Gaussian processes, Gaussian comparison inequalities

Def: $\{X_t\}_{t\in T}$ is a Gaussian process if, for any finite subset $T_0 \subseteq T$, $\{X_t\}_{t\in T_0}$ has a multivariate Gaussian law.

Its covariance function is $\Sigma(t,s) = \mathrm{cov}(X_t, X_s)$.

We will assume $\{X_t\}_{t\in T}$ is mean-0, so the law of $\{X_t\}_{t\in T_0}$ for any finite $T_0 \subseteq T$ is uniquely determined by $\Sigma(t,s)$.

Goal: Upper and lower bounds on $\sup_{t\in T} X_t$ using additional tools of Gaussian comparison inequalities.

Comparison inequalities

Thm (Sudakov-Fernique): If $\{X_t\}_{t\in T}$, $\{Y_t\}_{t\in T}$ are mean-zero, separable Gaussian processes such that

$$\mathbb{E}(X_t - X_s)^2 \leq \mathbb{E}(Y_t - Y_s)^2 \quad \forall s,t\in T$$

then

$$\mathbb{E}\sup_{t\in T} X_t \leq \mathbb{E}\sup_{t\in T} Y_t.$$

Thm (Slepian): If, in addition to the above, $\mathbb{E}X_t^2 = \mathbb{E}Y_t^2$ $\forall t\in T$, then for every $\tau\in\mathbb{R}$,

$$\mathbb{P}\left[\sup_{t\in T} X_t \geq \tau\right] \leq \mathbb{P}\left[\sup_{t\in T} Y_t \geq \tau\right].$$

To prove these comparison results, we use an interpolation argument + Gaussian integration-by-parts.

Lemma (integration by parts): If $X \sim \mathcal{N}(0,\Sigma) \in \mathbb{R}^n$, $f: \mathbb{R}^n \to \mathbb{R}$ cont. differentiable, $\mathbb{E}|X_i f(x)| < \infty$, $\mathbb{E}|\partial_i f(x)| < \infty$ $\forall i$ then

$$\mathbb{E}\, X f(X) = \Sigma \cdot \mathbb{E}\, \nabla f(X)$$

[For $n=1$ and $\Sigma = 1$: $\mathbb{E}\, X f(X) = \mathbb{E} f'(X)$.]

Proof: For $n=1$: If $X \sim \mathcal{N}(0,1)$, f compactly supported, then

$$\mathbb{E}\, X f(X) = \int x f(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int f'(x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \mathbb{E} f'(X).$$

For general f, approximate by sequence of compactly supported f.

For general n and Σ: Let $X = \Sigma^{1/2} Z$, $Z \sim \mathcal{N}(0, I)$.

Applying result for $n=1$:

$$\begin{aligned}
\mathbb{E}\, X f(X) &= \Sigma^{1/2}\, \mathbb{E}\, Z f(\Sigma^{1/2} Z) \\
&= \Sigma^{1/2} \left(\mathbb{E}\, Z_k f(\Sigma^{1/2} Z)\right)_{k=1}^{n} \\
&= \Sigma^{1/2} \left(e_k^T \Sigma^{1/2}\, \mathbb{E}\, \nabla f(\Sigma^{1/2} Z)\right)_{k=1}^{n} = \Sigma \cdot \mathbb{E}\, \nabla f(X).
\end{aligned}$$

∎

Lemma (Gaussian interpolation): Let $X \sim \mathcal{N}(0, \Sigma^X)$, $Y \sim \mathcal{N}(0, \Sigma^Y)$, in $\mathbb{R}^n$, $f: \mathbb{R}^n \to \mathbb{R}$ twice cont. differentiable. Define

$$Z(u) = \sqrt{u} \cdot X + \sqrt{1-u}\, Y, \qquad u \in [0,1].$$

Then $$\frac{d}{du} \mathbb{E} f(Z(u)) = \frac{1}{2} \sum_{i,j=1}^n \left(\Sigma^X_{ij} - \Sigma^Y_{ij} \right) \cdot \mathbb{E}\left[\partial_{ij} f(Z(u)) \right].$$

