Supplementary Material to "Machine Learning for Online Algorithm Selection under Censored Feedback"

We are aware that many of the details we provide in the appendix in the following are important. In case the reviewers find this paper worthy of publication, we plan to buy additional pages in order to include as much of the information as possible.

A Table of Symbols

The following table contains a list of symbols that are frequently used in the main paper as well as in the following supplementary material.

Basics	
$1_{[\![\cdot]\!]}$	indicator function
$A^\intercal$	transpose of a matrix A (possibly non-quadratic)
A^{-1}	inverse of an invertible (quadratic) matrix A
$\|\boldsymbol{x}\|$	$\sqrt{\boldsymbol{x}^\intercal \boldsymbol{x}}$ for any $\boldsymbol{x} \in \mathbb{R}^d$ (Euclidean norm)
$\|\boldsymbol{x}\|_A$	$\sqrt{\boldsymbol{x}^\intercal A^{-1} \boldsymbol{x}}$ for any $\boldsymbol{x} \in \mathbb{R}^d$ and semi-positive definite matrix $A \in \mathbb{R}^{d \times d}$
$\boldsymbol{x}[i]$	ith component of a vector $\boldsymbol{x} \in \mathbb{R}^d$ for $i = 1, \ldots, d$
I_d	identity matrix of size $d \times d$
$\mathrm{N}(\boldsymbol{\mu}, \Sigma)$	multivariate Gaussian distribution with mean vector $\boldsymbol{\mu} \in \mathbb{R}^d$ and covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$
$\mathrm{LN}(\mu, \sigma^2)$	log-normal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$
Φ_{μ,σ^2}	univariate Gaussian distribution with mean $\mu \in \mathbb{R}$ and standard deviation $\sigma > 0$
$X \sim P$	X is distributed as P **or** X is sampled from P
$\mathbb{P}, \mathbb{E}[\cdot]$	probability, expected value
OAS related	
$\mathcal{A}$	(finite) set of algorithms
$\mathcal{I}$	problem instance space
$f : \mathcal{I} \to \mathbb{R}^d$	feature function ($d \in \mathbb{N}$ feature dimensionality)
i_t	problem instance at timestep t
$f(i), \boldsymbol{f}_{i_t}$	features of problem instance i, $i_t \in \mathcal{I}$
$s : \mathcal{H} \times \mathcal{I} \to \mathcal{A}$	algorithm selector
a_t	algorithm chosen by algorithm selector at time step t, i.e., $a_t = s_t(i_t)$
$l : \mathcal{I} \times \mathcal{A} \to \mathbb{R}$	loss function
$l_{i,a}, l_{t,a}$	loss of algorithm a (or a_t) used on problem instance i (or i_t)
$s^* : \mathcal{H} \times \mathcal{I} \to \mathcal{A}$	oracle or virtual best solver $s_t^*(h_t, i_t) := \arg\min_{a \in \mathcal{A}} \mathbb{E}[l(i_t, a) \mid h_t]$
T	number of overall timesteps (element of $\mathbb{N}$)
$\mathcal{L}(s)$	averaged cumulative loss of algorithm selector s, i.e., $T^{-1} \sum_{t=1}^{T} l(i_t, a_t)$ if T finite, else $\lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} l(i_t, a_t)$
Runtime related	
$\mathcal{P} : \mathbb{R} \to \mathbb{R}$	penalty function (for penalizing unsolved instance)
C	cutoff time (element of $\mathbb{R}_+$)
$\boldsymbol{\theta}_a^*$	unknown weight vector (element of $\mathbb{R}^d$) of algorithm a according to the model in (5)
$m : \mathcal{I} \times \mathcal{A} \to \mathbb{R}_+$	runtime function, i.e., $m(i, a)$ is the (noisy) runtime of algorithm a on problem instance i
$m_{i,a}$	abbreviation for $m(i, a)$
$y_{i,a}$	logarithmic (noisy) runtime of algorithm a on problem instance i, i.e., $\log(m_{i,a})$
$\tilde{y}_{i,a}$	$\min(y_{i,a}, \log(C))$, i.e., possibly imputed logarithmic (noisy) runtime of algorithm a on problem instance i
$\epsilon_{i,a}$	stochastic noise variable in model (5)
Bandit related	
$X_{t,a}$	the design matrix corresponding to algorithm a at timestep t (stacking row by row the feature vectors whenever a is chosen)
$A_{t,a}$	$\lambda I_d + X_{t,a}^\intercal X_{t,a}$, i.e., the regularized (through $\lambda > 0$) Gram matrix corresponding to algorithm a at timestep t
$\widehat{\boldsymbol{\theta}}_{t,a}$	ridge regression (RR) estimate for θ_a^* at timestep t for linearized model (cf. (6)) (non-imputed version defined in (7), imputed version defined in (9))
$w_{t,a} : \mathbb{R}^d \to \mathbb{R}_+$	confidence width (function) of non-imputed RR estimate $\widehat{\boldsymbol{\theta}}_{t,a}$ for θ_a^* at timestep t for linearized model (cf. text below (8))
$w_{t,a}^{(bc)} : \mathbb{R}^d \to \mathbb{R}_+$	confidence width (function) of imputed RR estimate $\widehat{\boldsymbol{\theta}}_{t,a}$ for θ_a^* at timestep t for linearized model (cf. Sec. 4.1)
$\tilde{w}_{t,a} : \mathbb{R}^d \to \mathbb{R}_+$	randomized confidence width (function) of imputed RR estimate $\widehat{\boldsymbol{\theta}}_{t,a}$ for θ_a^* at timestep t for linearized model (cf. Sec. 4.2)
$N_{a,t}^{(C)}$	number of timeouts of algorithm a until t
$E_C, E_C(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*, \sigma)$	conditional expectation of $\mathrm{LN}(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*, \sigma^2)$ under a cutoff C (cf. (18))
$C_{i_t,a}^{(1)}$	$\frac{\log(C) - \boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^* - \sigma^2}{\sigma}$
$C_{i_t,a}^{(2)}$	$C_{i_t,a}^{(1)} + \sigma$

