SMOOTHING OF SAMPLES FOR MAXIMA

by

YASHASWINI MITTAL
STANFORD UNIVERSITY

TECHNICAL REPORT NO. 19

August 8, 1978

STUDY ON STATISTICS AND ENVIRONMENTAL
FACTORS IN HEALTH

PREPARED UNDER SUPPORT TO SIMS FROM
ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION (ERDA)
ROCKEFELLER FOUNDATION
SLOAN FOUNDATION
ENVIRONMENTAL PROTECTION AGENCY (EPA)
NATIONAL SCIENCE FOUNDATION (NSF)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

SMOOTHING OF SAMPLES FOR MAXIMA¹

by

Yashaswini Mittal
Stanford University

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

- ¹ AMS subject classification: Primary 60F05, 60F99; Secondary 60G15.
Key words and phrases: maxima; smoothed samples; central limit theorem; large deviations.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED *EB*

SMOOTHING OF SAMPLES FOR MAXIMA

by

Yashaswini Mittal

ABSTRACT

Smoothing of the data by averaging is suggested in order to study the maximum. The maximum of the smoothed data is approximated by that of a Gaussian sample and thus is more robust against outliers.

1. Introduction

Let $X_1, \dots, X_n$ be a sequence of random variables with mean zero and

$$Y_i = \frac{\sum_{j=(i-1)L(n)+1}^{iL(n)} X_j}{\text{var}^{1/2} \left(\sum_{j=(i-1)L(n)+1}^{iL(n)} X_j \right)}, \quad i = 1, 2, \dots, [n/L(n)]$$

where $L(n) \uparrow \infty$ with n and $[.]$ denotes the integral value. If the Central Limit Theorem is assumed to hold for $\{X_i\}$, then each of Y_i is asymptotically Gaussian with mean zero and variance one. With suitable assumptions on the dependence of $\{X_i\}$ and careful estimates of the error terms we show here that the asymptotic distribution of $\Pi_n = \max_{1 \leq i \leq [n/L(n)]} Y_i$

is the same as $\max_{1 \leq i \leq [n/L(n)]} Z_i$ where Z_i are independent standard normal variables.

The theory of maxima has found an increasing number of applications to real life situations over the past years. Gnedenko's [6] characterization of the extreme value distributions in the independent identically distributed case, besides being theoretically beautiful, opened up significant new possibilities for its applications. His results were extended to sequences with various degrees of dependence by a number of authors (see [7] for references). As far as the case of a general sequence $\{X_i\}$ of variables is concerned, one can hardly hope for any "weaker" conditions for finding the asymptotic distribution of $M_n = \max_{1 \leq i \leq n} X_i$. And yet in many cases - here we have the air pollution data in mind - these conditions are neither verifiable nor believable. Additional problems are caused by outliers and rounding errors in the data. The maximum Π_n defined above is much

more robust against these problems and, at least in certain situations, may be offered as a reasonable alternative to M_n . For example, even though concentrations of a pollutant may be recorded as often as every five minutes, the EPA clean air standard may be in terms of average concentrations over an 8 or 24 hour period. In this case, the actual maximum M_n of every-five-minute observations may be of no significant practical value. The study of Π_n offers the advantage that the X_i may be neither stationary nor identically distributed, both of these assumptions being very hard to verify and, in fact, are untrue in many cases of air pollution data. We are striving here to get the asymptotic distribution of Π_n under conditions that are as independent of the individual distributions of $\{X_i\}$ as possible. We can only hope to assume a very mild dependence between $\{X_i\}$. The dependence assumption we use is that of ϕ -mixing.

Naturally, the proofs of the theorems in Sections 2 and 3 are based on the large deviations type results for the Central Limit Theorem. In Section 2, we do the case of the i.i.d. random variables to illustrate the simple main idea of the proof. It also gives the "smallest value" of $L(n)$ for a given rate of convergence in Theorem 2.1. There are a few large deviations results available for dependent sequences. Any of these can be used to get a corresponding result of the type of Theorem 3.1. We choose the result of Babu, Ghosh and Singh [1] for our calculations. For other results on large deviations and an extensive bibliography see Statulevičius ([10], [11]). The result in [10] for strong mixing sequences has, in fact, faster convergence rates than those of

[1] and would lead to smaller block sizes in Theorem 3.1, but the conditions in Theorem 4 of [10] are rather complex. In Theorem 3.1 we have chosen an increasing sequence $\{L_i\}$ of block lengths rather than the fixed (for fixed n) length $L(n)$ used in Theorem 2.1. This allows one to easily incorporate new data and does away with the need to compute new averages at every stage n of the sample. In Theorem 3.2 we substitute a sample estimate for the unknown variances of the sums of $\{X_i\}$.

In Theorem 2.1 we discover that in a sample of size n we need to average over blocks of length no bigger than $(\ell n n)^{3+\delta}$ for some $\delta > 0$. In Theorems 3.1 and 3.2 we see that these blocks are approximately n^γ in length for some $\gamma > 0$. Even though n^γ is asymptotically much larger than $(\ell n n)^K$ for any K and any $\gamma > 0$, in practice the values may not differ significantly, especially for large values of K and small values of γ . A small value of γ can be chosen in Section 3 provided that sufficiently large order moments of $\{|X_i|\}$ are assumed to exist. Notice that a typical assumption in large deviations type result ([5], Theorem 1, page 517) used in Section 2 is that the moment generating function of X_i exists for all i .

It is natural to ask the question, "When can one predict M_n from Π_n ?" The question is impossible to answer unless one makes further assumptions about the distributions of $\{X_i\}$. In the last section we find a somewhat surprising joint distribution of M_n and Π_n in case $\{X_i\}$ is a stationary Gaussian sequence with a smooth covariance sequence.

2. The Case of i.i.d. Variables

Let $\{X_i, i \geq 1\}$ be a sequence of i.i.d. random variables with $EX_i = 0$; $EX_i^2 = \sigma^2$ and $E e^{tX_i} < \infty$ for all i and $|t| < t_0$. Let $L(n) \uparrow \infty$ be the smallest value for which

$$(2.1) \quad (2\ln[n/L(n)] - \ln(4\pi\ln[n/L(n)]))^{3+\delta} \leq L(n)$$

for some $\delta > 0$. Note that when no mention of an argument is made in a limiting statement, as in " $L(n) \uparrow \infty$," it is to be taken to read "as $n \rightarrow \infty$." Also, the symbols "0" and "o" will be used only when $n \rightarrow \infty$.

Remark. The choice of $L(n)$ in (2.1) needs to be made by trial and error. The "best" value of $L(n)$ is the one that barely satisfies the inequality. Asymptotically this choice of $L(n)$ becomes equal to about $(2\ln n)^{3+\delta}$. But for values of n even as large as 10,000, the choice in (2.1) leads to a considerably smaller value of $L(n)$ than $(2\ln n)^{3+\delta}$.