Proof: $\frac{d}{du} \mathbb{E} f(Z(u))$

$$= \sum_{i=1}^n \mathbb{E}\, \partial_i f(Z(u)) \cdot \frac{d}{du} Z(u)_i$$

$$= \sum_{i=1}^n \mathbb{E}\, \partial_i f(Z(u)) \cdot \left(\frac{1}{2\sqrt{u}} X_i - \frac{1}{2\sqrt{1-u}} Y_i \right)$$

$$= \sum_{i=1}^n \frac{1}{2\sqrt{u}} e_i^T \Sigma^X \cdot \mathbb{E} \nabla_X f(Z(u)) - \frac{1}{2\sqrt{1-u}} e_i^T \Sigma^Y \cdot \mathbb{E} \nabla_Y f(Z(u))$$

(by integration-by-parts lemma)

$$= \sum_{i=1}^n \frac{1}{2\sqrt{u}} \sum_{j=1}^n (\Sigma^X)_{ij} \sqrt{u}\, \mathbb{E} \partial_{ij} f(Z(u))$$
$$- \sum_{i=1}^n \frac{1}{2\sqrt{1-u}} \sum_{j=1}^n (\Sigma^Y)_{ij} \sqrt{1-u}\, \mathbb{E} \partial_{ij} f(Z(u))$$

$$= \frac{1}{2} \sum_{i,j=1}^n \left(\Sigma^X_{ij} - \Sigma^Y_{ij} \right) \cdot \mathbb{E}\left[\partial_{ij} f(Z(u)) \right].$$

$\blacksquare$

Proof of Sudakov-Fernique inequality: Suppose first $|T| < \infty$, i.e. $X \sim \mathcal{N}(0, \Sigma^X)$, $Y \sim \mathcal{N}(0, \Sigma^Y)$ in dimension $|T| = n$.

Fix any $\beta > 0$ and consider $f_\beta : \mathbb{R}^n \to \mathbb{R}$ given by

$$f_\beta(x) = \frac{1}{\beta} \log \sum_{i=1}^{n} e^{\beta x_i}$$

(For large β, this is a smooth approximation of $\max_{i=1}^{n} x_i$)

$$\Rightarrow \partial_i f_\beta(x) = \frac{e^{\beta x_i}}{\sum_k e^{\beta x_k}}, \quad \partial_{ij} f(x) = \beta \left[\underbrace{\frac{e^{\beta x_i}}{\sum_k e^{\beta x_k}}}_{:= p_i(x)} \mathbb{1}\{i = j\} - \underbrace{\frac{e^{\beta x_i} e^{\beta x_j}}{\left(\sum_k e^{\beta x_k}\right)^2}}_{:= p_i(x) p_j(x)} \right]$$

$$\Rightarrow \frac{d}{du} \mathbb{E} f_\beta(Z(u)) = \frac{\beta}{2} \sum_{i=1}^{n} \left(\Sigma^X_{ii} - \Sigma^Y_{ii}\right) \mathbb{E}\, p_i(Z(u)) \Big(\underbrace{1 - p_i(Z(u))}_{= \sum_{j \neq i} p_j(Z(u))}\Big)$$

$$- \frac{\beta}{2} \sum_{i \neq j} \left(\Sigma^X_{ij} - \Sigma^Y_{ij}\right) \mathbb{E}\, p_i(Z(u)) p_j(Z(u))$$

$$= \frac{\beta}{2} \sum_{i \neq j} \left(\Sigma^X_{ii} - \Sigma^Y_{ii} - \Sigma^X_{ij} + \Sigma^Y_{ij}\right) \mathbb{E}\, p_i(Z(u)) p_j(Z(u))$$

