LOCATING THE MAXIMUM OF SOME STATIONARY GAUSSIAN SEQUENCES

by

Michael T. Lacey

University of North Carolina, Chapel Hill
ABSTRACT

Let $X_j, j \in \mathbb{N}$ be a stationary Gaussian sequence with covariance $r_j$ satisfying $r_0 = 1$, and $r_j \to 0$, $j \to +\infty$. Let $j_n^\ast$ be the location of the maximum of $X_j$ on $\{1, 2, \ldots, n\}$. We consider the problem of locating $j_n^\ast$ by random sequential observation of $X_j$. If $j^\alpha r_j \downarrow 0$, as $j \to +\infty$, the number of observations that is both necessary and sufficient as $n \to \infty$ is essentially determined. This generalizes a result of B. Hajek.
1. Introduction

Let \( X_j, j \in \mathbb{N} \) be a stationary Gaussian sequence with covariance

\[
r_j = E X_0 X_j, \quad j \in \mathbb{N},
\]

and standardized so that \( r_0 = 1 \). We consider the problem of locating the maximum of \( X \) on \( \{1, 2, \ldots, n\} \) by a minimal number of random sequential observations. More precisely, let

\[
(1.1) \quad X_n^* = \max_{1 \leq j \leq n} X_j,
\]

and let \( j_n^* \) be the location of the maximum on \( \{1, 2, \ldots, n\} \). A sequential search of length \( v \) in \( \{1, 2, \ldots, n\} \) is a sequence of integers \( 1 \leq k_1, k_2, \ldots, k_v \leq n \), so that for all \( i \), \( k_{i+1} \) is a function of

\[
k_1, \ldots, k_i, X_{k_1}, \ldots, X_{k_i},
\]

and the covariance of \( X \). Notice that \( k_2, \ldots, k_v \) are random. Let

\[
(1.2) \quad s(n, v) = \max P(j_n^* \in \{k_1, \ldots, k_v\}),
\]

the maximum being taken over all sequential searches of length \( v \) in \( \{1, 2, \ldots, n\} \). Thus \( s(n, v) \) is the largest success probability one can have for locating \( j_n^* \) with \( v \) observations.

One of our main results is as follows. For \( a, \alpha > 0 \), and \(-\frac{1}{2} < \delta < \frac{1}{2}\), let

\[
e(n, \delta) = (a^{-1}(2 \log n)^{(\frac{1}{2} + \delta)})^{1/\alpha}
\]
(1.3) **Corollary** Assume that for some \( \alpha > 1 \) and \( a > 0 \), \( n^\alpha r_n \to a \), \( n \to +\infty \). Then for all \( 0 < p < 1 \), and \( 0 < \delta < \frac{1}{2} \) the following holds. For

\[
v(\delta) = v(n, p, \delta) = \left[ \frac{(1+\delta)p n}{\epsilon(n, -\delta)} \right].
\]

(1.4) \( \liminf_{n \to +\infty} s(n, v(\delta)) \geq p \),

while

(1.5) \( \limsup_{n \to +\infty} s(n, v(-\delta)) \leq p \).

That is, to locate \( j_n^* \) with success probability \( p \),

\[
\left[ \frac{p n}{2 \epsilon(n, 0)} \right]
\]

observations are (almost) necessary and sufficient. We learned of this problem from [H], who motivated his paper with the practical problem of maximizing an objective function about which little is known, and is costly to evaluate. Such functions arise in geological exploration, for example. He proved the theorem above when \( r_j = a^j \), \( 0 < a < 1 \), in which case \( \epsilon(n, 0) \sim -\log \log n/2 \log a \). Also observe that for \( r_j = j^{-1} \), we have \( \epsilon(n, 0) \sim (2 \log n)^{-\frac{1}{2}} \).

Note that if the \( X \)'s are independent, \( [pn] \) observations are both necessary and sufficient to locate \( j_n^* \) with success probability \( p \). Thus, the fraction \( (2 \epsilon(n, \cdot))^{-1} \) in the corollary represents the percentage savings one realizes due to the dependence structure of \( X_j \). The two examples above show that this percentage goes to zero as \( n \to +\infty \), but at a rate too slow to be of any practical use.

In view of this difficulty, one might wonder if the problem posed above is too hard. That is, is it significantly easier to locate a large, but not
the largest value of $X$? In this regard, we make the following conjecture, stated, for brevity's sake, in less than precise language. For $0 < p, q < 1$, and $\delta > 0$, in order to locate $k_n^* \in \{1, 2, \ldots, n\}$ so that

$$X_{k_n^*} \geq \sqrt{q} X_n^*$$

with probability tending to $p$ as $n \to +\infty$.

$$\left[ \begin{array}{c} p n^q \\ 2 \epsilon(n^q, -\delta) \end{array} \right]$$

observations are sufficient. In particular, to locate a point which exceeds 70% of the true maximum, $o(n^{0.49})$ observations suffice. A plausible argument for this would proceed this way. In a very strong sense, ([LLR]),

(1.6) $X_n^* \sim (2 \log n)^{1/2}, n \to +\infty,$

so that $X_{n^q}^* \sim \sqrt{q} X_n^*, n \to +\infty$. Thus to locate $k_n^*$ as above, it suffices to locate $j_{n^q}^*$. One then applies Corollary 1.3, but note the reduction of this "easier" problem to the original one.