Hereafter the constants c_k and b_k are defined by

$$(2.2) \quad c_k = (2\ln k)^{\frac{1}{2}}; \quad b_k = c_k - \frac{\ln(4\pi\ln k)}{2c_k} \quad k = 1, 2, \dots$$

Let us define $Y_i = \sum_{j=(i-1)L(n)+1}^{iL(n)} X_j / (\sigma(L(n))^{\frac{1}{2}})$; $\Pi_n = \max_{1 \leq i \leq [n/L(n)]} Y_i$

and $u_n = b_{[n/L(n)]} + x/c_{[n/L(n)]}$.

Theorem 2.1. For the quantities defined above

$$(2.3) \quad \lim_{n \rightarrow \infty} P(\Pi_n \leq u_n) = \exp(-e^{-x})$$

for $-\infty < x < \infty$.

Proof. Write $K_n = [n/L(n)]$. Since Y_i are i.i.d. $P(\Pi_n \leq u_n) = \{1 - P(Y_i > u_n)\}^{K_n}$. Using Theorem 2, page 520, of Feller [5], we have $P(Y_i > u_n) = (1 - \Phi(u_n)) E_n$ where $\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^x \exp(-u^2/2) du$ and $E_n = \exp\{u_n^2 \lambda(u_n/L^{1/2}(n))\} [1 + o(u_n/L^{1/2}(n))]$. The function $\lambda(x)$ is defined by the equation

$$x^2 \lambda(x) = \lambda_1 x^3 + \lambda_2 x^4 + \dots,$$

λ_i being constants and $\lambda_1 = EX_1^3/6\sigma^3$. For the choice of $L(n)$ which gives asymptotic equality in (2.1), we have $u_n^3/L^{1/2}(n) \sim u_n^{-\delta}$. Thus $u_n^2 \lambda(u_n/L^{1/2}(n)) = \lambda_1 u_n^{-\delta} + o(u_n^{-2(1+\delta)})$ and $E_n = 1 + \lambda_1 u_n^{-\delta} + o(u_n^{-2(1+\delta)})$. In the following we use the fact that $K_n(1 - \Phi(u_n)) = \exp(-x)(1 + o(u_n^{-1}))$.

Now

$$\begin{aligned} P(\Pi_n \leq u_n) &= \{1 - (1 - \Phi(u_n))(1 + \lambda_1 u_n^{-\delta} + o(u_n^{-2(1+\delta)}))\}^{K_n} \\ &= \{\Phi(u_n) - \lambda_1 K_n^{-1} u_n^{-\delta} \exp(-x)(1 + o(u_n^{-1})) + o(u_n^{-2(1+\delta)} K_n^{-1})\}^{K_n} \\ &= \Phi^{K_n}(u_n) \{1 - (K_n \Phi(u_n))^{-1} \lambda_1 \exp(-x) u_n^{-\delta} + o(\exp(-x) u_n^{-(1+\delta)-1} K_n)\}^{K_n} \\ (2.4) \quad &= \Phi^{K_n}(u_n) - \lambda_1 \exp(-x) u_n^{-\delta} \Phi^{-1}(u_n) + o(\exp(-x) u_n^{-(1+\delta)}) . \end{aligned}$$

Remark. Simple arithmetic calculations show that $\Phi_n^K(u_n) = \exp(-e^{-x}) + O(\exp(-x)(\ln \ln K_n)^2 u_n^{-2})$. Thus theoretically the error rate for the convergence in (2.3) is $\max\{u_n^{-\delta}; (\ln \ln K_n)^2 u_n^{-2}\}$. In practice, the second term of the R.H.S. of (2.4) can be considered as the error term since $\Phi_n^K(u_n)$ is computable for a given sample.

This concludes the proof of Theorem 2.1 and illustrates the relationship between the choice of $L(n)$ and the error rate.

3. The Main Theorems

In this section we assume that the sequence of variables $\{X_i, i \geq 1\}$ is ϕ -mixing, i.e., there exists a nonincreasing sequence ϕ_n of numbers such that $\phi_n \rightarrow 0$ and

$$(E) \quad |P(A \cap B) - P(A)P(B)| \leq \phi_n P(A)$$

for all $A \in \mathcal{F}_1^k$ and $B \in \mathcal{F}_{k+n}^\infty$, $k \geq 1$. By $\mathcal{F}_1^k$ and $\mathcal{F}_{k+n}^\infty$ we denote the σ -fields generated by $(X_1, \dots, X_k)$ and $(X_{k+n}, X_{k+n+1}, \dots)$ respectively. Let $\{L_i\}$ be a nondecreasing sequence of integers.

Define $\Delta_i = \sum_{k=1}^i L_k$ and K_n is such that $\Delta_{K_n-1} < n \leq \Delta_{K_n}$.

We choose $\{L_i\}$ such that $\Delta_{i-1}^{\gamma_1} \leq L_i < n^{\gamma_2}$ for some $0 < \gamma_1 < \gamma_2 < 1$.

We notice that this allows the choice $L_i = n^\gamma$ for all $1 \leq i \leq K_n$ for some $0 < \gamma < 1$.

In the first theorem we prove the sums of $\{X_i\}$ over various blocks will be normalized using its variances which are unknown in practice.

In Theorem 3.2 we will replace these by sample estimates.

Let us define $Y_i = \sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j / V_i^{1/2}$, where $V_i = \text{Var} \left(\sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j \right)$,
 $i = 1, 2, \dots, K_n$ and $\Pi_n = \max_{1 \leq i \leq K_n} Y_i$. We recall the definitions of c_k
 and b_k given by (2.2) and let $u_n = b_{K_n} + x/c_{K_n}$.

Theorem 3.1. Let

$$(3.1) \quad EX_j = 0 \quad \text{for all } j \geq 1.$$

$$(3.2) \quad \sum_{n=1}^{\infty} \phi_n^{1/2} < \infty.$$

$$(3.3a) \quad A_1 \leq \sigma_n^2 \leq A_2 \quad \text{for large } n \text{ and some constants } A_2 > A_1 > 0$$

$$\text{where } \sigma_n^2 = \text{Var}(X_n).$$

$$(3.3b) \quad \liminf_i V_i / L_i > 0.$$

$$(3.4) \quad E|X_j|^{2+c} \leq M \quad \text{for all } j \geq 1 \text{ and some constants } M > 1$$

and $c > 2$. (We will take γ_1 in the choice of L_i

such that $\gamma_1 c/2 > 2$.)

Then

$$(3.5) \quad \lim_{n \rightarrow \infty} P(\Pi_n \leq u_n) = \exp(-e^{-x}) \quad \text{for } -\infty < x < \infty.$$

Proof of Theorem 3.1. As mentioned in the introduction, we make repeated use of the results of Babu, Ghosh and Singh [1]. For convenience we

reproduce their Theorems 2 and 5 below. Conditions (3.1) - (3.4) above are very similar to their conditions except somewhat stronger. The first theorem below is the large deviations result and Theorem B is stated in [1] as a moderate deviations result. We need Theorem B in case L_1 is quite large.

