A Nyström Estimator Error bound

Nyström estimator can easily approximate the kernel mean embedding ψ_{p_1}, ψ_{p_2} as well as the MMD distance between two distribution density p_1 and p_2. We need first assume the boundedness of the feature map to the kernel k:

Assumption 2. *There exists a positive constant $K \leq \infty$ such that* $\sup_{x \in \mathcal{X}} \|\phi(x)\| \leq K$

The true MMD distance between p_1 and p_2 is denoted as $\text{MMD}(p_1, p_2)$. The estimated MMD distance when using a Nyström sample size n_i, sub-sample size m_i for p_i respectively, is denoted as $\text{MMD}_{(p_i, m_i, n_i)}$. Then the error

$$\text{Err}_{(p_i, n_i, m_i)} := |\text{MMD}(p_1, p_2) - \text{MMD}_{(p_i, m_i, n_i)}|$$

and now we have the lemma from Theorem 5.1 in [8]

Lemma 1. *Let Assumption 2 hold. Furthermore, assume that for $i \in 1, 2$, the data points $X_1^i, \cdots, X_{n_i}^i$ are drawn i.i.d. from the distribution ρ_i and that $m_i \leq n_i$ sub-samples $\tilde{X}_1^i, \cdots, \tilde{X}_{m_i}^i$ are drawn uniformly with replacement from the dataset $\{X_1^i, \cdots, X_{n_i}^i\}$. Then, for any $\delta \in (0, 1)$, it holds with probability at least $1 - 2\delta$*

$$Err_{(p_i, n_i, m_i)} \leq \sum_{i=1,2} \left(\frac{c_1}{\sqrt{n_i}} + \frac{c_2}{m_i} + \frac{\sqrt{\log(m_i/\delta)}}{m_i} \sqrt{\mathcal{N}^{p_i}(\frac{12K^2 \log(m_i/\delta)}{m_i})} \right),$$

provided that, for $i \in \{1, 2\}$,

$$m_i \geq \max(67, 12K^2 \|C_i\|_{\mathcal{L}(\mathcal{H})}^{-1}) \log(m_i/\delta)$$

where $c_1 = 2K\sqrt{2\log(6/\delta)}$, $c_2 = 4\sqrt{3}K \log(12/\delta)$ and $c_4 = 6K\sqrt{\log(12/\delta)}$. The notation $\mathcal{N}^{p_i}$ denotes the effective dimension associated to the distribution p_k.

Specifically, when the effective dimension $\mathcal{N}$ satisfies, for some $c \geq 0$,

- *either $\mathcal{N}^{\rho_i}(\sigma^2) \leq c\sigma^{2-\gamma}$ for some $\gamma \in (0, 1)$,*
- *or $\mathcal{N}^{\rho_i}(\sigma^2) \leq \log(1 + c/\sigma^2)/\beta$, for some $\beta > 0$.*

Then, choosing the subsample size m to be

- *$m_i = n_i^{1/(2-\gamma)} \log(n_i/\delta)$ in the first case*
- *or $m_i = \sqrt{n_i} \log(\sqrt{n_i} \max(1/\delta, c/(6K^2))$ in the second case,*

we get $Err_{(\rho_i, n_i, m_i)} = O(1/\sqrt{n_i})$

B Proofs of Section 4

B.1 Exact kernel uncertainty $\mathcal{GP}$ formulating

Following the same notation in Section 4, now we can construct a Gaussian process $\mathcal{GP}(0, \hat{k})$ modelling functions over $\mathcal{P}$. This $\mathcal{GP}$ model can then be applied to learn $\hat{f}$ from a given set of observations $\mathcal{D}_n = \{(P_i, y_i)\}_{i=1}^n$. Under zero mean condition, the value of $\hat{f}(P_*)$ for a given $P_* \in \mathcal{P}$ follows a Gaussian posterior distribution with

$$\hat{\mu}_n(P_*) = \hat{\mathbf{k}}_n(P_*)^T (\hat{\mathbf{K}}_n + \sigma^2 \mathbf{I})^{-1} \mathbf{y}_n \tag{21}$$

$$\hat{\sigma}_n^2(P_*) = \hat{k}(P_*, P_*) - \hat{\mathbf{k}}_n(P_*)^T (\hat{\mathbf{K}}_n + \sigma^2 \mathbf{I})^{-1} \hat{\mathbf{k}}_n(P_*), \tag{22}$$

where $\mathbf{y}_n := [y_1, \cdots, y_n]^T$, $\hat{\mathbf{k}}_n(P_*) := [\hat{k}(P_*, P_1), \cdots, \hat{k}(P_*, P_n)]^T$ and $[\hat{\mathbf{K}}_n]_{ij} = \hat{k}(P_i, P_j)$.

