Metric Entropy

Jiuzhou Miao

April 9, 2025

Proof of Proposition 4. We first prove that $\mathcal{N}(\delta, \mathbb{T}, \varrho) \leq \mathcal{M}(\delta, \mathbb{T}, \varrho)$. Denote $\mathcal{M}(\delta, \mathbb{T}, \varrho) = m$. Let $\{t_1, \ldots, t_m\}$ be a δ-packing of $\mathbb{T}$. We claim that $\{t_1, \ldots, t_m\}$ can form a δ-cover of $\mathbb{T}$, i.e., for each $t \in \mathbb{T} \backslash \{t_1, \ldots, t_m\}$, there exist some $i \in \{1, \ldots, m\}$ such that $\varrho(t, t_i) \leq \delta$. Assume that this argument failed, i.e., there exist $t \in \mathbb{T} \backslash \{t_1, \ldots, t_m\}$ such that $\varrho(t, t_i) > \delta$ for each $i \in \{1, \ldots, m\}$. Then $\{t_1, \ldots, t_m\} \cup \{t\}$ can form a new δ-packing of $\mathbb{T}$, which is contradictory to the packing number of $\mathbb{T}$ is m. Note that the covering number $\mathcal{N}(\delta, \mathbb{T}, \varrho)$ is the cardinality of the smallest δ-cover, then this inequality holds.

We then prove that $m' = \mathcal{M}(2\delta, \mathbb{T}, \varrho) \leq \mathcal{N}(\delta, \mathbb{T}, \varrho) = n$. Assume that $m' \geq n+1$. Let $\{y_1, \ldots, y_{m'}\}$ and $\{x_1, \ldots, x_n\}$ be the 2δ-packing and δ-cover of $\mathbb{T}$ respectively. By the definition of δ-cover, one has that each y_i belongs to a closed ball centered at some x_k with radius δ. By pigeonhole, one has that there exist some y_i and y_j which belong to a closed ball centered at some x_k with radius δ, i.e., $\varrho(y_i, y_j) \leq 2\delta$, which is contradictory to that $\{y_1, \ldots, y_{m'}\}$ is the 2δ-packing of $\mathbb{T}$. Then this inequality holds. $\square$

Proof of Proposition 5. We first prove the first inequality. Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ be a δ-cover of B, then one has that

$$B \subseteq \bigcup_{i=1}^{N} \{\boldsymbol{x}_i + \delta B'\},$$

which implies that

$$\mathrm{Vol}(B) \leq \sum_{i=1}^{N} \mathrm{Vol}(x_i + \delta B') = N\delta^d \mathrm{Vol}(B').$$

Then the first inequality holds.

We then prove the second inequality. Let $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\}$ be a maximal $\delta/2$-packing of B in the $\|\cdot\|'$-norm. We claim that $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_M\}$ can form a δ-cover of B in the $\|\cdot\|'$-norm. The balls $\{\boldsymbol{y}_i + 2\delta^{-1}B'\}$ are disjoint and contained within $B + 2\delta^{-1}B'$. Then one has that

$$\begin{aligned}
(\delta/2)^d \mathrm{Vol}(2^{-1}\delta B + B') &= \mathrm{Vol}(B + 2\delta^{-1}B') \\
&\geq \sum_{i=1}^{M} \mathrm{Vol}(\boldsymbol{y}_i + 2\delta^{-1}B') \\
&= M(\delta/2)^d \mathrm{Vol}(B').
\end{aligned}$$

Then the second inequality holds. $\square$

Proof of Proposition 6. Let

$$\tilde{\varepsilon}_i = \begin{cases} 1, & Z_i \geq 0 \\ -1, & Z_i < 0 \end{cases}.$$

Then one has that $\{\tilde{\varepsilon}_i\}_{i=1}^d$ are independent Rademacher random variables and are independent with $\{|Z_i|\}_{i=1}^d$ and $Z_i \overset{d}{=} |Z_i|\tilde{\varepsilon}_i$. Note that

$$\mathbb{E}\Big(\sum_{i=1}^d t_i |Z_i| \tilde{\varepsilon}_i \Big| \tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_n\Big) \overset{a.s.}{=} \sum_{i=1}^d t_i \tilde{\varepsilon}_i \mathbb{E}\big(|Z_i| \big| \tilde{\varepsilon}_i\big) \overset{d}{=} \sqrt{2/\pi} \sum_{i=1}^d t_i \varepsilon_i,$$

one has that

$$\sup_{\boldsymbol{t}\in\mathbb{T}} \sum_{i=1}^d t_i \varepsilon_i \overset{d}{=} \sqrt{\pi/2} \sup_{\boldsymbol{t}\in\mathbb{T}} \mathbb{E}\Big(\sum_{i=1}^d t_i |Z_i| \tilde{\varepsilon}_i \Big| \tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_n\Big).$$

