Lecture 9

In this lecture we will try to understand some ideas from "rate adaptive" estimation of quadratic functionals of a density.

Recall: We proved the following, $\left(\chi(\eta) = \int \eta^2 d\mu\right)$

$$\inf_{T_n} \sup_{\eta \in \mathcal{P}(\alpha, M, B)} \mathbb{E}_\eta \left(T_n - \chi(\eta)\right)^2 \sim \begin{cases} n^{-8\alpha/4\alpha+1} & \text{if } \alpha \leq 1/4 \\ n^{-1} & \text{if } \alpha > 1/4 \end{cases}$$

Our estimator $\chi(\eta)$ was

$$\hat{\chi}_n = \frac{1}{n(n-1)} \sum_{i \neq j} \sum_{\ell=1}^{k} e_\ell(x_i) e_\ell(x_j) \qquad \text{for an o.n.b. } \{e_\ell\} \text{ of } L_2[0,1]$$

We showed that $\mathbb{E}_\eta \left(\hat{\chi}_n - \chi(\eta)\right)^2 \leq \left(\text{Bias}_\eta (\hat{\chi}_n)\right)^2 + \text{Var}_\eta (\hat{\chi}_n)$

$$\leq C(M) k^{-4\alpha} + C(B) \left(\frac{1}{n} + \frac{k}{n^2}\right)$$

The match between $(\text{bias})^2$ and variance was then obtained by choosing $k \sim n^{2/4\alpha+1}$ for $\alpha \leq 1/4$ and $k \sim n$ for $\alpha > 1/4$.

Therefore it was crucial that we knew α!

Question: Can we find an estimator $\hat{\chi}_n$ which does not depend on α s.t for any α,

$$\sup_{\eta \in \mathcal{P}(\alpha, M, B)} \mathbb{E}_\eta \left(\hat{\chi}_n - \chi(\eta)\right)^2 \sim \inf_{T_n} \sup_{\eta \in \mathcal{P}(\alpha, M, B)} \mathbb{E}_\eta \left(T_n - \chi(\eta)\right)^2 \;?$$

"Answer": YES if $\alpha \geq 1/4$
NO if $\alpha < 1/4$

we shall prove the following result.

Theorem: Let $0 < \alpha < 1$. (i) Then there exists an estimator $\hat{x}_n$ (free of α) such that

$$\sup_{\eta \in \mathcal{P}(\alpha, M, B)} \mathbb{E}_\eta \left(\hat{x}_n - x(\eta)\right)^2 \sim \begin{cases} \left(\frac{n}{\sqrt{\log n}}\right)^{-\frac{8\alpha}{4\alpha+1}} & \text{if } \alpha < 1/4 \\ n^{-1} & \text{if } \alpha \geq \frac{1}{4} \end{cases}$$

(ii) Further, if for some estimator T_n, one has for some $\alpha < \frac{1}{4}$

$$\sup_{\eta \in \mathcal{P}(\alpha, M, B)} \mathbb{E}_\eta \left(T_n - x(\eta)\right)^2 \lesssim \left(\frac{n}{\sqrt{\log n}}\right)^{-\frac{8\alpha}{4\alpha+1}}$$

Then $\exists\, \alpha' < \alpha$ s.t.

$$\sup_{\eta \in \mathcal{P}(\alpha', M, B)} \mathbb{E}_\eta \left(T_n - x(\eta)\right) \gtrsim \left(\frac{n}{\sqrt{\log n}}\right)^{-\frac{8\alpha'}{4\alpha'+1}}$$

Proof: The proof of (i) (Upper Bound) is based on an idea, often referred to as Lepski's Method.
The proof of (ii) (Lower Bound) is based on constrained Risk Inequalities.
We will take this opportunity learn a little bit about both of these ideas

(i) Upper Bound & Lepski's Method

Crucial to our construction is noting that proof of a non-adaptive estimator $\hat{x}_n$ for known α depends on a bias variance tradeoff.

Let $\hat{x}_n(j) = \frac{1}{n(n-1)} \sum_{i_1 \neq i_2} K_{V_j}(X_{i_1}, X_{i_2})$ $\quad K_{V_j}(x_1, x_2) = \sum_{l,k} \psi_{lk}(x_1)\psi_{lk}(x_2)$

Then $\mathbb{E}_\eta\left(\hat{x}_n(j) - x(\eta)\right)^2$

$$= \left\{Bias_\eta(\hat{x}_n)\right\}^2 + Var_\eta(\hat{x}_n)$$

$$= \|\Pi_{V_j^\perp}(\eta)\|_2^4 + Var_\eta(\hat{x}_n)$$

$$\leq \|\Pi_{V_j^\perp}(\eta)\|_2^2 + \frac{4}{n}\mathbb{E}_\eta\left[\mathbb{E}_\eta^2\left(K_{V_j}(X_1, X_2) \mid X_1\right)\right] + \frac{2}{n(n-1)}\mathbb{E}_\eta\left(K_{V_j}^2(X_1, X_2)\right)$$

$$\leq \|\Pi_{V_j^\perp}(\eta)\|_2^4 + \frac{4\|\eta\|_\infty^3}{n} + \frac{2\|\eta\|_\infty^2}{n(n-1)} \cdot 2^j$$

$$\leq \|f\|_\alpha 2^{-4j\alpha} + \frac{4\|\eta\|_\infty^3}{n} + \frac{2\|\eta\|_\infty^2}{n(n-1)} 2^j$$

$$\leq \|f\|_\alpha 2^{-4j\alpha} + \frac{4\|\eta\|_\infty^3}{n} + \frac{3\|\eta\|_\infty^2}{n^2} \cdot 2^j \quad \text{if } n \geq 3$$

where $\|f\|_\alpha = \sup_{x,y \in [0,1]} \frac{|f(x) - f(y)|}{|x-y|^\alpha}$ for $0 < \alpha < 1$.

Now when α is known, ~~the~~ we choose from the collection of $\{\hat{x}_n(j), j \geq \log_2 n\}$ the estimator which minimizes the mean squared error.