Now we restrict our Gaussian process in the subspace $\mathcal{P}_\mathcal{X} = \{P_x, x \in \mathcal{X}\} \subset \mathcal{P}$. We assume the observation $y_i = f(x_i) + \zeta_i$ with the noise ζ_i. The input-induced noise is defined as $\Delta f_{p_{x_i}} := f(x_i) - \mathbb{E}_{P_{x_i}}[f] = f(x_i) - \hat{f}(P_{x_i})$. Then the total noise is $y_i - \mathbb{E}_{P_{x_i}}[f] = \zeta_i + \Delta f_{p_{x_i}}$. To formulate the regret bounds, we introduce the information gain given any $\{P_t\}_{t=1}^n \subset \mathcal{P}$:

$$\hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{P_t\}_{t=1}^n) := \frac{1}{2} \ln \det(\mathbf{I} + \sigma^{-2} \hat{\mathbf{K}}_n), \tag{23}$$

and the maximum information gain is defined as $\hat{\gamma}_n := \sup_{\mathcal{R} \in \mathcal{P}_\mathcal{X}; |\mathcal{R}|=n} \hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \mathcal{R})$. Here $\hat{\mathbf{f}}_n := [\hat{f}(p_1), \cdots, \hat{f}(p_n)]^T$.

We define the sub-Gaussian condition as follows:

Definition 1. *For a given $\sigma_\xi > 0$, a real-valued random variable ξ is said to be σ_ξ-sub-Gaussian if:*

$$\forall \lambda \in \mathbb{R}, \mathbb{E}[e^{\lambda \xi}] \leq e^{\lambda^2 \sigma_\xi^2 / 2} \tag{24}$$

Now we can state the lemma for bounding the uncertain-inputs regret of exact kernel evaluations, which is originally stated in Theorem 5 in [25].

Lemma 2. *Let $\delta \in (0, 1)$, $f \in \mathcal{H}_k$, and the corresponding $\|\hat{f}\|_{\hat{k}} \leq b$. Suppose the observation noise $\zeta_i = y_i - f(x_i)$ is conditionally σ_ζ-sub-Gaussian. Assume that both k and P_x satisfy the conditions for Δf_{P_x} to be σ_E-sub-Gaussian, for a given $\sigma_E > 0$. Then, we have the following results:*

- *The following holds for all $x \in \mathcal{X}$ and $t \geq 1$:*

$$|\hat{\mu}_n(P_x) - \hat{f}(P_x)| \leq \left(b + \sqrt{\sigma_E^2 + \sigma_\zeta^2} \sqrt{2 \left(\hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{P_t\}_{t=1}^n) + 1 + \ln(1/\delta) \right)} \right) \hat{\sigma}_n(P_x) \tag{25}$$

- *Running with upper confidence bound (UCB) acquisition function $\alpha(x|\mathcal{D}_n) = \hat{\mu}_n(P_x) + \hat{\beta}_n \hat{\sigma}_n(P_x)$ where*

$$\hat{\beta}_n = b + \sqrt{\sigma_E^2 + \sigma_\zeta^2} \sqrt{2 \left(\hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{P_t\}_{t=1}^n) + 1 + \ln(1/\delta) \right)}, \tag{26}$$

  and set $\sigma^2 = 1 + 2/n$, the uncertain-inputs cumulative regret satisfies:

$$\hat{R}_n \in O(\sqrt{n \hat{\gamma}_n} (b + \sqrt{\hat{\gamma}_n + \ln(1/\delta)})) \tag{27}$$

  with probability at least $1 - \delta$.