$$= \frac{\beta}{4} \sum_{i \neq j} \left(\Sigma^X_{ii} + \Sigma^X_{jj} - \Sigma^Y_{ii} - \Sigma^Y_{jj} - 2\Sigma^X_{ij} + 2\Sigma^Y_{ij}\right) \mathbb{E}\, p_i(Z(u)) p_j(Z(u))$$

$$= \frac{\beta}{4} \sum_{i \neq j} \Big[\underbrace{\mathbb{E}(X_i - X_j)^2 - \mathbb{E}(Y_i - Y_j)^2}_{\leq 0 \ \forall i,j \text{ by assumption}} \Big] \cdot \mathbb{E}\, p_i(Z(u)) p_j(Z(u)) \leq 0.$$

$$\Rightarrow \mathbb{E} f_\beta(X) = \mathbb{E} f_\beta(Z(1)) \leq \mathbb{E} f_\beta(Z(0)) = \mathbb{E} f_\beta(Y) \qquad (**)$$

Note $F_\beta(x) \geq \max_{i=1}^n x_i$, $F_\beta(x) \leq \frac{\log n}{\beta} + \max_{i=1}^n x_i$.

Take $\beta \nearrow \infty$ on both sides of $(**)$ using dominated convergence thm:

$$\mathbb{E} \max_{i=1}^n X_i \leq \mathbb{E} \max_{i=1}^n Y_i.$$

For $|T| = \infty$, by separability there exist $t_1, t_2, t_3, \ldots \in T$ s.t.

$$\mathbb{E} \sup_{t \in T} X_t = \mathbb{E} \sup_{t \in \{t_1, t_2, \ldots\}} X_t = \sup_{k \geq 1} \mathbb{E} \sup_{t \in \{t_1, \ldots, t_k\}} X_t.$$

Apply result for each fixed k. $\blacksquare$

Proof of Slepian inequality: Again suppose $|T| = n < \infty$, $X \sim N(0, \Sigma^X)$, $Y \sim N(0, \Sigma^Y)$. Fix $\tau \in \mathbb{R}$, let $h_\beta : \mathbb{R} \to [0,1]$ be smooth, decreasing functions s.t. $h_\beta(x) \to \mathbb{1}\{x < \tau\}$ as $\beta \to \infty$. Consider

$$F_\beta(x) = \prod_{i=1}^n h_\beta(x_i).$$

$$\Rightarrow \partial_{ij} F_\beta(x) = h_\beta'(x_i) h_\beta'(x_j) \prod_{k \neq i,j} h_\beta(x_k) \geq 0 \quad \forall i \neq j.$$

$$\Rightarrow \frac{d}{du} \mathbb{E} F_\beta(Z(u)) = \frac{1}{2} \sum_{i,j=1}^n \underbrace{(\Sigma^X_{ij} - \Sigma^Y_{ij})}_{=0 \text{ if } i=j,\ \geq 0 \text{ if } i \neq j \text{ by assumption}} \mathbb{E} \partial_{ij} F_\beta(Z(u)) \geq 0$$

$$\Rightarrow \mathbb{E} F_\beta(X) = \mathbb{E} F_\beta(Z(1)) \geq \mathbb{E} F_\beta(Z(0)) = \mathbb{E} F_\beta(Y).$$

Take $\beta \to \infty$: $f_\beta(x) \to \mathbb{1}\{\sup_{i=1}^n x_i < \tau\}$

$$\Rightarrow \mathbb{P}\left[\sup_{i=1}^n X_i < \tau\right] \geq \mathbb{P}\left[\sup_{i=1}^n Y_i < \tau\right],$$

i.e. $$\mathbb{P}\left[\sup_{i=1}^n X_i \geq \tau\right] \leq \mathbb{P}\left[\sup_{i=1}^n Y_i \geq \tau\right].$$

If $|T| = \infty$, apply separability of T as in preceding proof. ■

An illustration/application of these inequalities is the following:

Thm: Let $X \in \mathbb{R}^{n \times m}$ have iid $\mathcal{N}(0,1)$ entries. Then

$$\mathbb{E}\|X\|_{op} \leq \sqrt{n} + \sqrt{m}.$$

[Compare from Lecture 7: If X has iid 1-subgaussian entries, then

$\mathbb{E}\|X\|_{op} \leq \sqrt{n} + C\sqrt{m}$ for a universal constant $C > 0$.]