The choice is given by $2^j \sim n^{\frac{2}{4\alpha+1}}$ if $0<\alpha<1/4$.

Now suppose we do not know α, but our guess of α is α' with $|\alpha-\alpha'| \leq \frac{c}{\log n}$ for some constant $c>0$.
We then will want $2^j \sim n^{\frac{2}{4\alpha'+1}}$

How much do we loose in MSE with this choice? $\longrightarrow$
We loose by a constant! This implies the following idea.

Idea (Consider adapting over $0<\alpha<1/4$)

(i) Discretize $(0, \frac{1}{4})$ into a grid $\alpha_0 > \alpha_1 > \dots > \alpha_N$ s.t. $|\alpha_i - \alpha_{i+1}| \leq \frac{c}{\log n}$.

(ii) ~~Test between~~ Perform a multiple hypothesis testing to choose $\hat{\alpha} \in \{\alpha_0, \alpha_1, \dots, \alpha_N\}$

(iii) Take candidate estimator as $\hat{x}_n(\hat{j})$ with $2^{\hat{j}} \sim n^{\frac{2}{4\hat{\alpha}+1}}$

Caveat: Note that we are ~~doing~~ same data do perform the test and construction of the estimator. ~~The~~ Also the price to be paid for adaptation also comes from this multiple hypothesis testing.

The I try to understand this stratigy is the following. Begin with a two point adaptation problem i.e you are given the knowledge that $\alpha \in \{\alpha_0, \alpha_1\}$ with $\alpha_0 > \alpha_1$. We want to construct an estimator $\hat{\chi}_n$ which ~~does not dep~~ attains the "right rate" over both α_0 and α_1.

For α_0 we want to use $\hat{\chi}_n(j_0)$ for some j_0 depending on α_0

For α_1 " " " " $\hat{\chi}_n(j_1)$ " " j_1 " " α_1

We will need to have $j_1 > j_0$ to kill the bias at α_1.

Note $\hat{\chi}_n(j) = \frac{1}{n(n-1)} \sum_{i_1 \neq i_2} K_{V_j}(X_{i_1}, X_{i_2})$ where $K_{V_j}(x_1, x_2) = \sum_{l=-1}^{j} \sum_{k=0}^{2^l - 1 \vee 0} \psi_{lk}(x_1)\psi_{lk}(x_2)$

Assume for now $2^j \ll n^2$ but $2^j \gg n$. Therefore

$Var(\hat{\chi}_n(j)) \leq C(M,B)\, 2^j/n^2$.

Now indeed $\dfrac{\hat{\chi}_n(j) - \mathbb{E}(\hat{\chi}_n(j))}{\sqrt{Var(\hat{\chi}_n(j))}}$ is a tight sequence of random variables by Chebychev's inequality. Assume for now

$$\frac{\hat{\chi}_n(j) - \mathbb{E}(\hat{\chi}_n(j))}{\sqrt{Var(\hat{\chi}_n(j))}} \xrightarrow{d} N(0,1) \quad \text{as } n \to \infty. \qquad (*)$$

Let $\hat{j} \in \{j_0, j_1\}$ be a data dependent choice of truncation level. How does one analyze the resulting estimator $\hat{\chi}_n(\hat{j})$?

First consider the truth to be $\alpha=\alpha_0$ i.e. $\eta \in \mathcal{P}(\alpha_0, M, B)$

$$\Rightarrow \mathbb{E}_\eta\left((\hat{x}_n(\hat{j}) - x(\eta))^2\right) = \mathbb{E}_\eta\left((\hat{x}_n(\mathring{j}_0) - x(\eta))^2 I(\hat{j}=\mathring{j}_0)\right)$$
$$+ \mathbb{E}_\eta\left((\hat{x}_n(\mathring{j}_1) - x(\eta))^2 I(\hat{j}=\mathring{j}_1)\right)$$
$$= T_0 + T_1$$

Now, T_0 is under control since the truth is α_0 and we chose $\mathring{j}_0$. The crucial part therefore is control of T_1. Here the truth is α_0 but your choice guides you to lower smoothness truncation i.e $\mathring{j}_1$.

Now, $(\hat{x}_n(\mathring{j}_1) - x(\eta))^2 = O_p\left(2^{-4\mathring{j}_1\alpha_0} + 2^{\mathring{j}_1}/n^2\right)$
$$\sim O_p\left(2^{-4\mathring{j}_1\alpha_0} + 2^{-4\mathring{j}_1\alpha_1}\right) \sim O_p\left(2^{-4\mathring{j}_1\alpha_1}\right) \quad \text{as } \alpha_1<\alpha_0.$$

$\therefore$ $(\hat{x}_n(\mathring{j}_1) - x(\eta))^2$ can be at most like $O_p(n^{-8\alpha_1/4\alpha_1+1})$.
But for T_1 to behave like $O_p(n^{-8\alpha_0/4\alpha_0+1})$, the event $\{\hat{j}=\mathring{j}_1\}$ must have very low probability under α_0. This probability therefore intuitively needs to $O(1/n)$ since for any $\alpha_1<\alpha_0<1/4$ we need to get $n^{-8\alpha_1/4\alpha_1+1}$ down to $n^{-8\alpha_0/4\alpha_0+1}$ by multiplying with this probability.

Consider the test between α_0 and α_1 given by:

Choose $\mathring{j}_0$ if $(\hat{x}_n(\mathring{j}_0) - \hat{x}_n(\mathring{j}_1))^2 \le a_n\, 2^{\mathring{j}_1}/n^2$

choose $\mathring{j}_1$ if $(\hat{x}_n(\mathring{j}_0) - \hat{x}_n(\mathring{j}_1))^2 > a_n\, 2^{\mathring{j}_1}/n^2$

This choice is guided by the philosophy that as j increases, the bias decreases but variance increases.
But what should be the choice of a_n?

