Majorizing Measure Bounds for Processes with Mixed Exponential Tails

Fuchang Gao
Department of Mathematics
University of Idaho
Moscow, ID 83843
fuchang@uidaho.edu

November 19, 1999
Abstract

If \( \{X_t\}_{t \in T} \) is a process with tail probability

\[
\Pr(|X_t - X_s| > u) \leq \exp \left( -\min \left\{ \frac{A_i u^{\alpha_i}}{\|t - s\|^{\alpha_i}} : 1 \leq i \leq d \right\} \right),
\]

where \( A_i \) and \( \alpha_i \) are positive constants, and \( \| \cdot \|_i, 1 \leq i \leq d \), represent \( d \) different norms. Suppose \( \alpha_1 = \min \{\alpha_i\} \). Then there exists a constant \( K \) depending only on \( d \) and \( \alpha_i \)'s, such that

\[
\left\| \sup_{t \in T} |X_t| \right\|_{\psi_{\alpha_1}} \leq \|X_{t_0}\|_{\psi_{\alpha_1}} + K \sum_{i=1}^{d} A_i^{-1/\alpha_i} \gamma_{\alpha_i}(T, \| \cdot \|_i),
\]

where \( t_0 \in T \), and \( \gamma_{\alpha_i} \) are the majorizing measures of \( T \).
This is an extension of the corresponding result in [T2], where \( d = 2 \) was considered. Some techniques we use here are similar to that of [LT], though the latter contains an error. After this paper had been written, the author learned from Talagrand that the result had appeared in [M]. The author have not got the access to [M]. But according to Talagrand, the method of [M] is a rewriting of [T2]. So, our methods should be different. Also, we deal with Orliz norm, which is a stronger result.

A few weeks after a preprint of this paper was sent to Talagrand, the author received a preprint [T4] from him, which contains a different (and simple !) proof of this result. The author noticed and informed Talagrand that by regrouping the summation in our approach, the main result of [T4] can be recovered. However, the approach in [T4] is superior. So the author decide not to submit this paper for publication.

Lemma 1 seems to be new. It gives a quick proof of the theorem for the case \( d = 1 \) with Orliz norm. (For example, the sub-Gaussian case.) This is the main reason to make this paper accessible to public.

\[ \text{0.1 Introduction} \]

Consider a process \( \{X_t\}_{t \in T} \) with tail probability

\[
\Pr(|X_t - X_s| > u) \leq \exp\left(-\min\left\{ \frac{A_i u^{\alpha_i}}{\|t - s\|_i^{\alpha_i}} : 1 \leq i \leq d \right\}\right),
\]

(1)

where \( A_i \) and \( \alpha_i \) are positive constants, and \( \| \cdot \|_i, 1 \leq i \leq d \), represent \( d \) different norms. It is of interest to study \( \sup_{t \in T} |X_t| \). In this note, we will prove

**Theorem 1** If \( \{X_t\}_{t \in T} \) is a process satisfying (1) with \( \alpha_1 = \min\{\alpha_i : 1 \leq i \leq d\} \), then there exists a constant \( K \) depending only on \( d \) and \( \alpha_i \)'s, such that

\[
\left\| \sup_{t \in T} |X_t| \right\|_{\psi_{\alpha_1}} \leq \|X_{t_0}\|_{\psi_{\alpha_1}} + K \sum_{i=1}^{d} A_i^{-1/\alpha_i} \gamma_{\alpha_i}(T, \| \cdot \|_i),
\]

1
where \( t_0 \in T \), and \( \gamma_{\alpha} \) are the majorizing measures of \( T \). (See the definition below.)

