Batch Bayesian Optimization on Permutations using the Acquisition Weighted Kernel - Supplementary Material -

Changyong Oh
QUvA lab, IvI
University of Amsterdam
changyong.oh0224@gmail.com

Roberto Bondesan
Qualcomm AI Research[‡‡]
rbondesa@qti.qualcomm.com

Efstratios Gavves
QUvA lab, IvI
University of Amsterdam
egavves@uva.nl

Max Welling
QUvA lab, IvI
University of Amsterdam
m.welling@uva.nl

A Regret Analysis

In this section, we show that LAW with GP-UCB or EST has the vanishing simple regret with high probability.

In Bayesian optimization (BO), the goal is to find a minimum for a given objective f

$$\mathbf{x}^\star = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x})$$

First, we introduce different types of regret. Our analysis on the vanishing simple regret of LAW only requires batch version of all regrets below. Therefore, the proof for the vanishing simple regret can be read without referring to sequential version of regrets below. The sequential version definitions are used when we contrast our regret analysis with the regret analysis in existing works [DKB14, KKSP18].

We begin with two equivalent round indexing in *the batch setting*, the sequential indexing, an 1-tuple and the batch indexing, an ordered 2-tuple which are related via following mappings.

$$\begin{aligned} \mathfrak{T}_{bat}^{(B)} &: \mathbb{N} \to \mathbb{N} \times [B] & t &\mapsto ([(t-1)\ \mathbf{mod}\ B] + 1, [(t-1)\ \mathbf{rem}\ B] + 1) \\ \mathfrak{T}_{seq}^{(B)} &: \mathbb{N} \times [B] \to \mathbb{N} & (t,b) &\mapsto (t-1) \times B + b \end{aligned}$$

The batch indexing is primarily used and the sequential indexing is expressed via $\mathfrak{T}_{seq}^{(B)}$.

With two indexing, we have batch and sequential versions of regret definitions with the instantaneous regret $r_{t,b} = f(\mathbf{x}^\star) - f(\mathbf{x}_{t,b})$ at a query point $\mathbf{x}_{t,b}$. In the case of a noisy objective, $y_{t,b} = f(\mathbf{x}_{t,b}) + \epsilon_{t,b}$ is a corresponding evaluation with a noise $\epsilon_{t,b}$.
Note that we use the definition of sequential simple/cumulative regret in the context of batch BO. Since sequential simple regret is equal to batch simple regret, we call both simple regret without prefixes. In [CBRV13], sequential cumulative regret is termed full cumulative regret to contrast with batch cumulative regret.

The simple regret is in accord with the goal of BO[KKSP18] while the cumulative regret is prevalent in bandit[LS20].

[‡‡]Qualcomm AI Research is an initiative of Qualcomm Technologies, Inc.

Type	Batch	Sequential
Instantaneous	$r_t^{(B)} = \min_{b=1,\cdots,B} r_{t,b}$	$r_{\mathfrak{T}_{seq}^{(B)}(t,b)} = r_{t,b}$
Simple	$S_T^{(B)} = \min_{t=1,\cdots,T} r_t^{(B)}$	$S_{\mathfrak{T}_{seq}^{(B)}(T,b)} = \min_{\mathfrak{T}_{seq}^{(B)}(t,b') \leq \mathfrak{T}_{seq}^{(B)}(T,b)} r_{T_{seq}^{(B)}(t,b')}$
Cumulative	$R_T^{(B)} = \sum_{t=1}^{T} r_t^{(B)}$	$R_{\mathfrak{T}_{seq}^{(B)}(T,b)} = \sum_{\mathfrak{T}_{seq}^{(B)}(t,b') \leq \mathfrak{T}_{seq}^{(B)}(T,b)} r_{T_{seq}^{(B)}(t,b')}$
Simple & Cumulative	$S_T^{(B)} \leq \frac{1}{T} R_T^{(B)}$	$S_{\mathfrak{T}_{seq}^{(B)}(T,b)} \leq \frac{1}{\mathfrak{T}_{seq}^{(B)}(T,b)} R_{\mathfrak{T}_{seq}^{(B)}(T,b)}$
Between Simple	$S_T^{(B)} = S_{\mathfrak{T}_{seq}^{(B)}(T,B)}$	
Between Cumulative	$R_T^{(B)} \leq \frac{1}{B} R_{\mathfrak{T}_{seq}^{(B)}(T,B)}$	

Table 3: Types of regrets

When an algorithm exhibits that cumulative regret averaged over rounds converges to zero, then the algorithm is called no regret. As stated in Table 3, simple regret is bounded above by cumulative regret averaged over rounds, therefore, vanishing simple regret is often proved by showing that the algorithm is no regret[KKSP18].

In batch BO, there are two types of query depending on the accessible information. *Non-delayed* query point uses all previous query points with all corresponding evaluations, e.g. $\{\mathbf{x}_{t,1}\}_{t\in[T]}$ in LAW while *delayed* query point uses all previous query points but some of evaluations for previous query points are not used, e.g. $\{\mathbf{x}_{t,b}\}_{t\in[T],b=2,\cdots,B}$ in LAW. For evaluation, instantaneous regret and posterior variance, we can say *non-delayed* and *delayed* according to the query point with which it is defined.

The our analysis consists of steps below

1. Bound batch cumulative regret with the sum of non-delayed instantaneous regrets, i.e, the regrets from the first points in each batch
$$R_T^{(B)} = \sum_{t=1}^{T} r_t^{(B)} \leq \sum_{t=1}^{T} r_{t,1}$$
2. Bound non-delayed instantaneous regrets with non-delayed posterior variance, i.e, posterior variance conditioned on all previous query points with their evaluations.
$$\sum_{t=1}^{T} r_{t,1} \leq \sum_{t=1}^{T} \eta_t \sigma_{t-1,1}(\mathbf{x}_{t,1})$$
η_t depends on the acquisition function and the details are given in Theorem A.18.
3. Bound non-delayed posterior variance with all posterior variance (Lemma A.4)
$$\sum_{t=1}^{T} \sigma_{t-1,1}(\mathbf{x}_{t,1}) \leq 1 + \frac{w_+}{w_-}\frac{1}{B}\sum_{t=1}^{T}\sum_{b=1}^{B} \sigma_{t-1,b}(\mathbf{x}_{t,b})$$

While the regret analysis on sequential cumulative regret [DKB14, KKSP18][10] requires high probability confidence interval for $r_{t,b}$ for all $t \in [T]$ and $b \in [B]$, our analysis on batch cumulative regret requires high probability confidence interval for $r_{t,1}$ for all $t \in [T]$. More detailed discussion on the differences between two approaches is given after the proof (see Subsection A.2).

A.1 Vanishing simple regret of LAW

In Bayesian optimization using LAW, the surrogate model is Gaussian processes with a kernel K. At t-th round of batch Bayesian optimization with the batch size of B, we have L_t^{AW} which defines L-ensembles of k-DPP. L_t^{AW} is obtained using the product of the predictive covariance function

[10] In [KKSP18], vanishing simple regret is proved by showing that a bound with sequential cumulative regret averaged over rounds converges to zero.

of K conditioned on $\mathcal{D}_{t-1} = \{(\mathbf{x}_{s,b}, y_{s,b})\}_{s\in[t-1],b\in[B]}$ and the acquisition function a_t using the evaluation data $\mathcal{D}_{t-1}$ as follows

$$L_t^{AW}(\mathbf{x}, \mathbf{x}') = w(a_t(\mathbf{x})) \cdot L_t(\mathbf{x}, \mathbf{x}') \cdot w(a_t(\mathbf{x}')). \tag{9}$$

where $L_t(\mathbf{x}, \mathbf{x}') = K(\mathbf{x}, \mathbf{x}' \,|\, \mathcal{D}_{t-1})$ is the *diversity gauge* and $w : \mathbb{R} \to \mathbb{R}$ is positive increasing, $w_- = \inf_{x\in\mathbb{R}} w(x) > 0$ and $w_+ = \sup_{x\in\mathbb{R}} w(x) < \infty$, which we call the *weight function*.

The batch with B points $\mathbf{x}_{t,1}, \cdots, \mathbf{x}_{t,B}$ are acquired by

$$\begin{aligned}
\mathbf{x}_{t,1} &= \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} a_t(\mathbf{x}) = \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} w(a_t(\mathbf{x})) = \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} \log w(a_t(\mathbf{x}))^2 \\
\mathbf{x}_{t,b} &= \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} \log\det([L_t^{AW}(\mathbf{x}, \mathbf{x})]_{\{\mathbf{x}_i\}_{i\in[b-1]}\cup\{\mathbf{x}\}}) \\
&= \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} \log(L_t(\mathbf{x}, \mathbf{x} \,|\, \{\mathbf{x}_{t,i}\}_{i\in[b-1]}) \cdot w(a_t(\mathbf{x}))^2)
\end{aligned} \tag{10}$$

where $L_t(\mathbf{x}, \mathbf{x} \,|\, \{\mathbf{x}_{t,i}\}_{i\in[b-1]})$ is the posterior variance of the kernel L_t conditioned on $\{\mathbf{x}_{t,i}\}_{i\in[b-1]}$.

Note that the posterior variance of the *posterior covariance function* K_t conditioned on $\{\mathbf{x}_{t,i}\}_{i\in[b-1]}$ is equal to the posterior variance of the *prior covariance function* K conditioned on $\mathcal{D}_{t-1} \cup \{\mathbf{x}_{t,i}\}_{i\in[b-1]}$.

In the rest of the section, we use below simpler notation

$$\sigma_{t-1,b}^2(\mathbf{x}) = \begin{cases} L_t(\mathbf{x}, \mathbf{x}) = K(\mathbf{x}, \mathbf{x} \,|\, \mathcal{D}_{t-1}) & b = 1 \\ L_t(\mathbf{x}, \mathbf{x} \,|\, \{\mathbf{x}_{t,i}\}_{i\in[b-1]}) = K(\mathbf{x}, \mathbf{x} \,|\, \mathcal{D}_{t-1} \cup \{\mathbf{x}_{t,i}\}_{i\in[b-1]}) & b = 2, \cdots, B \end{cases} \tag{11}$$

$$\mu_t(\mathbf{x}) \text{ is the predictive mean conditioned on } \mathcal{D}_{t-1}\,. \tag{12}$$

Note that $\sigma_{t-1,b}^2$ is well defined for $b = 2, \cdots, B$ since the posterior variance does not depend on output values while the predictive mean is defined only when $b = 1$ where evaluated output $y_{s,b}$ exists for each $\mathbf{x}_{s,b}$ in conditioning data.