Note that although the original theorem restricted to the case when $\hat{k}(p, q) = \langle \psi_P, \psi_Q \rangle_k$, the results can be easily generated to other kernels over $\mathcal{P}$, as long as its universal w.r.t $C(\mathcal{P})$ given that $\mathcal{X}$ is compact and the mean map ψ is injective [11, 21].

B.2 Error estimates for inexact kernel approximation

Now let us derivative the inference under the introduce of inexact kernel estimations.

Theorem 2. *Under the Assumption 1 for $\varepsilon > 0$, let $\tilde{\mu}_n, \tilde{\sigma}_n, \tilde{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{P_t\}_{t=1}^n)$ as defined in* (14),(15),(16) *respectively, and $\hat{\mu}_n, \hat{\sigma}_n, \hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{P_t\}_{t=1}^n)$ as defined in* (21),(22),(23). *Assume* $\max_{x \in \mathcal{X}} f(x) = M$, *and assume the observation error $\zeta_i = y_i - f(x_i)$ satisfies $|\zeta_i| < A$ for all i. Then we have the following error bound holds with probability at least $1 - \varepsilon$:*

$$|\hat{\mu}_n(P_*) - \tilde{\mu}_n(P_*)| < (\frac{n}{\sigma^2} + \frac{n^2}{\sigma^4})(M + A) e_\varepsilon + O(e_\varepsilon^2) \tag{28}$$

$$|\hat{\sigma}_n^2(P_*) - \tilde{\sigma}_n^2(P_*)| < (1 + \frac{n}{\sigma^2})^2 e_\varepsilon + O(e_\varepsilon^2) \tag{29}$$

$$\left| \tilde{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{P_t\}_{t=1}^n) - \hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{P_t\}_{t=1}^n) \right| < \frac{n^{3/2}}{2\sigma^2} e_\varepsilon + O(e_\varepsilon^2) \tag{30}$$

Proof. Denote $e(P_*, Q) = \tilde{k}(P_*, Q) - \hat{k}(P_*, Q)$, $\mathbf{e}_n(P_*) = [e(P_*, P_1), \cdots, e(P_*.P_n)]^T$, and $[\mathbf{E}_n]_{i,j} = e(P_i, P_j)$. Now according to the matrix inverse perturbation expansion,

$$(X + \delta X)^{-1} = X^{-1} - X^{-1}\delta X X^{-1} + O(\|\delta X\|^2),$$

we have

$$(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I} + \mathbf{E}_n)^{-1} = (\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1} - (\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{E}_n(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1} + O(\|\mathbf{E}_n\|^2),$$

thus

$$\tilde{\mu}_n(P_*) = (\hat{\mathbf{k}}_n(P_*) + \mathbf{e}_n(P_*))^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I} + \mathbf{E}_n)^{-1}\mathbf{y}_n \tag{31}$$

$$= \hat{\mu}_n(P_*) + \mathbf{e}_n(P_*)^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{y}_n - \hat{\mathbf{k}}_n(P_*)^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{E}_n(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{y}_n \tag{32}$$

$$+ O(\|\mathbf{E}_n\|^2) + O(\|\mathbf{e}_n(P_*))\| \cdot \|\mathbf{E}_n\|) \tag{33}$$

$$\tilde{\sigma}_n^2(P_*) = \hat{\sigma}_n^2(P_*) + e(P_*, P_*) - (\hat{\mathbf{k}}_n(P_*) + \mathbf{e}_n(P_*))^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I} + \mathbf{E}_n)^{-1}(\hat{\mathbf{k}}_n(P_*) + \mathbf{e}_n(P_*)) \tag{34}$$

$$= \hat{\sigma}_n^2(P_*) + e(P_*, P_*) - 2\mathbf{e}_n(P)^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\hat{\mathbf{k}}_n(P_*) + \hat{\mathbf{k}}_n(P)^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{E}_n(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\hat{\mathbf{k}}_n(P_*) \tag{35}$$

$$+ O(\|\mathbf{E}_n\|^2) + O(\|\mathbf{e}_n\| \cdot \|\mathbf{E}_n\|) + O(\|\mathbf{e}_n\|^2 \cdot \|\mathbf{E}_n\|) \tag{36}$$