For a bounded (not a singleton) set \( T \) in a normed space \((X, \| \cdot \|)\), the majorizing measure \( \gamma_\alpha(T, \| \cdot \|) \) is defined by

\[
\gamma_\alpha(T, \| \cdot \|) := \inf_{\mu} \sup_{t \in T} \int_0^{\infty} \left[ \log \frac{1}{\mu(B(t, \varepsilon))} \right]^{1/\alpha} d\varepsilon, \tag{2}
\]

where \( B(t, \varepsilon) \) is the ball in \( X \) with center \( t \in T \) and radius \( \varepsilon > 0 \); \( \mu \) is a probability measure on \( T \); and \( \mu(B(t, \varepsilon)) \) is understood as \( \mu(B(t, \varepsilon) \cap T) \). The infimum in (2) is taken over all probability measure \( \mu \). We note that the right side of (2) remains the same if we replace \( T \) by any dense subset \( T' \subset T \). It can easily be proved that

\[
\frac{1}{r} \gamma_\alpha(T) \leq \sup_{t \in T} \sum_{k=b}^{\infty} r^{-k} \left[ \log \frac{1}{\mu(B(t, r^{-k}))} \right]^{1/\alpha} \leq \frac{r}{r-1} \gamma_\alpha(T)
\]

for \( r > 1 \), where \( b \) is an integer satisfying \( r^{-b} \geq \text{diam}(T) \).

### 0.2 An Equivalent Definition

In this section, we introduce “net ordering” to explain majorizing measures. The idea of net-ordering was developed by analyzing [T1] when the author was assigned as a course project for a probability course in early 96. Later on, we learned that [T3] more or less contained a similar idea, but not the same. We believe this “net-ordering” enables one to understand majorizing measures better, at least in the author’s case. We also had realized that by regrouping the summands in the definition, one could obtain a simpler form as in [T4]. However, we will then lose some useful information about the set. Also, the current formulation itself well explains the relation between and Dudley’s metric entropy condition.

For \( \varepsilon > 0 \), an \( \varepsilon \)-net \( N(\varepsilon) \) of \( T \) is a set such that for every \( t \in T \), there exists a point \( s = s(t) \in N(\varepsilon) \) satisfying \( \| t - s \| \leq \varepsilon \).

An ordering of a finite set \( S \) is a one-to-one map from \( S \) onto \( \{1, 2, \ldots, |S|\} \), where \( |S| \) is the cardinality of \( S \).
For $r > 1$ and integer $k$, let $N_k$ be an $r^{-k}$-net of $T$, and let $I_k$ be an ordering of $N_k$. Because $T$ is bounded, there exists the largest integer $b$, such that $N_b$ contains only one element, say $t_0$.

Fix $n \geq b$. For each $t \in N_n$. We denote $\pi_n(t, n) = t$, and select $\pi_{n-1}(t, n), \pi_{n-2}(t, n), \ldots, \pi_k(t, n), \ldots$ in order, so that $\pi_k(t, n) \in N_k$,

$$\|\pi_k(t, n) - t\| \leq r^{-k},$$

and $I_k(\pi_k(t, n))$ is as small as possible. The procedure stops when we reach $\pi_b(t, n) = t_0$. Thus we have the expansion

$$t = t_0 + \sum_{k=b}^{n-1} \pi_{k+1}(t, n) - \pi_k(t, n). \quad (3)$$

In general, $\pi_k(t, n)$ may depend on $n$. When there is no confusion, we simply denote it by $\pi_k(t)$.

We also denote

$$M_\alpha(r, n) = \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} \left[ \log I_k(\pi_k(t)) \right]^{1/\alpha}.$$ 

Clearly, $M_\alpha$ depends on nets $N_k$, and their orderings $I_k$, $b \leq k \leq n$. Define

$$O_\alpha(T, r) = \inf_{n \geq b} \sup_{r} M_\alpha(r, n), \quad (4)$$

where the infimum is taken over all choices of nets $N_k$ and all the orderings of these nets.

The following equivalent condition will be used later.

**Theorem 2** There exist a constant $K$ depending only on $r$, $r \geq 3$, and $\alpha$ such that

$$\frac{1}{K} \gamma_\alpha(T) \leq O_\alpha(T, r) \leq K \gamma_\alpha(T).$$

**Proof:** (i) Without loss of generality, we assume $T$ be compact. Suppose $\gamma_\alpha(T) < \infty$. There exists a probability measure $\mu$ on $T$, such that

$$\sup_{t \in T} \sum_{k \geq b} r^{-k} \left[ \log \frac{1}{\mu(B(t, r^{-k-1}))} \right]^{1/\alpha} \leq r^2 \gamma_\alpha(T).$$
For each \( k \geq b \), we let \( N_k \) consist of elements \( s_1, s_2, \ldots, s_i, \ldots \) chosen successively, such that
\[
\|s_i - s_j\| \geq r^{-k}
\] (5)
for \( i \neq j \), and \( \mu(B(s_i, r^{-k-1})) \) is as large as possible. (The maximum can be reached by assuming \( T \) be compact.) Let \( I_k \) be the natural ordering of \( N_k \). That is \( I_k(s_i) = i \).