We have in fact proved (1.3) for non-decreasing covariance; this generality requires the notions of covering and packing numbers. Their appearance is not surprising: they are implicit in [H] and are well known, in some sense, to control stationary Gaussian processes. See e.g. [JM]. The positive or sufficient half is taken up in section 2. The proof therein holds for covariances with even a logarithmic decay to zero; see the end of section 2. The proof of the negative, or necessary half, is in section 3.

The strategy of the proofs below are due to Hajek. Our innovations are these: The aforementioned covering and packing numbers; carrying out the necessary half in a non-Markovian setting (Compare the proof of lemma 2, [H])
to lemma (3.10) below); and the use of Borell's inequality ([B]; (27) below) which should be of use in studying versions of this problem for which (1.6) does not hold.

2. Sufficiency

We define the notion of covering numbers. Let

\[(2.1) \quad B(i,u) = \{ j \in Z : r_{i-j} > u \}.\]

Observe that the sets \( B(\cdot, u) \) are translation invariant, and are balls of radius \( (2(1-u))^{\frac{1}{2}} \) in the metric \( d(1,j) = (E(X_i - X_j)^2)^{\frac{1}{2}} \). Let

\[N_n(u) = \text{least number of } B(\cdot, u) \text{ needed to cover } \{1, 2, \ldots, n\},\]

and let \( M_n \) be the median of \( X_n^* \): \( P(X_n^* > M_n) = \frac{1}{2} \).

\[(2.1) \quad \text{Theorem.} \quad \text{Assume that for all } \alpha > 0\]

\[(2.2) \quad (\log n)^\alpha r_n \to 0, \quad n \to +\infty.\]

Fix \( 0 < p, \delta < 1 \), and let

\[v = v(n, p, \delta) = \left[ (1+\delta) N_{[pn]}(M_n^{-1+\delta}) \right].\]

Then

\[(2.3) \quad \liminf_{n \to +\infty} s(n,v) \geq p.\]

Note that (2.2) is satisfied if \( r_n \sim n^{-\alpha}, \quad n \to +\infty, \quad \text{for any } \alpha > 0, \) or more generally \( r_n \sim \exp(-(\log n)^\alpha), \quad n \to +\infty. \) In fact the proof given below handles covariances which decay slower than those just given; see remarks at the end of this section.

Further note that (2.2) implies that \( r_n = o((\log n)^{-1}) \). Thus, we have the following (Theorem 4.3.3, [LLR]): \( a_n = (2 \log n)^{\frac{1}{2}} \) and
(2.4) \( b_n = (2 \log n)^{1/6} \) 

\[- \frac{3}{5} (2 \log n)^{-1/2} (\log \log n + \log 4\pi).\]

Under (2.1), we have \( M_n = b_n + O(a_n^{-1}) \), \( n \to +\infty \) and

(2.5) \( n \psi(M_n) \to \log 2, \ n \to +\infty. \)

where

\[ \psi(u) = \frac{1}{\sqrt{2\pi}} u^{\infty} e^{-x^2/2} \, dx. \]

Using the fact that \( M_n \sim (2 \log n)^{1/6}, \) (1.4) follows from (2.1) by an easy calculation.

For the proof, we recall Borell's inequality, [B], in a weak form, convenient for our needs.

(2.7) Theorem. Let \( Z_t, \ t \in T \) be a mean-zero Gaussian process with median \( M \)

in sup norm: \( P(\sup_{t \in T} Z_t > M) = \frac{3}{5}. \) Then for all \( \lambda > 0 \)

\[ P(\sup_{t \in T} |Z_t - M| > u) \leq 2 \psi(u/\sigma) \]

where \( \sigma^2 = \sup_{t \in T} \mathbb{E} Z_t^2. \)

Proof of (2.1). We prove this result by exhibiting a strategy with the desired properties. First, however, we reduce the result to the case \( p = 1. \) Assume the Theorem for \( p = 1. \) That is,

(2.8) \( s(n,v(n,1,\delta)) \to 1, \ n \to +\infty, \) for all \( 0 < \delta < \frac{1}{4}. \)

Now consider \( p < 1. \) As the law of \( n^{-1} j_n^{-1} \) is asymptotically uniform on \([0,1]\)

(Chapter 5, [LLR]) it suffices to locate \( j_{[pn]}^{\ast} \) by a sequential strategy of length \( v(n,p), \) with probability tending to 1, as \( n \to +\infty. \) But note that for large \( n, \)

\[ v([pn],1,\delta) = N_{[pn]}(M_{[pn]}^{1+\delta} \leq N_{[pn]}(M_n^{-1+2\delta}) = v(n,p,2\delta) \]
Therefore, by (2.8) the Theorem for all \( p \) follows.