Notic that the following holds with a probability at least $1 - \varepsilon$,

$$|\mathbf{e}_n(P_*)^T(\hat{K}_n + \sigma^2\mathbf{I})^{-1}\mathbf{y}_n| \le \|\mathbf{e}_n(P_*)\|_2\|(\hat{K}_n + \sigma^2\mathbf{I})^{-1}\|_2\|\mathbf{y}_n\|_2 \le \frac{n}{\sigma^2}(M + A)e_\varepsilon, \tag{37}$$

$$|\hat{\mathbf{k}}_n(P_*)^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{E}_n(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{y}_n| \le \|\hat{\mathbf{k}}_n(P_*)\|_2\|(\hat{K}_n + \sigma^2\mathbf{I})^{-1}\|_2^2\|\mathbf{E}_n\|_2\|\mathbf{y}_n\|_2 \tag{38}$$

$$\le \sqrt{n}\sigma^{-4}ne_\varepsilon\sqrt{n}(M + A) = \frac{n^2}{\sigma^4}(M + A), \tag{39}$$

here we use the fact that $\hat{K}_n$ semi-definite (which means $\|(\hat{K}_n + \sigma^2 I)^{-1}\|_2 \le \sigma^{-2}$), $\hat{k}(P_*, P_*) \le 1$, $|y_i| \le M + A$. Combining these results, we have that

$$|\tilde{\mu}_n(P_*) - \hat{\mu}_n(P_*)| < (\frac{n}{\sigma^2} + \frac{n^2}{\sigma^4})(M + A)e_\varepsilon + O(e_\varepsilon^2),$$

holds with a probability at least $1 - \varepsilon$.

Similarly, we can conduct the same estimation to $\mathbf{e}_n(P)^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\hat{\mathbf{k}}_n(P_*)$ and $\hat{\mathbf{k}}_n(P)^T(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\mathbf{E}_n(\hat{\mathbf{K}}_n + \sigma^2\mathbf{I})^{-1}\hat{\mathbf{k}}_n(P_*)$, and get

$$|\tilde{\sigma}_n^2(P_*) - \hat{\sigma}_n^2(P_*)| < (1 + \frac{n}{\sigma^2})^2 e_\varepsilon + O(e_\varepsilon^2)$$

holds with a probability at least $1 - \varepsilon$.

It remains to estimate the error for estimating the information gain. Notice that, with a probability at least $1 - \varepsilon$,

$$\left|\tilde{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{p_t\}_{t=1}^n) - \hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n | \{p_t\}_{t=1}^n)\right| = \left|\frac{1}{2}\log\frac{\det(\mathbf{I} + \sigma^{-2}\tilde{\mathbf{K}}_n)}{\det(\mathbf{I} + \sigma^{-2}\hat{\mathbf{K}}_n)}\right| \tag{40}$$

$$= \left|\frac{1}{2}\log\det(\mathbf{I} - (\sigma^2\mathbf{I} + \hat{\mathbf{K}}_n)^{-1}\mathbf{E}_n)\right| \tag{41}$$

$$= \left|\frac{1}{2}\text{Tr}(\log(\mathbf{I} - (\sigma^2\mathbf{I} + \hat{\mathbf{K}}_n)^{-1}\mathbf{E}_n))\right| \tag{42}$$

$$= \left|\frac{1}{2}\text{Tr}(-(\sigma^2\mathbf{I} + \hat{\mathbf{K}}_n)^{-1}\mathbf{E}_n) + O(\|\mathbf{E}_n\|^2)\right| \tag{43}$$

$$\le \frac{n^{3/2}}{2\sigma^2}e_\varepsilon + O(\|\mathbf{E}_n\|^2), \tag{44}$$

here the second equation uses the fact that $\det(AB^{-1}) = \det(A)\det(B)^{-1}$, and the third and fourth equations use $\log\det(I+A) = \mathrm{Tr}\log(I+A) = \mathrm{Tr}(A - \frac{A^2}{2} + \cdots)$. The last inequality follows from the fact

$$\mathrm{Tr}(\sigma^2\mathbf{I} + \hat{\mathbf{K}}_n)^{-1}\mathbf{E}_n) \le \|(\sigma^2\mathbf{I} + \hat{\mathbf{K}}_n)^{-1}\|_F \|\mathbf{E}_n\|_F \le n^{3/2}\sigma^{-2}e_\varepsilon$$

and $\hat{\mathbf{K}}_n$ is semi-definite. □

With the uncertainty bound given by Lemma 2, let us prove that under inexact kernel estimations, the posterior mean is concentrated around the unknown reward function $\hat{f}$