B Pseudo Codes

In this section, we provide pseudo codes for the algorithms introduced in Sec. 4 and 5. However, we abstain from defining the pseudo code of the plain Thompson Sampling algorithm in Sec. 4.3 as it should be clear from its definition. Moreover, note that it is straightforward to see that the solution of (7) is $\widehat{\boldsymbol{\theta}}_{t,a} = (A_{t,a})^{-1}\boldsymbol{b}_{t,a}$, where $\boldsymbol{b}_{t,a} = X_{t,a}^{\intercal}\boldsymbol{y}_{t,a}$ and $\boldsymbol{y}_{t,a}$ is the (column) vector storing all observed non-censored log-runtimes until t whenever a has been chosen. The solution of (9) is $\widehat{\boldsymbol{\theta}}_{t,a} = (A_{t,a})^{-1}\boldsymbol{b}_{t,a}$, where $\boldsymbol{b}_{t,a} = X_{t,a}^{\intercal}\tilde{\boldsymbol{y}}_{t,a}$ and $\tilde{\boldsymbol{y}}_{t,a}$ is the (column) vector storing all observed and possibly imputed log-runtimes until t whenever a has been chosen. Further, both $A_{t,a}$ and $\boldsymbol{b}_{t,a}$ can be updated in an iterative fashion without actually storing all seen samples: If algorithm a is chosen at timestep t, then

$$A_{t+1,a} = \begin{cases} A_{t,a} + \boldsymbol{f}_{i_t}\boldsymbol{f}_{i_t}^{\intercal}, \\ A_{t,a} + 1_{[\![m(i_t,a_t)\leqslant C]\!]}\boldsymbol{f}_{i_t}\boldsymbol{f}_{i_t}^{\intercal}, \end{cases} \quad \text{and} \quad \boldsymbol{b}_{t+1,a} = \begin{cases} \boldsymbol{b}_{t,a} + \tilde{y}_{i_t,a}\boldsymbol{f}_{i_t}, & \text{for (9)}, \\ \boldsymbol{b}_{t,a} + 1_{[\![m(i_t,a_t)\leqslant C]\!]}y_{i_t,a}\boldsymbol{f}_{i_t}, & \text{for (7)}. \end{cases}$$

As the updates of the matrices $A_{t+1,a}$ are via a rank-one update, one can use the well-known Sherman-Morrison formula to compute their inverse in a sequential manner as well (similarly for $A_{t+1,a}$ based on (7)):

$$(A_{t+1,a})^{-1} = (A_{t,a} + \boldsymbol{f}_{i_t}\boldsymbol{f}_{i_t}^{\intercal})^{-1} = A_{t,a}^{-1} - \frac{A_{t,a}^{-1}\boldsymbol{f}_{i_t}\boldsymbol{f}_{i_t}^{\intercal}A_{t,a}^{-1}}{1 + \boldsymbol{f}_{i_t}^{\intercal}A_{t,a}^{-1}\boldsymbol{f}_{i_t}}.$$

B.1 LinUCB

Alg. 1 provides the pseudo code for the LinUCB variants introduced in Sec. 4, that is, BlindUCB, BClinUCB and RandUCB. For sake of convenience, we recall the role of each input parameter in the following.