For each \( n \geq b \) and \( t \in N_n \), by (3) we have the expansion
\[
t = t_0 + \sum_{k=b}^{n-1} \pi_{k+1}(t) - \pi_k(t),
\]
where \( \pi_0(t) = t_0, \ldots, \pi_k(t), \ldots, \pi_n(t) = t \) satisfy \( \pi_k(t) \in N_k \),
\[
\|\pi_k(t) - t\| \leq r^{-k}
\]
and \( I_k(\pi_k(t)) \) is as small as possible.

Suppose \( I_k(\pi_k(t)) = m \). By the selection of \( \pi_k(t) \), if \( s \in N_k \) satisfies \( I_k(s) < m \), then
\[
\|s - t\| > r^{-k}.
\] (6)

Therefore, by the selection of \( \pi_k(t) \), we have
\[
\mu(B(t, r^{-k-1})) \leq \mu(B(\pi_k(t), r^{-k-1})).
\] (7)

We also note that by (5), the sets \( B(s, r^{-k-1}), I_k(s) \leq m \), are pairwise disjoint, and each of them satisfies
\[
\mu(B(s, r^{-k-1})) \geq \mu(B(\pi_k(t), r^{-k-1})).
\] (8)

Therefore
\[
\mu(B(\pi_k(t), r^{-k-1})) \leq \frac{1}{m} = \frac{1}{I_k(\pi_k(t))}.
\]

Plug into (7), we obtain
\[
I_k(\pi_k(t)) \leq \frac{1}{\mu(B(t, r^{-k-1}))}.
\]
Therefore

\[ M_\alpha(r, n) = \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} [\log I_k(\pi_k(t))]^{1/\alpha} \]

\[ \leq \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} \left[ \log \frac{1}{\mu(B(t, r^{-k-1}))} \right]^{1/\alpha} \]

\[ \leq r^2 \gamma_\alpha(T), \]

which implies \( O_\alpha(T, r) \leq K \gamma_\alpha(T) \) for \( K = r^2 \).

(ii) Suppose \( O_\alpha(T, r) < \infty \). There exist nets \( N_k \) and their orderings \( I_k \), such that for all \( n > b \),

\[ \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} [\log I_k(\pi_k(t))]^{1/\alpha} \leq 2O_\alpha(T, r). \]

We take a discrete probability measure \( \mu \) such that for each \( n \geq b \) and \( t \in N_n \),

\[ \mu(\{\pi_k(t)\}) \geq \left( \frac{6}{\pi^2} \right)^2 \left[ \frac{1}{(k-b+1)I_k(\pi_k(t))} \right]^2. \]

Because \( \|t - \pi_k(t)\| \leq r^{-k} \), we have \( \pi_k(t) \in B(t, r^{-k}) \). Therefore

\[ \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} \left[ \log \frac{1}{\mu(B(t, r^{-k}))} \right]^{1/\alpha} \]

\[ \leq \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} \left[ \log \frac{1}{\mu(\{\pi_k(t)\})} \right]^{1/\alpha} \]

\[ \leq \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} \left[ 2 \log \left( \frac{6}{\pi^2} \right) + 2 \log(k-b+1) + 2 \log I_k(\pi_k(t)) \right]^{1/\alpha} \]

\[ \leq K_1 r^{-b} + K_2 \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} [\log I_k(\pi_k(t))]^{1/\alpha} \]

\[ \leq K O_\alpha(T, r), \]

where in the last inequality we used the assumption that \( T \) is not a singleton.

Because \( \bigcup_{n \geq b} N_n \) is dense in \( T \), we have

\[ \sup_{t \in T} \sum_{k=b}^{n-1} r^{-k} \left[ \frac{1}{\mu(B(t, r^{-k}))} \right]^{1/\alpha} = \sup_{n \geq b} \sup_{t \in N_n} \sum_{k=b}^{n-1} r^{-k} \left[ \log \frac{1}{\mu(B(t, r^{-k}))} \right]^{1/\alpha}. \]
Thus $\gamma_\alpha(T) \leq KO_\alpha(T, r)$. 