Theorem A. Let $\{X_j\}$ be a ϕ -mixing sequence for which (3.1) and (3.2) hold. Also let

$$(3.3)' \quad \inf_{n \geq 1} n \operatorname{Var} \left(\sum_{j=1}^n X_j \right) > 0$$

$$(3.4)' \quad E|X_j|^{2+c} \leq M \text{ for some } c > 0 \text{ and } M > 1.$$

Then for all $t^2 > (c+1) \ln n$ and some $K > 0$,

$$(3.6) \quad P|P(\eta_n \leq t) - \Phi(t)| \leq K n^{-c/2} t^{-2-c} (\ln t)^{2+2c}$$

$$\text{where } \eta_n = \sum_{j=1}^n X_j / \operatorname{Var}^{1/2} \left(\sum_{j=1}^n X_j \right).$$

Theorem B. For $\{X_j\}$ as in Theorem A and $t^2 \leq (c+1) \ln n$,

$$(3.7) \quad |P(\eta_n \leq t) - \Phi(t)| \leq K n^{-\lambda} \exp(-t^2/2) + O(n^{-c/2} (1+|t|)^{-2-c}) \text{ for}$$

some $\lambda > 0$.

We turn now to the proof of Theorem 3.1. We use standard arguments that have been used in the literature before. Divide $Y_1, \dots, Y_{K_n}$ into m blocks of length K_n/m each. Clip a small portion from the righthand end of each to "separate" these intervals. On each block, find upper and lower bounds for the probability that the maximum is bounded by u_n . Finally, let K_n/m tend to infinity first and then take a limit as $m \rightarrow \infty$.

For any integer $k \geq 1$, let $I_j = \{(j-1)K_n + 1, (j-1)K_n + 2, \dots, jK_n - [K_n^{1/2}]\}$, $I_j^* = \{jK_n - [K_n^{1/2}] + 1, \dots, jK_n\}$ for $j = 1, 2, \dots, m$

and $I = \bigcup_{j=1}^m I_j$; $I^* = \bigcup_{j=1}^m I_j^*$. Define $\Pi_n(m) = \max_{1 \leq i \leq mK_n} Y_i$,

$\Pi_{I_j} = \max_{i \in I_j} Y_i$ and $\Pi_I = \max_{i \in I} Y_i$. Also $u_{nm} = b_{mK_n} + x/c_{mK_n}$. Now

$$(3.8) \quad P\{\Pi_n(m) \leq u_{nm}\} = \prod_{j=1}^m P\{\Pi_{I_j} \leq u_{nm}\} + P_1 + P_2$$

where $P_1 = P\{\Pi_n(m) \leq u_{nm}\} - P\{\Pi_I \leq u_{nm}\}$ and $P_2 = P\{\Pi_I \leq u_{nm}\}$

$- \prod_{j=1}^m P(\Pi_{I_j} \leq u_{n,m})$. For the first step of the proof, we will show that

upper bounds on both $|P_1|$ and $|P_2|$ are $o(1)$.

$$|P_1| \leq P(Y_i > u_{nm} \text{ for at least one } i \in I^*)$$

$$(3.9) \quad \leq \sum_{i \in I^*} P(Y_i > u_{nm}).$$

We can use Theorems A and B to estimate the probability in the R.H.S. of (3.9) if $L_k \uparrow \infty$. The smallest value of i for $i \in I^*$ is $K_n - [K_n^{1/2}]$. The choice of L_i gives $L_i \geq \Delta_{i-1}^{\gamma_1} > (i-1)^{\gamma_1}$ for some $\gamma_1 > 0$. Thus Theorems A and B together give

$$(3.10) \quad P(Y_i > u_{nm}) \leq 1 - \Phi(u_{nm}) + K \max\{K_n^{-(1+\delta)}; L_i^{-c/2} u_{nm}^{-2-c} (\ln u_{nm})^{2+2c}\}$$

for some constants $\delta > 0$ and $K > 0$. The second term in the R.H.S. of (3.10) is maximum of the error terms in Theorems A and B. We need Theorem B in case L_i is so large that $u_{nm}^2 \leq (c+1) \ln L_i$. We notice that if t is too small, Theorem B can give significantly bigger error term than that of Theorem A. However, in our case u_{nm} is large enough that it does not make much difference.

We choose γ_1 such that $\gamma_1 c/2 > 1$ and replace L_i by $|i-1|^{\gamma_1}$ in the R.H.S. of (3.10). The total number of i in I^* is $m[K_n^{1/2}]$ and $i \geq K_n - [K_n^{1/2}]$. Substituting the value of u_{nm} and taking the sum over i in (3.10) we get that

$$(3.11) \quad |P_1| = o(K_n^{-\delta})$$

for some $0 < \delta < 1/2$. To find an upper bound for $|P_2|$, we use the ϕ -mixing property of $\{X_j\}$. Each I_j is separated by a distance of at least $[K_n^{1/2}]$ from I_ℓ for $\ell \neq j$. Applying (E) m times we get that

$$(3.12) \quad |P_2| \leq m \phi_{[K_n^{1/2}]} = o(1).$$

Thus for any integer $m \geq 1$,

$$(3.13) \quad P(\Pi_n(m) \leq u_{nm}) = \prod_{j=1}^m P(\Pi_{I_j} \leq u_{nm}) + o(1).$$

Now, for any $z > 0$,

$$\begin{aligned} o(1) + \sum_{j=1}^{[z]+1} P(\Pi_{I_j} \leq u_{nz}) &= P(\Pi_n([z]+1) \leq u_{nz}) \\ &\leq P(\Pi_n(z) \leq u_{nz}) \\ &\leq P(\Pi_n([z] \leq u_{nz}) = \prod_{j=1}^{[z]} P(\Pi_{I_j} \leq u_{nz}) + o(1). \end{aligned}$$

We proceed now to find upper and lower bounds for $P(\Pi_{I_j} \leq u_{nz})$. For a lower bound, Boole's inequality gives $P(\Pi_{I_j} \leq u_{nz}) \geq 1 - \sum_{i \in I_j} P(Y_i > u_{nz})$ for all $j = 1, 2, \dots, [z]+1$. Let us look at the case when $j=1$. Let n_1 be some integer such that $n_1 = o(\ln K_n)^{1/2}$. For $1 \leq i \leq n_1$, we use Chebychev's inequality to get

$$P(Y_i > u_{nz}) \leq EY_i^2 / u_{nz}^2 = u_{nz}^{-2} = O((\ln K_n)^{-1}).$$

Thus $\sum_{i=1}^{n_1} P(Y_i > u_{nz}) = o(1)$. For $i > n_1$, we use Theorems A and B as

in (3.10). Using the same kinds of arguments as after (3.10), we get

$$\sum_{i=n_1}^{K_n - [K_n^{1/2}]} P(Y_i > u_{nz}) \leq K_n \frac{\phi(u_{nz})}{u_{nz}} + K \max\{u_{nz}^{-2-c} (\ln u_{nz})^{2+2c} \sum_{i=n_1}^{K_n} |i-1|^{-\gamma_1 c/2}; K_n^{-\delta}\}$$

for some $\delta > 0$. We use the traditional notation $\phi(x) = (2\pi)^{-1/2} \exp(-x^2/2)$.