- $\lambda > 0$ is a regularization parameter due to the considered ridge regression.
- $\alpha > 0$ essentially controls the degree of exploration as a multiplicative term of the confidence width (see (8)). The higher (lower) it is chosen, the more (less) exploration is conducted.
- $C > 0$ is the cutoff time and depends on the considered problem scenario (specified by the ASlib library).
- $\tilde{\sigma} > 0$ is (only) used for RandUCB in order to specify the random sample's variance within the confidence width (see Sec. 4.2). The higher (lower) its choice, the larger (smaller) the effective exploration term, i.e., $|r| \cdot w_{t,a}^{(bc)}(\boldsymbol{x}_t)$.

Algorithm 1 `LinUCB` variants

Input parameters $\lambda \geqslant 0, \alpha, C > 0$ half-normal parameter $\tilde{\sigma}$ for RandUCB
Initialization
for all $a \in \mathcal{A}$ **do**
  $A_{t,a} = \lambda I_{d\times d}, \boldsymbol{b}_{t,a} = 0_{d\times 1}, \widehat{\boldsymbol{\theta}}_{t,a} = 0_{d\times 1}, \widehat{l}_{t,a} = 0, N_{t,a}^{(C)} = 0$
end for
Main Algorithm
for time steps $t = 1 \ldots, T$ **do**
  Observe instance i_t and its features $\boldsymbol{x}_t = f(i_t) \in \mathbb{R}^d$
  if $t \leqslant |\mathcal{A}|$ **then**
    Take algorithm $a_t \in \mathcal{A}$ and obtain $y_t = \min(\log(m(i_t, a_t)), \log(C))$
  else
    for all $a \in \mathcal{A}$ **do**
      $\widehat{\boldsymbol{\theta}}_{t,a} \leftarrow (A_{t,a})^{-1}\boldsymbol{b}_{t,a}$
      $\widehat{l}_{t,a} \leftarrow \begin{cases} \boldsymbol{x}_t^{\intercal}\widehat{\boldsymbol{\theta}}_{t,a} - \alpha \cdot w_{t,a}(\boldsymbol{x}_t), & \text{BlindUCB} \\ \boldsymbol{x}_t^{\intercal}\widehat{\boldsymbol{\theta}}_{t,a} - \alpha \cdot w_{t,a}^{(bc)}(\boldsymbol{x}_t), & \text{BClinUCB} \\ \boldsymbol{x}_t^{\intercal}\widehat{\boldsymbol{\theta}}_{t,a} - \alpha \cdot |r| \cdot w_{t,a}^{(bc)}(\boldsymbol{x}_t), \quad r \sim N(0, \tilde{\sigma}^2) & \text{RandUCB} \end{cases}$
    end for
    Take algorithm $a_t = \arg\min_{a\in\mathcal{A}} \hat{l}_{t,a}$ and obtain $y_t = \min(\log(m(i_t, a_t)), \log(C))$
  end if
  Updates:
  $N_{t,a}^{(C)} \leftarrow N_{t,a}^{(C)} + 1_{[\![m(i_t,a_t)>C]\!]}$
  $A_{t,a} \leftarrow \begin{cases} A_{t,a} + \boldsymbol{x}_t\boldsymbol{x}_t^{\intercal}, \\ A_{t,a} + 1_{[\![m(i_t,a_t)\leqslant C]\!]}\boldsymbol{x}_t\boldsymbol{x}_t^{\intercal}, \end{cases} \quad \boldsymbol{b}_{t,a} \leftarrow \begin{cases} \boldsymbol{b}_{t,a} + y_t\boldsymbol{x}_t, & \text{(BClinUCB,RandUCB)} \\ \boldsymbol{b}_{t,a} + 1_{[\![m(i_t,a_t)\leqslant C]\!]}y_t\boldsymbol{x}_t, & \text{(BlindUCB)} \end{cases}$
end for

B.2 LinUCB Revisited

Alg. 2 provides the pseudo code for the LinUCB variants introduced in Sec. 4 adapted to the refined loss decomposition in (11) by incorporating the selections strategy in (12). Compared to Alg. 1 there are two additional parameters:

- $\sigma > 0$ which ideally should correspond to the standard deviation of the noise variables in the model assumption (5).
- $\mathcal{P} : \mathbb{R} \to \mathbb{R}$ the penalty function for penalizing unsolved problem instances (we used $\mathcal{P}(z) = 10z$ corresponding to the $PAR10$ score).