Theorem 3. *Under the former setting as in Theorem 2, with probability at least $1 - \delta - \varepsilon$, let $\sigma_\nu = \sqrt{\sigma_\zeta^2 + \sigma_E^2}$, taking $\sigma = 1 + \frac{2}{n}$, the following holds for all $x \in \mathcal{X}$:*

$$|\tilde{\mu}_n(P_x) - \hat{f}(P_x)| \le \tilde{\beta}_n\tilde{\sigma}_n(P_x) + \left(\tilde{\beta}_n(1+n) + \tilde{\sigma}_n(P_x)\sigma_\nu n^{3/4}\right)e_\varepsilon^{1/2} + \left(n + n^2\right)(M + A)e_\varepsilon, \tag{45}$$

$$\text{where } \tilde{\beta}_n = \left(b + \sigma_\nu\sqrt{2(\tilde{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n|\{P_t\}_{t=1}^n) - \ln(\delta) + 1)}\right) \tag{46}$$

Proof. According to Lemma 2, equation (25), we have

$$|\hat{\mu}_n(P_x) - \hat{f}(P_x)| \le \hat{\beta}_n\hat{\sigma}_n(P_x) \tag{47}$$

with

$$\hat{\beta}_n = b + \sigma_\nu\sqrt{2\left(\hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n|\{P_t\}_{t=1}^n) + 1 + \ln(1/\delta)\right)}. \tag{48}$$

Notice that

$$|\tilde{\mu}_n(P_x) - \hat{f}(P_x)| \le |\tilde{\mu}_n(P_x) - \hat{\mu}_n(P_x)| + |\hat{\mu}_n(P_x) - \hat{f}(P_x)|, \tag{49}$$

$$\hat{\beta}_n = b + \sigma_\nu\sqrt{2\left(\hat{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n|\{P_t\}_{t=1}^n) + 1 + \ln(1/\delta)\right)} \tag{50}$$

$$\le b + \sigma_\nu\sqrt{2\left(\tilde{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n|\{P_t\}_{t=1}^n) + \frac{n^{3/2}}{2}e_\varepsilon + 1 + \ln(1/\delta)\right)} \tag{51}$$

$$\le b + \sigma_\nu\sqrt{2\left(\tilde{I}(\mathbf{y}_n; \hat{\mathbf{f}}_n|\{P_t\}_{t=1}^n) + 1 + \ln(1/\delta)\right)} + \sigma_\nu n^{3/4}e_\varepsilon^{1/2} \tag{52}$$

where the second inequality follows from Theorem 2, (30), and the third inequality follows from the inequality $\sqrt{a_1 + a_2} \le \sqrt{a_1} + \sqrt{a_2}, a_1 > 0, a_2 > 0$.

We also have (29), which means

$$\hat{\sigma}_n(P_x) = \sqrt{\hat{\sigma}_n(P_x)^2} \le \sqrt{\tilde{\sigma}_n(P_x)^2 + (1+n)^2 e_\varepsilon} \le \tilde{\sigma}_n(P_x) + (1+n)e_\varepsilon^{1/2}, \tag{53}$$

combining (28), (49), (50) and (53), we finally get the result in (45).

□

B.3 Proofs for Theorem 1

Now we can prove our main theorem 1.