We start with lemmas used in the regret bound analysis.

Lemma A.1. *Assume a kernel such that $K(\cdot, \cdot) \leq 1$. For each $t \in [T]$, LAW acquires a batch using the evaluation data $\mathcal{D}_{t-1}$, the diversity measure $L_t(\cdot, \cdot) = K(\cdot, \cdot \,|\, \mathcal{D}_{t-1})$, an acquisition function $a_t(\cdot)$ and a weight function $w(\cdot)$ (as defined below Eq. 9). The posterior variance defined as Eq. 11. has the following relation*

$$\sigma_{t,1}(\mathbf{x}_{t+1,1}) \leq \frac{w_+}{w_-} \sigma_{t-1,b}(\mathbf{x}_{t,b}) \quad 1 \leq t \leq T, 2 \leq b \leq B$$

Proof. By the definition of $\mathbf{x}_{t,b}$

$$\mathbf{x}_{t,b} = \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} \log(L_t(\mathbf{x} \,|\, \{\mathbf{x}_{t,i}\}_{i\in[b-1]}) \cdot w(a_t(\mathbf{x}))^2) = \operatorname*{argmax}_{\mathbf{x}\in\mathcal{X}} \log(\sigma_{t-1,b}^2(\mathbf{x}) \cdot w(a_t(\mathbf{x}))) \tag{13}$$

we have

$$w(a_t(\mathbf{x}))\sigma_{t-1,b}(\mathbf{x}) \leq w(a_t(\mathbf{x}_{t,b}))\sigma_{t-1,b}(\mathbf{x}_{t,b}) \quad \forall\, \mathbf{x} \in \mathcal{X} \tag{14}$$

thus

$$\sigma_{t-1,b}(\mathbf{x}) \leq \frac{w(a_t(\mathbf{x}))}{w(a_t(\mathbf{x}_{t,b}))} \sigma_{t-1,b}(\mathbf{x}_{t,b}) \leq \frac{w_+}{w_-} \sigma_{t-1,b}(\mathbf{x}_{t,b}) \quad \forall\, \mathbf{x} \in \mathcal{X} \tag{15}$$

By the "Information never hurts" principle[KSG08], i.e. the posterior variance decreases as the conditioning set increases, we have

$$\sigma_{t,1}(\mathbf{x}) \leq \sigma_{t-1,b}(\mathbf{x}) \quad \forall\, \mathbf{x} \in \mathcal{X}$$

since σ_t is conditioned by $\mathcal{D}_t = \mathcal{D}_{t-1} \cup \{\mathbf{x}_{t,i}\}_{i\in[B]}$ while $\sigma_{t,b}$ is conditioned by $\mathcal{D}_{t-1} \cup \{\mathbf{x}_{t,i}\}_{i\in[b-1]}$. Combining these two inequalities, we have

$$\sigma_{t,1}(\mathbf{x}) \leq \sigma_{t-1,b}(\mathbf{x}) \leq \frac{w_+}{w_-} \sigma_{t-1,b}(\mathbf{x}_{t,b}) \quad \forall\, \mathbf{x} \in \mathcal{X}$$

which also applies when $\mathbf{x} = \mathbf{x}_{t+1,1}$.

Q.E.D. □

Remark A.2. LAW does not use the heuristic called the relevant region[CBRV13, KDK16], which makes the proof simpler compared with the Lemma 6.5 in [KDK16].

Remark A.3. The Lemma 6.5 in [KDK16] claims that the inequality similar to Eq. 14 and Eq. 15 holds for sampling(DPP-SAMPLE). However, such inequality relies on fact that $\mathbf{x}_{t,b}$ is the maximum of an objective which is not guaranteed to hold for sampling(DPP-SAMPLE). The regret analysis of DPP-SAMPLE in [KDK16] appears to need a revision. In our version, we do not make any claim in the case of sampling.

Lemma A.4. *Assume a kernel such that $K(\cdot,\cdot) \le 1$. For each $t \in [T]$, LAW acquires a batch using the evaluation data $\mathcal{D}_{t-1}$, the diversity measure $L_t(\cdot,\cdot) = K(\cdot,\cdot|\mathcal{D}_{t-1})$, an acquisition function $a_t(\cdot)$ and a weight function $w(\cdot)$ (as defined below Eq. 9). The posterior variance defined as Eq. 11. has the following relation*

$$\sum_{t=1}^{T} \sigma_{t-1,1}(\mathbf{x}_{t,1}) \le 1 + \frac{w_+}{w_-}\frac{1}{B}\sum_{t=1}^{T}\sum_{b=1}^{B} \sigma_{t-1,b}(\mathbf{x}_{t,b}). \tag{16}$$

Proof. From Lemma A.1, for $b = 2, \cdots, B$, we have

$$\sigma_{t,1}(\mathbf{x}_{t+1,1}) = \sigma_t(\mathbf{x}_{t+1,1}) \le \frac{w_+}{w_-}\sigma_{t-1,b}(\mathbf{x}_{t,b})$$

Summing these for $b = 2, \cdots, B$ and $\sigma_{t-1,1}(\mathbf{x}_{t,b})$

$$\sigma_{t-1,1}(\mathbf{x}_{t,b}) + (B-1)\sigma_{t,1}(\mathbf{x}_{t+1,1}) \le \frac{w_+}{w_-}\sum_{b=1}^{B}\sigma_{t-1,b}(\mathbf{x}_{t,b})$$

since $w_- \le w_+$. Summing this with respect to t, we have

$$\sum_{t=1}^{T}\sigma_{t-1,1}(\mathbf{x}_{t,b}) + (B-1)\sum_{t=1}^{T}\sigma_{t,1}(\mathbf{x}_{t+1,1}) \le \frac{w_+}{w_-}\sum_{t=1}^{T}\sum_{b=1}^{B}\sigma_{t-1,b}(\mathbf{x}_{t,b})$$

The term on the left hand side can be rewritten

$$B\sum_{t=1}^{T}\sigma_{t-1,1}(\mathbf{x}_{t,b}) + (B-1)(\sigma_{T,1}(\mathbf{x}_{T+1,1}) - \sigma_{0,1}(\mathbf{x}_{1,1})) \tag{17}$$

Since $(B-1)(\sigma_{0,1}(\mathbf{x}_{1,1}) - \sigma_{T,1}(\mathbf{x}_{T+1,1})) \le B\sigma_{0,1}(\mathbf{x}_{1,1})$

$$\sum_{t=1}^{T}\sigma_{t-1,1}(\mathbf{x}_{t,b}) \le \sigma_{0,1}(\mathbf{x}_{1,1}) + \frac{w_+}{w_-}\frac{1}{B}\sum_{t=1}^{T}\sum_{b=1}^{B}\sigma_{t-1,b}(\mathbf{x}_{t,b})$$

Q.E.D. □

Remark A.5. In Lemma 3 in [CBRV13] and Lemma 6.5 in [KDK16], the second term in Eq. 17 is ignored. However, $\sigma_{T,1}(\mathbf{x}_{T+1,1}) - \sigma_{0,1}(\mathbf{x}_{1,1})$ can be negative, which should not be ignored. Nevertheless, this error does not change the regret analysis in [CBRV13] because constant terms divided by T vanishes. Our version has the additional constant 1 on the right hand side of Eq. 16.

Definition A.6. The maximum information gain γ_T is defined as below

$$\gamma_T = \gamma(T; \mathcal{X}) = \max_{X \subset \mathcal{X}, |X|=T} \mathbf{I}(Y_X; \mathbf{f}_X) = \max_{X \subset \mathcal{X}, |X|=T} H(Y_X) - H(Y_X|\,\mathbf{f}_X)$$

where Y is the observation at X and H is the differential entropy.

For Gaussian processes with the kernel K and the variance of observation noise σ^2

$$\gamma_T = \gamma(T; \mathcal{X}, K, \sigma^2) = \max_{X \subset \mathcal{X}, |X|=T} \frac{1}{2}\log\det(I + \sigma^{-2}K(X,X))$$

We rephrase lemmas from previous works with the batch indexing for notational ease and discuss the noteworthy points in the rephrased version compared with the original ones.

Lemma A.7 (Lemma 3 [SKKS09], Lemma 4 [CBRV13], Theorem 3.1 [WZJ16]). *Assume a kernel such that $K(\cdot,\cdot) \leq 1$. For each $t \in [T]$, LAW acquires a batch using the evaluation data $\mathcal{D}_{t-1}$, the diversity measure $L_t(\cdot,\cdot) = K(\cdot,\cdot|\mathcal{D}_{t-1})$, an acquisition function $a_t(\cdot)$ and a weight function $w(\cdot)$ (as defined below Eq. 9). The posterior variance defined as Eq. 11. has the following relation*

$$\sum_{t=1}^{T}\sum_{b=1}^{B}\sigma^2_{t-1,b}(\mathbf{x}_{t,b}) \leq C_1\gamma_{TB}$$

where $C_1 = \frac{2}{\log(1+\sigma^{-2})}$ and γ_{TB} is the maximum information gain at TB

Proof. Following the trick used in the proof of Lemma 5.4 in [SKKS09],

$$\sigma^2_{t-1,b}(\mathbf{x}) = \sigma^2\sigma^{-2}\sigma^2_{t-1,b}(\mathbf{x}) \leq \frac{1}{\log(1+\sigma^{-2})}\log(1+\sigma^{-2}\sigma^2_{t-1,b}(\mathbf{x})). \tag{18}$$

In LAW, $\mathbf{x}_{t,1}$ and $\mathbf{x}_{t,b}$ deterministic conditioned respectively on $\mathcal{D}_{t-1} = \{(\mathbf{x}_{s,b}, y_{s,b})\}_{s\in[t-1],b\in[B]}$ and $\mathcal{D}_{t-1} \cup \{\mathbf{x}_{t,c}\}_{c=2,\cdots,b-1}$ for $b = 2,\cdots,B$. Also, $\mathbf{x}_{t,b}$ does not depend on $\{y_{s,c}\}_{\mathfrak{T}^{(B)}_{seq}(s,c)<\mathfrak{T}^{(B)}_{seq}(t,b)}$ as long as $\{\mathbf{x}_{s,c}\}_{\mathfrak{T}^{(B)}_{seq}(s,c)\leq\mathfrak{T}^{(B)}_{seq}(t,b)}$. Therefore, the proof of Lemma 5.3 in [SKKS09] can be applied

$$\begin{aligned}\sum_{t=1}^{T}\sum_{b=1}^{B}\sigma^2_{t-1,b}(\mathbf{x}_{t,b}) &\leq \frac{1}{\log(1+\sigma^{-2})}\sum_{t=1}^{T}\sum_{b=1}^{B}\log(1+\sigma^{-2}\sigma^2_{t-1,b}(\mathbf{x}_{t,b}))\\ &= \frac{2}{\log(1+\sigma^{-2})}\mathbf{I}(Y_{\{\mathbf{x}_{t,b}\}_{t\in[T],b\in[B]}};\mathbf{f}_{\{\mathbf{x}_{t,b}\}_{t\in[T],b\in[B]}}) \leq \frac{2}{\log(1+\sigma^{-2})}\gamma_{TB}\end{aligned}$$

□

Remark A.8. In contrast to Lemma 5.4 in [SKKS09] which bounds the sum of square of regrets, Lemma A.7 bounds the sum of the posterior variances. The delayed evaluation $\{y_{t,b}\}_{t\in[T],b\in[B]}$ does not cause any impediment in the proof.