0.3 Nets and Their Orderings

Fix $r$ so that $r \geq 3^{\alpha_i}$ for all $1 \leq i \leq d$. By homogeneity, we can assume that $A_i = 1$ and $\gamma_{\alpha_i}(T, \| \cdot \|_i) \leq r^{-3/\alpha_i}$ for all $1 \leq i \leq d$. By Theorem 2, for any $1 \leq i \leq d$ and $k \geq 0$, there exists a $r^{-k/\alpha_i}$-net $N_k^{(i)}$ of $(T, \| \cdot \|_i)$ and an ordering $I_k^{(i)}$, such that, for all $1 \leq i \leq d$ and $n > 0$

$$\sup_{t \in N_n^{(i)}} \sum_{k=0}^{n-1} r^{-k/\alpha_i} \left[ \log I_k^{(i)}(\pi_k^{(i)}(t)) \right]^{1/\alpha_i} \leq r^{2/\alpha_i} \gamma_\alpha(Y, \| \cdot \|_i).$$

(9)

Denote $$\beta_k^{(i)}(t) = \sum_{j=1}^{k+1} \left( \log I_j^{(i)}(\pi_j^{(i)}(t)) \right)^{1/\alpha_i}.$$ Then, (9) implies that

$$\sup_{t \in N_n^{(i)}} \sum_{k=0}^{n-1} r^{-k/\alpha_i} \beta_k^{(i)}(t) \leq r^{3/\alpha_i} \gamma_\alpha(T, \| \cdot \|) \leq 1$$

(10)

for all $n > 0$ and $1 \leq i \leq d$.

Fix $n \geq 0$. For each $1 \leq i \leq d$, consider $N_n^{(i)}$. For each $u_i \in N_n^{(i)}$ and $0 \leq k \leq n$, let $m_i(k, u_i)$ be an integer depending on $k$ and $u_i$, and to be fixed later, such that

$$k \leq m_i(k, u_i) \leq 2k,$$

and

$$m_i(k, u_i) \geq m_i(k-1, u_i) \text{ for } 0 < k \leq n.$$

Denote

$$B_k^{(i)}(u_i) = \left\{ t \in T : \| t - \pi_{m_i(u_i,k)}(u_i) \|_i \leq 2r^{-m_i(k,u_i)/\alpha_i} \right\}.$$ Because

$$B_k^{(i)}(u_i) \supset \left\{ t \in T : \| t - u_i \|_i \leq r^{-m_i(k,u_i)/\alpha_i} \right\} \supset \left\{ t \in T : \| t - u_i \|_i \leq r^{-2n/\alpha_i} \right\},$$
we have \( \bigcup_{u_i \in N_{2n}^{(i)}} B_k^{(i)}(u_i) \) covers \( T \).

For each \( u_i \in N_{2n}^{(i)}, 1 \leq i \leq d \). If the set \( \bigcap_{i \leq d} B_k^{(i)}(u_i) \) is non-empty, we randomly select a point from this set, and denote it by \( \pi_k(u_1, u_2, \ldots, u_d) \). We further assume that if

\[
\bigcap_{i \leq d} B_k^{(i)}(u_i) = \bigcap_{i \leq d} B_k^{(i)}(u'_i)
\]

then

\[
\pi_k(u_1, u_2, \ldots, u_d) = \pi_k(u'_1, u'_2, \ldots, u'_d).
\]

Thus, \( \pi_k(u_1, u_2, \ldots, u_d) \) is determined by \( m_i(k, u_i) \) and \( \pi_{m_i(u_i, k)}(u_i), 1 \leq i \leq d \).