By using Jensen's inequality, one has that

$$\begin{aligned}
\mathcal{R}(\mathbb{T}) &= \mathbb{E}\Big\{\sup_{\boldsymbol{t}\in\mathbb{T}} \sum_{i=1}^d t_i \varepsilon_i\Big\} \\
&= \sqrt{\pi/2}\mathbb{E}\Big\{\sup_{\boldsymbol{t}\in\mathbb{T}} \mathbb{E}\Big(\sum_{i=1}^d t_i |Z_i| \tilde{\varepsilon}_i \Big| \tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_n\Big)\Big\} \\
&\leq \sqrt{\pi/2}\mathbb{E}\Big\{\mathbb{E}\Big(\sup_{\boldsymbol{t}\in\mathbb{T}} \sum_{i=1}^d t_i |Z_i| \tilde{\varepsilon}_i \Big| \tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_n\Big)\Big\} \\
&= \sqrt{\pi/2}\mathbb{E}\Big(\sup_{\boldsymbol{t}\in\mathbb{T}} \sum_{i=1}^d t_i |Z_i| \tilde{\varepsilon}_i\Big) \\
&= \sqrt{\pi/2}\mathcal{G}(\mathbb{T}).
\end{aligned}$$

□

Lemma 1 (Upper bounds for sub-Gaussian maxima). *Let $\{X_i\}_{i=1}^n$ be a sequence of zero mean sub-Gaussian random variables with parameter σ. Then one has that for $n \geq 2$*

$$\mathbb{E}\big(\max_{1\leq i\leq n} |X_i|\big) \leq 2\sqrt{\sigma^2 \log n}.$$

Proof. Note that $\max_{1\leq i\leq n} |X_i| = \max_{1\leq i\leq 2n} X_i'$, where $X_i' = X_i$ for $1 \leq i \leq n$ and $X_i' = -X_i$ for $n+1 \leq i \leq 2n$, we only to bound $\mathbb{E}\big(\max_{1\leq i\leq n} X_i\big)$. By Jensen's inequality, one has that for all $\lambda \geq 0$,

$$\begin{aligned}
\exp\Big\{\lambda \mathbb{E}\big(\max_{1\leq i\leq n} X_i\big)\Big\} &\leq \mathbb{E}(e^{\lambda \max_{1\leq i\leq n} X_i}) \\
&= \mathbb{E}\big(\max_{1\leq i\leq n} e^{\lambda X_i}\big) \\
&\leq n e^{\sigma^2\lambda^2/2}.
\end{aligned}$$

By taking logarithm, one has that

$$\mathbb{E}\big(\max_{1\leq i\leq n} X_i\big) \leq \inf_{\lambda\geq 0}\{\lambda^{-1}\log n + 2^{-1}\lambda\sigma^2\} = \sqrt{2\sigma^2 \log n}.$$

Then one has that

$$\mathbb{E}\big(\max_{1\le i\le n}|X_i|\big) = \mathbb{E}\big(\max_{1\le i\le 2n} X_i'\big) \le \sqrt{2\sigma^2\log 2n} \le 2\sqrt{\sigma^2\log n}$$

for all $n \ge 2$. $\square$

Proof of Proposition 8. For a given $\delta \ge 0$ and covering number $N = \mathcal{N}(\delta, \mathbb{T}, \varrho_X)$, let $\{t_1, \dots, t_N\}$ be a δ-cover of $\mathbb{T}$. For any $t \in \mathbb{T}$, one can find some t_i such that $\varrho_X(t, t_i) \le \delta$. Note that