Algorithm 2 `LinUCB` variants based on (11) (_rev) versions

1: *Input parameters* $\lambda \geqslant 0, \sigma > 0, \alpha > 0, C > 0, \mathcal{P} : \mathbb{R} \to \mathbb{R}$, half-normal parameter $\tilde{\sigma}$ for RandUCB
2: *Initialization*
3: **for** all $a \in \mathcal{A}$ **do**
4: $\quad A_{t,a} = \lambda I_{d\times d}, \boldsymbol{b}_{t,a} = 0_{d\times 1}, \widehat{\boldsymbol{\theta}}_{t,a} = 0_{d\times 1}, \widehat{l}_{t,a} = 0, N^{(C)}_{t,a} = 0$
5: **end for**
6: *Main Algorithm*
7: **for** time steps $t = 1 \ldots, T$ **do**
8: $\quad$ Observe instance i_t and its features $\boldsymbol{x}_t = f(i_t) \in \mathbb{R}^d$
9: $\quad$ **if** $t \leqslant |\mathcal{A}|$ **then**
10: $\quad\quad$ Take algorithm $a_t \in \mathcal{A}$ and obtain $y_t = \min(\log(m(i_t, a_t)), \log(C))$
11: $\quad$ **else**
12: $\quad\quad$ **for** all $a \in \mathcal{A}$ **do**
13: $\quad\quad\quad \widehat{\boldsymbol{\theta}}_{t,a} \leftarrow (A_{t,a})^{-1}\boldsymbol{b}_{t,a}$
14: $\quad\quad\quad r \sim \mathrm{N}(0, \tilde{\sigma}^2)$ (only for RandUCB)
15: $\quad\quad\quad p_{t,a} \leftarrow \begin{cases} \boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} + \alpha \cdot w_{t,a}(\boldsymbol{x}_t), & \\ \boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} + \alpha \cdot w^{(bc)}_{t,a}(\boldsymbol{x}_t), & \\ \boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} + \alpha \cdot |r| \cdot w^{(bc)}_{t,a}(\boldsymbol{x}_t), & \end{cases} \quad o_{t,a} \leftarrow \begin{cases} \boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} - \alpha \cdot w_{t,a}(\boldsymbol{x}_t), & \text{BlindUCB} \\ \boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} - \alpha \cdot w^{(bc)}_{t,a}(\boldsymbol{x}_t), & \text{BClinUCB} \\ \boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} - \alpha \cdot |r| \cdot w^{(bc)}_{t,a}(\boldsymbol{x}_t), & \text{RandUCB} \end{cases}$
16: $\quad\quad\quad \hat{C}^{(1,p)}_{i_t,a} \leftarrow \frac{\log(C) - p_{t,a} - \sigma^2}{\sigma}, \quad \hat{C}^{(1,o)}_{i_t,a} \leftarrow \frac{\log(C) - o_{t,a} - \sigma^2}{\sigma}, \quad \hat{C}^{(2,o)}_{i_t,a} \leftarrow \frac{\log(C) - o_{t,a}}{\sigma}, \quad \text{and} \quad \hat{C}^{(2,p)}_{i_t,a} \leftarrow \frac{\log(C) - p_{t,a}}{\sigma}.$
17: $\quad\quad\quad \widehat{l}_{t,a} \leftarrow \exp(o_{t,a} + \sigma^2/2) \cdot \frac{\Phi_{0,1}(\hat{C}^{(1,p)}_{i_t,a})}{\Phi_{0,1}(\hat{C}^{(2,o)}_{i_t,a})} + \big(1 - \Phi_{p_{t,a},\sigma}(\log(C))\big) \cdot \left(\mathcal{P}(C) - \exp(p_{t,a} + \sigma^2/2) \cdot \frac{\Phi_{0,1}(\hat{C}^{(1,o)}_{i_t,a})}{\Phi_{0,1}(\hat{C}^{(2,p)}_{i_t,a})}\right)$
18: $\quad\quad$ **end for**
19: $\quad\quad$ Take algorithm $a_t = \arg\min_{a\in\mathcal{A}} \hat{l}_{t,a}$ and obtain $y_t = \min(\log(m(i_t, a_t)), \log(C))$
20: $\quad$ **end if**
21: $\quad$ *Updates:* Same as lines 19–20 of Alg. 1
22: **end for**

B.3 Thompson Sampling Revisited

Alg. 3 provides the pseudo code for the revisited Thompson algorithm and its variant inspired by the Buckley-James estimate introduced in Sec. 5.2. As before, we discuss the role each input parameter plays in the behavior of the algorithm:

- the role of $\lambda, \mathcal{P}, C$ is the same as in Alg. 2, respectively.
- $\sigma > 0$ specifies the magnitude of the posterior distribution's variance and is therefore slightly different to σ in Alg. 2.
- BJ specifies whether the Buckley-James inspired imputation strategy described at the end of Sec. 5.2 (BJ = TRUE) or the näive imputation strategy (BJ = FALSE) should be deployed.