Lemma A.9 (Lemma 6.1 [KDK16], Lemma 3.2 [WZJ16]). *For $\zeta_t = \left(2\log\left(\frac{\pi_t^2}{2\delta}\right)\right)^{1/2}$ with $\delta \in (0,1)$ and $\pi_t > 0$ such that $\sum_{t=1}^{\infty}\pi_t \leq 1$, an arbitrary sequence of actions $\mathbf{x}_{1,1},\cdots,\mathbf{x}_{T,1} \in \mathcal{X}$*

$$P\left(\bigcap_{t\in[T]}\left\{f \,\middle|\, |f(\mathbf{x}_{t,1}) - \mu_{t-1}(\mathbf{x}_{t,1})| \leq \zeta_t\cdot\sigma_{t-1,1}(\mathbf{x}_{t,1})\right\}\right) \geq 1-\delta.$$

Remark A.10. ζ_t only depends on the number of batch round t and is independent with the batch size B. Therefore, ζ_t in batch BO is the same as one in the sequential BO.

Lemma A.11 (Lemma 3.3 [WZJ16]). *If $|f(\mathbf{x}_{t,1}) - \mu_{t-1}(\mathbf{x}_{t,1})| \leq \zeta_t\sigma_{t-1,1}(\mathbf{x}_{t,1})$*

$$r_{t,1} = f(\mathbf{x}_{t,1}) - f(\mathbf{x}^\star) \leq (\nu_t + \zeta_t)\sigma_{t-1,1}(\mathbf{x}_{t,1})$$

where $\nu_t = \left(\min_{\mathbf{x}\in\mathcal{X}}\frac{\mu_{t-1}(\mathbf{x})-\hat{m}}{\sigma_{t-1,1}(\mathbf{x})}\right)$, $\hat{m}$ is the estimate of the optimum[WZJ16] and $\zeta_t = \left(2\log\left(\frac{\pi_t^2}{2\delta}\right)\right)^{1/2}$ with $\delta \in (0,1)$ and $\pi_t > 0$ such that $\sum_{t=1}^{\infty}\pi_t^{-1} \leq 1$.

Remark A.12. In Lemma A.11, we only bound regrets in $(t,1)$-th round where there is no delayed evaluation.

Remark A.13. In contrast to the original condition $\sum_{t=1}^{T}\pi_t \leq 1$, we use $\sum_{t=1}^{\infty}\pi_t \leq 1$ so that π_ts are T independent as the recommendation of the choice $\pi_t = \frac{1}{6}\pi^2t^2$ in [WZJ16]. When the number of rounds T is known in advance, T dependent π_t is possible, e.g, $\pi_t = T$ [WZJ16]. By making π_t independent with T, EST becomes anytime, i.e. not requiring that the number of rounds is known in advance.

Lemma A.14 (Lemma 5.1 [SKKS09]). *For $\beta^{(B)UCB}_{t,1} = 2\log\left(\frac{|\mathcal{X}|\pi^2(\mathfrak{T}^{(B)}_{seq}(t,1))^2}{6\delta}\right)$ with $\delta \in (0,1)$,*

$$P\left(\bigcap_{\mathbf{x}\in\mathcal{X}}\left\{f \,\middle|\, |f(\mathbf{x}) - \mu_{t-1}(\mathbf{x})| \leq (\beta^{(B)UCB}_{t,1})^{1/2}\cdot\sigma_{t-1,1}(\mathbf{x})\right\}\right) \geq 1-\delta.$$

Remark A.15. Note that $\beta_{t,1}^{(B)UCB} = \beta_{\mathfrak{T}_{seq}^{(B)}(t,1)}^{UCB}$. in batch BO with the batch size of B, β is set as if there is B times more rounds.

Lemma A.16 (Lemma 5.2 [SKKS09], Lemma 1 [CBRV13]). *If* $|f(\mathbf{x}) - \mu_{t-1}(\mathbf{x})| \leq (\beta_{t,1}^{(B)UCB})^{1/2}\sigma_{t-1,1}(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{X}$*, then*

$$r_{t,1} = f(\mathbf{x}_{t,1}) - f(\mathbf{x}^\star) \leq 2(\beta_{t,1}^{(B)UCB})^{1/2}\sigma_{t-1,1}(\mathbf{x}_{t,1}).$$

Remark A.17. In Lemma A.16, we only bound regrets in $(t,1)$-th round where there is no delayed evaluation.

Theorem A.18. *Assume a kernel such that $K(\cdot,\cdot) \leq 1$, $|\,\mathcal{X}\,| < \infty$ and $f : \mathcal{X} \to \mathbb{R}$ is sampled from $\mathcal{GP}(\mathbf{0}, K)$. In each round $t \in [T]$ of batch Bayesian optimization, LAW acquires a batch using the evaluation data $\mathcal{D}_{t-1}$, the diversity measure $L_t(\cdot,\cdot) = K(\cdot,\cdot|\,\mathcal{D}_{t-1})$, an acquisition function $a_t(\cdot)$ and a weight function $w(\cdot)$ (as defined below Eq. 9).*

Let $C_1 = \frac{36}{\log(1+\sigma^{-2})}$ where σ^2 is the variance of the observation noise and $\delta \in (0,1)$.

For GP-UCB, define $\beta_{t,1}^{(B)UCB} = 2\log\left(\frac{|\,\mathcal{X}\,|\pi^2(\mathfrak{T}_{seq}^{(B)}(t,1))^2}{6\delta}\right)$ and let

$$\eta_t^{(B)} = 2(\beta_{t,1}^{(B)UCB})^{1/2}$$

For EST, define $\nu_t = \min_{\mathbf{x}}\left(\frac{\mu_{t-1}(\mathbf{x}) - \hat{m}_t}{\sigma_{t-1,1}(\mathbf{x})}\right)$ where $\hat{m}_t$ is the estimate of the optimum [WZJ16], $\zeta_t = \left(2\log\left(\frac{\pi_t^2}{2\delta}\right)\right)^{1/2}$, $\pi_t > 0$ such that $\sum_{t=1}^{\infty}\pi_t^{-1} \leq 1$ and let

$$\eta_t^{(B)} = \nu_{t^*} + \zeta_t$$

where $t^* = \operatorname*{argmax}_{s\in[t]} \nu_s$.

Then batch cumulative regret satisfies the following bound

$$P\left(\left\{\frac{R_T^{(B)}}{T} \leq \frac{\eta_T^{(B)}}{T} + \eta_T\frac{w_+}{w_-}\sqrt{C_1\frac{\gamma_{TB}}{TB}}\right\}\right) \geq 1-\delta.$$

Proof. Let $\eta_t^{(B)} = \begin{cases} \nu_{t^*} + \zeta_t & \text{EST} \\ 2(\beta_{t,1}^{(B)UCB})^{1/2} & \text{UCB} \end{cases}$.

For the batch cumulative regret case, we use Lemma A.1.

$$\begin{aligned}
R_T^{(B)} &= \sum_{t=1}^{T} r_t^{(B)} = \sum_{t=1}^{T}\min_{b=1,\cdots,B} r_{t,b} && (19)\\
&\leq \sum_{t=1}^{T} r_{t,1} && (20)\\
&\leq \eta_T^{(B)}\cdot\sum_{t=1}^{T}\sigma_{t-1,1}(\mathbf{x}_{t,1}) \quad \text{by} && \begin{cases}\text{Lemma A.11} & \text{EST}\\ \text{Lemma A.16} & \text{UCB}\end{cases}\\
&\leq \eta_T^{(B)}\cdot\left(1+\frac{w_+}{w_-}\frac{1}{B}\sum_{t=1}^{T}\sum_{b=1}^{B}\sigma_{t-1,b}(\mathbf{x}_{t,b})\right) && \text{by Lemma A.1}\\
&\leq \eta_T^{(B)}\cdot\left(1+\frac{w_+}{w_-}\sqrt{\frac{T}{B}\sum_{t=1}^{T}\sum_{b=1}^{B}\sigma_{t-1,b}^2(\mathbf{x}_{t,b})}\right) && \text{by Cauchy-Schwarz}\\
&\leq \eta_T^{(B)}\cdot\left(1+\frac{w_+}{w_-}\sqrt{\frac{T}{B}C_1\gamma_{TB}}\right) && \text{by Lemma A.7}
\end{aligned}$$

By Lemma A.9 for EST and Lemma A.14 for UCB, above two inequalities hold with the probability at least $1 - \delta$. $\square$

Remark A.19. Due to the difference of the statements in Lemma A.1, the batch cumulative regret bound additionally has the term $\frac{w_+}{w_-}$. Even with this additional term, it shows that the bound of the batch cumulative regret of LAW enjoys the same asymptotic behavior as existing methods [CBRV13, DKB14, KDK16].

Remark A.20. This theorem provides a rough guideline how to choose a weight function, that is, bounded below by a positive value and bounded above, which is the condition we specify for the weight function. Even though this shows that simple regret vanishes, this regret bound for LAW is loose because not much specific structure of the weight function other than the bound is used. We expect that, using other properties of the weight function along with the boundedness, the bound can be improved.

A.2 Difference to analysis of sequential cumulative regret

In our regret analysis, we analyze batch cumulative regret. In existing works, sequential cumulative regret is analyzed as an end goal [DKB14] and as a medium to show vanishing simple regret in [KKSP18]. In [CBRV13],[11] both batch cumulative regret and sequential cumulative regret are analyzed.[12] We discuss the differences between these two approaches and the technical details in their proofs.