$$\begin{aligned} X_t - X_{t'} &= X_t - X_{t_i} + X_{t_i} - X_{t_1} + X_{t_1} - X_{t_i} + X_{t_i} - X_{t'} \\ &\le 2\sup_{\gamma,\gamma'\in\mathbb{T},\varrho_X(\gamma,\gamma')\le\delta}(X_\gamma - X_{\gamma'}) + 2\max_{1\le i\le N}|X_{t_i} - X_{t_1}|, \end{aligned}$$

then one has that

$$\sup_{t,t'\in\mathbb{T}}(X_t - X_{t'}) \le 2\sup_{\gamma,\gamma'\in\mathbb{T},\varrho_X(\gamma,\gamma')\le\delta}(X_\gamma - X_{\gamma'}) + 2\max_{1\le i\le N}|X_{t_i} - X_{t_1}|$$

and

$$\mathbb{E}\big\{\sup_{\gamma,\gamma'\in\mathbb{T}}(X_t - X_{t'})\big\} \le 2\mathbb{E}\big\{\sup_{\gamma,\gamma'\in\mathbb{T},\varrho_X(\gamma,\gamma')\le\delta}(X_\gamma - X_{\gamma'})\big\} + 2\mathbb{E}\big(\max_{1\le i\le N}|X_{t_i} - X_{t_1}|\big).$$

Note that $\{X_t, t \in \mathbb{T}\}$ be a zero-mean sub-Gaussian process with respect to the metric ϱ_X, then for each i, the random variable $X_{t_i} - X_{t_1}$ is zero-mean and sub-Gaussian with parameter $\varrho_X(t_i, t_1) \le D$. Then by Lemma 1, one has that

$$\mathbb{E}\big(\max_{1\le i\le N}|X_{t_i} - X_{t_1}|\big) \le 2\sqrt{D^2\log N} = 2D\sqrt{\log\mathcal{N}(\delta, \mathbb{T}, \varrho_X)}.$$

$\square$

Proof of Theorem 9. Recall that

$$\sup_{t,t'\in\mathbb{T}}(X_t - X_{t'}) \le 2\sup_{\gamma,\gamma'\in\mathbb{T},\varrho_X(\gamma,\gamma')\le\delta}(X_\gamma - X_{\gamma'}) + 2\max_{1\le i\le N}|X_{t_i} - X_{t_1}|.$$

Define $\mathbb{U} = \{t_1, \dots, t_N\}$ where $N = \mathcal{N}(\delta, \mathbb{T}, \varrho_X)$. Then one has that

$$\max_{1\le i\le N}|X_{t_i} - X_{t_1}| \le \max_{t,\tilde{t}\in\mathbb{U}}|X_t - X_{\tilde{t}}|.$$

For each integer $m = 1, \dots, L$, let $\mathbb{U}_m$ be a minimal $2^{-m}D$-cover set of $\mathbb{U}$ in the metric ϱ_X, where we allow for any element of $\mathbb{T}$ to be used in forming the cover. Since each $\mathbb{U}_m$ is a subset of $\mathbb{T}$, each set has cardinality $N_m = \mathrm{Card}(\mathbb{U}_m) \le \mathcal{N}(2^{-m}D, \mathbb{T}, \varrho_X)$. Since $\mathbb{U}_m$ is finite, then there is some finite integer L such that $\mathbb{U}_L = \mathbb{U}$. For each m, define the mapping $\pi_m : \mathbb{U} \to \mathbb{U}_m$ via

$$\pi_m(t) = \arg\min_{\beta\in\mathbb{U}_m}\varrho_X(t, \beta).$$

Using this notation, we can decompose X_t into a sum of increments in terms of an associated sequence $\{\gamma_m\}_{m=1}^L$ with $\gamma_L = t$ and $\gamma_{m-1} = \pi_{m-1}(\gamma_m)$ for $m = 2, \dots, L$. Then one has the chaining relation

$$X_t - X_{\gamma_1} = \sum_{m=2}^{L}(X_{\gamma_m} - X_{\gamma_{m-1}}),$$

which can be bounded by

$$|X_t - X_{\gamma_1}| \le \sum_{m=2}^{L} \max_{\beta\in\mathbb{U}_m} |X_\beta - X_{\pi_{m-1}(\beta)}|.$$