(There does not seem to be any danger of confusion between the function ϕ_n in (E) and the standard normal density $\phi(x)$.) Substituting value for u_{nz} and taking limits we observe that

$$(3.15) \quad \lim_{n \rightarrow \infty} \sum_{i \in I_j} P(Y_i > u_{nz}) \leq \tau/z$$

where $\tau = \exp(-x)$ and $j = 1$. For $j = 2, 3, \dots, [z]+1$, we can use Theorems A and B directly since $i \geq K_n$. Thus (3.15) holds for all $1 \leq j \leq [z]+1$. Substituting in (3.14) we get

$$(3.16) \quad \liminf_{n \rightarrow \infty} P(\Pi_n(z) \leq u_{nz}) \geq \liminf_{j=1}^{[z]+1} P(\Pi_{I_j} \leq u_{nz}) \\ \geq (1 - \tau/z)^{[z]+1}.$$

Next, we will find an upper bound for $P(\Pi_{I_j} \leq u_{nz})$. To do this we first need to "separate" $Y_i, i \in I_j$. Without loss of generality, we will assume that $i > n_0$ where $n_0 = [K_n^\delta]$ for some $\delta > 0$, since excluding $Y_1, \dots, Y_{n_0}$ from consideration only increases the probability

$P(\Pi_{I_1} \leq u_{nz})$. Thus for all i under consideration $L_i \geq i^{\gamma_1} \geq K_n^{\gamma_1 \delta}$. Let us fix $j = 1$ and notice that here, $I_1 = \{n_0, n_0 + 1, \dots, K_n - [K_n^{1/2}]\}$. The computations for $j = 2, 3, \dots, [z]+1$ are very similar.

For z fixed and n large enough, define $V_i' = \text{Var} \left(\sum_{\ell=\Delta_{i-1}+1}^{\Delta_i-[z]} X_\ell \right)$ and $Y_i' = \sum_{\ell=\Delta_{i-1}+1}^{\Delta_i-[z]} X_\ell / (V_i')^{1/2}$. Also let $Y_i'' = \sum_{\ell=\Delta_i-[z]+1}^{\Delta_i} X_\ell$. Then

we can write

$$(3.17) \quad \Pi_{I_1} = \max_{i \in I_1} \{ (V_i'/V_i)^{1/2} Y_i' + V_i^{-1/2} Y_i'' \}.$$

Because $\sigma_j^2 \leq A_2$ for all j , $\text{Var}(Y_j'') \leq [z]^2 A_2$ and

$$V_i - V_i' \leq [z] A_2 \left\{ 1 + \sum_{\ell=1}^z \phi_\ell \right\} \leq [z] (\text{const.}) \text{ using (E) and Lemma 1, page 170,}$$

of Billingsley [4]. Thus we can write $(V_i'/V_i)^{1/2} = 1 + O(K_n^{-\delta \gamma_1})$ by

(3.3b). This gives that

$$(3.18) \quad P \left(\max_{i \in I_1} (V_i'/V_i)^{1/2} Y_i' \leq u_{nz} \right) = P \left(\max_{i \in I_1} Y_i' \leq u_{nz}' \right)$$

where $u_{nz}' = u_{nz} + o((\ln K_n)^{-1})$. Since Y_i'' is a sum of only $[z]$ number of X_ℓ , each one of which has finite absolute moments up to order $2 + c$, we must have $E|Y_i''|^{2+c} < \infty$ for all i . Using Chebychev's inequality,

$$\begin{aligned} P \left\{ \max_{i \in I_1} (Y_i''/V_i^{1/2}) > (\ln K_n)^{-1} \right\} &\leq \sum_{i \in I_1} P \left(Y_i'' > \frac{V_i^{1/2}}{\ln K_n} \right) \\ &\leq (\text{const.}) (\ln K_n)^{2+c} \sum_{i \in I_1} V_i^{-(2+c)/2}. \end{aligned}$$

Substituting i^{Y_1} for L_i and taking the sum we see that the R.H.S. above is $o(K_n^{-\delta})$ for some $\delta > 0$ and

$$(3.19) \quad P(\Pi_{I_1} \leq u_{nz}) \leq P\left(\max_{i \in I_1} Y'_i \leq u'_{nz}\right) + o(K_n^{-\delta}).$$

By Boole's inequality,

$$(3.20) \quad P\left(\max_{i \in I_1} Y'_i \leq u'_{nz}\right) \leq 1 - \sum_{i \in I_1} P(Y'_i > u'_{nz}) + 2 \sum_{\substack{i, \ell \in I_1 \\ i \neq \ell}} P(Y'_i > u'_{nz}; Y'_\ell > u'_{nz}).$$

The second term in the R.H.S. of (3.20) is computed in very similar manner to that of (3.15) which gives $\sum_{i \in I_1} P(Y'_i > u'_{nz}) = (\tau/z)(1 + o(1))$.

Events in the third term of the R.H.S. of (3.20) are separated by a distance at least equal to $(\ell-i)[z]$. Using (E) we get an upper bound for this last sum in (3.20) to be

$$\begin{aligned} & 2 \sum_{\substack{i, \ell \in I_1 \\ i \neq \ell}} P(Y'_i > u'_{nz}) P(Y'_\ell > u'_{nz}) + 2 \sum_{i \in I_1} P(Y'_i > u'_{nz}) \sum_{\substack{\ell > i \\ \ell \in I_1}} \phi_{(\ell-i)[z]} \\ & \leq \left(\sum_{i \in I_1} P(Y'_i > u'_{nz}) \right)^2 + 2g(z)(\tau/z)(1 + o(1)) \end{aligned}$$

where $g(z) = \sum_{\ell=1}^{\infty} \phi_{\ell}[z]$ is a function of z . Because of (E), it is easy to see that $\lim_{z \rightarrow \infty} g(z) = 0$.

Substituting in (3.14) we get

$$(3.21) \quad \limsup_{n \rightarrow \infty} P(\Pi_n(z) \leq u_{nz}) \leq (1 - \tau/z + \tau^2/z^2 + 2g(z)\tau/z)^{[z]}.$$

Combining (3.16), (3.21) and taking limit as $z \rightarrow \infty$ we get

$$(3.22) \quad \lim_{z \rightarrow \infty} \lim_{n \rightarrow \infty} P(\Pi_n(z) \leq u_{nz}) = \exp(-e^{-x})$$

for $-\infty < x < \infty$. Simple arithmetic then leads to (3.5). This concludes the proof of Theorem 3.1.