Proof of Theorem 1. Let x^* maximize $\hat{f}(P_x)$ over $\mathcal{X}$. Observing that at each round $n \ge 1$, by the choice of x_n to maximize the aquisition function $\tilde{\alpha}(x|\mathcal{D}_{n-1}) = \tilde{\mu}_{n-1}(P_x) + \tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_x)$, we have

$$\tilde{r}_n = \hat{f}(P_{x^*}) - \hat{f}(P_{x_n}) \tag{54}$$
$$\leq \tilde{\mu}_{n-1}(P_{x^*}) + \tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_{x^*}) - \tilde{\mu}_{n-1}(P_{x_n}) + \tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_{x_n}) + 2Err(n-1, e_\varepsilon) \tag{55}$$
$$\leq 2\tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_{x_n}) + 2Err(n-1, e_\varepsilon). \tag{56}$$

Here we denote $Err(n, e_\varepsilon) := \left(\tilde{\beta}_n(1+n) + \tilde{\sigma}_n(P_x)\sigma_\nu n^{3/4}\right) e_\varepsilon^{1/2} + \left(n + n^2\right)(M + A)e_\varepsilon$. The second inequality follows from (45),

$$\hat{f}(P_{x^*}) - \tilde{\mu}_{n-1}(P_{x^*}) \leq \tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_{x^*}) + Err(n-1, e_\varepsilon) \tag{57}$$
$$\tilde{\mu}_{n-1}(P_{x_n}) - \hat{f}(P_{x_n}) \leq \tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_{x_n}) + Err(n-1, e_\varepsilon), \tag{58}$$

and the third inequality follows from the choice of x_n:

$$\tilde{\mu}_{n-1}(P_{x^*}) + \tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_{x^*}) \leq \tilde{\mu}_{n-1}(P_{x_n}) + \tilde{\beta}_{n-1}\tilde{\sigma}_{n-1}(P_{x_n}).$$

Thus we have

$$\tilde{R}_n = \sum_{t=1}^{n} \tilde{r}_t \leq 2\tilde{\beta}_n \sum_{t=1}^{n} \tilde{\sigma}_{t-1}(P_{x_t}) + \sum_{t=1}^{T} Err(t-1, e_\varepsilon). \tag{59}$$

From Lemma 4 in [9], we have that

$$\sum_{t=1}^{n} \tilde{\sigma}_{t-1}(P_{x_t}) \leq \sqrt{4(n+2)\ln\det(I + \sigma^{-2}\tilde{K}_n)} \leq \sqrt{4(n+2)\tilde{\gamma}_n},$$

and thus

$$2\tilde{\beta}_n \sum_{t=1}^{n} \tilde{\sigma}_{t-1}(P_{x_t}) = O\left(\sqrt{n\tilde{\gamma}_n} + \sqrt{n\tilde{\gamma}_n(\tilde{\gamma}_n - \ln\delta)}\right).$$

On the other hand, notice that

$$\sum_{t=1}^{n} Err(t-1, e_\varepsilon) = O\left((\sqrt{\tilde{\gamma}_n}n^2 + n^{7/4})e_\varepsilon + (n^2 + n^3)e_\epsilon\right),$$

we immediately get the result. □

C Evaluation Details

C.1 Implementation

In our implementation of AIRBO, we design the kernel k used for MMD estimation to be a linear combination of multiple Rational Quadratic kernels as its long tail behavior circumvents the fast decay issue of kernel [6]:

$$k(x, x') = \sum_{a_i \in \{0.2, 0.5, 1, 2, 5\}} \left(1 + \frac{(x - x')^2}{2a_i l_i^2}\right)^{-a_i}, \tag{60}$$

where l_i is a learnable lengthscale and a_i determines the relative weighting of large-scale and small-scale variations.

Depending on the form of input distributions, the sampling and sub-sampling sizes for Nyström MMD estimator are empirically selected via experiments. Moreover, as the input uncertainty is already modeled in the surrogate, we employ a classic UCB-based acquisition as Eq. 5 with $\beta = 2.0$ and maximize it via an L-BFGS-B optimizer.