Indeed by the above heuristics, we want that under the smoothness α_0, $\mathbb{P}\left(\left(\hat{x}_n(j_0) - \hat{x}_n(j_1)\right)^2 > a_n\, 2^{j_1}/n^2\right) = O(1/n)$

Guided by the fact that $\hat{x}_n(j_0)$, $\hat{x}_n(j_1)$ are asymptotically normal under right centering & scaling, $j_1 > j_0$, we have $a_n \sim \log n$. This is because we know that $\mathbb{P}\left(|N(0,1)| > \sqrt{2\log n}\right) = O(1/n)$.

But then this choice of ~~a_n~~ the list might not guarantee that when the truth is α_1, we will choose ~~smoothing~~ $\hat{j} = j_1$ with high probability. This can be heuristically understood by the calculations below.

If α_1 is the truth, then

$$\left(\hat{x}_n(j_0) - \hat{x}_n(j_1)\right)^2 \sim \left(\hat{x}_n(j_0) - x(\eta)\right)^2 + \left(\hat{x}_n(j_1) - x(\eta)\right)^2$$

$$\sim 2^{-4j_0\alpha_1} + \frac{2^{j_0}}{n^2} + 2^{-4j_1\alpha_1} + 2^{j_1}/n^2 \qquad (**)$$

~~we have~~ We have already decided to choose j_0 to be the right truncation corresponding to α_0 i.e. $2^{j_0} \sim n^{2/4\alpha_0+1}$. If we also choose $2^{j_1} \sim n^{2/4\alpha_1+1}$ ~~then $(**) \sim n^{-8\alpha_1/4\alpha_0+1} + n^{-8\alpha_1/4\alpha_1+1}$~~ we get suboptimal result ~~So this leads to that $\log n\, 2^{j_1}/n^2 \sim \log n\, n^{-8\alpha_1/4\alpha_1+1}$ which cannot happen~~

The idea is to change the choice of $2^{j_1} \sim \left(\frac{n^2}{\log n}\right)^{1/4\alpha_1+1}$.

Then $(**)$ can be of the order of $\log n\; 2^{j_1}/n^2$. But this means that we pay a price at the mean squared error for α_1 due to oversmoothing.

Let's do the calculations in detail now to make things concrete.

Let j_0 and j_1 be such that $2^{j_0} \sim n^{2/(4\alpha_0+1)}$ and $2^{j_1} \sim \left(n^2/\log n\right)^{1/(4\alpha_1+1)}$

Let $\hat{j} = j_0$ if $\left(\hat{x}_n(j_0) - \hat{x}_n(j_1)\right)^2 \le C^* \log n \; 2^{j_1}/n^2$

$\hat{j} = j_1$ otherwise. (C^* will be chosen later)

Consider the estimator $\hat{x}_n = \hat{x}_n(\hat{j})$. We now study the mean squared error properties of this estimator over $\mathcal{P}(\alpha_0, M, B)$ and $\mathcal{P}(\alpha_1, M, B)$.

First let $\eta \in \mathcal{P}(\alpha_1, M, B)$. Then

$$\mathbb{E}_\eta\left(\hat{x}_n(\hat{j}) - x(\eta)\right)^2 = \mathbb{E}_\eta\left(\left(\hat{x}_n(\hat{j}) - x(\eta)\right)^2 I(\hat{j} = j_0)\right) + \mathbb{E}_\eta\left(\left(\hat{x}_n(\hat{j}) - x(\eta)\right)^2 I(\hat{j} = j_1)\right)$$
$$= T_0 + T_1$$

$$T_1 \le \mathbb{E}_\eta\left(\left(\hat{x}_n(j_1) - x(\eta)\right)^2\right) \le c(M,B)\left[2^{-4j_1\alpha_1} + 2^{j_1}/n^2\right] \quad \text{as } 2^{j_1} \gg n$$

$$\le 2c(M,B) \left(\frac{n}{\sqrt{\log n}}\right)^{-\frac{8\alpha_1}{4\alpha_1+1}}$$

matches the claimed adaptive bound.

$$T_0 = \mathbb{E}_\eta\left(\left(\hat{x}_n(\hat{j}) - x(\eta)\right)^2 I(\hat{j} = j_0)\right)$$

$$\le 2\,\mathbb{E}_\eta\left(\left(\hat{x}_n(j_0) - \hat{x}_n(j_1)\right)^2 I(\hat{j} = j_0)\right) + 2\,\mathbb{E}\left(\left(\hat{x}_n(j_1) - x(\eta)\right)^2\right)$$

$$\le 2C^* \log n \,\frac{2^{j_1}}{n^2} + 2c(M,B)\left(\frac{n}{\sqrt{\log n}}\right)^{-\frac{8\alpha_1}{4\alpha_1+1}}$$

$$\le \left(2C^* + 2c(M,B)\right)\left(\frac{n}{\sqrt{\log n}}\right)^{-\frac{8\alpha_1}{4\alpha_1+1}}$$ $\Rightarrow$ $T_0 + T_1$ is under desired control.

Now let $\eta \in \mathcal{P}(\alpha_0, M, B)$. Then

$$\mathbb{E}_\eta\left((\hat{x}_n(\hat{j}) - x(\eta))^2\right) = \mathbb{E}_\eta\left((\hat{x}_n(\hat{j}) - x(\eta))^2 I(\hat{j} = j_0)\right)$$
$$+ \mathbb{E}_\eta\left((\hat{x}_n(\hat{j}) - x(\eta))^2 I(\hat{j} = j_1)\right)$$
$$= T_0 + T_1$$

This time, $T_0 \leq \mathbb{E}_\eta\left((\hat{x}_n(j_0) - x(\eta))^2\right) \leq c(M,B)\, n^{-8\alpha_0/4\alpha_0+1}$

The crucial term is T_1, because this corresponds to choosing a lower smoothness when the truth is higher.

$$T_1 = \mathbb{E}_\eta\left((\hat{x}_n(\hat{j}) - x(\eta))^2 I(\hat{j} = j_1)\right)$$

$$\leq \left\{\mathbb{E}_\eta\left((\hat{x}_n(j_1) - x(\eta))^{2q}\right)\right\}^{1/q} \left\{\mathbb{P}_\eta(\hat{j} = j_1)\right\}^{1/p} = T_{11} \times T_{12} \quad \left(\frac{1}{p} + \frac{1}{q} = 1\right)$$

We show that T_{11} is appropriately bounded and $\mathbb{P}_\eta(\hat{j} = j_1)$ is small.