Similarly, for another $\tilde{t} \in \mathbb{U}$, one can define $\{\tilde{\gamma}_m\}_{m=1}^{L}$ to obtain

$$X_{\tilde{t}} - X_{\tilde{\gamma}_1} = \sum_{m=2}^{L} (X_{\tilde{\gamma}_m} - X_{\tilde{\gamma}_{m-1}}),$$

which can be bounded by

$$|X_{\tilde{t}} - X_{\tilde{\gamma}_1}| \le \sum_{m=2}^{L} \max_{\beta\in\mathbb{U}_m} |X_\beta - X_{\pi_{m-1}(\beta)}|.$$

Then for $t, t' \in \mathbb{U}_m$, one has that

$$\begin{aligned}|X_t - X_{t'}| &= |X_t - X_{\gamma_1} + X_{\gamma_1} - X_{\tilde{\gamma}_1} - X_{\tilde{\gamma}_1} - X_{t'}| \\ &\le \max_{\gamma,\tilde{\gamma}\in\mathbb{U}_1} |X_\gamma - X_{\tilde{\gamma}}| + 2\sum_{m=2}^{L} \max_{\beta\in\mathbb{U}_m} |X_\beta - X_{\pi_{m-1}(\beta)}|.\end{aligned}$$

We first upper bound the finite maximum over $\mathbb{U}_1$, which has $\mathcal{N}(D/2, \mathbb{T}, \varrho_X)$ elements. Note that $X_\gamma - X_{\tilde{\gamma}}$ is sub-Gaussian with parameter $\varrho_X(\gamma, \tilde{\gamma}) \le D$. By Lemma 1, one has that

$$\mathbb{E}\big\{\max_{\gamma,\tilde{\gamma}\in\mathbb{U}_1} |X_\gamma - X_{\tilde{\gamma}}|\big\} \le 2D\sqrt{\log \mathcal{N}(D/2, \mathbb{T}, \varrho_X)}.$$

Similarly, for $m = 2, \ldots, L$, $\mathbb{U}_m$ has $\mathcal{N}(2^{-m}D, \mathbb{T}, \varrho_X)$ elements and $\varrho_X(\beta, \pi_{m-1}(\beta)) \le 2^{-(m-1)}D$. Then one has that

$$\mathbb{E}\big\{\max_{\beta\in\mathbb{U}_m} |X_\beta - X_{\pi_{m-1}(\beta)}|\big\} \le 2 \times 2^{-(m-1)}D\sqrt{\log \mathcal{N}(2^{-m}D, \mathbb{T}, \varrho_X)}.$$

Combining the pieces, we conclude that

$$\begin{aligned}\mathbb{E}\big\{\max_{t,\tilde{t}\in\mathbb{U}} |X_t - X_{t'}|\big\} &\le 2D\sqrt{\log \mathcal{N}(D/2, \mathbb{T}, \varrho_X)} + 4\sum_{m=2}^{L} 2^{-(m-1)}D\sqrt{\log \mathcal{N}(2^{-m}D, \mathbb{T}, \varrho_X)} \\ &\le 4\sum_{m=1}^{L} 2^{-(m-1)}D\sqrt{\log \mathcal{N}(2^{-m}D, \mathbb{T}, \varrho_X)} \\ &= 16\sum_{m=1}^{L} 2^{-(m+1)}D\sqrt{\log \mathcal{N}(2^{-m}D, \mathbb{T}, \varrho_X)} \\ &\le 16\sum_{m=1}^{L} \int_{2^{-(m+1)}D}^{2^{-m}D} \sqrt{\log \mathcal{N}(u, \mathbb{T}, \varrho_X)}\mathrm{d}u \\ &\le 16\int_{\delta/4}^{D} \sqrt{\log \mathcal{N}(u, \mathbb{T}, \varrho_X)}\mathrm{d}u.\end{aligned}$$

□

Lemma 2. *Let $\{Y_i\}_{i=1}^N$ be non-negative random variables such that $\|Y_i\|_{\psi_q} \le 1$. Then for any measurable set A, one has that*

$$\mathbb{E}_A(Y_i) \le \mathbb{P}(A)\psi_q^{-1}\big(1/\mathbb{P}(A)\big)$$

for all $i = 1, \dots, N$ and

$$\mathbb{E}_A\big(\max_{1\le i\le N} Y_i\big) \le \mathbb{P}(A)\psi_q^{-1}\big(N/\mathbb{P}(A)\big),$$

where $\mathbb{E}_A(Y) = \int_A Y \mathrm{d}\mathbb{P}$.