For the proof below, it is convenient to extend \( X_j \) to a stationary sequence of \( Z \). Then let

\[
B_n = B(0, M_n^{-1+\delta});
\]

\[
b_n = \#B_n;
\]

\( I_n \) = set of centers of a minimal covering of \( \{1, 2, \ldots, n\} \) by translates of \( B_n \);

and

\[
i_n = \#I_n = N_n(M_n^{-1+\delta}).
\]

Our strategy is then:

1) Observe \( X \) at each \( j \in I_n \). Order these observations by height. That is \( I_n = \{j_1, j_2, \ldots\} \) where \( X_{j_1} > X_{j_2} > X_{j_3} \ldots \)

2) Completely search \( j_1 + B_n, j_2 + B_n, \ldots \) until a total of \( \lceil (1+\delta)i_n \rceil \) observations have been made. Take as \( j_n^* \) the location of the largest value observed.

Let

\[
S_n = \{j \in I_n : j + B_n \text{ is searched}\};
\]

\[
s_n = \#S_n \leq \delta i_n / b_n;
\]

and

\[
A_n = \text{median of } b_n \text{ standard independent Gaussian r.v.}
\]

Observe that by (2.5),

\[
(2.9) \quad b_n \psi(A_n) \to \log 2, \quad n \to +\infty.
\]

Referring to the strategy above, we have that
\[ P(J_n^* \text{ not located}) \]
\[ = P(j_n^* \in j + B_n \text{ for some } j \in I_n/S_n), \]
\[ \leq P(X_n^* < M_n - \Lambda_n) \]
\[ + P(X_j > 2 \Lambda_n \text{ for some } j \in I_n/S_n) \]
\[ + P(\text{for some } j \in I_n/S_n \text{ and } k \in j + B_n, X_j < 2 \Lambda_n \text{ and } X_k > M_n - \Lambda_n) \]
\[ = w_1 + w_2 + w_3. \]

It is enough to show that each \( j, w_j \to 0 \) as \( n \to +\infty \). We do this for each term in order. As \( n \to +\infty \), \( b_n \to +\infty \), hence \( \Lambda_n \to +\infty \). Thus by Borell's inequality (2.7),
\[ w_1 \to 0, \quad n \to +\infty. \]

For the second term, use (2.9) and Chebyshev's inequality.
\[ w_2 \leq P(\#\{j \in I_n : X_j > 2 \Lambda_n\} \geq s_n) \]
\[ \leq s_n^{-1} \sum_{I_n} 1\{X_j > 2 \Lambda_n\} \]
\[ \leq s_n^{-1} \sum_{I_n} 1 \]
\[ \leq d_n^{-1} \psi(2 \Lambda_n) \]
\[ \leq d_n^{-1} \log 2 \psi(2 \Lambda_n)/\psi(\Lambda_n) \]
\[ \to 0. \]

For the third term,
(2.10) \[ w_3 \leq \inf_n P(X_0 < 2 \Lambda_n \text{ and } \sup_{j \in B_n} X_j > M_n - \Lambda_n ) \]

\[ \leq \inf_n P(\sup_{j \in B_n} X_j > M_n - \Lambda_n \mid X_0 < 2 \Lambda_n ). \]

Given \( X_0 = x < 2 \Lambda_n \), and \( j \in B_n \), \( X_j \) is Gaussian with mean

\[ m_j = x < 2 \Lambda_n \]

and variance

\[ 1 - r_j \leq 1 - M_n^{-1+\delta} \]

Moreover, the median of

\[ \sup_{j \in B_n} X_j - m_j \mid X_0 = x \]

is dominated by \( 2 \Lambda_n \). Under (2.2) we have \( b_n = 0(\exp((\log n)^{\delta/2})) \), so that \( \Lambda_n = 0((\log n)^{\delta/4}) \). Hence

\[ \frac{M_n^{1+\delta/2}}{M_n^{1+\delta/2} - 10 M_n \Lambda_n} \to +\infty, \quad n \to +\infty \]

Therefore, again using Borell's inequality (2.7), and continuing from (2.10).

(2.11) \[ w_3 \leq \inf_n \psi \left( \frac{M_n - 5 \Lambda_n}{1 - M_n^{-1+\delta}} \right) \leq \inf_n \psi (M_n) \exp \left( -\frac{M_n^{-1+\delta/2}}{2} + 10 M_n \Lambda_n \right) \]

\[ \to 0, \quad n \to +\infty. \]

This finishes the proof.

We remark that the theorem above holds in the case where

\( (\log n)^{\alpha} r_n \to a, \quad n \to +\infty \), for some \( \alpha > 1 \), and \( a > 0 \). For then (2.5) continues to hold, and the proof just given only needs to be modified in (2.11), by using the better estimate...
\[ i_n \sim \frac{1}{n} \exp\left(-\alpha^{-1}(2 \log n)^{\frac{1}{2-\delta}}\right)^{1/\alpha}. \]

This proof fails for \( 0 < \alpha < 1 \), for then (2.5) fails. See Chapter 6, [LLR].