Algorithm 3 (bj_) Thompson_rev

Input parameters $\sigma > 0, \lambda \geqslant 0, \mathcal{P} : \mathbb{R} \to \mathbb{R}, C$, BJ $\in \{\text{TRUE}, \text{FALSE}\}$
Initialization
for all $a \in \mathcal{A}$ **do**
  $A_{t,a} = \lambda \cdot I_{d\times d}, \boldsymbol{b}_{t,a} = 0_{d\times 1}, \widehat{\boldsymbol{\theta}}_{t,a} = 0_{d\times 1}, \tilde{\sigma}^2_{t,a} \leftarrow 0$ and $\widehat{l}_{t,a} = 0$
end for
Main Algorithm
for time steps $t = 1 \ldots, T$ **do**
  Observe instance i_t and its features $\boldsymbol{x}_t = f(i_t) \in \mathbb{R}^d$
  if $t \leqslant |\mathcal{A}|$ **then**
    Take algorithm $a_t \in \mathcal{A}$ and obtain $y_t = \min(\log(m(i_t, a_t)), \log(C))$
  else
    for all $a \in \mathcal{A}$ **do**
      $\widehat{\boldsymbol{\theta}}_{t,a} \leftarrow (A_{t,a})^{-1}\boldsymbol{b}_{t,a} \quad \tilde{\sigma}^2_{t,a} \leftarrow \sigma\|\boldsymbol{f}_{i_t}\|^2_{A_{t,a}}$
      Sample $\widetilde{\boldsymbol{\theta}}_a \sim \mathrm{N}\big(\widehat{\boldsymbol{\theta}}_{t,a}, \sigma(A_{t,a})^{-1}\big)$
      $\widehat{l}_{t,a} \leftarrow \big(1 - \Phi_{\boldsymbol{f}_{i_t}^\intercal \tilde{\boldsymbol{\theta}}_a, \tilde{\sigma}^2_{t,a}}(\log(C))\big) \cdot \big(\mathcal{P}(C) - E_C(\boldsymbol{f}_{i_t}^\intercal \widetilde{\boldsymbol{\theta}}_a, \tilde{\sigma}_{t,a})\big) + E_C(\boldsymbol{f}_{i_t}^\intercal \widetilde{\boldsymbol{\theta}}_a, \tilde{\sigma}_{t,a})$ (RHS of (13))
    end for
    Take algorithm $a_t = \arg\min_{a\in\mathcal{A}} \hat{l}_{t,a}$ and obtain $y_t = \min(\log(m(i_t, a_t)), \log(C))$
  end if
  Updates:
  if $y_t = \log(C)$ and BJ = TRUE **then**
    Sample $\breve{\boldsymbol{\theta}}_a \sim \mathrm{N}\big(\widehat{\boldsymbol{\theta}}_{t,a}, \sigma(A_{t,a})^{-1}\big)$ (if $\exp(\boldsymbol{x}_t^\intercal \breve{\boldsymbol{\theta}}_a) \leqslant C$ sample again)
    $y_t \leftarrow \log(\boldsymbol{x}_t^\intercal \breve{\boldsymbol{\theta}}_a)$
  end if
  $A_{t,a} \leftarrow A_{t,a} + \boldsymbol{x}_t\boldsymbol{x}_t^\intercal$
  $\boldsymbol{b}_{t,a} \leftarrow \boldsymbol{b}_{t,a} + y_t\boldsymbol{x}_t$
end for

C Deriving the Bias-corrected Confidence Bounds

This section is concerned with giving the details for deriving the bias-corrected confidence bounds in Sec. 4.1. In particular, we focus on the solution of (9) which is given by $\widehat{\boldsymbol{\theta}}_{t,a} = (A_{t,a})^{-1}\boldsymbol{b}_{t,a}$, where $\boldsymbol{b}_{t,a} = X_{t,a}^\intercal \tilde{\boldsymbol{y}}_{t,a}$ and $\tilde{\boldsymbol{y}}_{t,a}$ is the (column) vector storing all observed and possibly imputed log-runtimes until t whenever a has been chosen. We follow the approach by (Chu et al. 2011) and analyze the deviation of $\boldsymbol{x}_t^\intercal\hat{\theta}_{t,a}$ and $\boldsymbol{x}_t^\intercal\boldsymbol{\theta}^*_a$, where we write for sake of convenience $\boldsymbol{x}_t = f(i_t)$. First, note that