By definition, the analysis of sequential cumulative regret takes into account all instantaneous regrets incurred while batch cumulative regret considers minimum instantaneous regrets in each batch. Therefore, bounds on sequential cumulative regret are stronger than ones on batch cumulative regret in this sense (as shown in Table 3 Relation between two Cumulative). However, each has its own more appropriate scenario to use. The sequential cumulative regret is often appropriate in the situation where the optimization objective represents the cost of evaluations. For example, in multi-armed bandit, each instantaneous regret represents the cost of evaluation (playing arm-pulling) and the goal is to minimize the incurred cost in finding the best bandit machine. On the other hand, batch cumulative regret is often reasonable when the optimization objective is different to the cost of evaluations. For example, in hyperparameter optimization, the cost of evaluations can be wall-clock time and the objective is the cross-validation error. In this case, we want to find a good hyperparameter no matter how bad hyperparameters are evaluated, which possibly acts as exploratory query points.

In proofs, each analysis takes a slightly different route. As argued in [DKB14], to bound all instantaneous regret, a wider confidence bound is needed to bound instantaneous regret with the corresponding posterior variance. While the posterior mean is not update in the batch acquisition until all query points are evaluated, the posterior variance is updated whenever a new query point is given no matter whether it is evaluated or not. To guarantee high probability bound for all instantaneous regrets, an additional kernel-dependent constant is introduced and the constant is controlled with an initialization scheme[DKB14]. In [KKSP18], the analysis relies on such kernel-dependent constant and the initialization scheme but it is empirically shown that the algorithm performs well without the initialization scheme. The necessity of the kernel-dependent constant suggests that the analysis of sequential cumulative regret in [CBRV13] requires a revision.

For the purpose to show vanishing simple regret, batch cumulative regret can be used circumventing the additional constant and the initialization scheme proposed in [DKB14]. In the analysis using batch cumulative regret, only non-delayed regret is considered and bounded by non-delayed posterior variance (Eq. 20). Then non-delayed posterior variance is bounded by the average of non-delayed posterior variance and delayed posterior variances in the same batch (Lemma A.4). Therefore, the effect of the batch size influences the bound in this posterior variance bounding step. However, in the analysis using sequential cumulative regret[DKB14, KKSP18], both non-delayed and delayed instantaneous regrets need to be bounded. The bound is the corresponding posterior variance multiplied by a specially design number to handle delayed cases. In response to this, the additional kernel-dependent constant and the initialization scheme are introduced in [DKB14].

[11] Sequential cumulative regret is termed full cumulative regret in [CBRV13].

[12] The analysis of sequential cumulative regret in [CBRV13] may need modification and not be correct, see following paragraph for a brief explanation and for more elaborated explanation, refer to [DKB14].

Batch cumulative regret is enough in showing vanishing simple regret. The proof only considers non-delayed instantaneous regrets in batches. Therefore, the analysis of batch cumulative regret reveals how delayed query points in a batch explore effectively and help to reduce future non-delayed instantaneous regrets. We admit that some may argue that a tighter bound is possible by taking into account delayed evaluations with smaller instantaneous regrets. Still, this is aligned with the intuition of many batch acquisition methods promoting diversity in batches. In practice, it is not unlikely to observe a delayed evaluation is better than the non-delayed evaluation in the same batch.

A.3 Growth Rate of UCB/EST hyperparameter

$$\eta_t^{(B)} = \begin{cases} \nu_{t^*} + \zeta_t & \text{EST} \\ 2(\beta_{t,1}^{(B)UCB})^{1/2} & \text{UCB} \end{cases}$$

where $\beta_{t,1}^{(B)UCB} = 2\log\left(\frac{|\mathcal{X}|\pi^2((t-1)B+1)^2}{6\delta}\right)$, $t^* = \underset{s\in[t]}{\operatorname{argmax}}\,\nu_s$, $\nu_t = \min_{\mathbf{x}}\left(\frac{\mu_{t-1}(\mathbf{x})-\hat{m}_t}{\sigma_{t-1,1}(\mathbf{x})}\right)$ where $\hat{m}_t$ is the estimate of the optimum [WZJ16], $\zeta_t = \left(2\log\left(\frac{\pi_t^2}{2\delta}\right)\right)^{1/2}$, $\pi_t > 0$ such that $\sum_{t=1}^{\infty}\pi_t^{-1} \leq 1$.

For UCB, it is clear that $2(\beta_{t,1}^{(B)UCB})^{1/2} = \mathcal{O}((\log(tB))^{1/2})$.

For EST, we first look into ζ_t. If we choose $\pi_t = \frac{\pi^2 t^2}{6}$ as suggested in [WZJ16], then $\zeta_t = \mathcal{O}((\log(tB))^{1/2})$. Since $\hat{m}_t = E_{f\sim GP(\mu_{t-1}(\cdot),\sigma_{t-1}^2(\cdot))}[\inf_{\mathbf{x}} f(\mathbf{x})]$ [WZJ16], from Lemma 5.1 in [SKKS09], we have

$$|f(\mathbf{x}) - \mu_{t-1}(\mathbf{x})| \leq \tau_t\sigma_{t-1}(\cdot) \quad \forall \mathbf{x} \in \mathcal{X}$$

where $\tau_t^{1/2} = 2\log\left(\frac{|\mathcal{X}|\pi^2 t^2}{6\delta}\right)$.

Then

$$\hat{m}_t \geq \mu_{t-1}(\mathbf{x}_{lb}) - \tau_t\sigma_{t-1}(\mathbf{x}_{lb})$$

with $\mathbf{x}_{lb} = \underset{\mathbf{x}}{\operatorname{argmin}}\,\mu_{t-1}(\mathbf{x}) + \tau_t\sigma_{t-1}(\mathbf{x})$,

$$\min_{\mathbf{x}}\left(\frac{\mu_{t-1}(\mathbf{x}) - \hat{m}_t}{\sigma_{t-1,1}(\mathbf{x})}\right) \leq \frac{\mu_{t-1}(\mathbf{x}_{lb}) - \mu_{t-1}(\mathbf{x}_{lb}) + \tau_t\sigma_{t-1}(\mathbf{x}_{lb})}{\sigma_{t-1,1}(\mathbf{x}_{lb})} = \tau_t$$

Since, τ_t is increasing, we have $\nu_{t^*} \leq \tau_t$ and $\nu_{t^*} = \mathcal{O}((\log(tB))^{1/2})$.

Therefore, for EST, $\eta_t^{(B)} = \nu_{t^*} + \zeta_t = \mathcal{O}((\log(tB))^{1/2})$.

B Information Gain

In this section, we present results of our analysis on the information gain and the position kernel.

In Theorem. B.1 in Subsection B.1, we show that for a kernel on a finite space, the information gain grows $\mathcal{O}(\log(T))$. In combination with Section A, this shows that an arbitrary kernel on a finite space including the position kernel achieves sublinear regret. To our knowledge, the positive definiteness of the position kernel has not been shown rigorously, [ZSBB14] used randomly generated data to empirically check that whether the position kernel is positive definite and [ZYLW19] argued that the exponential of a metric is positive definite, which is not true in general. Therefore, we show that the position kernel is positive definite and further provide a lower and upper bound of the eigenvalues of the position kernel in Subsection B.2. In Theorem B.4 in Subsection B.3, we show that by using the properties of the position kernel, a tighter bound on the maximum information gain is achievable.

B.1 Information gain of kernels on a finite space

In this subsection, we show a bound of the information gain of kernels defined on a finite set.

Theorem B.1. *K is a kernel on a finite set $\mathcal{X}$ ($|\mathcal{X}| < \infty$), σ^2 is the variance of the observation noise and $\Lambda = \{\lambda_n\}_{1,\cdots,|\mathcal{X}|}$ ($\lambda_n \geq \lambda_{n+1} \geq 0$) is the set of eigenvalues of the gram matrix $K(\mathcal{X},\mathcal{X})$.*

The number of elements in a set A is denoted by N_A, so $N_{\mathcal{X}}$ is the number of elements of $\mathcal{X}$ and is equal to the number of eigenvalues of $K(\mathcal{X}, \mathcal{X})$.

Then

$$\gamma(T; K, \mathcal{X}, \sigma^2) \leq \frac{1}{2} \min \left\{ T \cdot \log\det(1 + \sigma^{-2} \max_{x \in \mathcal{X}} K(x, x)), N_{\mathcal{X}} \cdot \log(1 + \sigma^{-2} \lambda_{max} \cdot T) \right\}$$

Proof. Let us consider the eigenvalues and the eigenvectors of the gram matrix $K(\mathcal{X}, \mathcal{X})$.

$$K(\mathcal{X}, \mathcal{X}) = U \Lambda U^T$$

with $\Lambda = diag(\lambda_1, \cdots, \lambda_{N_{\mathcal{X}}})$, $U = [u_1, \cdots, u_{N_{\mathcal{X}}}] \in \mathbb{R}^{N_{\mathcal{X}} \times N_{\mathcal{X}}}$ where λ_i is an eigenvalue and u_i is the corresponding eigenvector.

Since

$$K(x, x') = \sum_{i=1}^{N_{\mathcal{X}}} \lambda_i [u_i]_x [u_i]_{x'},$$

the map

$$\phi(x) = [\sqrt{\lambda_1}[u_1]_x, \cdots, \sqrt{\lambda_{N_{\mathcal{X}}}}[u_{N_{\mathcal{X}}}]_x]^T$$

is a $N_{\mathcal{X}}$ dimensional feature map

$$K(x, x') = \phi(x)^T \cdot \phi(x').$$

For a sequence $A = \{a_1, \cdots, a_{N_A}\}$ of $a_i \in \mathcal{X} = x_{i}{}_{i=1}^{N_{\mathcal{X}}}$, the gram matrix $K(A, A)$ can be expressed with the projection matrix $P_A^{\mathcal{X}} \in \{0, 1\}^{N_A \times N_{\mathcal{X}}}$ from X to A such that $[P_A^{\mathcal{X}}]_{ij} = 1$ if $a_i = x_j$

$$K(A, A) = P_A^{\mathcal{X}} U \Lambda U^T (P_A^{\mathcal{X}})^T. \tag{21}$$

Remark B.2. Note that $(P_A^{\mathcal{X}})^T \cdot P_A^{\mathcal{X}}$ is $N_{\mathcal{X}} \times N_{\mathcal{X}}$ diagonal matrix and $[(P_A^{\mathcal{X}})^T \cdot P_A^{\mathcal{X}}]_{ii}$ is how many times x_i appears in the sequence A.