3. Necessity

Recall the definition (2.1), and define packing numbers by

\[ P(n, u) = \text{greatest number of disjoint balls } B(\cdot, u) \text{ that can be placed} \]
\[ \text{in } \{1, 2, \ldots, n\}. \]

We prove that

(3.1) **Theorem** Assume that for some \( \alpha > 1 \), \( n^{\alpha} r_n \to 0, \ n \to +\infty \). Fix \( 0 < \delta \), \( p < 1 \) and let

\[ v = v(n, p, \delta) = P([\lfloor (1-\delta)pn \rfloor], (2 \log n)^{-\delta}) \]
\[ + P([\lfloor pn \rfloor], (2 \log n)^{-1-\delta}). \]

Then

\[ \limsup_{n \to +\infty} \delta(v, n) \leq p. \]

(1.5) follows from this by a routine calculation. Concerning the relationships between packing and covering numbers, we have

\[ P_n(u) \leq n(\#B(\cdot, u))^{-1} \leq N_n(u). \]

For the range of \( u \) we considering, \( P_n(u) \) and \( N_n(u) \) need not even be of the same order; their values could also be quite difficult to compute.

Throughout the proof below C denotes a positive numerical constant that may change from line to line.
Proof. Let
\[ \alpha_n = (2 \log n)^{\delta/2}, \]
\[ \beta_n = (2 \log n)^{\delta+\delta}, \]
and
\[ \gamma_n = (2 \log n)^{1+\delta}. \]
Let \( O = O_n \) be a strategy of length (3.2); let
\[ (3.3) \ M = M_n = \{ i \in O_n : |X_i| < \alpha_n \}. \]
\[ N = N_n = \{ i \in O_n : |X_i| > \alpha_n \}. \]

\[ (3.4) \ B_i = \begin{cases} B(i, \beta_n^{-1}), & i \in M \\ B(i, \gamma_n^{-1}), & i \in N, \end{cases} \]
and \( U = U_n = \bigcup_{i \in O} B_i \).

We shall prove that
\[ (3.5) \ \limsup_n P(j_n^* \in U_n) \leq p. \]
This allows us to assume without loss of generality that for \( i, j \in O \).

\[ (3.6) \ r_{i-j} \leq \begin{cases} \beta_n, & |X_j|, |X_j| < \alpha_n \\ \gamma_n, & \text{otherwise} \end{cases} \]

Moreover, writing \( O = \{ i_1 \leq i_2 < \ldots \} \) we have
\[ (3.7) \ r(i_u - i_v) \leq C \beta_n^{-1} |u-v|^{-\alpha}. \]

Let \( \theta \) denote all information obtained from \( O_n \); that is,
\( \theta = \{ O_n, X_j, j \in O_n \} \). Set
\[ A = \{ \theta : \max_{1 \leq i \leq n} |X_{ij}| < (2 \log n)^{\frac{2}{3}} \}; \]

\[ \#\{i \in O_n : |X_i| > \alpha_n < P([\delta_n], \gamma_n) \}. \]

We prove that

(3.8) \[ P(A) \to 1, \ n \to \infty. \]

and for \( G = G_n = \{1, 2, \ldots, n\}/U_n \).

(3.9) \[ \theta \in A \Rightarrow \#G \geq [(1-p)n]. \]

This last claim is immediate from the definitions of \( A \) and \( U \), and the following inequality: \( P_n(u) \leq n(B(\cdot, u))^{-1} \). (Note that \( O_n \), and hence \( G_n \) are random sets of integers.)

(3.8) is seen in two steps. First,

\[ P\left( \max_{1 \leq i \leq n} |X_{ij}| > (2 \log n)^{\frac{2}{3}} \right) \]

\[ \leq 2n \psi((2 \log n)^{\frac{2}{3}}) \]

\[ \to 0, \ n \to +\infty. \]

\( \psi(\cdot) \) is defined in (2.6) above. Second, let \( \rho_n = P([p\delta_n], \gamma_n^{-1}) \). Then

\[ P(\#\{i \in O_n : |X_i| > \alpha_n \} \geq \rho_n) \]

\[ \leq P(X_n \{X_i > \alpha_n \} \geq \rho_n) \]

\[ \leq \rho_n^{-1} n \psi(\alpha_n). \]

\[ \leq (2 \log n)^{(1+\delta)/\alpha} \psi(\alpha_n) \to 0, \ n \to +\infty. \]

The last inequality holds for large \( n \), and follows from the observation that

\[ B(0, \gamma_n^{-1}) \supset \{j : |j| < (2 \log n)^{(1+\delta)/\alpha} \} \]

for large \( n \). These last two calculations finish the proof of (3.8).
A given outcome $\theta$, induces a Gaussian conditional probability $P_\theta$ on $X_u$, $u \not\in 0$. Denoting expectation with respect to $P_\theta$ by $E_\theta$, let $m_u = E_\theta X_u$, and for $v \not\in 0$,
\[ c_{uv} = E_\theta(X_u - m_u)(X_v - m_v), \]
be the conditional mean and covariance, respectively.