Proof: $$\|X\|_{op} = \sup_{t \in S^{n-1}, u \in S^{m-1}} \underbrace{t^T X u}_{:= X_{ut}}.$$

We have $$\mathbb{E}(X_{ut} - X_{vs})^2 = \mathbb{E}\Big(\sum_{ij} X_{ij}(t_i u_j - s_i v_j)\Big)^2$$

$$= \sum_{ij}(t_i u_j - s_i v_j)^2 = \|tu^T - sv^T\|_F^2 = 2 - 2 s^T t \cdot u^T v.$$

Define $Y_{ut} = g^T t + h^T u$ where $g \sim \mathcal{N}(0, I_n)$, $h \sim \mathcal{N}(0, I_m)$ independent.

Then $$\mathbb{E}(Y_{ut} - Y_{vs})^2 = \mathbb{E}\big(g^T(t-s) + h^T(u-v)\big)^2$$

$= \|t-s\|_2^2 + \|u-v\|_2^2 = 4 - s^T t - u^T v.$

Observe $(4 - s^T t - u^T v) - (2 - 2 s^T t \cdot u^T v) = 2(1 - s^T t)(1 - u^T v) \geq 0$

$$\Rightarrow \mathbb{E}(X_{ut} - X_{vs})^2 \leq \mathbb{E}(Y_{ut} - Y_{vs})^2 \quad \forall u, v \in S^{m-1},\ t, s \in S^{n-1}.$$

By Sudakov-Fernique,

$$\mathbb{E}\|X\|_{op} = \mathbb{E} \sup_{u \in S^{n-1},\, t \in S^{m-1}} X_{ut}$$

$$\leq \mathbb{E} \sup_{u \in S^{n-1},\, t \in S^{m-1}} Y_{ut} = \mathbb{E} \sup_{t \in S^{m-1}} g^T t + \mathbb{E} \sup_{u \in S^{n-1}} h^T u$$

$$= \mathbb{E}\|g\|_2 + \mathbb{E}\|h\|_2 \leq \sqrt{n} + \sqrt{m}. \quad \blacksquare$$

Gaussian process lower bounds

Def: The canonical metric (or natural distance) associated to a mean-0 Gaussian process $\{X_t\}_{t \in T}$ is $d(t,s) = \left(\mathbb{E}(X_t - X_s)^2\right)^{1/2}$.

Remarks: ① $\{X_t\}_{t \in T}$ is subgaussian wrt its canonical metric, because

$$\mathbb{E} e^{\lambda(X_t - X_s)} = e^{\frac{\lambda}{2}\mathbb{E}(X_t - X_s)^2} = e^{\frac{\lambda}{2} d(t,s)^2}$$

② For the process $X_t = g^T t$, $g \sim N(0, I)$, $T \subseteq \mathbb{R}^n$, we have $\mathbb{E}(X_t - X_s)^2 = \|t-s\|_2^2$ so $d(t,s)$ is the usual Euclidean distance on T.

③ Let $X_1, \dots, X_n \overset{iid}{\sim} P$, $\mathcal{F}$ a function class s.t. $\mathbb{E}_P f(X) = 0 \ \forall f \in \mathcal{F}$, Consider the (renormalized) empirical process

$$Z_f = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} f(X_i), \quad f \in \mathcal{F}.$$

By CLT, the finite-dimensional marginals of $\{Z_f\}_{f \in \mathcal{F}}$ converge to a Gaussian process w/ $\Sigma(f,g) = \mathrm{Cov}_P[f(X), g(X)]$.