Proof. By definition and Jensen's inequality, one has that

$$\begin{aligned}
\mathbb{E}_A(Y) &= \mathbb{P}(A)\mathbb{E}(Y|Y\in A)\\
&= \mathbb{P}(A)\mathbb{E}\Big\{\psi_q^{-1}\big(\psi_q(Y)\big)\Big|Y\in A\Big\}\\
&\le \mathbb{P}(A)\psi_q^{-1}\Big(\mathbb{E}\big\{\psi_q(Y)\big|Y\in A\big\}\Big)\\
&= \mathbb{P}(A)\psi_q^{-1}\Big(\mathbb{E}_A\big\{\psi_q(Y)\big\}/\mathbb{P}(A)\Big)\\
&\le \mathbb{P}(A)\psi_q^{-1}\big(1/\mathbb{P}(A)\big).
\end{aligned}$$

Any measurable set A can be partitioned into a disjoint union of sets $\{A_i\}_{i=1}^N$ such that $Y_i = \max_{1\le j\le N} Y_j$ on A_i. Then

$$\begin{aligned}
\mathbb{E}_A\big(\max_{1\le i\le N} Y_i\big) &= \sum_{i=1}^N \mathbb{E}_{A_i}(Y_i)\\
&\le \sum_{i=1}^N \mathbb{P}(A_i)\psi_q^{-1}\big(1/\mathbb{P}(A_i)\big)\\
&= \mathbb{P}(A)\sum_{i=1}^N \mathbb{P}(A_i)\psi_q^{-1}\big(1/\mathbb{P}(A_i)\big)/\mathbb{P}(A)\\
&\le \mathbb{P}(A)\psi_q^{-1}\big(N/\mathbb{P}(A)\big).
\end{aligned}$$

□

Proof of Theorem 12. Denote $Z = \sup_{t,t'\in\mathbb{T}} |X_t - X_{t'}|$. If we have shown that

$$\mathbb{E}_A(Z) \le 16\mathbb{P}(A)\int_0^D \psi_q^{-1}\Big(\frac{\mathcal{N}(u,\mathbb{T},\varrho_X)}{\mathbb{P}(A)}\Big)\mathrm{d}u.$$

Then by taking $A = \{Z \ge z\}$, one has that

$$\mathbb{P}(Z\ge z) \le z^{-1}\mathbb{E}_{\{Z\ge z\}}(Z) \le 16t^{-1}\mathbb{P}(Z\ge z)\int_0^D \psi_q^{-1}\Big(\frac{\mathcal{N}(u,\mathbb{T},\varrho_X)}{\mathbb{P}(Z\ge z)}\Big)\mathrm{d}u.$$

Then by canceling out $\mathbb{P}(Z\ge z)$ and using $\psi_q^{-1}(ab) \le c\big\{\psi_q^{-1}(a) + \psi_q^{-1}(b)\big\}$, one has that

$$\begin{aligned}
t &\le 16\int_0^D \psi_q^{-1}\Big(\frac{\mathcal{N}(u,\mathbb{T},\varrho_X)}{\mathbb{P}(Z\ge z)}\Big)\mathrm{d}u\\
&\le 16c\Big\{\mathcal{J}_q(0,D) + D\psi_q^{-1}\big(1/\mathbb{P}(Z\ge z)\big)\Big\}.
\end{aligned}$$

Let $\epsilon > 0$, set $z = 16c\mathcal{J}_q(0, D) + 16c\epsilon$. Some algebra then yields the inequality

$$\mathbb{P}\Big[Z \geq c\big\{\mathcal{J}_q(0, D) + \epsilon\big\}\Big] \leq 1/\psi_q(\epsilon/D).$$