We obtain two bounds. The first one is

$$\log\det(I + \sigma^{-2} K(A, A)) \leq \sum_{a \in A} \log\det(1 + \sigma^{-2} K(a, a)) \tag{22}$$

using Hadamard's inequality.[13]

Adopting the proof for the information gain of the linear kernel from [SKKS09], the second one is

$$\begin{aligned}
&\log\det(I + \sigma^{-2} K(A, A)) && (23)\\
&= \log\det(I + \sigma^{-2} P_A^{\mathcal{X}} K(X, X) (P_A^{\mathcal{X}})^T) \\
&= \log\det(I + \sigma^{-2} P_A^{\mathcal{X}} U \Lambda U^T (P_A^{\mathcal{X}})^T) && \text{by Eq. 21} \\
&= \log\det(I + \sigma^{-2} \Lambda^{\frac{1}{2}} U^T (P_A^{\mathcal{X}})^T P_A^{\mathcal{X}} U \Lambda^{\frac{1}{2}}) && \text{by Weinstein-Aronszajn identity} \\
&\leq \sum_{i=1}^{N_{\mathcal{X}}} \log\det(1 + \sigma^{-2} \lambda_i [U^T (P_A^{\mathcal{X}})^T P_A^{\mathcal{X}} U]_{ii}) && \text{by Hadamard's inequality} \\
&\leq \sum_{i=1}^{N_{\mathcal{X}}} \log\det(1 + \sigma^{-2} \lambda_i T) && \text{by Eq. 25} \qquad (24)
\end{aligned}$$

[13]https://en.wikipedia.org/wiki/Hadamard%27s_inequality

using Weinstein-Aronszajn identity[14], Hadamard's inequality[15] and below

$$
\begin{aligned}
&[U^T(P_A^{\mathcal{X}})^T P_A^{\mathcal{X}} U]_{ii} \\
=&\sum_{k=1}^{N_{\mathcal{X}}}\sum_{l=1}^{N_{\mathcal{X}}}([U]_{ki})[(P_A^{\mathcal{X}})^T P_A^{\mathcal{X}}]_{kl}([U]_{li}) \\
=&\sum_{k=1}^{N_{\mathcal{X}}}[(P_A^{\mathcal{X}})^T P_A^{\mathcal{X}}]_{kk}([U]_{ki})^2 && \text{by Rmk. B.2} \\
\leq&\underbrace{\sum_{k=1}^{N_{\mathcal{X}}}[(P_A^{\mathcal{X}})^T P_A^{\mathcal{X}}]_{kk}}_{=N_A=T}\cdot\underbrace{\sum_{k=1}^{N_{\mathcal{X}}}([U]_{ki})^2}_{=1} && \sum_i a_i b_i \leq (\sum_i a_i)(\sum_i b_i) \quad \text{if} \quad a_i, b_i \geq 0
\end{aligned} \tag{25}
$$

where the second equality comes from the fact that $(P_A^{\mathcal{X}})^T \cdot P_A^{\mathcal{X}}$ is $N_{\mathcal{X}} \times N_{\mathcal{X}}$ diagonal matrix and the last inequality is possible because every numbers are non-negative since $[(P_A^{\mathcal{X}})^T \cdot P_A^{\mathcal{X}}]_{ii}$ is how many times x_i appears in the sequence A.

Putting Eq. 22 and Eq. 23 together,

$$\log\det(I+\sigma^{-2}K(A,A)) \leq \min\left\{T\cdot\log\det(1+\sigma^{-2}\max_{x\in\mathcal{X}}K(x,x)),\ N_{\mathcal{X}}\cdot\log\det(1+\sigma^{-2}\lambda_{max}T)\right\}$$

Q.E.D. $\square$

B.2 Positive definiteness of the position kernel

In this subsection, we show that the positive definiteness of the position kernel and the bound of its eigenvalues.

Theorem B.3. *The position kernel $K(\cdot,\cdot|\tau)$ defined on S_N is positive definite and the eigenvalues of the $K(A,A)$ where $A \subset \mathcal{X}$ lie between $\left(\frac{1-\rho}{1+\rho}\right)^N$ and $\left(\frac{1+\rho}{1-\rho}\right)^N$ where $\rho = \exp(-\tau)$.*

Proof. We show that the kernel is positive definite on a larger set

$$\mathcal{X} = \prod_{i=1}^{N}\{1,\cdots,N\}.$$

Since $S_N \subset \mathcal{X}$, $K(S_N, S_N)$ is a principal submatrix of $K(\mathcal{X},\mathcal{X})$ With Poincaré seperation theorem (or Cauchy interlacing theorem), we show that the position kernel is positive definite and that the eigenvalues of $K(S_N, S_N)$ lie between the smallest eigenvalue and the largest eigenvalue of $K(\mathcal{X},\mathcal{X})$.

On $\mathcal{X}$, the position kernel is a product kernel of N kernels defined $\{1,\cdots,N\}$ as below

$$K(\pi_1,\pi_2|\tau) = \exp\Big(-\tau\cdot\sum_i|\pi_1^{-1}(i) - \pi_2^{-1}(i)|\Big).$$

and its gram matrix on each component has following form

$$[\rho^{|i-j|}]_{ij} = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-2} & \rho^{N-1} \\ \rho & 1 & \rho & \cdots & \rho^{N-3} & \rho^{N-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \cdots & \rho & 1 \end{bmatrix}$$

where $\rho = \exp(-\tau)$.

[14] https://en.wikipedia.org/wiki/Weinstein%E2%80%93Aronszajn_identity

[15] https://en.wikipedia.org/wiki/Hadamard%27s_inequality

This form of matrix is known as Kac-Murdock-Szegö (KMS) matrix [GS58, Tre99], which we denote by $KMS(\rho)$ $(0 < \rho < 1)$.

Their eigenvalues λ_n are bounded as below [GS58, Tre99]

$$\lambda_n = \frac{1-\rho^2}{1+\rho^2-2\rho\cos(\theta_n)}$$

where

$$\frac{n-1}{N+1}\pi < \theta_n < \frac{n}{N+1}\pi$$

Therefore

$$\frac{1-\rho}{1+\rho} < \frac{1-\rho^2}{1+\rho^2-2\rho\cos(\frac{n}{N+1}\pi)} < \lambda_n < \frac{1-\rho^2}{1+\rho^2-2\rho\cos(\frac{n-1}{N+1}\pi)} < \frac{1+\rho}{1-\rho}$$

We observe that the each component kernel is positive definite with above bounds on the eigenvalues.

Since

$$K(\mathcal{X},\mathcal{X}) = \bigotimes_{i=1}^{N} K(\{1,\cdots,N\},\{1,\cdots,N\})$$

where $\bigotimes$ is the Kronecker product, the lower bound and the upper bound of the eigenvalues of $K(\mathcal{X},\mathcal{X})$ are $\left(\frac{1-\rho}{1+\rho}\right)^N$ and $\left(\frac{1+\rho}{1-\rho}\right)^N$, respectively.

For $A \in S_N \in \mathcal{X}$, these bounds also apply to the eigenvalues of $K(A,A)$ by Poincaré seperation theorem (or Cauchy interlacing theorem).

Q.E.D. □

B.3 Information gain of the position kernel

Theorem B.4. *$K(\cdot,\cdot|\tau)$ is the position kernel defined on S_N, σ_{obs}^2 is the variance of the observation noise, $\rho = \exp(-\tau)$ and*

$$D_{max} = \begin{cases} \frac{N^2}{2} & N \text{ mod } 2 = 0 \\ \frac{N^2-1}{2} & N \text{ mod } 2 = 1 \end{cases}$$

Then

$$\gamma_T \leq \frac{1}{2}\min\{A(T), N_{\mathcal{X}} \cdot \log(1+\sigma_{obs}^{-2}\lambda_{max}\cdot T)\}$$

where

$$A(T) = \log(1+\sigma_{obs}^{-2}(1+(T-1)\rho^{D_{max}})) + (T-1)\log(1+\sigma_{obs}^{-2}(1-\rho^{D_{max}}))$$

which is smaller than $T \cdot \log\det(1+\sigma_{obs}^{-2}\max_{x\in\mathcal{X}} K(x,x))$.

Proof. γ_T is defined as

$$\frac{1}{2}\max_{A\subset\mathcal{X},|A|=T} \log\det(I+\sigma_{obs}^{-2}K(A,A))$$

By Lem. B.6 in Supp.Subsec. B.3, for $i,,j = 1,\cdots,T$, $\rho^{D_{max}} \leq [K(A,A)]_{ij} \leq 1$.

By Perron-Frobenius theorem, the largest eigenvalue of $K(A,A)$ is bounded below by

$$1+(T-1)\rho^{D_{max}}$$

When $\lambda_i^{(A)}$ is the i-th eigenvalue of $K(A,A)$, with the constraint $\lambda_1^{(A)} \geq 1+(T-1)\rho^{D_{max}}$,

$$\prod_{i=1}^{T}(1+\sigma_{obs}^{-2}\lambda_i^{(A)})$$

is bounded above by

$$(1+\sigma_{obs}^{-2}(1+(T-1)\rho^{D_{max}}))\prod_{i=2}^{T}\Big(1+\sigma_{obs}^{-2}\frac{T-(1+(T-1)\rho^{D_{max}})}{T-1}\Big)$$
$$(1+\sigma_{obs}^{-2}(1+(T-1)\rho^{D_{max}}))\prod_{i=2}^{T}(1+\sigma_{obs}^{-2}(1-\rho^{D_{max}}))$$

Here, we use the fact that for $\sum_i x_i = C$, $x_i > 0$ if there are p and q such that $x_p < x_q$, then for x'_i defined as $x'_i = x_i$ for $i \neq p, 1$ and $x'_p = x_p + d$, $x'_q = x_q - d$ where $d \leq (x_q - x_p)/2$

$$\prod_i x_i \leq \prod_i x'_i.$$

Note that without the constraint on the lower bound of the largest eigenvalue,

$$\prod_{i=1}^{T}(1+\sigma_{obs}^{-2}\lambda_i^{(A)}) \leq \Big(1+\sigma_{obs}^{-2}\frac{trace(K(A,A))}{T}\Big)^T$$

where $trace(K(A,A)) = T$ for kernels such that $K(x,x)$ is a constant independent of $x \in \mathcal{X}$ as is for the position kernel.

This shows that the bound in this theorem is tighter than that of Thm. B.1.

Q.E.D. □

Remark B.5. specially when σ_{obs}^2 and/or ρ is large, i.e. $\log(1+\sigma_{obs}^{-2}(1-\rho^{D_{max}})) \approx 0$, we can observe that even in the finite-time regime, the regret is almost sublinear since it is dominated by $\log(1+\sigma_{obs}^{-2}(1+(T-1)\rho^{D_{max}}))$. In this case, the theorem provide a bound which is significantly tighter than the bound in Thm. 3.10 even in the finite-time regime. Even though both are the same in the asymptotic regime, they may differ significantly in the finite-time regime.