<u>Analysis of T_{12}</u>

$$\mathbb{P}_\eta(\hat{j} = j_1) = \mathbb{P}_\eta\left(|\hat{x}_n(j_0) - \hat{x}_n(j_1)| > \sqrt{c^* \log n}\sqrt{2^{j_1}/n^2}\right)$$

$$\leq \mathbb{P}_\eta\Big(|\hat{x}_n(j_0) - \mathbb{E}_\eta(\hat{x}_n(j_0))| + |\hat{x}_n(j_1) - \mathbb{E}_\eta(\hat{x}_n(j_1))|$$
$$> \sqrt{c^* \log n}\sqrt{2^{j_1}/n^2} - |\mathbb{E}_\eta(\hat{x}_n(j_1)) - \mathbb{E}_\eta(\hat{x}_n(j_0))|\Big)$$

Now, $|\mathbb{E}_\eta(\hat{x}_n(j_1)) - \mathbb{E}_\eta(\hat{x}_n(j_0))| \leq |\mathrm{Bias}_\eta(\hat{x}_n(j_1))| + |\mathrm{Bias}_\eta(\hat{x}_n(j_0))|$

$$\leq M\left(2^{-2j_1\alpha_0} + 2^{-2j_0\alpha_0}\right)$$
$$\leq 2M\, 2^{-2j_0\alpha_0} \quad \text{as } j_1 \geq j_0$$
$$\leq 2M\sqrt{2^{j_1}/n^2} \quad \forall n$$

$$\Rightarrow \mathbb{P}_\eta\left(\hat{j}=j_1\right) \le \mathbb{P}_\eta\Big[\left|\hat{x}_n(j_0)-\mathbb{E}_\eta(\hat{x}_n(j_0))\right| + \left|\hat{x}_n(j_1)-\mathbb{E}_\eta(\hat{x}_n(j_1))\right| > \sqrt{\tfrac{C^*}{2}\log n}\sqrt{2^{j_1}/n^2}\Big]$$

$$\le \mathbb{P}_\eta\Big[\left|\hat{x}_n(j_0)-\mathbb{E}_\eta(\hat{x}_n(j_0))\right| > \sqrt{\tfrac{C^*}{4}\log n}\sqrt{2^{j_1}/n^2}\Big] + \mathbb{P}_\eta\Big[\left|\hat{x}_n(j_1)-\mathbb{E}_\eta(\hat{x}_n(j_1))\right| > \sqrt{\tfrac{C^*}{4}\log n}\sqrt{\tfrac{2^{j_1}}{n^2}}\Big]$$

$$\le \mathbb{P}_\eta\Big[\left|\hat{x}_n(j_0)-\mathbb{E}_\eta(\hat{x}_n(j_0))\right| > \sqrt{\tfrac{C^*}{4}\log n}\sqrt{\tfrac{2^{j_0}}{n^2}}\Big] + \mathbb{P}_\eta\Big[\left|\hat{x}_n(j_1)-\mathbb{E}_\eta(\hat{x}_n(j_1))\right| > \sqrt{\tfrac{C^*}{4}\log n}\sqrt{\tfrac{2^{j_1}}{n^2}}\Big]$$

Therefore the analysis depends on

$$\mathbb{P}_\eta\Big[\left|\hat{x}_n(j)-\mathbb{E}_\eta(\hat{x}_n(j))\right| > \sqrt{C^{**}\log n}\sqrt{2^j/n^2}\Big] \quad \text{for } 2^j \gg n$$

$\quad\&\ 2^j \ll n^2$.

$$\hat{x}_n(j)-\mathbb{E}_\eta(\hat{x}_n(j)) = \frac{1}{n(n-1)}\sum_{i_1\ne i_2} K_{V_j}(x_{i_1},x_{i_2}) - \mathbb{E}_\eta\left(K_{V_j}(x_{i_1},x_{i_2})\right)$$

$$= \frac{2}{n}\sum_{i=1}^{n}\sum_{l=-1}^{j}\sum_{k=0}^{2^l-1\vee 0}\left(\psi_{lk}(x_i)-\langle\psi_{lk},\eta\rangle\right)\langle\psi_{lk},\eta\rangle$$

$$+ \frac{1}{n(n-1)}\sum_{i_1\ne i_2}\sum_{l,k}\left(\psi_{lk}(x_{i_1})-\langle\psi_{lk},\eta\rangle\right)\left(\psi_{lk}(x_{i_2})-\langle\psi_{lk},\eta\rangle\right)$$

$$= I_1 + I_2$$

$$\Rightarrow \mathbb{P}_\eta\Big[\left|\hat{x}_n(j)-\mathbb{E}_\eta(\hat{x}_n(j))\right| > \sqrt{C^{**}\log n}\sqrt{2^j/n^2}\Big]$$

$$\le \mathbb{P}_\eta\Big[|I_1| > \sqrt{\tfrac{C^{**}}{2}\log n}\sqrt{2^j/n^2}\Big] + \mathbb{P}_\eta\Big[|I_2| > \sqrt{\tfrac{C^{**}}{2}\log n}\sqrt{2^j/n^2}\Big]$$

Now $I_1 = \frac{2}{n}\sum_{i=1}^{n} R(x_i)$ where $R(x) = \sum_{l=-1}^{j}\sum_{k=0}^{(2^l-1)\vee 0}\left(\psi_{lk}(x)-\langle\psi_{lk},\eta\rangle\right)\langle\psi_{lk},\eta\rangle$

To control tail of I_1 we will use standard Hoeffding's inequality.