(3.10) **Lemma.** For $\theta \in A$, $u \in G$, and $v \not\in 0$,\n
(3.11) $|m_u| < C (\log n)^{-\frac{1}{2} - \delta/2}$.

(3.12) $|c_{uu}| > 1 - C \gamma_n^{-1}$.

and

(3.13) $|c_{uv} - r_{u-v}| \leq C(\beta_n^{-1} \wedge |u-v|^{-\theta})$.

We observe here a corollary to (the proof of) this lemma. The only facts about $O$ used in the proof are (3.6) and (3.7). Hence, for $j \in O$, and $u \in B_j$ (Recall (3.4)), let $\tilde{m}_u = r_{i-u} X_i$, and $\tilde{c}_{uu} = 1 - r_{i-u}^2$, which are respectively the mean and variance of $X_u$ given $X_i$. Then by (3.11) and (3.12),

(3.14) $|m_u - \tilde{m}_u|, |c_{uu} - \tilde{c}_{uu}| \leq C(\log n)^{-\frac{1}{2} - \delta/2}$.

**Proof.** Fix $u$ and $v$ as in the lemma. Let $o = \#_0$, and write the covariance matrix of the Gaussian vector $((X_u, X_v), (X_i, i \in O))$ as

\[
\begin{bmatrix}
A & \Lambda \\
\Lambda^t & \Gamma
\end{bmatrix}
\]

where

\[
\Lambda = \begin{bmatrix} 1 & r_{u-v} \\ r_{u-v} & 1 \end{bmatrix};
\]

$A = (A_{w1})$ is a $2 \times 0$ matrix; and $\Gamma = (\Gamma_{ij})$ is a $o \times o$ matrix. Given $(X_i, i \in O) = y$, $(X_u, X_v)$ is a bivariate Gaussian vector with mean vector
(3.15) \( \Lambda \Gamma^{-1} y \)

and covariance matrix

(3.16) \( \Lambda = \Lambda \Gamma^{-1} \Lambda^t \).

Notice that \( \Gamma_{ii} = 1 \), and by (3.7), \( |\Gamma_{ij}| < C \beta_n^{-1} |i-j|^{-\alpha} \). That is \( \Gamma \) is nearly the identity, but standard techniques for estimating the terms of \( \Gamma^{-1} \) directly fail here. (We shall see in a moment that \( \Gamma^{-1} \) exists.) Thus the technique we employ is to define various norms on \( \mathbb{R}^o \), and estimate norm of \( \Gamma : \mathbb{R}^o \to \mathbb{R}^o \) below. This gives upper bounds on \( \Gamma^{-1} : \mathbb{R}^o \to \mathbb{R}^o \).

Recalling (3.3) and (2.4), for \( y \in \mathbb{R}^o \) let

\[
\|y\| = (\max_{u \in N} |y_u|) \vee (a_n \max_{u \in M} |y_u|).
\]

and let \( \ell = (\mathbb{R}^o, \| \cdot \|) \). Observe that \( \ell^* \), the dual of \( \ell \), has norm

\[
\|y\|_{\ell^*} = \sum_{N} |y_u| + a_n^{-1} \sum_{M} |y_u|.
\]

Consider the \( i \)th coordinate of \( y = \Gamma x \). By (3.7),

\[
|y_i| = |\sum_{j} \Gamma_{ij} x_j| \\
\geq |x_i| - \sum_{M} |x_j r_{i-j}| \\
- \sum_{N} |x_j r_{i-j}| \\
\geq |x_i| - C a_n^{-1} \beta_n^{-1} \|x\| \sum_{j \geq 1} |j|^{-\alpha} \\
- C \tau_n^{-1} \|x\| \sum_{j \geq 1} |j|^{-\alpha} \\
\geq |x_i| - C \tau_n^{-1} \|x\|.
\]

Thus, for large \( n \), the image of the unit ball of \( \ell \) under \( \Gamma \) is an open set.
so $r^{-1}$ exists. Moreover, \( \| \Gamma \|_{\ell \rightarrow \ell} > \frac{1}{4} \). This, and the fact that $\Gamma$ is self-adjoint imply that
\[
\|r^{-1}\|_{\ell \rightarrow \ell} < 2.
\]

Let $A_u = (A_{ui} : i \in \mathbb{N}) = (r_{u,i} : i \in \mathbb{N})$. As $u \in \mathbb{C}$,

\[
\|A_u\| \leq C; \quad \|A_u\|_* \leq C r_n^{-1}.
\]

