# Meta-Learning Hypothesis Spaces for Sequential Decision-making

Parnian Kassraie, Jonas Rothfuss, Andreas Krause

ETH zürich
Learning &
Adaptive Systems

# Confidence Sequences for Sequential Decision-making

noisy observ. actions

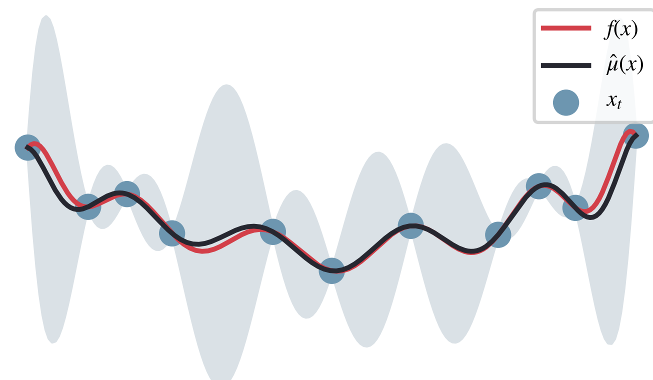
$$y_t = f^*(\boldsymbol{x}_t) + \varepsilon_t$$

unknown

Repeat

explore

exploit




Confidence sequences are great for guiding explorations!

E-processes are even better?

width $\longleftrightarrow$ current uncertainty

center $\longleftrightarrow$ current knowledge

# Problem Setting I (facts)

$$y_t = f^*(\boldsymbol{x}_t) + \varepsilon_t$$

$\boldsymbol{x}_t \in \mathcal{X}$, depends on the history $(\boldsymbol{x}_{1:t-1}, y_{1:t-1})$

$\mathcal{X} \subset \mathbb{R}^{d_0}$, compact

$\varepsilon_t$: $\sigma^2$ sub-Gussian, i.i.d.

$f^* : \mathcal{X} \to R$, $f^* \in \mathcal{H}_{k^*}$, $\|f^*\|_{k^*} \leq B$

$k^*$ unknown

We commit to confidence sets of the form,

$$\mathcal{C}_{t-1}(k; \boldsymbol{x}) = [\mu_{t-1}(k; \boldsymbol{x}) \pm \nu_t \sigma_{t-1}(k; \boldsymbol{x})]$$

$$\mu_{t-1}(k; \boldsymbol{x}) = \boldsymbol{k}_{t-1}^T(\boldsymbol{x})(\boldsymbol{K}_{t-1} + \bar{\sigma}^2 \boldsymbol{I})^{-1} \boldsymbol{y}_{t-1}$$

$$\sigma_{t-1}^2(k; \boldsymbol{x}) = k(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}_{t-1}^T(\boldsymbol{x})(\boldsymbol{K}_{t-1} + \bar{\sigma}^2 \boldsymbol{I})^{-1} \boldsymbol{k}_{t-1}(\boldsymbol{x})$$

Find $\hat{k}$ such that these sets are valid,

$$\mathbb{P}\left(\forall \boldsymbol{x} \in \mathcal{X},\, \forall t \geq 1 :\ f^*(\boldsymbol{x}) \in \mathcal{C}_{t-1}(\hat{k}; \boldsymbol{x})\right) \geq 1 - \delta$$

# Problem Setting II (luxuries)

$k^*$ unknown

$$k^*(\boldsymbol{x}, \boldsymbol{x}') = \sum_{j=1}^{p} \eta_j^* k_j(\boldsymbol{x}, \boldsymbol{x}')$$

$p < \infty$

$\eta_j^*$: unknown, non-negative

Better choice is $\eta_j^* \in \{0, 1\}$

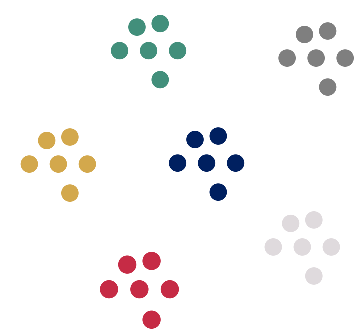
Data from similar tasks is available (fixed design)

$$y_{s,i} = f_s(\boldsymbol{x}_{s,i}) + \varepsilon_{s,i}$$

$1 \leq i \leq n$ and $1 \leq s \leq m$

$\varepsilon_{s,i}$: also $\sigma^2$ sub-Gussian, i.i.d.