Its canonical metric is

$$d(f,g) = \sqrt{\Sigma(f,f) + \Sigma(g,g) - 2\Sigma(f,g)} = \sqrt{\mathrm{Var}_P[f-g]} = \|f-g\|_{L^2(P)}.$$

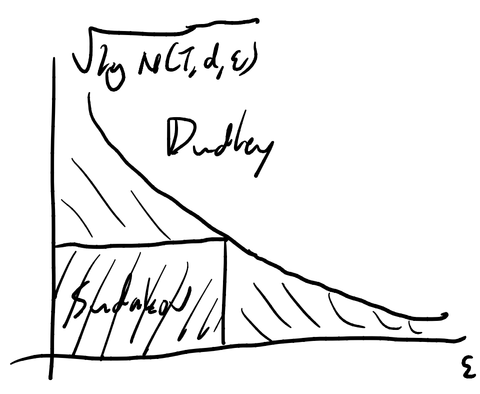
By Dudley's inequality,

$$\mathbb{E} \sup_{t \in T} X_t \leq C \cdot \int_0^\infty \sqrt{\log N(T,d,\varepsilon)}\, d\varepsilon$$

Using the Sudakov-Fernique inequality, we will derive corresponding lower bounds, to understand when Dudley's inequality is tight.

Thm (Sudakov): Let $\{X_t\}_{t \in T}$ be a mean-zero Gaussian process, $d(t,s)$ the canonical metric. Then for a universal constant $c > 0$,

$$\mathbb{E} \sup_{t \in T} X_t \geq c \cdot \sup_{\varepsilon > 0} \varepsilon \sqrt{\log N(T,d,\varepsilon)}$$



Proof: Let $\mathcal{D}$ be a maximum-size ε-packing, so

$$|\mathcal{D}| = D(T, d, \varepsilon) \geq N(T, d, \varepsilon).$$

Idea: Compare $\sup_{t \in \mathcal{D}} X_t$ with the maximum of iid Gaussians.

$$\mathbb{E}(X_t - X_s)^2 = d(t,s)^2 \geq \varepsilon^2 \quad \forall s \neq t \in \mathcal{D}.$$

Let $\{Y_t\}_{t \in \mathcal{D}}$ be iid $\mathcal{N}(0, \frac{\varepsilon^2}{2}) \Rightarrow \mathbb{E}(Y_t - Y_s)^2 = \varepsilon^2 \; \forall s \neq t \in \mathcal{D}$.

By Sudakov-Fernique,

$$\mathbb{E} \sup_{t \in T} X_t \geq \mathbb{E} \sup_{t \in \mathcal{D}} X_t \geq \mathbb{E} \sup_{t \in \mathcal{D}} Y_t \geq c\varepsilon \sqrt{\log |\mathcal{D}|}.$$

$\blacksquare$

Example: $X_t = g^T t$, $t \in T \subseteq \mathbb{R}^n$, $g \sim \mathcal{N}(0, I)$ in $\mathbb{R}^n$, $d(t,s) = \|t-s\|_2$.

① Consider $T = B^n$, the unit ball. $N(T, d, \varepsilon) \in \left[\left(\frac{1}{\varepsilon}\right)^n, \left(\frac{3}{\varepsilon}\right)^n\right]$

So $$\sup_{\varepsilon \in (0,1)} \varepsilon \sqrt{\log N(T,d,\varepsilon)} \asymp \sqrt{n} \cdot \sup_{\varepsilon \in (0,1)} \varepsilon \sqrt{\log 1/\varepsilon} \asymp \sqrt{n}.$$

Also $$\int_0^1 \sqrt{\log N(T,d,\varepsilon)}\, d\varepsilon \asymp \sqrt{n} \cdot \int_0^1 \sqrt{\log 1/\varepsilon}\, d\varepsilon \asymp \sqrt{n}.$$