By following the one-step discretization, one has that

$$\begin{aligned}
\mathbb{E}_A(Z) &\leq 2\mathbb{E}_A\big\{\sup_{\gamma,\gamma'\in\mathbb{T},\varrho_X(\gamma,\gamma')\leq\delta}(X_\gamma - X_{\gamma'})\big\} + 2\mathbb{E}_A\big\{\max_{1\leq i\leq N}|X_{t_i} - X_{t_1}|\big\} \\
&\leq 2\mathbb{E}_A\big\{\sup_{\gamma,\gamma'\in\mathbb{T},\varrho_X(\gamma,\gamma')\leq\delta}(X_\gamma - X_{\gamma'})\big\} + 2\mathbb{E}_A\big\{\max_{t,\tilde{t}\in\mathbb{U}}|X_t - X_{\tilde{t}}|\big\},
\end{aligned}$$

where $\mathbb{U} = \{t_1, \ldots, t_N\}$ is a δ-cover of $\mathbb{T}$. For $m = 1, \ldots, L$, let $\mathbb{U}_m$ be a minimal $2^{-m}D$-cover of $\mathbb{U}$ with $N_m = \mathcal{N}(2^{-m}D, \mathbb{U}, \varrho_X)$ elements. For each m, define the mapping $\pi_m : \mathbb{U} \to \mathbb{U}_m$ via

$$\pi_m(t) = \arg\min\nolimits_{\beta\in\mathbb{U}_m}\varrho_X(t, \beta).$$

Then one has that

$$\mathbb{E}_A\big\{\max_{t,\tilde{t}\in\mathbb{U}}|X_t - X_{\tilde{t}}|\big\} \leq 2\sum_{m=1}^{L}\mathbb{E}_A\big\{\max_{\beta\in\mathbb{U}_m}|X_\beta - X_{\pi_{m-1}(\beta)}|\big\}.$$

For each $\beta \in \mathbb{U}_m$, one has that

$$\|X_\beta - X_{\pi_{m-1}(\beta)}\|_{\psi_q} \leq \varrho_X\big(\beta, \pi_{m-1}(\beta)\big) \leq 2^{-(m-1)}D.$$

By applying Lemma 2, one has that

$$\mathbb{E}_A\big\{\max_{\beta\in\mathbb{U}_m}|X_\beta - X_{\pi_{m-1}(\beta)}|\big\} \leq \mathbb{P}(A)2^{-(m-1)}D\psi_q^{-1}\Big(\frac{\mathcal{N}(2^{-m}D, \mathbb{U}, \varrho_X)}{\mathbb{P}(A)}\Big).$$

Then one has that

$$\begin{aligned}
2\mathbb{E}_A\big\{\max_{t,\tilde{t}\in\mathbb{U}}|X_t - X_{\tilde{t}}|\big\} &\leq 4\mathbb{P}(A)\sum_{m=1}^{L}2^{-(m-1)}D\psi_q^{-1}\Big(\frac{\mathcal{N}(2^{-m}D, \mathbb{U}, \varrho_X)}{\mathbb{P}(A)}\Big) \\
&= 16\mathbb{P}(A)\sum_{m=1}^{L}2^{-(m+1)}D\psi_q^{-1}\Big(\frac{\mathcal{N}(2^{-m}D, \mathbb{U}, \varrho_X)}{\mathbb{P}(A)}\Big) \\
&\leq 16\mathbb{P}(A)\int_0^D\psi_q^{-1}\Big(\frac{\mathcal{N}(u, \mathbb{T}, \varrho_X)}{\mathbb{P}(A)}\Big)\mathrm{d}u.
\end{aligned}$$

Let $\delta \to 0^+$,

$$\mathbb{E}_A(Z) \leq 16\mathbb{P}(A)\int_0^D\psi_q^{-1}\Big(\frac{\mathcal{N}(u, \mathbb{T}, \varrho_X)}{\mathbb{P}(A)}\Big)\mathrm{d}u.$$

□

Lemma 3 (Gaussian interpolation)**.** *Consider two independent Gaussian N-dimensional random vectors $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_1)$ and $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_2)$. Define*

$$\mathbf{Z}(t) = \sqrt{1-t}\mathbf{X} + \sqrt{t}\mathbf{Y}.$$

Then for any twice-differentiable function $F : \mathbb{R}^N \to \mathbb{R}$, one has that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\Big\{F\big(\mathbf{Z}(t)\big)\Big\} = 2^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}(\mathbf{\Sigma}_{2,ij} - \mathbf{\Sigma}_{1,ij})\mathbb{E}\Big\{\frac{\partial^2F}{\partial u_i\partial u_j}\big(\mathbf{Z}(t)\big)\Big\}.$$