$f_s : \mathcal{X} \to R$, $f_s \in \mathcal{H}_{k^*}$, $\|f_s\|_{k^*} \leq B$



# Meta-learning $\mathcal{H}_{k^*}$

Let $\boxed{\hat{k}}$ be the minimizer of

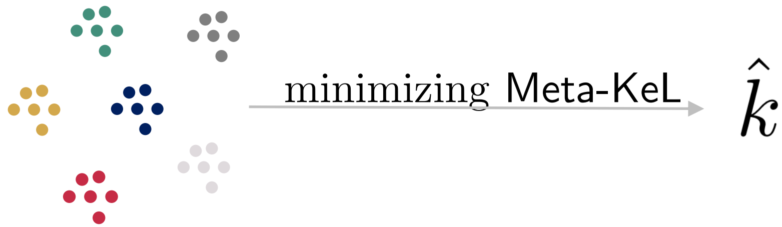
$$\min_{\boldsymbol{\eta}, f_1, \ldots, f_m} \frac{1}{m} \sum_{s=1}^{m} \left[ \frac{1}{n} \sum_{i=1}^{n} \left( y_{s,i} - f_s(\boldsymbol{x}_{s,i}) \right)^2 \right] + \frac{\lambda}{2} \sum_{s=1}^{m} \|f_s\|_k^2 + \frac{\lambda}{2} \|\boldsymbol{\eta}\|_1$$

$$\text{s.t. } \forall s : f_s \in \mathcal{H}_k,\ k = \sum_{j=1}^{p} \eta_j k_j,\ 0 \leq \boldsymbol{\eta} \qquad \text{(META-KEL)}$$

**Proposition**

*Meta-KeL is convex, has a solution and optimizing it is as difficult as the Group Lasso.*

# Properties of the meta-learned kernel

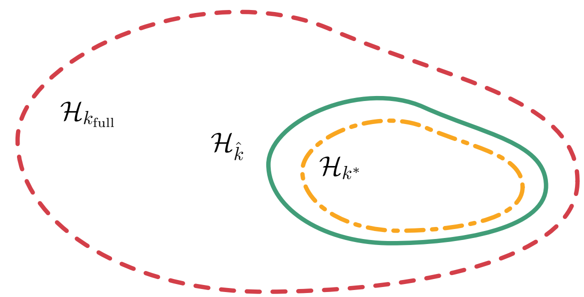



**Theorem (Hypothesis Space Recovery, Informal)**

*Let $0 < \delta < 1$. Assume meta-data satisfies the Restricted Eigenvalue Assumption and (a relaxed) Beta-min Condition. For the appropriate choice of $\lambda$, with probability greater than $1 - \delta$,*

- $\mathcal{H}_{k^*} \subseteq \mathcal{H}_{\hat{k}}$,
- $\|f\|_{\hat{k}} \leq \|f\|_{k^*}\Big(1 + \epsilon(n, m)\Big)$,
- *and $k^*$ and $\hat{k}$ have the same sparsity pattern, if $mn$ large.*

$$\epsilon(n, m) = O\left(1/\sqrt{mn}\left(\log p/\delta + \sqrt{md_{\max}\log p/\delta}\right)\right)$$




$$k_{\text{full}}(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{p}\sum_{j=1}^{p} k_j(\boldsymbol{x}, \boldsymbol{x}')$$

# The "meta-learned" confidence sequences

$$\mathcal{C}_{t-1}(\hat{k};\boldsymbol{x}) = [\mu_{t-1}(\hat{k};\boldsymbol{x}) \pm \nu_t \sigma_{t-1}(\hat{k};\boldsymbol{x})]$$

$$\nu_t = B\Big(1+\epsilon(n,m)\Big) + \sigma\sqrt{\hat{d}\log\left(1+\bar{\sigma}^{-2}t\right)+2+2\log(1/\delta)}$$

$$\bar{\sigma} = 1 + 2/t$$

**Theorem (Confidence Bounds with Meta-KeL)**

*Let $f \in \mathcal{H}_{k^*}$ with $\|f\|_{k^*} \leq B$, where $k^*$ is unknown. Under the assumptions of the previous theorem,*

$$\mathbb{P}\left(\forall \boldsymbol{x} \in \mathcal{X},\, \forall t \geq 1:\; f(\boldsymbol{x}) \in \mathcal{C}_{t-1}(\hat{k};\boldsymbol{x})\right) \geq 1-\delta.$$

The meta-learned confidence bounds approach the oracle bounds, as amounts of offline data grows.

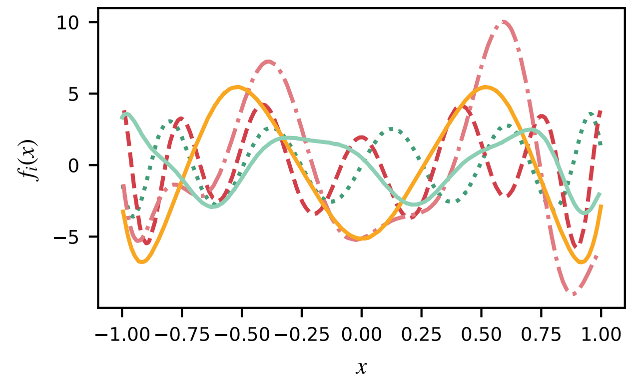
# How do the confidence sets look?