② Consider $T = \left\{ \frac{e_k}{\sqrt{1 + \log k}} \right\}_{k=1}^n$. Here

$$M(n) := \int_0^\infty \sqrt{\log N(T,d,\varepsilon)}\, d\varepsilon \to 0 \text{ as } n \to \infty \quad \text{(Homework 8)}.$$

However, $$\sup_{\varepsilon > 0} \varepsilon \sqrt{\log N(T,d,\varepsilon)} \asymp \sup_{\varepsilon > 0} \varepsilon \cdot \frac{1}{\varepsilon} \asymp 1.$$

(In this case, the lower bound is tight, upper bound is not.)

Def: A Gaussian process $\{X_t\}_{t\in T}$ is <u>stationary</u> if there is a group G acting on T s.t.

- For any $g \in G$, $\{X_t\}_{t\in T} \overset{L}{=} \{X_{g\cdot t}\}_{t\in T}$
- For any $t, s \in T$, there exists $g \in G$ s.t. $g \cdot t = s$.

Example: $T = S^{n-1} = \{t \in \mathbb{R}^n : \|t\|_2 = 1\}$, G = group of rotations of T.

$\{X_t\}_{t\in T}$ (mean-zero) is stationary if $\Sigma(t,s)$ depends only on the spherical distance between t and s.

Thm (Fernique): Let $\{X_t\}_{t\in T}$ be a separable, mean-zero, <u>stationary</u> Gaussian process. Then for a universal constant $c > 0$,

$$\mathbb{E} \sup_{t\in T} X_t \geq c \int_0^\infty \sqrt{\log N(T, d, \varepsilon)}\, d\varepsilon.$$

Lemma: If $\{X_t\}_{t\in T}$ is a separable Gaussian process, then

$\sup_{t\in T} X_t - \mathbb{E} \sup_{t\in T} X_t$ is $\sup_{t\in T} \mathrm{Var}[X_t]$-subgaussian.

Proof: By separability, suffices to consider $|T| \equiv n$ finite.

Write $X = \Sigma^{1/2} Z$, $Z \sim N(0, I)$ in $\mathbb{R}^n$.

Then each X_i is Lipschitz in Z, w/ Lipschitz constant

$$\|e_i^T \Sigma^{1/2}\|_2 = \sqrt{\sum_j (\Sigma^{1/2})_{ij}^2} = \sqrt{(\Sigma^{1/2}\cdot\Sigma^{1/2})_{ii}} = \sqrt{\operatorname{Var} X_i}.$$

$\Rightarrow \sup_{i=1}^n X_i$ is $\sqrt{\sup_{i=1}^n \operatorname{Var} X_i}$-Lipschitz. Apply Gaussian concentration. ∎

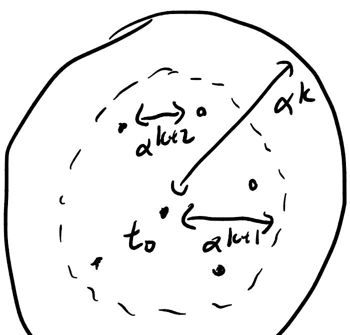
Proof: Idea: chaining + packing to construct multi-scale lower bound.

Fix any $t_0 \in T$. Let $B(t,\varepsilon) = \{s \in T : d(s,t) \le \varepsilon\}$, $B(\varepsilon) \equiv B(t_0, \varepsilon)$.

Fix any $\alpha \in (0, 1/2)$ and consider $G(k) := \mathbb{E} \sup_{t \in B(\alpha^k)} X_t$.

Let $\mathcal{D}_k$ be a maximal α^{k+2}-packing of $B(\alpha^{k+1})$.