For that we need to bound $\|R\|_\infty$. Fix $x \in [0,1]$.

First note that $|\langle \psi_{\ell k}, \eta \rangle| = \left|\int \psi_{\ell k}(x)\,\eta(x)\,dx\right| \leq 2^{\ell/2}\|\eta\|_\infty \cdot 1/2^{\ell}$

$$\leq B/2^{\ell/2}$$

Also, for any l, $\exists$ a fixed k_x s.t. $x \in \text{support}(\psi_{\ell k_x})$.

By disjointness of the supports of $\psi_{\ell k}$'s as $k = 0, 1, \ldots, 2^{\ell}-1$, we have for each $l = 0, 1, \ldots, j$, $\left|\sum_{\ell=0}^{(2^{\ell}-1)\vee 0} (\psi_{\ell k}(x) - \langle \psi_{\ell k}, \eta\rangle)\langle \psi_{\ell k}, \eta\rangle\right|$

$$\leq 2^{\ell/2} \cdot 2B/2^{\ell/2} \leq 2B$$

$\Rightarrow |R(x)| \leq 2(j+1)B \leq 4jB$ for all x.

$\Rightarrow$ By Hoeffding's inequality,

$$\mathbb{P}_\eta\left[|I_1| > \sqrt{\frac{c^{**}\log n}{2}}\sqrt{\frac{2^j}{n^2}}\right] \leq \exp\left[-2n\,\frac{c^{**}\log n}{2}\cdot\frac{2^j}{n^2}\cdot\frac{1}{4jB}\right]$$

$$= \exp\left[-\frac{c^{**}\log n \times 2^j}{4B\,jn}\right]$$

Now $2^j \leq n^2$ $\Rightarrow j \leq \log n/2$

$$\leq \exp\left[-\frac{c^{**}}{2B}\times\frac{2^j}{n}\right]$$

Now we need to control the tail of I_2. Now

$I_2 = \frac{1}{n(n-1)}\sum_{i_1 \neq i_2} H(X_{i_1}, X_{i_2})$ where H is symmetric & degenerate under $\mathbb{P}_\eta$.

We shall use the following very important tail bound for second order degenerate U-statistics.

Lemma (Giné, Latala, Zinn (2000), Houdré, Renaud-Bouret (2003))

If $R(x_1, x_2)$ is degenerate under $\mathbb{P}$ and symmetric, then

$$\mathbb{P}\left[\left|\sum_{i\neq j} R(x_i, x_j)\right| \geq C_0\left(\Lambda_1\sqrt{u} + \Lambda_2 u + \Lambda_3 u^{3/2} + \Lambda_4 u^2\right)\right] \leq 6\exp(-u)$$

where $\Lambda_1^2 = \frac{n(n-1)}{2}\,\mathbb{E}\left(R^2(X_1, X_2)\right)$

$$\Lambda_2 = n \sup\left\{\mathbb{E}\left(R(X_1, X_2)\zeta(X_1)\xi(X_2)\right) : \mathbb{E}(\zeta^2(X_1)) \leq 1,\ \mathbb{E}(\xi^2(X_2)) \leq 1\right\}$$

$$\Lambda_3 = \left\| n\,\mathbb{E}\left(R^2(X_1, \cdot)\right)\right\|_\infty^{\frac{1}{2}}$$

$$\Lambda_4 = \sup_{x_1, x_2} |R(x_1, x_2)|.$$

We apply the above result with $R(x_1, x_2) = \sum_{l,k}\left(\psi_{lk}(x_1) - \langle\psi_{lk}, \eta\rangle\right)\left(\psi_{lk}(x_2) - \langle\psi_{lk}, \eta\rangle\right)$ which symmetric and degenerate under $\mathbb{P}_\eta$. We begin by bounding $\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4$ for our problem.

$$\Lambda_1^2 = \frac{n(n-1)}{2}\mathbb{E}_\eta\left(R^2(X_1, X_2)\right) \leq \frac{n(n-1)}{2}\|\eta\|_\infty^2\, 2^j$$
$$\leq B^2 \cdot n^2 \cdot 2^j$$

(We have done this calculation many times before)

To control Λ_2 consider any $\zeta(x_1)$ and $\xi^2(x_2)$ s.t. $\mathbb{E}_\eta\left(\zeta^2(X_1)\right) \leq 1$ and $\mathbb{E}_\eta\left(\xi^2(X_2)\right) \leq 1$.

$$\Rightarrow \mathbb{E}_\eta\left(R(X_1, X_2)\zeta(X_1)\xi(X_2)\right) = \mathbb{E}_\eta\Big(\big(K_{V_j}(X_1, X_2) - \mathbb{E}_\eta(K_{V_j}(X_1, X_2)|X_1) - \mathbb{E}_\eta(K_{V_j}(X_1, X_2)|X_2) + \mathbb{E}_\eta(K_{V_j}(X_1, X_2))\big)\zeta(X_1)\xi(X_2)\Big)$$

First note

$$\left|\mathbb{E}_\eta\left(K_{V_j}(X_1, X_2)\zeta(X_1)\xi(X_2)\right)\right| = \left|\iint K_{V_j}(x_1, x_2)\zeta(x_1)\eta(x_1)\,\xi(x_2)\eta(x_2)\,dx_1 dx_2\right|$$
$$= \left|\int \Pi_{V_j}(\zeta\eta)(x_2)\,\xi(x_2)\eta(x_2)\,dx_2\right|$$

$$\bullet = \left| \int \Pi_{V_j}(\zeta\eta)(x_2)\, \Pi_{V_j}(\xi\eta)(x_2)\, dx_2 \right|$$

$$\leq \sqrt{\int \zeta^2(x_2)\eta^2(x_2)\,dx_2}\,\sqrt{\int \xi^2(x_1)\eta^2(x_1)\,dx_1} \leq B$$