Proof. By using the chain rule, one has that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\Big\{F\big(\mathbf{Z}(t)\big)\Big\} = 2^{-1}\sum_{i=1}^{N}\mathbb{E}\Big[\frac{\partial F}{\partial u_i}\big(\mathbf{Z}(t)\big)\big\{-(1-t)^{-1/2}X_i + t^{-1/2}Y_i\big\}\Big]$$
$$= -2^{-1}\sum_{i=1}^{N}(1-t)^{-1/2}\mathbb{E}\Big[X_i\frac{\partial F}{\partial u_i}\big(\mathbf{Z}(t)\big)\Big] + 2^{-1}\sum_{i=1}^{N}t^{-1/2}\mathbb{E}\Big[Y_i\frac{\partial F}{\partial u_i}\big(\mathbf{Z}(t)\big)\Big].$$

By using Gaussian integration by parts, one has that

$$\mathbb{E}\Big[X_i\frac{\partial F}{\partial u_i}\big(\mathbf{Z}(t)\big)\Big] = \sum_{j=1}^{N}\mathbf{\Sigma}_{1,ij}\mathbb{E}\Big[\frac{\partial^2 F}{\partial u_i\partial u_j}\big(\mathbf{Z}(t)\big)\Big] \times \sqrt{1-t}$$

and

$$\mathbb{E}\Big[Y_i\frac{\partial F}{\partial u_i}\big(\mathbf{Z}(t)\big)\Big] = \sum_{j=1}^{N}\mathbf{\Sigma}_{2,ij}\mathbb{E}\Big[\frac{\partial^2 F}{\partial u_i\partial u_j}\big(\mathbf{Z}(t)\big)\Big] \times \sqrt{t}.$$

Combining the two sums we complete the proof. □

Proof of Theorem 13. This is trivial by using Lemma 3. □

Proof of Theorem 14. Let $F_\beta(x_1,\dots,x_N) = \beta^{-1}\log\big(\sum_{j=1}^{N}e^{\beta x_j}\big)$. One has that

$$\max_{1\le j\le N} x_j \le F_\beta(x_1,\dots,x_N) \le \max_{1\le j\le N} x_j + \beta^{-1}\log N$$

for all $\beta > 0$. Then use Theorem 13 by taking $A = \emptyset$ and $B = \{(i,j) : i \ne j\}$, one has that

$$\mathbb{E}\big\{\max_{1\le j\le N} X_i\big\} \le \mathbb{E}\big\{F_\beta(X_i)\big\} \le \mathbb{E}\big\{F_\beta(Y_i)\big\} \le \mathbb{E}\big\{\max_{1\le j\le N} Y_i\big\} + \beta^{-1}\log N.$$

Let $\beta \to \infty$ yield the result. □

Proof of Theorem 17. Denote $M = \mathcal{M}(\delta,\mathbb{T},\varrho_X)$. For any $\delta > 0$, let $\{t_1,\dots,t_M\}$ be a δ-packing of $\mathbb{T}$. Then one has that

$$\mathbb{E}(X_{t_i} - X_{t_j})^2 = \varrho_X^2(t_i,t_j) > \delta^2$$

for all $i \ne j$. Let $\{Y_i\}_{i=1}^{M}$ be independent $\mathcal{N}(0,\delta^2/2)$ random variables. Then one has that

$$\mathbb{E}(X_{t_i} - X_{t_j})^2 \ge \mathbb{E}(Y_i - Y_j)^2$$

for all $i,j \in \{1,\dots,M\}$. Then by Theorem 15, one has that

$$\begin{aligned}
\mathbb{E}\big(\sup_{t\in\mathbb{T}} X_t\big) &\ge \mathbb{E}\big(\max_{1\le i\le M} X_{t_i}\big)\\
&\ge \mathbb{E}\big(\max_{1\le i\le M} Y_i\big)\\
&\ge \sqrt{\delta^2\log M/2} \times \mathbb{P}\big(\max_{1\le i\le M} Y_i \ge \sqrt{\delta^2\log M/2}\big)\\
&= \sqrt{\delta^2\log M/2}\big\{1 - \Phi^M\big(\sqrt{\log M}\big)\big\}\\
&\ge \sqrt{\delta^2\log M/2}\big\{1 - (2\pi)^{-1/2}M^{-1/2}\big\}\\
&\ge 2^{-1}\delta\sqrt{\log M}.
\end{aligned}$$

□