Then $\{B(s, \alpha^{k+3}) : s \in \mathcal{D}_k\}$ are disjoint balls in $B(\alpha^k)$.



$$G(k) \ge \mathbb{E} \sup_{s \in \mathcal{D}_k} \sup_{t \in B(s,\alpha^{k+3})} X_t$$

$$= \mathbb{E}\Big[\sup_{s \in \mathcal{D}_k} X_s + \underbrace{\sup_{t \in B(s,\alpha^{k+3})} X_t - X_s}_{:= Y_s}\Big].$$

$$= \mathbb{E}\Big[\sup_{s \in \mathcal{D}_k} X_s + (Y_s - \mathbb{E}Y_s) + \mathbb{E}Y_s\Big]$$

- $\mathbb{E} \sup_{s \in \mathcal{D}_k} X_s \ge c\,\alpha^{k+2} \sqrt{\log |\mathcal{D}_k|}$ by Sudakov lower bound.
- $\mathbb{E} Y_s = \mathbb{E} \sup_{t \in B(s,\alpha^{k+3})} X_t = G(k+3)$, same for all s by stationarity.

• Here $\mathrm{Var}(X_t - X_s) = d(t,s)^2 \le (\alpha^{k+3})^2$ so $Y_s - \mathbb{E}Y_s$ is $(\alpha^{k+3})^2$-subgaussian by previous lemma, for any $s \in \mathcal{D}_k$.

$\Rightarrow \mathbb{E} \sup_{s \in \mathcal{D}_k} |Y_s - \mathbb{E}Y_s| \le C\alpha^{k+3}\sqrt{\log |\mathcal{D}_k|}$ by maximal inequality

Combining the above, for $\alpha \in (0, 1/2)$ small enough and some $c' > 0$,

$$G(k) \ge c'\alpha^{k+2}\sqrt{\log |\mathcal{D}_k|} + G(k+3)$$

$$\ge c'\alpha^{k+2}\sqrt{\log N(B(\alpha^{k+1}), d, \alpha^{k+2})} + G(k+3).$$

Summing over $k, k+1, k+2$ and iterating this bound,

$$G(k) + G(k+1) + G(k+2) \ge c' \sum_{j \ge k+1} \alpha^{j+1}\sqrt{\log N(B(\alpha^j), d, \alpha^{j+1})}$$

There exists $K \in \mathbb{Z}$ s.t. $\alpha^{K+2} \ge \mathrm{diam}(T)$. (Otherwise $\mathbb{E} \sup_{t \in T} X_t = \infty$ by Sudakov lower bound, so the theorem is trivial.)

$\Rightarrow G(K), G(K+1), G(K+2) = \sup_{t \in T} X_t$, and

$\log N(B(\alpha^j), d, \alpha^{j+1}) = 0$ for all $j \le K$.

$$\Rightarrow \mathbb{E} \sup_{t \in T} X_t \ge \frac{c'}{3} \sum_{j \in \mathbb{Z}} \alpha^{j+1}\sqrt{\log N(B(\alpha^j), d, \alpha^{j+1})}$$

Finally, apply

$$N(T, d, \alpha^{k+1}) \le N(T, d, \alpha^k) \cdot N(B(\alpha^k), d, \alpha^{k+1})$$

and $\sqrt{a-b} \ge \sqrt{a} - \sqrt{b}$ for $a \ge b \ge 0$:

$$\mathbb{E} \sup_{t \in T} X_t \ge \frac{c'}{3} \sum_{j \in \mathbb{Z}} \alpha^{j+1} \left(\sqrt{\lg N(T, d, \alpha^{j+1})} - \sqrt{\lg N(T, d, \alpha^j)} \right)$$

$$= \frac{c'}{3}(1-\alpha) \sum_{j \in \mathbb{Z}} \alpha^j \sqrt{\lg N(T, d, \alpha^j)}$$

$$\ge c'' \int_0^\infty \sqrt{\log N(T, d, \varepsilon)}\, d\varepsilon.$$

$\blacksquare$