The averages Y_i computed for Theorem 3.1 are based on the variances V_i . If the dependence between $\{X_i\}$ is weak enough then

$$S_i = \sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j^2 \text{ should serve as an adequate estimate for } V_i. \text{ We}$$

can prove this in Theorem 3.2 under slightly stronger assumptions than ϕ -mixing assumed in (E). In order to make the new averages Y_i^S well defined, we truncate S_i away from zero. Let $S_i^\theta = \max\{\theta, S_i\}$ for some

$$\theta > 0 \text{ and } Y_i^S = \sum_{j=\Delta_{i-1}+1}^{\Delta_i} X_j / (S_i^\theta)^{1/2} = Y_i (V_i/S_i^\theta)^{1/2} \text{ for } i = 1, 2, \dots, K_n.$$

We define $\Pi_n^S = \max_{1 \leq i \leq K_n} Y_i^S$.

Theorem 3.2. We assume that (3.1) - (3.4) hold. In addition, if

$$(3.23) \quad V_i - ES_i = o(L_i / \ln L_i)$$

and

$$(3.24) \quad \inf_i (\text{Var } S_i / L_i) \geq c_1$$

for some constant $c_1 > 0$. Then

$$(3.25) \quad \lim_{n \rightarrow \infty} P(\Pi_n^S \leq u_n) = \exp(-e^{-x})$$

for $-\infty < x < \infty$.

Remark. Conditions (3.23) and (3.24) are on the correlations of (X_j, X_k) and (X_j^2, X_k^2) . Using (E) and Lemma 1, page 170, of [4], it can easily be shown that the L.H.S. of (3.23) is $o(L_i)$. Property (E) is related to the rate of decay of dependence between X_j and X_k . In (3.23) we assume that there is enough cancellation due to positive and negative correlations so that the L.H.S. of (3.23) grows slower than $L_i / \ln L_i$. For (3.24) it is sufficient to assume that $\liminf_j \text{Var } (X_j^2) > 0$ and

$$\sum_{\substack{j, k \in [\Delta_{i-1}+1, \Delta_i] \\ j \neq k}} E(X_j^2 - \sigma_j^2)(X_k^2 - \sigma_k^2) = o(L_i).$$

Proof of Theorem 3.2. The use of S_i^θ instead of S_i is simply to avoid embarrassing situations when S_i may be too small. We first show in Lemma 3.1 below that this does not occur if L_i is assumed to increase with n .

Lemma 3.1. If there exists a function $f(n) \uparrow \infty$ such that $L_i / f(n) \uparrow \infty$ for all $1 \leq i \leq K_n$, then

$$(3.26) \quad \sum_{i=1}^{K_n} P(S_i \neq S_i^\theta) = o(P).$$

Proof of Lemma 3.1. We know that

$$P(S_i \neq S_i^\theta) = P(S_i < \theta)$$

$$= P\left(\frac{S_i - ES_i}{(\text{Var } S_i)^{1/2}} < \frac{\theta - ES_i}{(\text{Var } S_i)^{1/2}}\right).$$

We see that $(\text{Var } S_i)^{-1/2}(\theta - ES_i) \leq -d_1 L_i^{1/2}$ for some constant $d_1 > 0$ because of (3.23) and (3.24). Applying Theorems A and B we have

$$(3.27) \quad \sum_{i=1}^{K_n} P(S_i \neq S_i^\theta) \leq \sum_{i=1}^{K_n} \left\{ 1 - \Phi(d_1 L_i^{1/2}) + K \max \left(L_i^{-(2+3c)/2} \left(\frac{\ell n L_i}{2} \right)^{2+2c}; \frac{L_i^{-\lambda-1/2} \phi(d_1 L_i^{1/2})}{d_1} \right) \right\}$$

for some constants $K > 0$ and $\lambda > 0$. We will split the sum in the R.H.S. of (3.27) into two parts, $1 \leq i \leq f(n)$ and $f(n) < i \leq K_n$. For the first part we write

$$L_i^{-(2+3c)/2} (\ell n L_i)^{2+2c} \leq \left(\max_{1 \leq x < \infty} (x^{-1/2} \ell n x) \right) L_i^{-c/2} < d_2 L_i^{-c/2}$$

for some constant $d_2 > 0$. Thus the first part of the sum in the R.H.S. of (3.27) is at most

$$\sum_{i=1}^{f(n)} \left\{ \frac{\Phi(d_1 f(n))}{d_1 f(n)} + K \max \left(d_2 (f(n))^{-c/2}, (f(n))^{-\lambda-1/2} \phi(d_1 f(n)^{1/2}) d_1^{-1} \right) \right\} = o(1).$$

For $i > f(n)$, we will substitute i^{γ_1} for L_i . The second part of the sum under consideration is at most

$$(\text{const.}) \sum_{i=f(n)}^{K_n} \left\{ i^{-\gamma_1/2} \exp(-d_1^2 i^{\gamma_1/2}) + (\ln i)^{2+2c} i^{-\gamma_1(2+3c)/2} \right\}.$$

Because $\gamma_1 c/2 > 1$, the above is $o(1)$. This proves Lemma 3.1.

The proof of Theorem 3.2 is slightly tricky. It is based on the basic fact that $(S_i/V_i - 1)$ goes to zero sufficiently fast (shown at 3.32). But for this we need to assume that $i > n_o = [K_n^\delta]$ for some $\delta > 0$. Thus, as a first step of the proof of Theorem 3.2, we exclude $Y_1^S, \dots, Y_{n_o}^S$ from consideration in (3.25), and this is where the somewhat confusing part of the proof occurs. Assuming that $L_i/f(n) \uparrow \infty$ for some $f(n) \uparrow \infty$, we can exclude $Y_1^S, \dots, Y_{n_o}^S$ by using a weaker version of (3.32) proved at (3.31). This brings us to the assumption that $L_i/f(n) \uparrow \infty$ which may be already true by choice of L_i . If not, then the first part of the proof below excludes enough Y_i^S to ensure the existence of some $f(n)$ as above for the remaining block lengths L_i under consideration.

We know that

$$(3.28) \quad P\left(\max_{n_o < i \leq K_n} Y_i^S \leq u_n\right) = P(\Pi_n^S \leq u_n) \leq \sum_{i=1}^{n_o} P(Y_i^S > u_n).$$

Ultimately, we want to show that the R.H.S. of (3.28) is $o(1)$. But first define $D = \{1 \leq i \leq n_1 \mid L_i \leq u_n\}$ where n_1 is some integer such that $n_1 = o((\ln K_n)^{1/2})$. Notice that D could be empty, in which case

$L_i > u_n$ for all $1 \leq i \leq K_n$. In any case, $L_i > f(n) = \min(n_1^{Y_1}, u_n)$ for all $i \notin D$. Thus we will first exclude Y_i^S for $i \in D$. We can substitute $Y_i(V_i/\theta)^{1/2}$ for Y_i^S since θ is the smallest value of S_i^θ . Also $V_i \sim (\text{const.}) L_i$ by (3.3a), (3.3b) and (3.23). Now

$$(3.29) \quad \sum_{i \in D} P(Y_i^S > u_n) \leq \sum_{i \in D} P(Y_i > d_3 u_n^{1/2})$$

$$\leq \sum_{i \in D} d_3^{-2} u_n^{-1}$$

for some constant $d_3 > 0$. The last statement follows because of Chebychev's inequality and the fact that $EY_i^2 = 1$ for all i . The total number of i in D is at most $o((\ln K_n)^{1/2})$ and the R.H.S. of (3.29) is $o(1)$. Now assuming that $L_i \uparrow \infty$, we will show that $(S_i/V_i - 1)$ is small. This will help us exclude the remaining Y_i^S .