The proof of (3.11) is then easy. As $\theta \in A$, $y = (X_i : i \in \mathbb{N})$ satisfies $\|y\| \leq (2 \log n)^{1/4-\delta/2}$. Thus by (3.15), (3.17) and (3.18),
\[
|m_u| = \left| \langle A_u, r^{-1} y \rangle \right|
\leq \|A_u\|_* \|r^{-1}\|_{\ell \rightarrow \ell} \|y\|
\leq C r_n^{-1} (2 \log u)^{1/4+\delta/2}
\leq C (2 \log n)^{-1/4-\delta/2}.
\]

The proof of (3.12) is similar. By (3.16), (3.17) and (3.18),
\[
|1 - c_{uu}| < \left| \langle A_u, r^{-1} A_u \rangle \right| < C r_n^{-1}.
\]

Likewise the proof of the first half of (3.13). Observe that $\|A_u\| \leq a_n$. Hence
\[
|c_{uv} - r_{u-v}| \leq \left| \langle A_u, r^{-1} A_v \rangle \right|
\leq \|A_u\|_* \|r^{-1}\|_{\ell \rightarrow \ell} \|A_v\|
\leq C r_n^{-1} a_n = C \beta_n^{-1}.
\]

Proving the second half of (3.13) requires a new norm. For $v \in U$, $y \in \mathbb{R}^0$, set
\[ \|y\|_v = \sup_{j \leq 0} |i-v|^\alpha |y_i| \]

Then, for \( z = \Gamma y \), consider the \( i \)th coordinate of \( z \), where without loss of generality we assume that \( i < v \). Then

\[
(3.19) \quad |z_i| > |y_i| - \|y\|_v \left[ \sum_{j \leq 0} \sum_{j > 0} \sum_{j > 0} r_{i-j} |j-v|^{-\alpha} \right]_{j < i, i < j < v, v < j} \\
= y_i - I - II - III.
\]

As \( n^\alpha r_n \to 0 \), and (3.6) holds,

\[
(3.20) \quad I, II \leq C \|y\|^2 v^{-1} |i-v|^{-\alpha}.
\]

Concerning II, a similar argument shows that

\[
\sum_{j \leq 0} \sum_{j > 0} r_{i-j} |j-v|^{-\alpha} < C \|y\|_v \beta_n^{-1} |i-v|^{-\alpha} \quad i < j < (i+v)/2
\]

and likewise for the sum over \((i+v)/2 \leq j < v\). Thus (3.20) holds for II as well. Denoting \((R^0, \|\cdot\|_v)\) by \( \ell_v \), (3.19) and (3.20) show that for large \( n \),

\[
(3.21) \quad \|\Gamma^{-1}\|_{\ell_v \to \ell_v} \leq 2.
\]

To see (3.13), note that \( \|A_v\|_v \leq C \). (For the definition of \( A_v \) see (3.18).) Thus by (3.16) and (3.21) it is enough to estimate

\[
|(A_u, \Gamma^{-1} A_v)| \leq C \sum_{j \leq 0} |r_{j-u} |j-v|^{-\alpha}|.
\]

But this last sum can be dominated in a manner very close to (3.19) and (3.20). This concludes the proof of the lemma. \( \blacksquare \)
Recall the definitions (2.3), and given \( x \in \mathbb{R} \), set \( x_n = b_n + a_n^{-1} x \). Let \( \tilde{x}_n = \max_{1 \leq j \leq n} X_j \). Then (Theorem 4.3.3., [LLR])
\[
(3.22) \quad P(\tilde{x}_n \leq x_n) \to e^{-e^{-x}}, \quad x \in \mathbb{R}.
\]

For an arbitrary \( F \subset \{1, 2, \ldots, n\} \) let \( \tilde{F} = \max_{i \in F} X_i \).

(3.23) Lemma. The following limit holds uniformly over \( \theta \in A \), and \( y \) in compact subsets of \( \mathbb{R} \).
\[
\lim \inf \frac{\log n}{n} P(\tilde{G}_n \leq y_n) \geq \exp(-(1-p)e^{-y}).
\]

Proof. By (3.9), \( g(n) = \#C_n \geq [(1-p)n] \). A routine calculation then shows that the lemma follows from this result: Uniformly in \( y \in \mathbb{R} \) and \( \theta \in A \),
\[
\lim_{n \to \infty} P_{\theta}(\tilde{G}_n \leq y_{g(n)}) \to e^{-e^{-y}}
\]

This last fact can be seen as an immediate corollary to (3.10) and (the obvious triangular array version of) Theorem 6.2.1. [LLR]. (It can also be seen by a direct computation involving the normal comparison lemma, see below.)

(3.8), (3.22) and the last lemma can be used to prove that
\[
\lim \sup_{n} P(J_n^* \in O_n) \leq p.
\]
(see [H]), but not (3.5), which we need to complete the proof. The next lemma provides the last ingredient we need.
(3.24) Lemma. For all $x \in \mathbb{R}$ uniformly in $\theta \in A$,

$$\sup_{u,v \geq x} \left| P_\theta(\tilde{U}_n \leq u_n; \tilde{G}_n \leq v_n) - P_\theta(\tilde{U}_n \leq u_n) P_\theta(\tilde{G}_n \leq v_n) \right|$$

$$\to 0, \quad n \to +\infty.$$ 

Furthermore,

$$\limsup_{n} P_\theta(\tilde{U}_n \leq x_n | \theta \in A) \leq e^{-p} e^{-x}$$

uniformly in $\theta \in A$, and $x$ in compact subsets of $\mathbb{R}$.