Denote

\[
N_k = \{ \pi_k(u_1, u_2, \ldots, u_d) : u_i \in N_{2n}^{(i)}, 1 \leq i \leq d \}.
\]

Next, we give an ordering to \( N_k \). We define \( I_k \) so that if \( \pi_k(u_1, u_2, \ldots, u_d) \in N_k \) and \( \pi_k(u'_1, u'_2, \ldots, u'_d) \in N_k \) satisfy

\[
\prod_{i=1}^d I_{m_i(k, u_i)}(\pi_{m_i(k, u_i)}(u_i)) > \prod_{i=1}^d I_{m_i(k, u'_i)}(\pi_{m_i(k, u'_i)}(u'_i)),
\]

then

\[
I_k(\pi_k(u_1, u_2, \ldots, u_d)) > I_k(\pi_k(u'_1, u'_2, \ldots, u'_d)).
\]

We prove that under such an ordering, we have

\[
I_k(\pi_k(u_1, u_2, \ldots, u_d)) \leq \left( (k + 1) \prod_{i=1}^d I_{m_i(k, u_i)}(\pi_{m_i(k, u_i)}(u_i)) \right)^d.
\]  \( (11) \)

To see this, let

\[
U = \prod_{i=1}^d I_{m_i(k, u_i)}(\pi_{m_i(k, u_i)}(u_i)),
\]

and suppose

\[
\prod_{i=1}^d I_{m_i(k, u'_i)}(\pi_{m_i(k, u'_i)}(u'_i)) \leq U.
\]  \( (12) \)
For each \( i \leq d \), \( m_i(k, u'_i) \) has no more than \( k + 1 \) choices; and when \( m_i(k, u'_i) \) fixed, \( \pi_{m_i(k, u'_i)}(u'_i) \) has no more than \( U \) choices. Because \( \pi_k(u'_1, u'_2, \ldots, u'_d) \) is determined by \( m_i(k, u'_i) \) and \( \pi_{i}(i) \), \( 1 \leq i \leq d \), we conclude that there are no more than \((k + 1)^dU^d\) different \( \pi_k(u'_1, u'_2, \ldots, u'_d) \) satisfying (12). This implies (11).

Fix \( n \geq 0 \). For each \( t = \pi_n(u_1, u_2, \ldots, u_d) \in N_n \), denote
\[
\pi_k(t) = \pi_k(u_1, u_2, \ldots, u_d) \quad \text{for} \quad 0 \leq k \leq n.
\]

Because we can assume \( N_0(t) \) be a singleton for all \( 1 \leq i \leq d \), we can assume that \( \pi_0(t) \) does not depend on \( t \). We denote \( \pi_0(t) = t_0 \). Now we can write
\[
\left\| \sup_{t \in N_n} |X_t| \right\|_{\psi_{\alpha}} 
\leq \|X_{t_0}\|_{\psi_{\alpha}} + \left\| \sup_{t \in N_n} \sum_{k=0}^{n-1} |X_{\pi_{k+1}(t)} - X_{\pi_k(t)}| \right\|_{\psi_{\alpha}} 
\leq \|X_{t_0}\|_{\psi_{\alpha}} + \sup_{t \in N_n} \sum_{k=0}^{n-1} \varepsilon_k(t) + \left\| \sup_{t \in N_n} \sum_{k=0}^{n-1} |Y_k(t)| \right\|_{\psi_{\alpha}}
\]
where
\[
Y_k(t) = \left( |X_{\pi_{k+1}(t)} - X_{\pi_k(t)}| - \varepsilon_k(t) \right)^+.
\]

To estimate the last term in (13), we use the following lemma.

### 0.4 A Lemma

**Lemma 1** Let \( \psi \) be a Young function satisfying
\[
\psi^{-1}(xy) \leq K\psi^{-1}(x)\psi^{-1}(y)
\]
and
\[
\psi^{-1}(x^2) \leq K\psi^{-1}(x)
\]
for some constant \( K \). Let \( F_k \) be a finite set of random variables of \( \psi \)-norm bounded by \( 2^{-k} \), and let \( J_k \) be an ordering of \( F_k \). Suppose \( S \) is a subset of
\( \left\{ \sum_{k \geq 0} Y_k : Y_k \in F_k, k \geq 0 \right\} \). Then

\[
\left\| \sup_{S} \sum_{k \geq 0} |Y_k| \right\|_{\psi} \leq C_\psi \sup_{S} \sum_{k \geq 0} \psi^{-1}(J_k(Y_k)) \|Y_k\|_{\psi},
\]

where \( C_\psi \) is a constant depending only on \( \psi \).