In view of (3.26) we write $Y_i^S \approx Y_i(V_i/S_i)^{1/2}$ for $i \in [1, 2, \dots, n_0] \cap D^c$. For any $\varepsilon > 0$,

$$(3.30) \quad P(Y_i^S > u_n) \leq P(Y_i > (1-\varepsilon) u_n) + P(|(V_i/S_i)^{1/2} - 1| > \varepsilon).$$

Using Theorems A and B as in (3.10), we see that the sum for $i \in [1, \dots, n_0]$ of the first term in the R.H.S. of (3.30) is $o(1)$ for sufficiently small values of $\delta > 0$ (i.e., of $n_0 = [K_n^\delta]$). We will now show that

$$(3.31) \quad \sum_{i=1}^{n_0} P(S_i > (1+2\varepsilon) V_i) = o(1).$$

Note that $(\text{Var } S_i)^{-1/2}(V_i - ES_i) = O(L_i^{1/2}/\ln L_i)$ due to (3.23) and (3.24). Thus $P(S_i > (1+2\varepsilon)V_i) \approx P((\text{Var } S_i)^{-1/2}(S_i - ES_i) > (\text{const.}) \varepsilon L_i^{1/2})$. Using the same estimates as in (3.27), (3.31) follows easily. The argument

for $\sum_{i=1}^{n_0} P(S_i < (1-2\varepsilon)V_i) = o(1)$ are very similar. A few arithmetic calculations show that this is sufficient to conclude that the R.H.S. of (3.28) is $o(1)$.

To complete the proof of Theorem 3.2, we write $Y_i^S \approx Y_i + Y_i\{(V_i/S_i)^{1/2} - 1\}$ (having substituted S_i for S_i^θ in view of 3.26). Theorem 3.1 implies that it is sufficient to show that

$(\ln K_n) \max_{n_0 \leq i \leq K_n} \{(V_i/S_i)^{1/2} - 1\} \xrightarrow{P} 0$ as $n \rightarrow \infty$. For this we will show

that

$$(3.32) \quad \lim_{n \rightarrow \infty} \sum_{i=n_0}^{K_n} P(S_i/V_i > 1 + \varepsilon/\ln K_n) = 0.$$

The corresponding result for $P(S_i/V_i < 1 - \varepsilon/\ln K_n)$ will then follow easily. Now

$$(3.33) \quad P(S_i/V_i > 1 + \varepsilon/\ln K_n) = P\left(\frac{S_i - ES_i}{(\text{Var } S_i)^{1/2}} > \frac{V_i - ES_i}{(\text{Var } S_i)^{1/2}} + \frac{\varepsilon V_i}{(\ln K_n)(\text{Var } S_i)^{1/2}}\right).$$

Because $i > [K_n^\delta]$, $\ln L_i > \gamma_1 \delta \ln K_n$ and (3.23) implies that

$(V_i - ES_i)(\text{Var } S_i)^{-1/2} = o(L_i^{1/2}(\ln K_n)^{-1})$. Thus the R.H.S. of (3.33) is at most $P\{(\text{Var } S_i)^{-1/2}(S_i - ES_i) > d_4 L_i^{1/2}(\ln K_n)^{-1}\}$ for some constant $d_4 > 0$.

The remaining steps are by now routine applications of Theorems A and B.

The sum in (3.33) is $o(1)$ and this completes the proof of Theorem 3.2.

4. Joint Distribution of M_n and Π_n

In this section we will assume that $\{X_j; j \geq 0\}$ is a stationary Gaussian sequence with mean zero, variance one and $r_j = EX_0 X_j$. Let $M_n = \max_{0 \leq j \leq n} X_j$. The limiting distribution of M_n is extensively studied under various conditions on $\{r_j\}$ (see, e.g., [2] and [8]). A possible motivation for looking at Π_n even under these circumstances could be the presence of the outliers in the data. If the predicted value of M_n based on Π_n is significantly different from the observed M_n , it can be construed as indicating the presence of the outliers.

As shown in Theorem 4.1 below, mild conditions on the covariance function $\{r_j\}$ imply asymptotic independence of M_n and Π_n . This conclusion is disappointing from the point of view of predicting M_n based on Π_n .

The conditions used in Theorem 4.1 are those of the moderately dependent case in Mittal ([9], Theorem 2.1). The joint distribution of M_n and Π_n in the "strongly dependent" case (see Theorem 3.1 of [9]) is conjectured to be degenerate on the diagonal. The normalizing constants in this case will involve the covariance function. This would be contrary to the considerations towards applications used in this paper. Thus we choose to omit the calculations for the "strongly dependent" case.

The sequence $\{L_i\}$ described in the beginning of Section 3 will have now even a greater freedom of choice. To keep the choice sensible, we will assume that $L_i < n^{\gamma_2}$ for some $0 < \gamma_2 < 1$ and

$L_i/L_{i+1}^{1-\gamma} \rightarrow \infty$ as $i \rightarrow \infty$ for all $0 < \gamma < 1$. In addition, assume that $\ln n / \ln K_n$ is bounded. (This allows the choice $L_i = (\ln n)^k$, $k \geq 1$ for all $1 \leq i \leq K_n$ which was not possible in Section 3. As before, K_n is such that $\sum_{i=1}^{K_n-1} L_i < n \leq \sum_{i=1}^{K_n} L_i$.)

Theorem 4.1. Suppose that for sufficiently large k ,

$$(4.1) \quad \sum_{i=-k}^k r_i = k^\delta G(k) > 0,$$

$-1 < \delta < 1$ and $G(k)$ is a slowly varying function satisfying the following conditions.

(4.2) For sufficiently large ℓ , $G(k)/G(\ell)$ is bounded if $k \geq \ell^\theta$,
 $\theta > 0$ and

(4.3) If $\ell \geq k$, $k \rightarrow \infty$ such that $(\ell-k)/k^{1-\gamma} \rightarrow \infty$ for all $\gamma > 0$ and
 and $\frac{\ell-k}{k} \rightarrow c$ for $0 \leq c \leq \infty$, then

$$\lim_{k \rightarrow \infty} \frac{\ell}{k} (G(\ell)/G(\ell-k) - 1) = 0.$$

Then

$$(4.4) \quad \lim_{n \rightarrow \infty} P(M_n \leq u_n(x); \Pi_n \leq u_{K_n}(y)) = \exp\{-e^{-x} - e^{-y}\}$$

for all $-\infty < x, y < \infty$ where $u_m(x) = b_m + x/c_m$ and b_m and c_m are defined in (2.2).