$$\begin{aligned}
\boldsymbol{x}_t^\intercal\widehat{\boldsymbol{\theta}}_{t,a} - \boldsymbol{x}_t^\intercal\boldsymbol{\theta}^*_a &= \boldsymbol{x}_t^\intercal A_{t,a}^{-1}\boldsymbol{b}_{t,a} - \boldsymbol{x}_t^\intercal A_{t,a}^{-1}A_{t,a}\boldsymbol{\theta}^*_a \\
&= \boldsymbol{x}_t^\intercal A_{t,a}^{-1}X_{t,a}^\intercal\tilde{\boldsymbol{y}}_{t,a} - \boldsymbol{x}_t^\intercal A_{t,a}^{-1}\Big(\lambda I_d + X_{t,a}^\intercal X_{t,a}\Big)\boldsymbol{\theta}^*_a \\
&= \boldsymbol{x}_t^\intercal A_{t,a}^{-1}X_{t,a}^\intercal\Big(\tilde{\boldsymbol{y}}_{t,a} - X_{t,a}\boldsymbol{\theta}^*_a\Big) - \lambda\boldsymbol{x}_t^\intercal A_{t,a}^{-1}\boldsymbol{\theta}^*_a \\
&= \boldsymbol{z}_{t,a}(\tilde{\boldsymbol{y}}_{t,a} - X_{t,a}\boldsymbol{\theta}^*_a) - \lambda\boldsymbol{x}_t^\intercal A_{t,a}^{-1}\boldsymbol{\theta}^*_a \\
&=: (A1) - (A2),
\end{aligned} \tag{14}$$

where we abbreviated $\boldsymbol{x}_t^\intercal A_{t,a}^{-1}X_{t,a}^\intercal$ by $\boldsymbol{z}_{t,a}$, which is a row vector with $N_a(t)$ components, i.e., $N_a(t)$ is the total number of times a has been chosen till t. We can write the term $(A1)$ as

$$\sum_{j:m_{i_j,a}\leqslant C} \boldsymbol{z}_{t,a}[j]\Big(\log(m(i_j, a)) - \boldsymbol{x}_{i_j}^\intercal\boldsymbol{\theta}^*_a\Big) + \sum_{j:m_{i_j,a}>C} \boldsymbol{z}_{t,a}[s]\Big(\log(C) - \boldsymbol{x}_{i_j}^\intercal\boldsymbol{\theta}^*_a\Big),$$

i.e., we split it into the sum over all uncensored observations and the sum of all censored observations. This is necessary as we want to apply Azuma's inequality and this can only be done for the terms in the first sum, since $\mathbb{E}\left[\log(m(i_j,a))\right] = \boldsymbol{x}_{i_j}^\intercal \theta_a^*$ due to our assumptions made in Sec. 3. By applying Azuma's inequality we get for any $\tilde{\alpha} > 0$ that

$$\mathbb{P}\Big(\Big|\sum\nolimits_{j:m_{i_j,a}\leqslant C} \boldsymbol{z}_{t,a}[j]\big(\log(m(i_j,a)) - \boldsymbol{x}_{i_j}^\intercal \boldsymbol{\theta}_a^*\big)\Big| > \tilde{\alpha}\, w_{t,a}(\boldsymbol{x}_t)\Big) \leqslant 2\exp\left(-\frac{2\,\tilde{\alpha}^2\, w_{t,a}^2(\boldsymbol{x}_t)}{\|\boldsymbol{z}_{t,a}\|^2}\right),$$

where $w_{t,a}$ is as in Sec. 4 defined by $w_{t,a}(\boldsymbol{x}_t) = \|\boldsymbol{x}_t\|_{A_{t,a}}$. Note that

$$\|\boldsymbol{z}_{t,a}\|^2 = \boldsymbol{x}_t^\intercal A_{t,a}^{-1} X_{t,a}^T X_{t,a} A_{t,a}^{-1} \boldsymbol{x}_t \leqslant \boldsymbol{x}_t^\intercal A_{t,a}^{-1}(\lambda I_d + X_{t,a}^T X_{t,a}) A_{t,a}^{-1}\boldsymbol{x}_t = \boldsymbol{x}_t^\intercal A_{t,a}^{-1}\boldsymbol{x}_t = w_{t,a}^2(\boldsymbol{x}_t),$$

since $\lambda A_{t,a}^{-1} I_d A_{t,a}^{-1}$ is semi-positive definite. Thus, by choosing $\tilde{\alpha}$ appropriately the latter probability is small. In particular, we obtain that

$$\Big|\sum\nolimits_{j:m_{i_j,a}\leqslant C} \boldsymbol{z}_{t,a}[j]\big(\log(m(i_j,a)) - \boldsymbol{x}_{i_j}^\intercal \boldsymbol{\theta}_a^*\big)\Big| \leqslant \tilde{\alpha}\, w_{t,a}(\boldsymbol{x}_t) \tag{15}$$