Similarly, $\left| \mathbb{E}_\eta \left(\mathbb{E}_\eta(K_{V_j}(X_1,X_2) \,\cancel{\zeta(X_1)\,\xi(X_2)}\,|X_1)\, \zeta(X_1)\,\xi(X_2) \right) \right|$

$$= \left| \iiint \left(\int K_{V_j}(x_1,x')\, \eta(x')\,dx' \right) \zeta(x_1)\,\xi(x_2)\,\eta(x_1)\,\eta(x_2)\,dx_1\,dx_2 \right|$$

$$= \left| \iint \Pi_{V_j}(\eta)(x_1)\,\zeta(x_1)\,\eta(x_1)\;\; \xi(x_2)\,\eta(x_2)\,dx_2 \right|$$

$$= \left| \int \Pi_{V_j}(\eta)(x_1)\,\Pi_{V_j}(\zeta\eta)(x_1)\,dx_1 \right| \left| \int \xi(x_2)\,\eta(x_2)\,dx_2 \right|$$

$$\leq \sqrt{\int \eta^2(x)\,dx \int \zeta^2(x)\eta^2(x)\,dx}\,\sqrt{\int \xi^2(x)\eta^2(x)\,dx} \leq B^2$$

Similarly $\quad \mathbb{E}_\eta\left(\mathbb{E}_\eta(K_{V_j}(X_1,X_2))\, \zeta(X_1)\,\xi(X_2) \right) \leq B^2$

$\Rightarrow \Lambda_2 \lesssim 4n B^2.$

To control Λ_3, fix $x_2 \in [0,1]$ and consider

$$\mathbb{E}_\eta(R^2(X_1,x_2)) = \mathbb{E}_\eta\Big[K_{V_j}(X_1,x_2) - \mathbb{E}_\eta[K_{V_j}(X_1,X_2)\,|X_1] - \mathbb{E}_\eta[K_{V_j}(X_1,X_2)\,|X_2 = x_2] + \mathbb{E}_\eta[K_{V_j}(X_1,X_2)] \Big]^2$$

$$\cancel{\leq 4\,\mathbb{E}(K^2_{V_j}(X_1,X_2)) + \mathbb{E}_\eta(\mathbb{E}_\eta[K_{V_j}(X_1,X_2)|X_1])^2}$$

$$\leq \mathbb{E}_\eta\left[K_{V_j}(X_1,x_2) - \mathbb{E}_\eta(K_{V_j}(X_1,x_2)) \right]^2 + \mathbb{E}_\eta\left[\mathbb{E}_\eta(K_{V_j}(X_1,X_2)|X_1) - \mathbb{E}_\eta(K_{V_j}(X_1,X_2)) \right]^2$$

Now, $\mathbb{E}_\eta \left[K_{v_j}(X_1, X_2) - \mathbb{E}_\eta (K_{v_j}(X_1, X_2)) \right]^2$

$$\le 2\, \mathbb{E}_\eta \Big(\sum_{l,k} \psi_{lk}(x_1) \psi_{lk}(x_2) \Big)^2 + 2 \int \Big(\sum_{l,k} \psi_{lk}(x_1) \psi_{lk}(x_2) \Big)^2 \eta(x_1)\, dx_1$$

$$\le 2B^2 \cdot 2^j + 2B \sum_{l,k} \psi_{lk}^2(x_2) \le 8B\, 2^j \quad \Big(\text{for each } x_2 \sum_{l,k} \psi_{lk}^2(x_2) \le 1 + 2 + 2^2 + \cdots 2^j \le 2^{j+1} - 1 \Big)$$

as

Similarly one can also show,

$$\mathbb{E}_\eta \left[\mathbb{E}_\eta (K_{v_j}(X_1, X_2) \mid X_1) - \mathbb{E}_\eta (K_{v_j}(X_1, X_2)) \right]^2 \le 8 B \cdot 2^j$$

$\Rightarrow$ For each $x_2 \in [0,1]$, $\mathbb{E}_\eta (R^2(X_1, x_2)) \le 16 B 2^j$

$\Rightarrow \Lambda_3 \le (16 B n 2^j)^{1/2}$

$$\Lambda_4 = \sup_{x_1, x_2} |R(x_1, x_2)| \qquad R(x_1,x_2) = K_{v_j}(x_1,x_2) - \mathbb{E}_\eta (K_{v_j}(X_1, X_2) \mid X_1 = x_1) - \mathbb{E}_\eta (K_{v_j}(X_1, X_2) \mid X_2 = x_2) + \mathbb{E}_\eta (K_{v_j}(X_1, X_2))$$

$$|K_{v_j}(x_1, x_2)| \le \sum_{l=-1}^{j} \sum_{k=0}^{2^l \vee 0} |\psi_{lk}(x_1)| |\psi_{lk}(x_2)|$$

$$\le 2^{j+1}$$

$\Rightarrow \Lambda_4 \le 4\, 2^{j+1}$

Collecting our bounds:

$$\Lambda_1^2 \le B^2 \cdot n^2 \cdot 2^j$$
$$\Lambda_2 \le 4 n B^2$$
$$\Lambda_3 \le 4 \sqrt{B n 2^j}$$
$$\Lambda_4 \le 4\, 2^{j+1} = 8\, 2^j$$

Therefore, plugging these bounds in the lemma we get

$$\mathbb{P}_\eta\left(\left|\sum_{i_1\neq i_2} R(X_{i_1}, X_{i_2})\right| \geq C_0\left[Bn2^{j/2}\sqrt{u} + 4nB^2u + 4\sqrt{Bn2^j}\,u^{3/2} + 8\,2^j u^2\right]\right) \leq 6e^{-u}$$

$$\Rightarrow \mathbb{P}_\eta\left(\left|\sum_{i_1\neq i_2} R(X_{i_1}, X_{i_2})\right| \geq 8C_0B^2\left[n2^{j/2}\sqrt{u} + nu + \sqrt{n}\,2^{j/2}u^{3/2} + 2^j u^2\right]\right) \leq 6e^{-u}$$

$$\Rightarrow \mathbb{P}_\eta\left(\frac{1}{n(n-1)}\left|\sum_{i_1\neq i_2} R(X_{i_1}, X_{i_2})\right| \geq 16C_0B^2\left[\frac{2^{j/2}}{n}\sqrt{u} + \frac{u}{n} + \frac{2^{j/2}}{n^{3/2}}2u^{3/2} + \frac{2^j}{n^2}u^2\right]\right) \leq 6e^{-u}$$