Proof. The proof consists of showing that M_n and Π_n are asymptotically independent and that Γ_n properly normalized converges to the double exponential distribution ($\exp(-e^{-y})$). Notice that conditions (4.1) - (4.3) are intuitively much weaker than the ϕ -mixing assumed in Section 3. This is to be expected since here we are dealing with Gaussian rather than general sequences.

We will first find the limiting behavior of Π_n . Without loss of generality, we can replace Π_n in (4.4) by $\max_{K_n - \gamma \leq i \leq K_n} Y_i$ for any

$$0 < \gamma < 1 \quad \text{since} \quad \sum_{i=1}^{K_n} P(Y_i > u_{K_n}^\gamma(y)) \leq K_n^\gamma \phi(u_{K_n}^\gamma(y)/u_{K_n}^\gamma(y)) = o(1).$$

Thus from now on let us assume $\Pi_n = \max_{K_n - \gamma \leq i \leq K_n} Y_i$. The result follows

from Theorem 3.1 of Berman [2] if we show that $|EY_i Y_j| < 1$ (see 4.9 below) and that $|EY_i Y_j| \ln|j-i| \rightarrow 0$ as $j-i \rightarrow \infty$ (shown at (4.11) and (4.12)).

Define $D_{ij} = \sum_{k=i}^{j-1} L_k$ and $B_i = \{\Delta_{i-1}+1, \Delta_{i-1}+2, \dots, \Delta_i\}$ where Δ_i is the same as in Section 3. Then

$$\begin{aligned} EY_i Y_j &= \sum_{\ell \in B_i} \sum_{k \in B_j} E(X_\ell X_k) / (V_i V_j)^{1/2} \\ (4.5) \quad &= \sum_{\ell=0}^{L_i-1} \sum_{k=D_{ij}}^{D_{ij}+L_j-1} r_{k-\ell} / (V_i V_j)^{1/2} \end{aligned}$$

Using condition (4.1), we get the numerator of the R.H.S. of (4.5) to be equal to

$$(4.6) \quad \sum_{\ell=0}^{L_i-1} \{ (D_{ij} + L_j - \ell)^\delta G(D_{ij} + L_j - \ell) - (D_{ij} - \ell)^\delta G(D_{ij} - \ell) \}.$$

First, let us assume $j = i + 1$. Then we can rewrite (4.6) as

$$(4.7) \quad \sum_{\ell=0}^{L_i-1} (L_i + L_j - \ell)^\delta G(L_i + L_j - \ell) - V_i.$$

Here we have $L_j < L_i + L_j - \ell \leq L_i + L_j$. Now $(L_i + L_j - \ell)/(L_j - \ell)^{1-\theta} \geq L_i^\theta \rightarrow \infty$ for all $\theta > 0$. Thus condition (4.3) implies that $G(L_i + L_j - \ell) \leq (1+\epsilon) G(L_i)$ for large n and $0 \leq \ell \leq L_i - \ell$. The variance $V_i \sim (\delta+1)^{-1} L_i^{\delta+1} G(L_i)$. (See (2.5) in [9].) The following upper bound for (4.7) ignores the second order terms. For every $\epsilon > 0$, we can find n large enough that (4.7) is at most

$$(4.8) \quad G(L_i)(1+\delta)^{-1} \{ (L_i + L_j)^{\delta+1} - L_i^{\delta+1} - L_j^{\delta+1} + \epsilon((L_i + L_j)^{\delta+1} - L_j^{\delta+1}) \}.$$

We recall the algebraic inequality $(a+b)^\theta - a^\theta - b^\theta < a^{\theta/2} b^{\theta/2}$ for all $a, b > 0$ and $0 < \theta < 2$. Also notice that $G(L_i)/G(L_j) \rightarrow 1$ as $n \rightarrow \infty$ because of (4.3) and the choice of $L_{i+1} = L_j$. Substituting in (4.5), we get the following strict upper bound for it.

$$(4.9) \quad (1 + \epsilon')(\rho + 2\epsilon)$$

where $\rho < 1$. We can choose ϵ and ϵ' small enough for n large so that (4.9) is strictly less than one.

Next, let us assume $j \geq i + 2$. Then (4.6) is equal to

$$(4.10) \sum_{\ell=0}^{L_i-1} \{ (D_{ij}+L_j-\ell)^\delta (G(D_{ij}+L_j-\ell)-G(D_{ij}-\ell)) + G(D_{ij}-\ell) ((D_{ij}+L_j-\ell)^\delta (D_{ij}-\ell)^\delta) \}.$$

Now, consider the first term in the sum of (4.10). We know that

$$(D_{ij}-\ell)/L_j^{1-\theta} \geq L_{j-1}/L_j^{1-\theta} \rightarrow \infty \text{ for all } \theta > 0. \text{ Hence applying (4.3) we get}$$

$$(D_{ij}+L_j-\ell)^\delta (G(D_{ij}+L_j-\ell) - G(D_{ij}-\ell)) \leq \varepsilon (D_{ij}-\ell) L_j (D_{ij}+L_j-\ell)^{-1+\delta}$$

for $\varepsilon > 0$ and n large. Also

$$\begin{aligned} (D_{ij}+L_j-\ell)^\delta - (D_{ij}-\ell)^\delta &= (D_{ij}+L_j-\ell) / (D_{ij}+L_j-\ell)^{1-\delta} - (D_{ij}-\ell)^\delta \\ &< L_j (D_{ij}+L_j-\ell)^{-1+\delta}. \end{aligned}$$

Thus (4.10) is less than

$$(1+\varepsilon) L_j \sum_{\ell=0}^{L_i-1} (D_{ij}+L_j-\ell)^{-1+\delta} G(D_{ij}-\ell).$$

Suppose that for all $0 < \theta < 1$, $L_i/(D_{ij} - 2L_i + 1)^{1-\theta}$ and

$L_j/(D_{ij} - L_i - L_j + 1)^{1-\theta}$ both tend to infinity with n , then using

(4.3) we know that $G(D_{ij}-\ell)/G(L_i)$ and $G(D_{ij}-\ell)/G(L_j)$ both tend to one

with $n \rightarrow \infty$. Substituting in (4.5) we get

$$(4.11) |EY_i Y_j| \leq (1+\varepsilon') (L_i L_j)^{(1-\delta)/2} / ((j-i-1)L_i + L_j)^{1-\delta} < (1+\varepsilon') (j-i)^{-(1-\delta)/2}.$$

Now, if there exists $0 < \theta < 1$ such that $L_i / (D_{ij} - 2L_i + 1)^{1-\theta}$ is bounded as $n \rightarrow \infty$, then

$$|EY_i Y_j| \leq (\text{const.}) \frac{\max_{0 \leq \ell \leq L_i - 1} G(D_{ij} - \ell)}{G^{1/2}(L_i) G^{1/2}(L_j)} \left(\frac{L_i}{(D_{ij} + L_j - L_i)^{(1-\theta)}} \right)^{(1-\delta)/2} \\ \cdot \left(\frac{L_j}{(D_{ij} + L_j - L_i)} \right)^{(1-\delta)/2} (D_{ij} + L_j - L_i)^{-\theta(1-\delta)/2}$$

$$(4.12) \quad < \varepsilon(j-i)^{\theta(1-\delta)/4}$$

for large n . Similarly for the case when $L_j / (D_{ij} - L_i - L_j + 1)^{1-\theta}$ is bounded. Thus (4.11) and (4.12) imply that $|EY_i Y_j| \ln |j-i| \rightarrow 0$ as $n \rightarrow \infty$ for all $i, j > K_n^Y$. This together with (4.9) and Theorem 3.1 of [2] implies that $\lim_{n \rightarrow \infty} P(\Pi_n \leq u_{K_n}(y)) = \exp(-e^{-y})$ for $-\infty < y < \infty$.