**Remark 1** The following weaker result was proved in [AG]:

\[
\left\| \sup_{k \geq 1} |Z_k| \right\|_{\psi} \leq C_\psi \sup_{k \geq 1} \|\psi^{-1}(k) \cdot Z_k\|_{\psi}.
\]  \hfill (14)

**Proof:** Without loss of generality, we assume

\[
C_\psi \sup_{S} \sum_{k \geq 0} \psi^{-1}(J_k(Y_k) + k + 1) \|Y_k\|_{\psi} \leq 1,
\]

where \( C_\psi \) is a constant to be determined later. For \( u \geq 1/2 \),

\[
A : = \left\{ \psi \left( \sup_{S} \sum_{k \geq 0} |Y_k| \right) > u \right\} = \bigcup_{S} \left\{ \sum_{k \geq 0} |Y_k| > \psi^{-1}(u) \right\} \subset \bigcup_{S} \bigcup_{k} \left\{ |Y_k| > C_\psi \psi^{-1}(J_k(Y_k) + k + 1) \|Y_k\|_{\psi} \psi^{-1}(u) \right\} = \bigcup_{k} \bigcup_{F_k} \left\{ |Y_k| > C_\psi \psi^{-1}(J_k(Y_k) + k + 1) \|Y_k\|_{\psi} \psi^{-1}(u) \right\}.
\]

Therefore

\[
\Pr(A) \leq \sum_{k \geq 0} \sum_{F_k} \Pr \left( |Y_k| > C_\psi \psi^{-1}(J_k(Y_k) + k + 1) \|Y_k\|_{\psi} \psi^{-1}(u) \right)
\]

\[
\leq \sum_{k \geq 0} \sum_{F_k} \Pr \left( \psi \left( \frac{|Y_k|}{\|Y_k\|_{\psi}} \right) > \psi \left( C_\psi \psi^{-1}(J_k(Y_k) + k + 1) \psi^{-1}(u) \right) \right)
\]

\[
\leq \sum_{k \geq 0} \sum_{F_k} \psi \left( C_\psi \psi^{-1}(J_k(Y_k) + k + 1) \psi^{-1}(u) \right)
\]
\[
\leq \sum_{k \geq 0} \sum_{F_k} \frac{1}{16(J_k(Y_k) + k + 1)^4u^2} \quad \text{(providing } C_\psi \text{ large enough)}
\]
\[
\leq \frac{1}{4u^2}.
\]

Thus,
\[
E_\psi \left( \sup_{k \geq 0} \sum_{Y_k} \right) \leq \frac{1}{2} + \int_{\frac{1}{2}}^{\infty} \frac{1}{4u^2} du = 1
\]
proving the lemma. \hfill \blacksquare

## 0.5 Proof of the Theorem

Now we turn back to (13). For each \( t = \pi_n(u_1, \ldots, u_d) \in N_n \). We choose a small \( \varepsilon_k(t) \), such that \( \|Y_k(t)\|_{\psi_{\alpha_i}} \leq Kr^{-k/\alpha_1} \) for some constant \( K \) depending only on \( \alpha_i \)'s. A proper choice is
\[
\varepsilon_k(t) = \sum_{i=2}^{d} r^{-\left(m_i(k,u_i) - k\right)/(\alpha_i - \alpha_1)}.
\]
(When \( \alpha_i = \alpha_1 \), we omit the corresponding term in the summation.) In fact, under such a choice
\[
\Pr \left( |Y_k(t)| \geq s \right)
\]
\[
\leq \Pr \left( |X_{\pi_{k+1}}(t) - X_{\pi_k}(t)| \geq s + \sum_{i=2}^{d} r^{-\left(m_i(k,u_i) - k\right)/(\alpha_i - \alpha_1)} \right)
\]
\[
\leq \exp \left( -\min \left\{ \sum_{i=1}^{s^{\alpha_1}} \left( \|\pi_{k+1}(t) - \pi_k(t)\|_{\alpha_i} \right)^{\alpha_i}, \sum_{i=2}^{d} r^{-m_i(k,u_i)/\alpha_i} \right\} \right).
\]
Because
\[
\|\pi_k(t) - u_i\|_i \leq \|\pi_k(t) - \pi_{m_i(k,u_i)}(u_i)\|_i + \|\pi_{m_i(k,u_i)}(u_i) - u_i\|_i
\]
\[
\leq 2r^{-m_i(k,u_i)/\alpha_i} + r^{-m_i(k,u_i)/\alpha_i} = 3r^{-m_i(k,u_i)/\alpha_i},
\]
and similarly,
\[
\|\pi_{k+1}(t) - u_i\|_i \leq 3r^{-m_i(k+1,u_i)/\alpha_i} \leq 3r^{-m_i(k,u_i)/\alpha_i},
\]