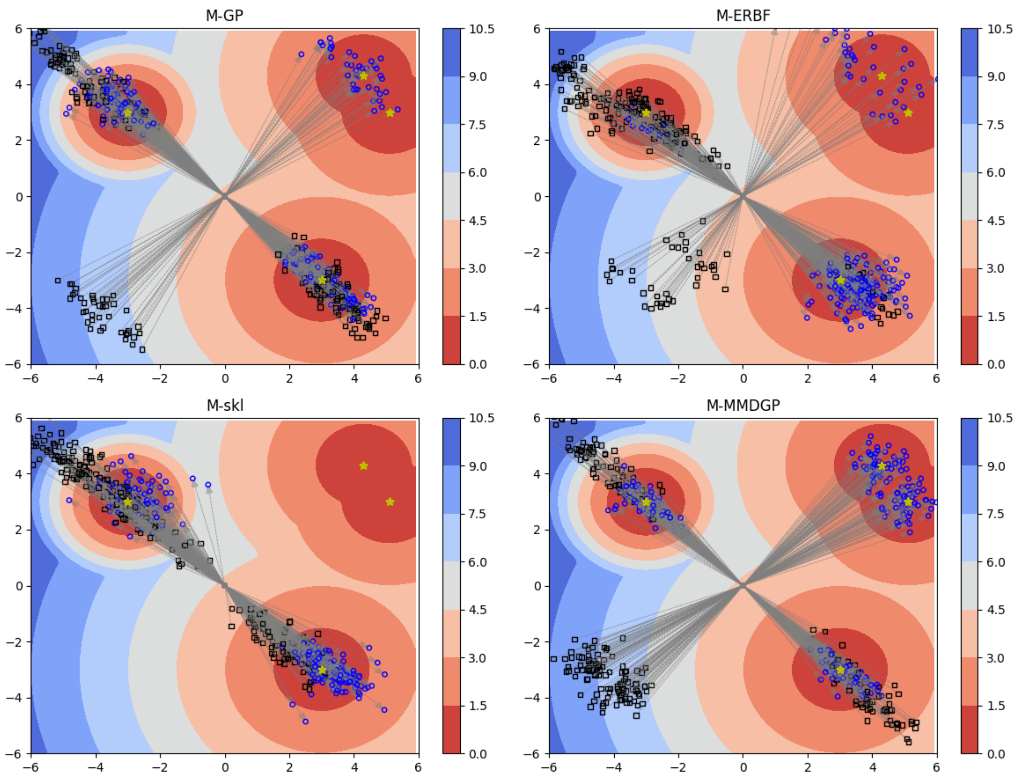

Figure 7: Simulation results of the push configurations found by different algorithms.

C.2 Supplementary Experiments

Robust Robot Pushing: This benchmark is based on a Box2D simulator from [30], where our objective is to identify a robust push configuration, enabling a robot to push a ball to predetermined targets under input randomness. In our experiment, we simplify the task by setting the push angle to $r_a = \arctan \frac{r_y}{r_x}$, ensuring the robot is always facing the ball. Also, we intentionally define the input distribution as a two-component Gaussian Mixture Model as follows:

$$(r_x, r_y, r_t) \sim GMM\Big(\mu = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} 0.1^2 & -0.3^2 & 1e-6 \\ -0.3^2 & 0.1^2 & 1e-6 \\ 1e-6 & 1e-6 & 1.0^2 \end{bmatrix}, w = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}\Big),$$

where the covariance matrix Σ is shared among components and w is the weights of mixture components. Figure 5b shows some example samples from this GMM distribution. Meanwhile, as the SKL-UCB and ERBF-UCB surrogates can only accept Gaussian input distributions, we choose to approximate the true input distribution with a Gaussian. As shown in Figure 5b, the approximation error is obvious, which explains the performance gap among these algorithms in Figure 5c.

Apart from the statistics of the found pre-images in Figure 6, we also simulate the robot pushes according to the found configurations and visualize the results in Figure 7. In this figure, each black hollow square represents an instance of the robot's initial location, the grey arrow indicates the push direction and duration, and the blue circle marks the ball's ending position after the push. We can find that, as the GP-UCB ignores the input uncertainty, it randomly pushes to these targets and the ball ending positions fluctuate. Also, due to the incorrect assumption of the input distribution, the SKL-UCB and ERBF-UCB fail to control the ball's ending position under input randomness. On the contrary, AIRBO successfully recognizes the twin targets in quadrant I as an optimal choice and frequently pushes to this area. Moreover, all the ball's ending positions are well controlled and centralized around the targets under input randomness.