Now, $2u^{3/2} \leq u + u^2$ by AM GM inequality.

$$\Rightarrow \mathbb{P}_\eta\left(\left|\frac{1}{n(n-1)}\sum_{i_1\neq i_2} R(X_{i_1}, X_{i_2})\right| \geq 16C_0B^2\left[\frac{2^{j/2}}{n}\sqrt{u} + \frac{u}{n}\left(\frac{2^{j/2}}{\sqrt{n}}+1\right) + \frac{2^{j/2}}{n^{3/2}}u^2\left(1+\frac{2^{j/2}}{\sqrt{n}}\right)\right]\right) \leq 6e^{-u}$$

$$\Rightarrow \mathbb{P}_\eta\left(|\hat{\chi}_n(j) - \mathbb{E}_\eta\hat{\chi}_n(j)| \geq a_1\sqrt{u} + a_2u + a_3u^2\right) \leq 6e^{-u}$$ where a_1, a_2, a_3 are as above.

If $h(u)$ is such that $a_1\sqrt{h(u)} + a_2h(u) + a_3h^2(u) \leq u$ then,

$$\mathbb{P}_\eta\left(|\hat{\chi}_n(j) - \mathbb{E}_\eta\hat{\chi}_n(j)| \geq u\right) \leq \mathbb{P}_\eta\left(|\hat{\chi}_n(j) - \mathbb{E}_\eta\hat{\chi}_n(j)| \geq a_1\sqrt{h(u)} + a_2h(u) + a_3h^2(u)\right)$$

$$\leq 6e^{-h(u)}$$

Letting $h(u) = \frac{1}{3}\min\{b_1u^2, b_2u, b_3\sqrt{u}\}$ with $b_1 = 1/a_1^2$, $b_2 = 1/a_2$ and $b_3 = 1/\sqrt{a_3}$

we have,

$$\mathbb{P}_\eta\left(|\hat{\chi}_n(j) - \mathbb{E}_\eta\hat{\chi}_n(j)| \geq u\right) \leq 6\exp\left\{-\frac{1}{3}\min\left(\frac{u^2}{a_1^2}, \frac{u}{a_2}, \frac{\sqrt{u}}{\sqrt{a_3}}\right)\right\}$$

where $a_1 = 16C_0B^2\sqrt{\frac{2^j}{n^2}}$, $a_2 = \frac{16C_0B^2}{n}\left(\sqrt{\frac{2^j}{n}}+1\right)$, $a_3 = 16C_0B^2\cdot\frac{2^{j/2}}{n^{3/2}}\left(1+\sqrt{\frac{2^j}{n}}\right)$

This also implies that

$$\mathbb{E}_\eta\left(|\hat{x}_n(j) - \mathbb{E}_\eta \hat{x}_n(j)|^{2q}\right) = 2q\int_0^\infty u^{2q-1}\,\mathbb{P}_\eta\left(|\hat{x}_n(j) - \mathbb{E}_\eta(\hat{x}_n(j))| \geq u\right) du$$

$$\leq 12q\left[\int_0^\infty u^{2q-1} e^{-b_1u^2} du + \int_0^\infty u^{2q-1} e^{-b_2 u} du + \int_0^\infty u^{2q-1} e^{-b_3\sqrt{u}} du\right]$$

$$\leq 12q\left(\frac{\Gamma(q)}{2b_1^q} + \frac{\Gamma(2q)}{b_2^{2q}} + \frac{2\Gamma(4q)}{b_3^{4q}}\right)$$

<u>Going Back to our Actual Problem</u>

$\eta \in \mathcal{P}(\alpha_0, M, B)$ and $n << 2^{j_0}, 2^{j_1} << n^2$

$$\mathbb{E}_\eta\left((\hat{x}_n(\hat{j}) - x(\eta))^2\right) \leq c(M,B)\, n^{-8\alpha_0/4\alpha_0+1} + T_1$$

$$T_1 \leq \left\{\mathbb{E}_\eta\left((\hat{x}_n(j_1) - x(\eta))^{2q}\right\}^{1/q} \left\{\mathbb{P}_\eta(\hat{j} = j_1)\right\}^{1/p}$$ for any $\frac{1}{p} + \frac{1}{q} = 1$, $p, q \geq 1$

Then we showed, for any $n << 2^j << n^2$,

$$\mathbb{P}_\eta(\hat{j} = j_1) \leq \mathbb{P}_\eta\left[|\hat{x}_n(j_0) - \mathbb{E}_\eta \hat{x}_n(j_0)| > \sqrt{\frac{C^*}{4}\log n}\sqrt{\frac{2^{j_0}}{n^2}}\right] + \mathbb{P}_\eta\left[|\hat{x}_n(j_1) - \mathbb{E}_\eta(\hat{x}_n(j_1))| > \sqrt{\frac{C^*}{4}\log n}\sqrt{\frac{2^{j_1}}{n^2}}\right]$$

We then showed

$$\mathbb{P}_\eta\left[|\hat{x}_n(j) - \mathbb{E}_\eta(\hat{x}_n(j))| > \overbrace{\sqrt{C^{**}\log n}\sqrt{2^j/n^2}}^{u}\right]$$

$$\leq \exp\left(-\frac{C^{**}}{2B}\times\frac{2^j}{n}\right) + 6\exp\left(-\frac{1}{3}\min\left\{\frac{u^2}{a_1^2}, \frac{u}{a_2}, \frac{\sqrt{u}}{a_3}\right\}\right)$$ with a_1, a_2, a_3 as on page 15

We also have a bound on $\mathbb{E}_\eta\left(|\hat{x}_n(j) - \mathbb{E}_\eta(\hat{x}_n(j))|^{2q}\right)$.