10
we have
\[ \|\pi_{k+1}(t) - \pi_k(t)\|_i \leq 6r^{-m_i(k,u_i)/\alpha_i}. \]

Therefore
\[
\Pr (|Y_k(t)| \geq s) \leq \exp \left( - \min_{1 \leq i \leq d} \left\{ 6^{-\alpha_i} r^{k} s^{\alpha_1} \right\} \right),
\]
which implies that
\[
\|Y_k(t)\|_{\psi_{\alpha_1}} \leq K r^{-k/\alpha_1}.
\] (15)

Next, we define the ordering of \( Y_k(t) \) so that if \( Y_k(t) \) and \( Y_k(t') \) satisfy
\[
I_{k+1}(\pi_{k+1}(t)) \cdot I_k(\pi_k(t)) > I_{k+1}(\pi_{k+1}(t')) \cdot I_k(\pi_k(t')),
\]
then \( J_k(Y_k(t)) > J_k(Y_k(t')) \). Under such definition, we have
\[
J_k(Y_k(t)) \leq [I_{k+1}(\pi_{k+1}(t)) \cdot I_k(\pi_k(t))]^2.
\] (16)

Applying Lemma 1, and then using (15), (16) and (11), we can bound the last two terms of (13) by
\[
\sum_{i=2}^{d} \sup_{t \in \mathbb{N}_n} \sum_{k=0}^{n-1} r^{-(m_i(k,u_i) - k)/(\alpha_i - \alpha)}
+ K_1 \sup_{t \in \mathbb{N}_n} \sum_{k=0}^{n-1} r^{-k/\alpha_1} (\log J_k(Y_k(t)))^{1/\alpha_1}
\leq \sum_{i=2}^{d} \sup_{t \in \mathbb{N}_n} \sum_{k=0}^{n-1} r^{-(m_i(k,u_i) - k)/(\alpha_i - \alpha)} + K_2 r^{n-1} \sum_{k=0}^{n-1} r^{-k/\alpha_1} [\log(k + 1)]^{1/\alpha_1}
+ K_2 \sup_{t \in \mathbb{N}_n} \sum_{k=0}^{n-1} r^{-k/\alpha_1} \left[ \log I_{m_i(k+1,u_i)}^{(i)}(\pi_{m_i(k+1,u_i)}^{(i)}(u_i)) \right]^{1/\alpha_1}
+ K_2 \sup_{t \in \mathbb{N}_n} \sum_{k=0}^{n-1} r^{-k/\alpha_1} \left[ \log I_{m_1(k+1,u_1)}^{(1)}(\pi_{m_1(k+1,u_1)}^{(1)}(u_1)) \right]^{1/\alpha_1}
\leq \sum_{i=2}^{d} \sup_{u_i \in \mathbb{N}_{2n}^{(i)}} \sum_{k=0}^{n-1} r^{-(m_i(k,u_i) - k)/(\alpha_i - \alpha)} + K_3
+ K_3 \sup_{i=2}^{d} \sup_{u_i \in \mathbb{N}_{2n}^{(i)}} \sum_{k=0}^{n-1} r^{-k/\alpha_1} \left[ \log I_{m_i(k+1,u_i)}^{(i)}(\pi_{m_i(k+1,u_i)}^{(i)}(u_i)) \right]^{1/\alpha_1}.
\[ + K_3 \sup_{u_i \in N^{(1)}_{2n}} \sum_{k=0}^{n-1} r^{-k/\alpha_1} \left[ \log f^{(1)}_{m_1(k+1,u_i)}(\pi^{(1)}_{m_1(k+1,u_i)}(u_i)) \right]^{1/\alpha_1} \]
\[ \leq \sum_{i=2}^{d} \sup_{u_i \in N^{(i)}_{2n}} \sum_{k=0}^{n-1} r^{-(m_i(k,u_i)-k)/(\alpha_i-\alpha_1)} + K_3 \tag{17} \]
\[ + K_3 \sum_{i=2}^{d} \sup_{u_i \in N^{(i)}_{2n}} \sum_{k=0}^{n-1} r^{-k/\alpha_1} \left( \beta^{(i)}_{m_i(k,u_i)}(u_i) \right)^{\alpha_i/\alpha_1} \tag{18} \]
\[ + K_3 \sup_{u_i \in N^{(1)}_{2n}} \sum_{k=0}^{n-1} r^{-k/\alpha_1} \left( \log f^{(1)}_{m_1(k+1,u_i)}(\pi^{(1)}_{m_1(k+1,u_i)}(u_i)) \right)^{1/\alpha_1}. \tag{19} \]