To complete the proof of the Theorem, fix $\delta > 0$, and choose $x$ so large that for large $n$,

$$P(A^c), P(\tilde{U}_n > x_n), P(\tilde{G}_n < (-x)_n) < \delta.$$ 

This is possible by (3.8), (3.23) and (3.26). Then by (3.24) and (3.25), for large $n$,

$$P(j \in U) = P(\tilde{U} > \tilde{G})$$

$$\leq 3 \delta + \int_{-x}^{\infty} P(\tilde{U} > y_n, \tilde{G} \in (dy)_n \setminus \theta \in A)$$

$$= 4 \delta + \int_{-x}^{\infty} e^{-p} e^{-y} P(\tilde{G} \in (dy)_n \setminus \theta \in A).$$

Now, integrating by parts and using (3.23), for large $n$ the last integral is less than

$$\delta + \int_{-x}^{\infty} e^{-(1-p)e^{-y}} d e^{-p} e^{-y}$$

$$\leq \delta + \int_{-\infty}^{\infty} e^{-(1-p)e^{-y}} d e^{-pe^{-y}} = \delta + p.$$ 

As $\delta > 0$ was arbitrary, this proves (3.5).
Proof of (3.24) (3.26) follows from (3.22), (3.23) and (3.25): For an arbitrary \( \delta > 0 \) and large \( n \),

\[
e^{-e^{-x}} \leq \delta + P(\tilde{X} \leq x_n) \]

\[
\leq 2\delta + P(\tilde{X} \leq x_n, \theta \in A) \]

\[
\leq 2\delta + P(\tilde{U}_n \leq x_n, \tilde{G}_n \leq x_n | \theta \in A) \]

\[
\leq 3\delta + P(\tilde{U}_n \leq x_n | \theta \in A) P(\tilde{G}_n \leq x_n | \theta \in A) \]

(3.23) then implies (3.26).

It remains to prove (3.25). To do this, given \( \theta \in A \), define Gaussian random variables \( Y_{i_1}^\theta, \ldots, Y_{i_n}^\theta \) so that \( E Y_{i_j}^\theta = m_{j_1} \), \( 1 \leq j \leq n \), and

\[
E(Y_{i_j}^\theta - m_{j_1})(Y_{j_j}^\theta - m_{j_1}) = \begin{cases} c_{ij} & (i, j) \in U \times U, \ G \times G \\ 0 & (i, j) \in U \times G, \ G \times U. \end{cases}
\]

Here \( m_{j_1} \) and \( c_{ij} \) are defined just before lemma (3.10).

Observe that

\[
P_{\theta}(\tilde{U}_n \leq u_n) P_{\theta}(\tilde{G}_n \leq v_n) = P(\max_{u} Y_{i_j}^\theta \leq u_n; \max_{G_n} Y_{i_j}^\theta \leq v_n).\]

hence (3.26) can be shown by an application of the normal comparison lemma.

Let \( \rho_{ij} = c_{ii}^{-1/2} c_{jj}^{-1/2} c_{ij} \), which is the normalized covariance of \( Y_{i_j}^\theta \) and \( Y_{j_j}^\theta \).

Further, let \( u_{nj} = c_{jj}^{-1/2} (u_n - m_j) \), likewise for \( v_{nj} \). Then by Theorem 4.2.1. [LLR], it is enough to show that for all \( x \in \mathbb{R} \),

\[
(3.27) \sup_{u,v2x} \sum_{j \in U} \sum_{k \in G} \sum_{j<k} \frac{\rho_{ik}}{\sqrt{1 - \rho_{ik}^2}} \exp \left[ -\frac{u_{nj}^2 + v_{nk}^2}{2(1 + \rho_{jk})} \right] \rightarrow 0, \quad n \rightarrow +\infty.
\]
Let $\epsilon > 0$ be a number to be fixed later. Then for $i \in 0$, let 
\[ E_i = \{ j : r_{i-j} > \epsilon \}. \]
We first show that
\[ \sup_{u,v \geq x} \sum_{i \in E_i} \sum_{j \in G} \frac{\rho_{jk}}{\sqrt{1 - \rho_{jk}^2}} \exp \left[ -\frac{u^2_{nj} + v^2_{nk}}{2(1 + \rho_{jk}^2)} \right] \to 0, \quad n \to +\infty. \]  

We have that
\[ \sup_{j \geq 1} r_j < \eta < 1. \]  