To show the asymptotic independence of M_n and Π_n , notice that

$Z^n = (Z_1^n, \dots, Z_{n+K_n}^n)^T = (X_1, \dots, X_n, Y_1, \dots, Y_{K_n})^T$ has a joint normal distribution for each n . Let $\Sigma(n)$ denote the covariance matrix of Z^n . That is, the $(i, j)^{\text{th}}$ element of $\Sigma(n)$ is $\sigma_{ij}(n) = EZ_i^n Z_j^n$. Let $\sigma_{ij}^0(n) = EZ_i^n Z_j^n$, if both i and $j \leq n$ or both i and j are $> n$. Take $\sigma_{ij}^0(n) = 0$ otherwise. That is, $\Sigma^0(n) = ((\sigma_{ij}^0(n)))$ is the covariance matrix of Z^n where X_ℓ and Y_m are assumed independent for each ℓ and m . Slight modification of (4.12) in Berman [3] (as used in [11]) gives that

$$\begin{aligned}
& |P(M_n \leq u_n(x); \Pi_n \leq u_{K_n}(y)) - P(M_n \leq u_n(x)) P(\Pi_n \leq u_{K_n}(y))| \\
& \leq (\text{const.}) \sum_{i=1}^{n+K_n} \sum_{j=1}^{n+K_n} |\sigma_{ij}(n) - \sigma_{ij}^0(n)| \exp\{-(u_n^2(x) - 2\sigma_{ij}(n)u_n(x)u_{K_n}(y) \\
& \quad + u_{K_n}^2(y)) / 2(1 - \sigma_{ij}^2(n))\}.
\end{aligned}
\tag{4.13}$$

We only need to compute $\sigma_{ij}(n) = \sigma_{ji}(n)$ for $i \leq n$; $j > n$, the sum in (4.13) being zero over the remaining terms. Now

$$EX_i Y_j = \sum_{\ell \in B_{j-n}} EX_i X_\ell / V_{j-n}^{1/2}.
\tag{4.14}$$

If $X_i \in B_{jn}$ then

$$EX_i Y_j \leq \frac{\ell n L_{j-n} + L_{j-n}^\delta G(L_{j-n})}{L_{j-n}^{(1+\delta)/2} G^{1/2}(L_{j-n})} = o(L_{j-n}^{-\epsilon})
\tag{4.15}$$

for some $\epsilon > 0$ of $X_i \notin B_{j-n}$, then let A_{ij} denote the "distance" between X_j and B_{j-n} , i.e., for all ℓ with $X_\ell \in B_{j-n}$, $\ell - i \geq A_{ij}$. Hence for $X_i \notin B_{j-n}$

$$\begin{aligned}
EX_i Y_j & \leq \frac{(A_{ij} + L_{j-n})^\delta G(A_{ij} + L_{j-n}) - A_{ij}^\delta G(A_{ij})}{L_{j-n}^{(1+\delta)/2} G^{1/2}(L_{j-n})} \\
& \leq L_{j-n}^{(1-\delta)/2} (A_{ij} + L_{j-n})^{-1+\delta+\epsilon}.
\end{aligned}
\tag{4.16}$$

The above follows by similar computations as before. The R. H. S. of (4.16) is $o(L_{j-n}^{-\epsilon})$ for some $\epsilon > 0$. However, if $A_{ij} \geq K_n^\gamma$ for some $\gamma > 0$, then we can write the R. H. S. of (4.16) as $o(K_n^{-\gamma'})$ for some $\gamma' > 0$. Substituting in (4.13) we get an upper bound for the R.H.S. there to be

$$(4.17) \quad n \sum_{j=n+1}^{n+K_n^\gamma} L_{j-n}^{-\epsilon} \exp \left\{ \frac{-u_n^2(x)}{2} - (1-2L_{j-n}^{-\epsilon} u_n(x)/u_n(y)) \frac{u_n^2(y)}{2} \right\} \\ + nK_n K_n^{-\gamma'} \exp \left\{ \frac{-u_n^2(x)}{2} - (1-2K_n^{-\gamma'} u_n(x)/u_n(y)) \frac{u_n^2(y)}{2} \right\}.$$

We have chosen K_n such that $u_n(x)/u_n(y)$ is bounded and

$L_{j-n}^{-\epsilon} \rightarrow 0$ as $n \rightarrow \infty$. Thus the terms in (4.17) are $o(1)$ and the result follows.

I would like to thank Professor Wassily Hoeffding for many references and helpful discussions.

REFERENCES

- [1] Babu, G., M. Ghosh and K. Singh (1978). On rates of convergence to normality for ϕ -mixing processes. Indian Statistical Institute, Calcutta. Technical Report No. 3.
- [2] Berman, S. M. (1964). Limit theorems for the maximum term in stationary sequences. Ann. Math. Stat. (35) 502-516.
- [3] Berman, S. M. (1971). Asymptotic independence of the numbers of high and low level crossing of stationary Gaussian processes. Ann. Math. Stat. (42) 927-945.
- [4] Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.
- [5] Feller, W. (1966). An Introduction to Probability Theory and Its Applications, Vol. II. Wiley, New York.
- [6] Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. Ann. Math. (44) 423-453.
- [7] Mittal, Y. (1974). Limiting behavior of maxima in stationary Gaussian sequences. Ann. Prob. (2) 231-242.
- [8] Mittal, Y. and D. Ylvisaker (1975). Limit distributions for the maxima of stationary Gaussian processes. Stoch. Proc. and Their Appln. (3) 1-18.
- [9] Mittal, Y. (1978). Extreme value distribution for normalized sums from stationary Gaussian sequences. Submitted to Stoch. Proc. and Their Appln.
- [10] Statulevičius, V. A. (1966). On large deviations. Z. Wahr. und Verw. Geb. (6) 133-144.
- [11] Statulevičius, V. A. (1974). Limit theorems for dependent random variables under various regularity conditions. Proc. of the International Congress of Math. Vancouver.