Now it is the time to choose \( m_i(k,u_i) \). We choose \( m_1(k,u_i) = k \). By (9), we can bound (19) by \( K \). For \( 2 \leq i \leq d \), we choose \( m_i(k,u_i) \) so as to “minimize”
\[ r^{-(m_i(k,u_i)-k)/(\alpha_i-\alpha_1)} + r^{-k/\alpha_1} \left( \beta^{(i)}_{m_i(k,u_i)}(u_i) \right)^{\alpha_i/\alpha_1}, \]
and keep the conditions \( k \leq m_i(k,u_i) \leq 2k \) and \( m_i(k,u_i) \geq m_i(k-1,u_i) \) in mind. A proper choice is to let \( m_i(k,u_i) \) be the largest integer, less than or equal to \( 2k \), such that for \( j \leq m_i(k,u_i) \),
\[ r^{-k/\alpha_1} \left( \beta^{(i)}_{j}(u_i) \right)^{\alpha_i/\alpha_1} \leq r^{-(j-k)/(\alpha_i-\alpha_1)}, \]
or equivalently,
\[ r^{-j/\alpha_i} \beta^{(i)}_{j}(u_i) \leq r^{-(j-k)/(\alpha_i-\alpha_1)}. \]

Thus, (17) + (18) can be bounded by
\[ \sum_{i=2}^{d} \sup_{u_i \in N^{(i)}_{2n}} \sum_{k=0}^{n-1} r^{-(m_i(k,u_i)-k)/(\alpha_i-\alpha_1)} \leq K \sum_{i=2}^{d} \sup_{u_i \in N^{(i)}_{2n}} \sum_{l=0}^{2n-2} \sum_{k \in D_i^{(i)}} r^{-(l-k)/(\alpha_i-\alpha_1)}, \tag{20} \]
where \( D_i^{(i)} = \{ k : m_i(k,u_i) = l \} \). Let \( k_0 = \min \{ k : k \in D_i^{(i)} \} \), we have
\[ \sum_{k \in D_i^{(i)}} r^{-(l-k)/(\alpha_i-\alpha_1)} \leq K r^{-(m_i(k_0,u_i)-k_0)/(\alpha_i-\alpha_1)}. \]
On the other hand,

\[
\begin{align*}
r^{-(m_i(k_0,u_i)+1-k_0)/(\alpha_i-\alpha_1)} &< r^{-(m_i(k_0,u_i)+1)/(\alpha_i-\alpha_1)} \beta^{(i)}_{m_i(k_0,u_i)+1}(u_i) + r^{-k_0/(\alpha_i-\alpha_1)} \\
&\leq r^{-(l+1)/(\alpha_i-\alpha_1)} \beta^{(i)}_{l+1}(u_i) + r^{-l/2(\alpha_i-\alpha_1)},
\end{align*}
\]

Thus, we can bound (20) by

\[
\sum_{i=2}^{d} \sup_{u_i \in N^{(i)}_{2n}} r^{1/(\alpha_i-\alpha_1)} \sum_{l=0}^{2n-2} \left[ r^{-(l+1)/(\alpha_i-\alpha_1)} \beta^{(i)}_{l+1}(u_i) + r^{-l/2(\alpha_i-\alpha_1)} \right].
\]

The theorem follows by applying (10).

**Acknowledgment** This work is mainly done under the supervise of Professor Ron Blei.

## 0.6 References


[M]