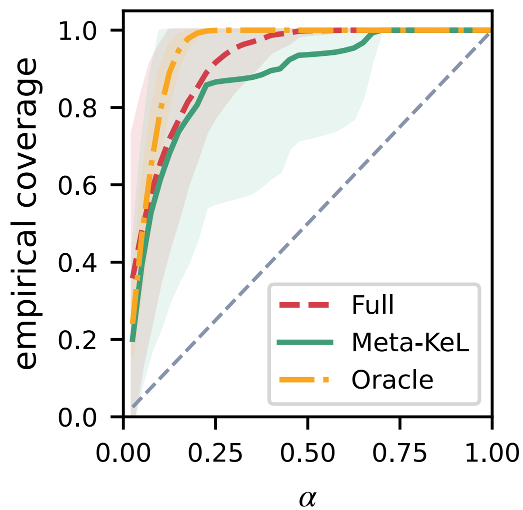
$$\mathcal{X} = [-1, 1]$$

$$k^*(x, x') = \sum_{j \in J_{k^*}} \eta_j^* P_j(x) P_j(x')$$

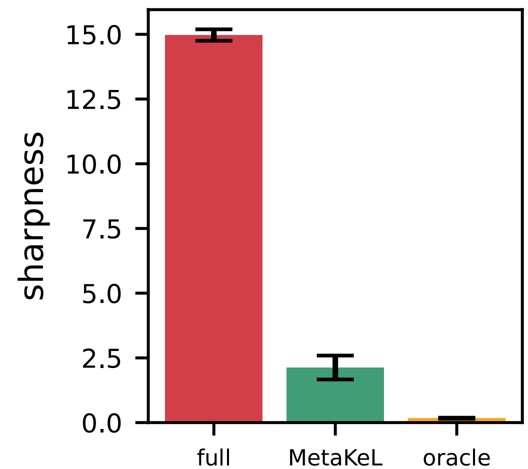
Legendre Polynomials

$$J_{k^*} \subset \{1, \cdots, 20\}, \ |J_{k^*}| = 5$$




Empirical coverage of the confidence band vs. the true $1 - \delta$



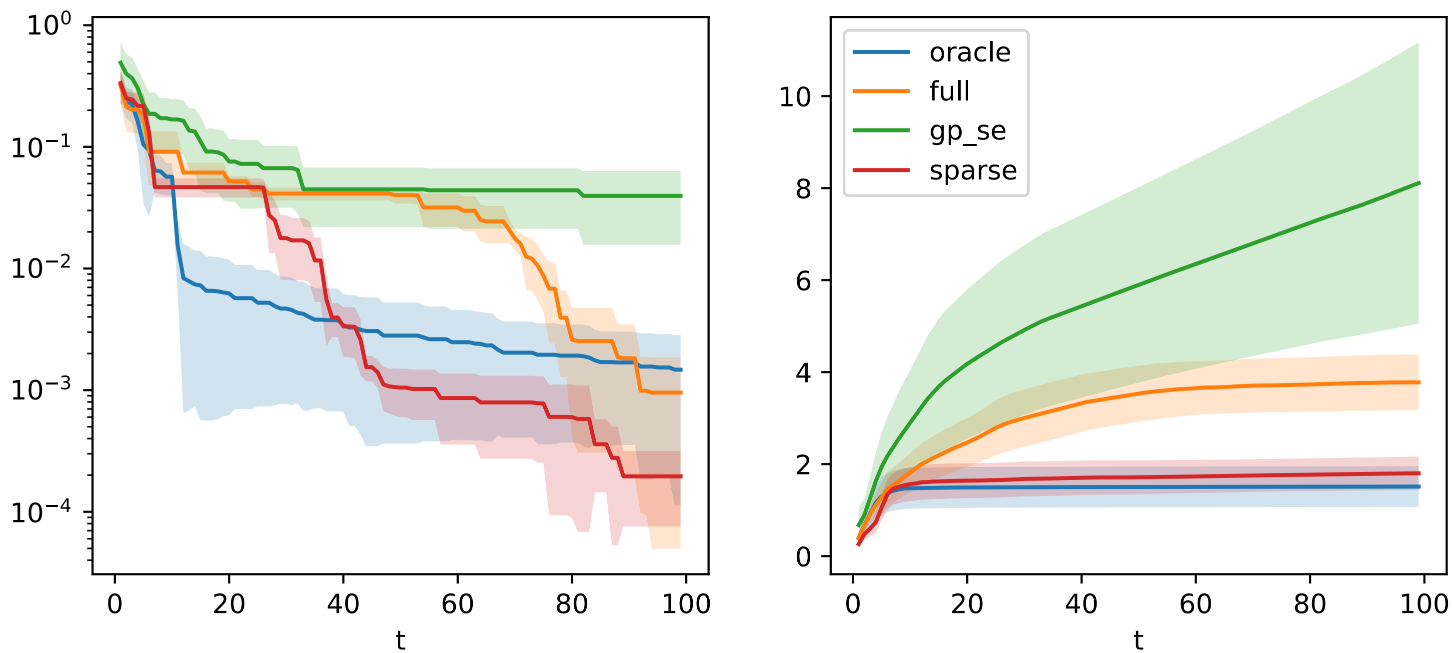

Empirical width of the confidence band

$$t = 4$$

# Thank you!

# Experiments

GLMNET data [Friedman et al 2010] + 2D Legendre features




$r_t(k) = f(x^*) - \max_{\tau \leq t} f(x_\tau)$ $\quad$ $R_T = \sum_{t=1}^{T} [f(\boldsymbol{x}^*) - f(\boldsymbol{x}_t)]$

# Experiments