holds with high probability by choosing $\tilde{\alpha} > 0$ appropriately. Now, we bound the (absolute values of the) censored sum. Using the Cauchy-Schwarz inequality we can infer

$$\begin{aligned}\Big|\sum\nolimits_{j:m_{i_j,a}>C} \boldsymbol{z}_{t,a}[s]\big(\log(C) - \boldsymbol{x}_{i_j}^\intercal \boldsymbol{\theta}_a^*\big)\Big| &\leqslant \|\boldsymbol{z}_{t,a}\| \Big\|\sum\nolimits_{j:m_{i_j,a}>C} \big(\log(C) - \boldsymbol{x}_{i_j}^\intercal \boldsymbol{\theta}_a^*\big)\Big\| \\ &\leqslant w_{t,a}(\boldsymbol{x}_t) \Big\|\sum\nolimits_{j:m_{i_j,a}>C} \big(\log(C) - \boldsymbol{x}_{i_j}^\intercal \boldsymbol{\theta}_a^*\big)\Big\| \\ &\leqslant 2\, w_{t,a}(\boldsymbol{x}_t)\sqrt{N_a^{(C)}(t)}\,\log(C),\end{aligned} \tag{16}$$

where we used for the last inequality that $\boldsymbol{x}_{i_j}^\intercal \boldsymbol{\theta}_a^* = \boldsymbol{f}_{i_j}^\intercal \boldsymbol{\theta}_a^* \leqslant \log(C)$ holds by our assumptions made in Sec. 3. Finally, the absolute value of the term $(A2)$ can be bounded as follows

$$|(A2)| \leqslant \lambda\|\boldsymbol{\theta}_a^*\|\|\boldsymbol{x}_t^\intercal A_{t,a}^{-1}\| \leqslant \lambda\|\boldsymbol{\theta}_a^*\| w_{t,a}(\boldsymbol{x}_t). \tag{17}$$

Combining (14)–(17), we have with high probability (depending on the choice of $\tilde{\alpha}$) that

$$|\boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} - \boldsymbol{x}_t^\intercal \boldsymbol{\theta}_a^*| \leqslant |(A1)| + |(A2)| \leqslant w_{t,a}(\boldsymbol{x}_t)\left(\tilde{\alpha} + \lambda\|\boldsymbol{\theta}_a^*\| + 2\log(C)\sqrt{N_a^{(C)}(t)}\right).$$

Thus, for some appropriate $\alpha > 0$ we have with high probability

$$|\boldsymbol{x}_t^\intercal \widehat{\boldsymbol{\theta}}_{t,a} - \boldsymbol{x}_t^\intercal \boldsymbol{\theta}_a^*| \leqslant \alpha w_{t,a}^{(bc)}(\boldsymbol{x}_t).$$

Bias issue. Note that from the definition of $w_{t,a}^{(bc)}(\boldsymbol{x}_t)$ we can see the bias issue of $\widehat{\boldsymbol{\theta}}_{t,a}$ due to the imputation employed. The term $w_{t,a}(\boldsymbol{x}_t)$ is asymptotically tending against zero with the rate $\approx \sqrt{1/N_a(t)}$. However, $w_{t,a}^{(bc)}(\boldsymbol{x}_t)$ does not tend to zero asymptotically for $t \to \infty$, if $\sqrt{N_a^{(C)}(t)/N_a(t)} \to C'$ for some constant $C' > 0$. The latter condition in turn seems to be satisfied if for any t it holds that $\mathbb{P}(m_{i_t,a} > C) > \epsilon > 0$, i.e., the probability of observing a censored runtime does not vanish in the course of time.

D Deriving the Refined Expected Loss Representation

In this section, we provide the details for showing (11). For this purpose, we need the following lemma showing the explicit form of the conditional expectation of a log-normal distribution under a certain cutoff.

Lemma 1. *Let $Y \sim \mathrm{LN}(\mu, \sigma^2)$, i.e., a log-normally distributed random variable with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$. Then, for any $C > 0$ it holds that*

$$\mathbb{E}\left[Y | Y \leqslant C\right] = \exp(\mu + {}^{\sigma^2}\!/_2) \cdot \frac{\Phi_{0,1}\left(\frac{\log(C)-\mu-\sigma^2}{\sigma}\right)}{\Phi_{0,1}\left(\frac{\log(C)-\mu}{\sigma}\right)}, \tag{18}$$

where $\Phi_{0,1}(\cdot)$ is the cumulative distribution function of a standard normal distribution.