This and (3.14) show that for large $n$, $0 < 1 - \eta - \epsilon < c_{jj} < 1 - 2\epsilon^2$, $j \in E_i$, $i \in 0$. Here we take $0 < \epsilon < 2^{-1/2}$ and $1 - \eta$, so that the inequalities above are meaningful. Also, by the definition of $E_i$, and $G$, and (3.13), $c_{jk} < C \beta_n^{-1}$ for $(j,k)$ as in (3.28). These facts and (3.12) imply that
\[ \rho_{jk} < C \beta_n^{-1}, \quad (j,k) \text{ as in (3.28)}. \]  

(3.29) and (3.14) show that for large $n$, $j \in E_j$, and $j \in 0$ we have
\[ m_i \leq c(\log n)^{-\frac{1}{2}-\frac{5}{2}} + r_{i-j} X_i \]
\[ \leq c(\log n)^{-\frac{1}{2}-\frac{5}{2}} + \eta a_n, \]
as $\theta \in A$. This, and (3.30) show that
\[ \frac{u_{nj}}{1 + \rho_{jk}} \geq (1 - \eta - \epsilon) a_n - \frac{x}{2} (2 \log n)^{-\frac{1}{2}} (\log \log n + \log 4\pi) \]
\[ = d_n, \]
where $u_{nj}$ is as in (3.28). Last of all, fix $y < x$ arbitrary. By (3.11) and (3.30), for large $n$ we have
\[ (3.32) \quad \frac{v_{nk}}{1 + \rho_{jk}} \geq y_n, \quad v, j, k \text{ as in (3.28)}. \]

We are now ready to estimate (3.28). By (3.30–32), (3.28) is dominated by a constant times
\[
\exp(-\frac{1}{2}(d_n^2 + y_n^2)) \sum_{j \in E_1} \sum_{j > k} \rho_{jk}.
\]
\[
\leq C \sum_{j=1}^{n_0} \exp(-\frac{1}{2}(d_n^2 + y_n^2))
\]
\[
\leq C \exp(-\frac{1}{2}(d_n^2 + y_n^2)) \rightarrow 0, \quad n \rightarrow +\infty.
\]

Here, we've used (3.13), and \( \sum_{j=1}^{n_0} j^{-\alpha} \leq C; \) \( \# E_1 \leq C, \) by stationarity of \( X; \) and \( n_0 \leq n. \) This proves (3.28).

It remains to prove that
\[
(3.33) \quad \sup_{u,v \geq x} \sum_{j \in B_1/E_1} \sum_{j < k} \rho_{jk} \exp\left(-\frac{u_{nj}^2 + v_{nk}^2}{2(1 + \rho_{jk})}\right) \rightarrow 0, \quad n \rightarrow +\infty.
\]

Here, \( B_1 \) is defined in (3.4).

In fact we will show that for \( \epsilon > 0 \) sufficiently small, and \( \delta > 0 \) depending only on \( \eta \) in (3.29),
\[
(3.34) \quad \sup_{u,v \geq x} \sup_{(j,k)} \exp\left(-\frac{u_{nj}^2 + v_{nk}^2}{2(1 - \rho_{jk})}\right) \leq n^{-1-\delta}.
\]

This and (3.13) easily imply (3.38), which finishes the proof of (3.27), concluding the proof of the lemma.
For any \( \epsilon > 0 \) fixed, and all \( j \in 0, j \in B_1/E_1, k \in G \) the following holds for sufficiently large \( n \). By the definition of \( E_1 \), (3.13), (3.14), and (3.29),

\[
\rho_{jk} = \frac{c_{jk}}{c_{jj}^{-\frac{1}{2}} c_{kk}^{-\frac{1}{2}}} = \frac{\eta + \epsilon}{(1-2\epsilon)^{\frac{1}{2}} (1-\epsilon)^{\frac{1}{2}}}.
\]

This last expression can be made strictly less than \( \mu \), \( \eta < \mu < 1 \), for \( \epsilon > 0 \) sufficiently small. Also, by (3.14), for \( i \in 0 \), and \( j \in B_1/E_1 \), and \( n \) sufficiently large,

\[
u_{nj} = b_n + a_n^{-1} u - m_j\]

\[
> b_n + a_n^{-1} u - 2\epsilon a_n
\]

\[
\geq (1-2\epsilon) a_n - \frac{1}{2} a_n^{-1} (\log \log n + \log 4\pi) + a_n^{-1} y
\]

\[
= e_n,
\]

where \( y < x \) is arbitrary. Similarly, we have \( v_{nk} \geq y_n \).

Fix \( \delta > 0 \) so that \( \frac{1}{2} + \delta < 1/(1+\mu) \). Then for \( \epsilon > 0 \) sufficiently small, the observations above show that

\[
\frac{u_{nj}^2 + v_{nk}^2}{2(1 + \rho_{jk})} > \frac{c_n^2 + y_n^2}{2(1 + \mu)}
\]

\[
> \frac{(2-3\epsilon)a_n^2}{2(1+\mu)}
\]

\[
> (1+\delta)(\log n).
\]

This implies (3.34).
REFERENCES


