Supplementary Material: Online Kernel Learning with a Near Optimal Sparsity Bound

L un Z an ZHANGLIJ MSU EDU

J n n Y YIJINFEN MSU EDU

on J n RONGJIN CSE MSU EDU

D p m n o Com p n n n n n n n n n A

n L n LIN M 8 MAILS TSINGHUA EDU CN

D p m n o A o m on n n B n C n

X ao H XIAOFEIHE CAD ZJU EDU CN

o o o CAD C Co o Com p n Z n n n o C n

A. Proof of Theorem 2

p o o o m on n m o ppo o o op m o m

F on o n m p $\mathcal{T}_1$ $\{ \mathbf{x} , y \}^{n}_{=1}$ $\langle \kappa \ \mathbf{x} , \cdot , \kappa \ \mathbf{x} , \cdot \rangle_{\mathcal{H}_\kappa}$ δ n $y \in \{ , - \}$ o m on on on d po no m n $\kappa \ \mathbf{x}, \mathbf{y}$ $\mathbf{x} \ \mathbf{y}$ n n n p $\mathbb{R}$ $m > n$ n $m > n$ n n o o ono m o $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ n $\mathbb{R}$ $\mathbf{x} \ \mathbf{x}$ n $i \ / \ j$ n $\mathbf{x} \ \mathbf{x}$ o on on o m p on $\kappa \ \mathbf{x} , \mathbf{x}$ δ Fo n n n n n x n $\mathbf{x}$ no o $\kappa \ \mathbf{x} , \mathbf{x} \approx \delta$

B on $\mathcal{T}_1$ on no $\mathcal{T}_2$ $, u \in \mathcal{T}_2$ n n $j \in n$ n n on $\xi \in \mathcal{H}$

$$\kappa \quad , \cdot \quad \kappa \ \mathbf{x} , \cdot \quad \xi, \ u \quad y \ , \quad \text{n} \quad \langle \xi, \kappa \ \mathbf{x} , \cdot \rangle_{\mathcal{H}_\kappa} \quad , \forall i \in \ n \ .$$

o $, u \in \mathcal{T}_2$ o pon n $\mathbf{x} , y \in \mathcal{T}_1$ on p n o n o $\mathcal{T}_2$ n p o n m o $\mathcal{T}_1$

on T n n n m p $_1, u_1 , \ldots , \quad , u$ o m $\mathcal{T}_2$ o $\mathbf{x}, y \in \mathcal{T}_1$ T/n m p on o m n o o $\ell \ y, z$ n p $-yz$ n m p Fo m o on on o

$$f_* \quad \frac{R}{\sqrt{n}} \sum_{=1}^{n} y \ \kappa \ \mathbf{x} , \cdot$$

m n m m o on T n n m p

$$f_* \quad \underset{\| \ \|_{\mathcal{H}_\kappa} \leq}{\text{m n}} \sum_{=1} \ell \ u , f \qquad \underset{\| \ \|_{\mathcal{H}_\kappa} \leq}{\text{m n}} \sum_{=1} \ell \ u , \langle f, \kappa \quad , \cdot \rangle_{\mathcal{H}_\kappa} \qquad \underset{\| \ \|_{\mathcal{H}_\kappa} \leq}{\text{m n}} \sum_{=1}^{n} \sum_{=1}^{n} \ell \ y \ , \langle f, \kappa \ \mathbf{x} \ , \cdot \quad \xi \quad \rangle_{\mathcal{H}_\kappa} \ ,$$

n m n m o n

$$\epsilon \quad T \ \text{n} \quad \text{p} \ -R/\sqrt{n} \ .$$

C oo n

$$R \quad \sqrt{n} \ \text{n} \frac{T}{n},$$

$$\epsilon \le T \quad p \; -R/\sqrt{n} \qquad n.$$

op o on f_* n ppo o n o o O n

n p o n o A o Fo o no o n n o ppo o o on o o

$$O \; \epsilon \; R^2 \qquad O \; n \;\; n T^{\;2} \; .$$

Fn on n o o p q n o n $f_1', \ldots, f'$ no o n $n -$ ppo o B o on on o T/n n n p n o $\sum_{=1} \ell \; u \, , f'$ *larger than*

$$\frac{T}{n} \; \mathrm{n} \qquad \left(\frac{T}{n}\right).$$

o o f_* O n o o o o n T/n n n o n O n nT ^{2} o T

B. Proof of Lemma 1

B n n n q o n C B n o o n on o n

or *(Bernstein's inequality for martingales). Let* $X_1, \ldots, X_n$ *be a bounded martingale difference sequence with respect to the filtration* $\mathcal{F}$ $\mathcal{F}$ $_{1 \le \; \le n}$ *and wit*

C. Proof of Lemma 2

m fn on o X K n ^{2} n App n B No A n pp o n o ^{2} n o m nno pp o m o n n m o p n p o n B n

$$\left(\geq \sqrt{GA \ \tau} \ -K\tau \right)$$
$$\left(\geq \sqrt{GA \ \tau} \ -K\tau, A \ \leq G_1 T \right)$$
$$\left(\geq \sqrt{GA \ \tau} \ -K\tau, \ ^2 \leq GA \ , A \ \leq G_1 T \right)$$
$$\leq \left(\geq \sqrt{GA \ \tau} \ -K\tau, \ ^2 \leq GA \ , A \ \leq \frac{}{T} \right)$$
$$\sum_{=1} \left(\geq \sqrt{GA \ \tau} \ -K\tau, \ ^2 \leq GA \ , \frac{-1}{T} < A \ \leq \frac{}{T} \right)$$
$$\leq \left(A \ \leq \frac{}{T} \right) \sum_{=1} \left(\geq \sqrt{\frac{G}{T}\tau} \ -K\tau, \ ^2 \leq \frac{G}{T} \right)$$
$$\leq \left(A \ \leq \frac{}{T} \right) me^{-} ,$$

m $\lceil$ o $_2$ $G_1 T^2$ $\rceil$ n p o o B n n n q o m n o p p oo n τ n m/δ n n p on A $>$ $/T$