Lemma B.6. *For π_1, π_2 in S_N,*

$$d_{pos}(\pi_1,\pi_2) = \sum_i |\pi_1^{-1}(i) - \pi_2^{-1}(i)| \geq D_{max}$$

where

$$D_{max} = \begin{cases} \frac{N^2}{2} & N \textit{ mod } 2 = 0 \\ \frac{N^2-1}{2} & N \textit{ mod } 2 = 1 \end{cases}$$

Proof. Note that d_{pos} is left-invariant, that is,

$$d_{pos}(\pi_1,\pi_2) = d_{pos}(\pi \circ \pi_1, \pi \circ \pi_2)$$

for $\pi \in S_N$, and thus

$$d_{pos}(\pi_1,\pi_2) = d_{pos}(\pi_{id}, (\pi_1)^{-1} \circ \pi_2)$$

where $\pi_{id} = (1, \cdots, N)$.

By induction on N, we show that

$$\max_{\pi \in S_N} d_{pos}(\pi_{id}, \pi) = d_{pos}((1,2,3,\cdots,N),(N,N-1,\cdots,2,1))$$

Base case($N = 2$) This is trivial.

Induction Step As the induction hypothesis, assume that above is true for $N = k$. When $N = k+1$, let us consider $\pi = (-,-,\cdots,, a) \in S_{k+1}$ an arbitrary permutation whose last element is $a \neq 1$,

$$d_{pos}(\pi_{id},\pi) = \sum_{i:\pi^{-1}(i)<a, i<k+1} |i-\pi^{-1}(i)| + \sum_{i:\pi^{-1}(i)>a, i<k+1} |i-\pi^{-1}(i)| + |k+1-a|$$

where $a \neq N$.

Then

$$
\begin{aligned}
&\sum_{i:\pi^{-1}(i)<a, i<k+1} |i-\pi^{-1}(i)| + \sum_{i:\pi^{-1}(i)>a, i<k+1} |i-\pi^{-1}(i)+1-1| \\
&\leq \sum_{i:\pi^{-1}(i)<a, i<k+1} |i-\pi^{-1}(i)| + \sum_{i:\pi^{-1}(i)>a, i<k+1} |i-(\pi^{-1}(i)-1)| + \sum_{i:\pi^{-1}(i)>a, i<k+1} 1 \\
&\leq d_{pos}((1,\cdots,k),(k,\cdots,1)) + (k-a) \\
&\leq d_{pos}((1,\cdots,k),(k,\cdots,1)) + k + (k \bmod 2) \\
&= d_{pos}((1,\cdots,k+1),(k+1,\cdots,1))
\end{aligned}
$$

where

$$
\sum_{i:\pi^{-1}(i)<a, i<k+1} |i-\pi^{-1}(i)| + \sum_{i:\pi^{-1}(i)>a, i<k+1} |i-(\pi^{-1}(i)-1)| \leq d_{pos}((1,\cdots,k),(k,\cdots,1))
$$

is from the induction hypothesis.

Therefore

$$
\max_{\pi \in S_N} d_{pos}(\pi_{id}, \pi) = D_{max} = \begin{cases} \frac{N^2}{2} & N \bmod 2 = 0 \\ \frac{N^2-1}{2} & N \bmod 2 = 1 \end{cases}
$$

Q.E.D. □

C Resemblance to the Local Penalization

Taken from [GDHL16], the local penalization strategy selects b-th point in a batch as follows

$$
\mathbf{x}_{t,b} = \operatorname*{argmax}_{\mathbf{x} \in \mathcal{X}} \left\{ g(a_t(\mathbf{x})) \prod_{i=1}^{b-1} \phi(\mathbf{x}, \mathbf{x}_{t,i}) \right\} \tag{26}
$$

where $\phi(\mathbf{x}, \mathbf{x}_{t,i})$ is a local penalizer which is non-decreasing function of Euclidean distance $\| \mathbf{x} - \mathbf{x}_{t,i} \|_2$ and $g(\cdot)$ is a positive increasing function similar to our weight function.

If we use the prior covariance function $K(\cdot,\cdot)$, which is the kernel of the GP surrogate model in place of the posterior covariance function $K_t(\cdot,\cdot)$ as the diversity gauge of L_t^{AW}, the greedy maximization objective becomes

$$
\begin{aligned}
\mathbf{x}_b = \operatorname*{argmax}_{\mathbf{x} \in \mathcal{X}} \Big[& w(a_t(\mathbf{x}))^2 \cdot \\
& \big(K(\mathbf{x},\mathbf{x}) - K(\mathbf{x}, \{\mathbf{x}_i\}_{|b-1|})(K(\{\mathbf{x}_i\}_{|b-1|}, \{\mathbf{x}_i\}_{|b-1|}) + \sigma^2 I)^{-1}(K(\{\mathbf{x}_i\}_{|b-1|}, \mathbf{x})\big)\Big]
\end{aligned} \tag{27}
$$

We call this LAW variant as LAW-prior-EST and LAW-prior-EI according to the acquisition function each uses.

Since the closer to the conditioning data it is, the smaller the predictive variance is, the predictive variance behaves exactly as what the local penalizer aims at. Another key difference is that, while the local penalizer Eq. 26 is heuristically designed, LAW-prior-EST\EI use the kernel whose hyperparameters are fitted in the surrogate model fitting step. Therefore, the diversity measured in LAW-prior-EST\EI is more guided by the collected evaluation data.

Additional comparison to these variants (Supp. Sec. G) reveals the contribution of the acquisition weights and thus further confirms the benefit of using acquisition weights in the optimization performance

D Score-based Structure Learning

In score-based structure learning, general-purpose optimization methods are utilized to optimize a score $S(\mathcal{G}, \mathcal{D})$ of DAG $\mathcal{G}$[SGG19] for a given data $\mathcal{D}$. Typically, $S(\cdot,\cdot)$ is a penalized likelihood score or an information theoretic criterion[DM17]. The prevalent choice of the optimization method is a local search which relies on the efficient computation of the DAG score to afford many score evaluations.[Chi02, KF09, SGG19]. For the efficient computation, it is critical for a score to be decomposable[SSS19], defined as below.

$$S(\mathcal{G}, \mathcal{D}) = \sum_{v \in \mathcal{V}} s(v | Pa^{\mathcal{G}}(v), \mathcal{D})$$

where $s(\cdot)$ is a score defined for a node $v \in \mathcal{V}$. In local search with a decomposable score, changes made by local modification of the DAG $\mathcal{G}$ can be reflected to the network score $S(\mathcal{G}, \mathcal{D})$ by updating corresponding components without calculating the network score from scratch. Despite of the computational benefit, the constraint on the decomposability of the score restricts the use of more suitable scores, such as, scores with non-factorized priors[CCD15] or sophisticated information criteria[GR19].

E Normalized Maximum Likelihood

E.1 Model Selection with Minimum Description Length

In minimum description length (MDL) principle[GG07], a distribution called a universal distribution is associated with each model class, for example, $\bar{p}_{\mathcal{G}}(\cdot)$ is associated with the $\mathcal{M}^{\mathcal{G}}$, BNs with a given DAG $\mathcal{G}$. For a given data $\mathcal{D}$, model selection can be performed by comparing the universal distribution relative to a model class

$$\bar{p}_{\mathcal{G}_1}(\mathcal{D}) \quad \text{VS} \quad \bar{p}_{\mathcal{G}_2}(\mathcal{D})$$

E.2 Normalized Maximum Likelihood

Normalized Maximum Likelihood (NML) is regarded as the most fundamental universal distribution[GR19]. For the discrete BNs with a DAG $\mathcal{G}$ with the data $\mathcal{D}$, NML is defined as

$$\bar{p}_{\mathcal{G}}(\mathcal{D}) = \frac{p_{BN}(\mathcal{D} \,|\, \mathcal{G}, \hat{\theta}_{ML}(\mathcal{G}, \mathcal{D}))}{\sum_{|\,\mathcal{D}\,| = |\,\mathcal{D}'\,|} p_{BN}(\mathcal{D}' \,|\, \mathcal{G}, \hat{\theta}_{ML}(\mathcal{G}, \mathcal{D}'))}$$

where $\hat{\theta}_{ML}(\mathcal{G}, \mathcal{D})$ is the maximum likelihood estimator of the parameters of the BN with the DAG $\mathcal{G}$ on the data $\mathcal{D}$. The summation over all possible data with the same cardinality is the computational bottleneck. The log of the denominator $REG_{NML}(\mathcal{G}, N) = \log(\sum_{|\,\mathcal{D}'\,| = N} p_{BN}(\mathcal{D}' \,|\, \mathcal{G}, \hat{\theta}_{ML}(\mathcal{G}, \mathcal{D}')))$ is called NML regret.[16]

E.3 NML regret estimation

Even though it is strongly principled, NML computation is restricted to certain classes of models, e.g, multinomial distribution[KM07], naive Bayes[MM07], which prevents its use in score-based structure learning. In Bayesian networks, efficient approximations were proposed and shown to perform better in model selection[RSKM08, SLAJR18].

Even though $REG_{NML}(\mathcal{G}, N)$ cannot be exactly computed, the summation can be estimated using Monte carlo with proper scaling when BN is discrete.

$$\begin{aligned} &\log\Big(\sum_{|\,\mathcal{D}\,| = |\,\mathcal{D}'\,|} p_{BN}(\mathcal{D}' \,|\, \mathcal{G}, \hat{\theta}_{ML}(\mathcal{G}, \mathcal{D}'))\Big) \\ &\approx \log\Big(\frac{\sum_{|\,\mathcal{D}\,| = |\,\mathcal{D}'\,|} 1}{|\,\mathcal{S}\,|}\Big) + \log\Big(\mathit{LSE}_{\mathcal{D}' \in \mathcal{S}} \log(p_{BN}(\mathcal{D}' \,|\, \mathcal{G}, \hat{\theta}_{ML}(\mathcal{G}, \mathcal{D}')))\Big) \end{aligned}$$

where *LSE* is the logsumexp whose implementation increases the numerical stability significantly[17].

[16] Originally, it is called regret but not to confuse with bandit regret, we prefix it with NML.

[17] `https://pytorch.org/docs/stable/generated/torch.logsumexp.html`

In our scaled MC estimate of $REG_{NML}(\cdot,\cdot)$, we observed that smaller samples tend to marginally underestimate the value. However, the estimation seems quickly saturated with respect to the sample size. We observed that a MC-estimate of NML regret using $|\,\mathcal{S}\,| = 10,000$ is a good compromise between the stability of the estimation and the time needed for the evaluation. With 10,000 samples, the estimation is stable and the difference made by using more samples is marginal to the difference made by the choice of different DAGs. On machines with Intel(R) Xeon(R) CPU E5-2630 v3 2.40GHz, the evaluation time of the objectives (Tab. 2) ranges from one minute to four minutes.