$\therefore$ We have all the ingredients to tackle T_1.

It is easy to check that with our choice of j_0, j_1, if $u = \sqrt{c^{**}\log n}\sqrt{2^j/n^2}$ and $j \in \{j_0, j_1\}$, then

$$\min\left\{\frac{u^2}{a_1^2}, \frac{u}{a_2}, \frac{\sqrt{u}}{a_3}\right\}$$

$$= \frac{1}{16C_0B^2}\left\{\frac{c^*\log n\ 2^j/n^2}{2^j/n^2}, \frac{(c^*\log n\ 2^j/n^2)^{\frac{1}{2}}(16C_0B)}{\frac{1}{n}\left(\sqrt{\frac{2^j}{n}}+1\right)}, \frac{(c^{**}\log n\ 2^j/n^2)^{\frac{1}{4}}(16C_0B)^{3/2}}{\left(\frac{2^{j/2}}{n^{3/2}}\left(1+\sqrt{\frac{2^j}{n}}\right)\right)^{\frac{1}{2}}}\right\}$$

$$= \frac{1}{(16C_0B^2)^2}\left\{c^{**}\log n, \frac{(n^2c^{**}\log n\cdot 2^j/n^2)^{\frac{1}{2}}(16C_0B)}{\left(\sqrt{\frac{2^j}{n}}+1\right)}, \frac{\left(\frac{n^3}{2^j}c^{**}\log n\ 2^j/n^2\right)^{\frac{1}{4}}\times(16C_0B)^{3/2}}{\left(1+\sqrt{\frac{2^j}{n}}\right)^{\frac{1}{2}}}\right\}$$

$$= \frac{1}{(16C_0B^2)^2}\left\{c^{**}\log n, \frac{(c^{**}\log n\cdot 2^j)^{\frac{1}{2}}(16C_0B)}{(1+\sqrt{2^j/n})}, \frac{(c^{**}\,n\log n)^{1/4}(16C_0B)^{3/2}}{(1+\sqrt{2^j/n})^{1/2}}\right\}$$

$$= \left(\frac{c^{**}\log n}{(16C_0B^2)^2}\right) \quad \forall n.$$

$$\Rightarrow \mathbb{P}_\eta\left(|\hat{x}_n(j) - \mathbb{E}_\eta(\hat{x}_n(j))| > \sqrt{c^{**}\log n}\sqrt{2^j/n^2}\right)$$

$$\leq \exp\left(-\frac{c^{**}}{2B}\times\frac{2^j}{n}\right) + 6\exp\left(-\frac{1}{3}\times\frac{c^{**}\log n}{(16C_0B^2)^2}\right) \qquad \text{if } j\in\{j_0, j_1\}$$

If $c^{**} > 6(16C_0B^2)^2$, then the above bound $\leq (7/n)$

$$\therefore \mathbb{P}_\eta\left[|\hat{x}_n(j) - \mathbb{E}_\eta(\hat{x}_n(j))| > \sqrt{\frac{c^*}{4}\log n}\sqrt{2^j/n^2}\right]$$

$$\leq \exp\left(-\frac{c^*}{8B}2^j/n\right) + 6\exp\left(-\frac{c^*\log n}{12(16C_0B)^2}\right)$$

Choose $c^* > 24(16C_0B)^2$

$$\Rightarrow \mathbb{P}_\eta\left[|\hat{x}_n(j) - \mathbb{E}_\eta(\hat{x}_n(j))| > \sqrt{\frac{c^*}{4}\log n}\sqrt{2^j/n^2}\right] \leq 7/n \quad \text{if } j\in\{j_0, j_1\}$$

We also have, if $j \in \{j_0, j_1\}$ (j satisfying $n \ll 2^j \ll n^2$)

$$\mathbb{E}_\eta\left(|\hat{x}_n(j) - \mathbb{E}_\eta(\hat{x}_n(j))|^{2q}\right)$$

$$\lesssim 12q\left(\frac{\Gamma(q)}{2b_1^q} + \frac{\Gamma(2q)}{b_2^{2q}} + \frac{2\Gamma(4q)}{b_3^{4q}}\right)$$

$$= 12q\left[\frac{\Gamma(q)}{2} \times \cancel{\frac{1}{(16 C_0 B^2)^q}} a_1^{2q} + \Gamma(2q) \times a_2^{2q} + 2\Gamma(4q)\, a_3^{2q}\right]$$

$$\lesssim 12q\left[\frac{\Gamma(q)}{2} + \Gamma(2q) + 2\Gamma(4q)\right] \max\{a_1^{2q}, a_2^{2q}, a_3^{2q}\}$$

$$\lesssim 12q\left[\frac{\Gamma(q)}{2} + \Gamma(2q) + 2\Gamma(4q)\right] (16 C_0 B^2)^{2q} \left(\frac{2^j}{n^2}\right)^q$$ (Check by direct calculations).

$$\Rightarrow T_1 \leqslant \left[\left\{12q\left(\frac{\Gamma(q)}{2} + \Gamma(2q) + 2\Gamma(4q)\right) \times (16 C_0 B^2)^2\right\}^{\frac{1}{q}} \frac{2^{j_1}}{n^2}\right] \times \left(\frac{7}{n}\right)^{\frac{1}{p}}$$

$$= (16 C_0 B^2) \frac{2^{j_1}}{n^2} \times \left(\frac{7}{n}\right)^{\frac{1}{p}} \times \left(12q\left(\frac{\Gamma(q)}{2} + \Gamma(2q) + 2\Gamma(4q)\right)\right)^{\frac{1}{q}}$$

We need to make this smaller than $2^{j_0}/n^2 = n^{-8\alpha_0/4\alpha_0+1}$

$\Rightarrow$ choose p sufficiently close to 1 (depending on α_0).

This completes the "proof" for adaptation over $0 < \alpha_0 \leq \alpha_1 < 1/4$.