F Additional Information on Experiments

F.1 Submodular Maximization

A set function $g : 2^{\Omega} \to \mathbb{R}$, where 2^{Ω} is the power set of Ω, is submodular when it has the diminishing returns property, that is, for all $P \subset Q \subset \Omega$ and $\mathbf{p} \in \Omega \setminus Q$

$$g(P \cup \{\mathbf{p}\}) - g(P) \geq g(Q \cup \{\mathbf{p}\}) - g(Q)$$

As a combinatorial version of convexity [Lov83], submodularity has been playing a critical role in combinatorial optimization [Fuj05].

One important property of the submodular function is that when it is positive ($g(\cdot) \geq 0$) and monotone ($P \subset Q \implies g(P) \leq g(Q)$), its maximization can be performed greedily with an approximation guarantee [NWF78] as given below. In the maximization of a positive monotone submodular function with the cardinality constraints, $g(P^*) = \max_{|P|=M} g(P)$, the solution $P^*_{greedy} = \{\mathbf{p}^*_1, \cdots, \mathbf{p}^*_M\}$ from the greedy strategy which sequentially solves $\mathbf{p}^*_m = \operatorname{argmax}_{\mathbf{p}\in\Omega} g(\{\mathbf{p}^*_1, \cdots, \mathbf{p}^*_{m-1}, \mathbf{p}\})$ has the following approximation guarantee

$$(1 - e^{-1})g(P^*) \leq g(P^*_{greedy}) \leq g(P^*) \tag{28}$$

In practice, this greedy method often provides almost optimum solutions [SKD15]. Moreover, it is possible to relax the conditions (positivity, monotonicity, and even submodularity) [FMV11, BBKT17, Sak20].

F.2 DPP-SAMPLE-EST implementation

In [KDK16], DPP-MAX-EST and DPP-SAMPLE-EST are compared on continuous spaces using the median of multiple runs, and the median of DPP-SAMPLE-EST outperforms the median of DPP-MAX-EST. On the permutation spaces of our experiments, DPP-SAMPLE-EST performs worse than DPP-MAX-EST in terms of the mean of multiple runs.

We attribute this to the sample size used in sampling, rather than the use of a different performance measure. Due to the large size of permutation spaces, it is infeasible to collect many samples as suggested in [AGR16], which is $\mathcal{O}(S\log(S/\epsilon))$ where S is the size of the permutation space and ϵ is the desired approximation level.

As suggested in [AGR16], we first use DPP-MAX-EST to pick the initial point of sampling and then we perform 100 MCMC sampling. Therefore, MCMC includes the maximization routine in its initialization. And we conjectured that not many sampling steps are necessary because it already starts from the most likely point expecting that being perturbed from the most likely point, it is still likely but retains reasonable diversity. However, from the results we have, it appears that more sampling steps are necessary. Since the sufficient condition given in [AGR16] requires an infeasibly huge number on permutation space and such huge number can make the batch acquisition time not negligible compared with evaluation time, LAW focuses on the maximization of the k-DPP density.

F.3 Combinatorial optimization problems on permutations

We consider three types of combinatorial optimization on permutations

Quadratic Assignment Problems [KB57] Given N facilities $\mathcal{P}$ and N locations $\mathcal{L}$, a distance $d(\cdot,\cdot)$ is given for each pair of locations and a weight $k(\cdot,\cdot)$ is given for each pair of facilities, for example, the cost of delivery between facilities. Then the goal is to find an assignment represented by a permutation π^* minimizing $f(\pi) = \sum_{a,b\in\mathcal{P}} k(a,b) \cdot d(\pi(a),\pi(b))$.

Data source (`https://www.opt.math.tugraz.at/qaplib/inst.html`): char12a[CB89], nug22[NVR68], esc32a[EW90]

Flowshop Scheduling Problems [Wik20] There are N machines and M jobs. Each job requires N operations to complete. The n-th operation of the job must be executed on the m-th machine. Each machine can process at most one operation at a time. Each operation in each job has its own execution specified. Even though jobs can be executed in any order, operations in each job should obey the given order. The problem is to find an optimal order of jobs to minimize execution time. For a formal description, please refer to [Ree95].

Data source (`http://people.brunel.ac.uk/~mastjjb/jeb/orlib/flowshopinfo.html`):car5[Car78], hel2[Hel60], reC19[Ree95]

Traveling Salesman Problems For given cities, a salesman visits each city exactly once while minimizing a given cost incurred in travelling. TSP is the most widely known example of combinatorial optimization on permutations.

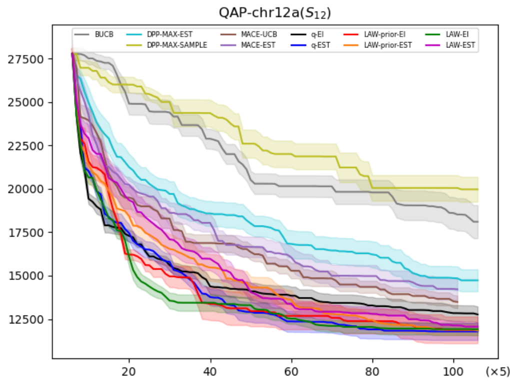
Data source (`http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/`)

G Experiment Results

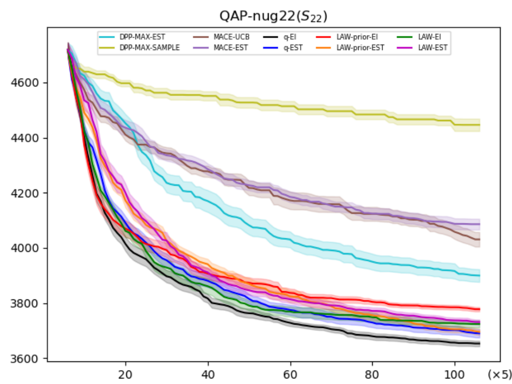
In this section, we provide the additional experimental results which we cannot present in the main text due to the page limit. Following results are presented from the next page

- Comparison with other LAW variants as combinatorial versions of the local penalization[GDHL16] (Subsec. 5.2.1)
- Figures of the structure learning experiments (Sec. 5.3)

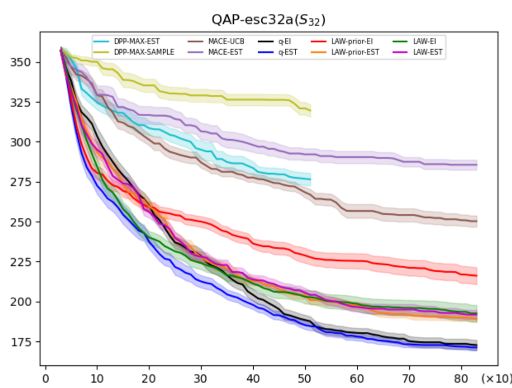
G.1 Quadratic Assignment Problems



Method	Mean±Std.Err.	#Eval
BUCB	$+18104.80 \pm 955.15$	530
DPP-MAX-EST	$+14731.60 \pm 633.79$	530
DPP-SAMPLE-EST	$+19969.60 \pm 718.90$	530
MACE-EST	$+14126.13 \pm 596.29$	530
MACE-UCB	$+13440.13 \pm 347.78$	530
q-EI	$+12769.20 \pm 457.11$	530
q-EST	$+11790.13 \pm 497.59$	530
LAW-EI	$+11914.40 \pm 345.21$	530
LAW-EST	$+12067.07 \pm 237.50$	530
LAW-prior-EI	$+11875.67 \pm 771.34$	530
LAW-prior-EST	$+11842.53 \pm 301.49$	530

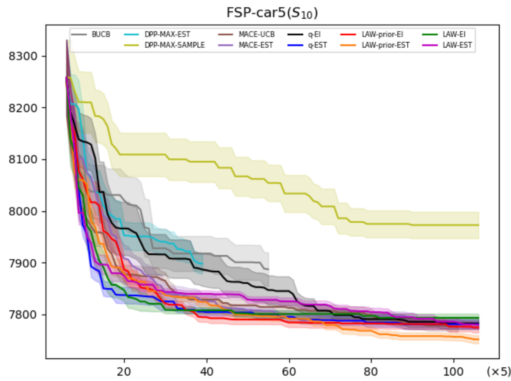


Method	Mean±Std.Err.	#Eval
DPP-MAX-EST	$+3899.60 \pm 23.04$	530
DPP-SAMPLE-EST	$+4446.13 \pm 22.45$	530
MACE-EST	$+4085.87 \pm 19.65$	530
MACE-UCB	$+4030.80 \pm 26.37$	530
q-EI	$+3653.07 \pm 10.06$	530
q-EST	$+3690.00 \pm 15.16$	530
LAW-EI	$+3724.00 \pm 12.71$	530
LAW-EST	$+3730.93 \pm 9.03$	530
LAW-prior-EI	$+3777.47 \pm 7.90$	530
LAW-prior-EST	$+3695.47 \pm 11.49$	530

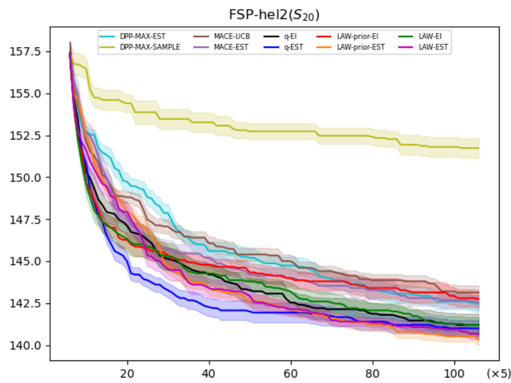


Method	Mean±Std.Err.	#Eval
DPP-MAX-EST	$+276.53 \pm 3.87$	510
DPP-SAMPLE-EST	$+319.60 \pm 3.78$	510
MACE-EST	$+285.60 \pm 3.13$	830
MACE-UCB	$+250.27 \pm 3.51$	830
q-EI	$+172.67 \pm 3.23$	830
q-EST	$+171.20 \pm 1.84$	830
LAW-EI	$+192.53 \pm 5.26$	830
LAW-EST	$+191.73 \pm 2.89$	830
LAW-prior-EI	$+216.13 \pm 5.30$	830
LAW-prior-EST	$+188.93 \pm 1.91$	830

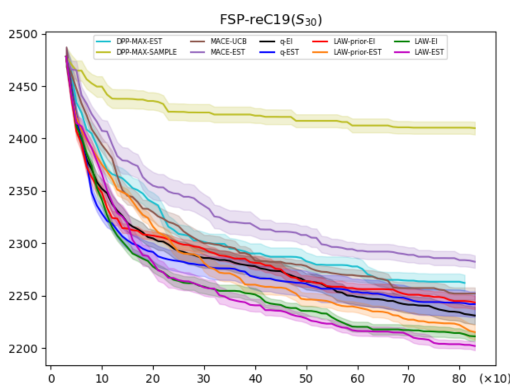
G.2 Flow-shop Scheduling Problems



Method	Mean±Std.Err.	#Eval
BUCB	$+7887.20 \pm 32.37$	275
DPP-MAX-EST	$+7795.67 \pm 11.11$	530
DPP-SAMPLE-EST	$+7972.73 \pm 25.60$	530
MACE-EST	$+7791.27 \pm 9.34$	530
MACE-UCB	$+7775.87 \pm 9.73$	530
q-EI	$+7782.67 \pm 10.76$	530
q-EST	$+7781.94 \pm 9.25$	530
LAW-EI	$+7793.53 \pm 7.89$	530
LAW-EST	$+7779.87 \pm 7.29$	530
LAW-prior-EI	$+7774.93 \pm 9.97$	530
LAW-prior-EST	$+7751.94 \pm 7.80$	530

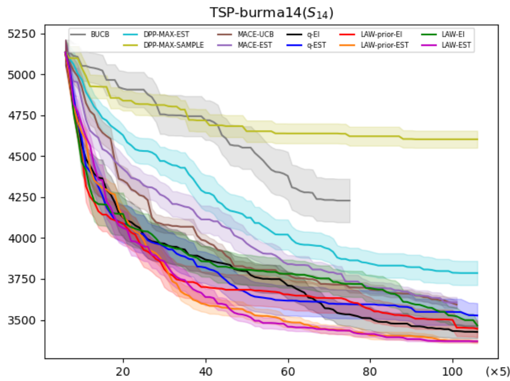


Method	Mean±Std.Err.	#Eval
DPP-MAX-EST	$+142.47 \pm 0.48$	530
DPP-SAMPLE-EST	$+151.73 \pm 0.58$	530
MACE-EST	$+142.53 \pm 0.45$	530
MACE-UCB	$+143.13 \pm 0.42$	530
q-EI	$+141.20 \pm 0.66$	530
q-EST	$+141.00 \pm 0.49$	530
LAW-EI	$+141.20 \pm 0.45$	530
LAW-EST	$+140.67 \pm 0.31$	530
LAW-prior-EI	$+142.73 \pm 0.49$	530
LAW-prior-EST	$+140.33 \pm 0.35$	530

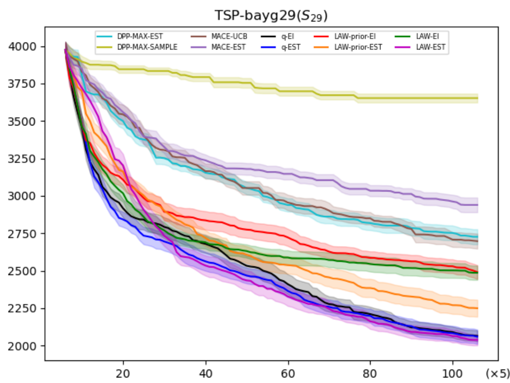


Method	Mean±Std.Err.	#Eval
DPP-MAX-EST	$+2262.13 \pm 7.66$	810
DPP-SAMPLE-EST	$+2409.87 \pm 6.09$	830
MACE-EST	$+2282.40 \pm 5.86$	830
MACE-UCB	$+2252.00 \pm 5.79$	830
q-EI	$+2231.07 \pm 8.39$	830
q-EST	$+2241.87 \pm 12.06$	830
LAW-EI	$+2211.20 \pm 4.47$	830
LAW-EST	$+2202.00 \pm 4.17$	830
LAW-prior-EI	$+2243.60 \pm 6.56$	830
LAW-prior-EST	$+2215.27 \pm 7.20$	830

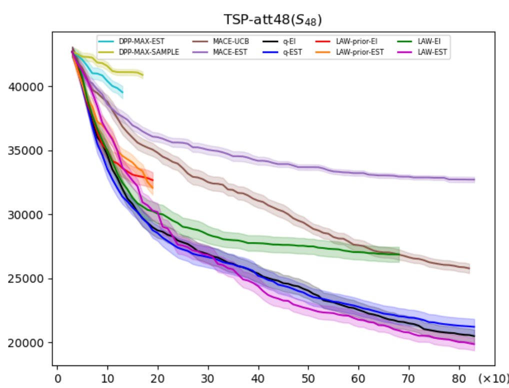
G.3 Traveling Salesman Problems



Method	Mean±Std.Err.	#Eval
BUCB	+4184.20 ± 132.13	405
DPP-MAX-EST	+3786.00 ± 73.76	530
DPP-SAMPLE-EST	+4602.93 ± 52.15	530
MACE-EST	+3575.53 ± 25.04	530
MACE-UCB	+3582.93 ± 20.93	530
q-EI	+3426.53 ± 39.93	530
q-EST	+3526.80 ± 75.02	530
LAW-EI	+3465.87 ± 25.69	530
LAW-EST	+3369.27 ± 7.20	530
LAW-prior-EI	+3445.53 ± 51.35	530
LAW-prior-EST	+3367.40 ± 10.66	530

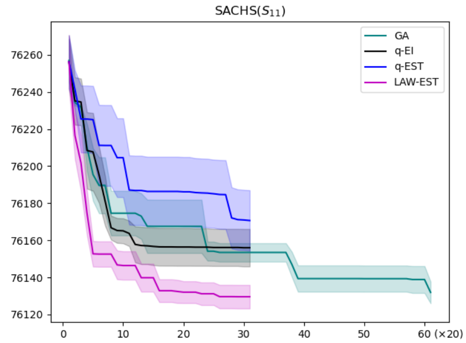


Method	Mean±Std.Err.	#Eval
DPP-MAX-EST	+2726.93 ± 50.37	530
DPP-SAMPLE-EST	+3652.87 ± 29.48	530
MACE-EST	+2939.67 ± 49.04	530
MACE-UCB	+2697.93 ± 50.18	530
q-EI	+2065.13 ± 36.48	530
q-EST	+2059.73 ± 47.93	530
LAW-EI	+2486.87 ± 47.36	530
LAW-EST	+2038.40 ± 36.28	530
LAW-prior-EI	+2491.27 ± 46.49	530
LAW-prior-EST	+2250.00 ± 56.72	530



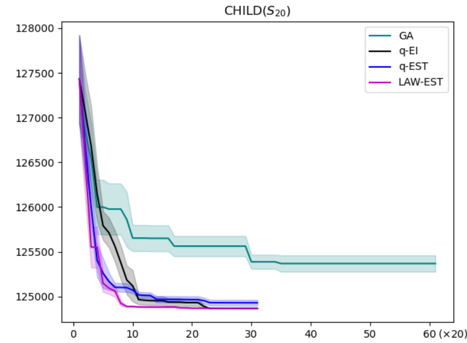
Method	Mean±Std.Err.	#Eval
DPP-MAX-EST	+39539.47 ± 486.85	130
DPP-SAMPLE-EST	+40893.30 ± 265.03	170
MACE-EST	+32710.55 ± 212.19	830
MACE-UCB	+25772.51 ± 370.62	820
q-EI	+20472.44 ± 502.39	830
q-EST	+21199.09 ± 619.65	830
LAW-EI	+26864.42 ± 589.32	680
LAW-EST	+19846.04 ± 484.86	830
LAW-prior-EI	+32670.35 ± 614.68	190
LAW-prior-EST	+32072.59 ± 544.12	190

G.4 Structure Learning



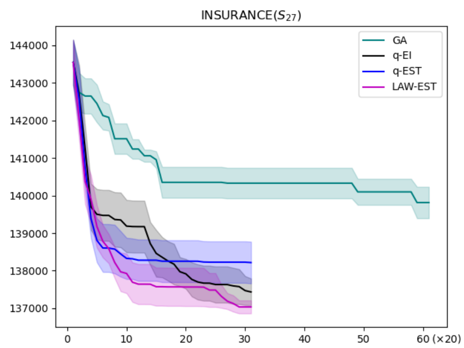
$C = 76100$

Method	#Eval	Mean±Std.Err.
GA	620	$(C + 53.46) \pm\ \ 4.99$
GA	1240	$(C + 31.90) \pm\ \ 5.86$
q-EI	620	$(C + 55.98) \pm 10.11$
q-EST	620	$(C + 70.67) \pm 16.31$
LAW-EST	620	$(C + 29.58) \pm\ \ 6.36$



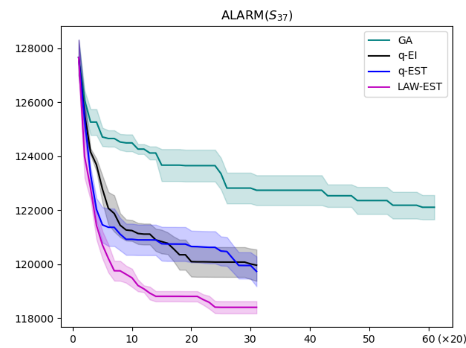
$C = 124000$

Method	#Eval	Mean±Std.Err.
GA	620	$(C + 1387.12) \pm 79.26$
GA	1240	$(C + 1368.07) \pm 92.26$
q-EI	620	$(C + \ \ 864.85) \pm\ \ 0.16$
q-EST	620	$(C + \ \ 928.83) \pm 32.97$
LAW-EST	620	$(C + \ \ 866.64) \pm\ \ 0.39$



$C = 135000$

Method	#Eval	Mean±Std.Err.
GA	620	$(C + 5330.60) \pm 406.92$
GA	1240	$(C + 4814.04) \pm 418.49$
q-EI	620	$(C + 2433.23) \pm 357.18$
q-EST	620	$(C + 3215.75) \pm 556.36$
LAW-EST	620	$(C + 2033.95) \pm 174.04$



$C = 117000$

Method	#Eval	Mean±Std.Err.
GA	620	$(C + 5825.19) \pm 570.55$
GA	1240	$(C + 5114.97) \pm 449.93$
q-EI	620	$(C + 2969.00) \pm 581.67$
q-EST	620	$(C + 2739.77) \pm 554.12$
LAW-EST	620	$(C + 1409.27) \pm 227.57$