Proof. The density function of Y is given by

$$f(x) = \frac{\exp\left(\frac{-(\log(x)-\mu)^2}{2\sigma^2}\right)}{x\sigma\sqrt{2\pi}}, \quad x > 0.$$

Thus, $f(x) = \frac{\phi_{\mu,\sigma}(\log(x))}{x}$, where $\phi_{\mu,\sigma}(\cdot)$ is the density function of a normal distribution with mean μ and standard deviation σ. Next, note that the density function of Y conditioned on $Y \leqslant C$ is $f(x|x \leqslant C) = \frac{f(x)}{F(C)}$, where $F(\cdot)$ is the cumulative distribution function of Y and given by $F(x) = \Phi_{0,1}\left(\frac{\log(x)-\mu}{\sigma}\right)$ for any $x \in \mathbb{R}$. With this,

$$\begin{aligned}
\mathbb{E}\left[Y|Y \leqslant C\right] &= \int_0^C x f(x|x \leqslant C)\,\mathrm{d}x \\
&= \frac{1}{\Phi_{0,1}\left(\frac{\log(C)-\mu}{\sigma}\right)} \int_0^C \phi_{\mu,\sigma}(\log(x))\,\mathrm{d}x \\
&= \frac{\exp(\mu + \sigma^2/2)}{\Phi_{0,1}\left(\frac{\log(C)-\mu}{\sigma}\right)} \int_{-\infty}^{\frac{\log(C)-\mu}{\sigma}} \phi_{0,1}(z-\sigma)\,\mathrm{d}z \\
&= \exp(\mu + \sigma^2/2) \frac{\Phi_{0,1}\left(\frac{\log(C)-\mu-\sigma^2}{\sigma}\right)}{\Phi_{0,1}\left(\frac{\log(C)-\mu}{\sigma}\right)}.
\end{aligned}$$

Here, we used for the third inequality the substitution $z = \frac{\log(x)-\mu}{\sigma}$, so that $\exp(\sigma z + \mu)\sigma\,\mathrm{d}z = \mathrm{d}x$ and

$$\exp\left(-\frac{(z-\sigma)^2}{2}\right)\exp\left(\frac{\sigma^2}{2}\right) = \exp\left(-\frac{z^2}{2}\right)\exp(\sigma z).$$

□

Recalling the modelling assumption on the runtimes made in (5) as well as assumption (**A2**), we obtain that $m_{i_t,a} \sim \mathrm{LN}(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*, \sigma^2)$. Using Lemma 1 and that $C_{i_t,a}^{(1)} = {}^{(\log(C)-\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^* - \sigma^2)}/_{\sigma}$, $C_{i_t,a}^{(2)} = {}^{(\log(C)-\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*)}/_{\sigma}$, we can derive (11) as follows:

$$\begin{aligned}
\mathbb{E}\left[l_{t,a}|\boldsymbol{f}_{i_t}\right] &= \mathbb{E}\left[l_{t,a}|m_{i_t,a} \leqslant C, \boldsymbol{f}_{i_t}\right] \cdot \mathbb{P}(m_{i_t,a} \leqslant C|\boldsymbol{f}_{i_t}) + \mathbb{E}\left[l_{t,a}|m_{i_t,a} > C, \boldsymbol{f}_{i_t}\right] \cdot \mathbb{P}(m_{i_t,a} > C|\boldsymbol{f}_{i_t}) \\
&= \exp\left(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^* + \frac{\sigma^2}{2}\right) \cdot \left(\frac{\Phi_{0,1}(C_{i_t,a}^{(1)})}{\Phi_{0,1}(C_{i_t,a}^{(2)})}\right) \cdot \mathbb{P}(m_{i_t,a} \leqslant C|\boldsymbol{f}_{i_t}) + \mathcal{P}(C) \cdot \mathbb{P}(m_{i_t,a} > C|\boldsymbol{f}_{i_t}) \\
&= \exp\left(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^* + \frac{\sigma^2}{2}\right) \cdot \left(\frac{\Phi_{0,1}(C_{i_t,a}^{(1)})}{\Phi_{0,1}(C_{i_t,a}^{(2)})}\right) + \mathbb{P}(m_{i_t,a} > C|\boldsymbol{f}_{i_t}) \cdot \left(\mathcal{P}(C) - \exp\left(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^* + \frac{\sigma^2}{2}\right) \cdot \frac{\Phi_{0,1}(C_{i_t,a}^{(1)})}{\Phi_{0,1}(C_{i_t,a}^{(2)})}\right) \\
&= E_C(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*, \sigma) + \left(1 - \Phi_{\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*,\sigma}(\log(C))\right) \cdot \left(\mathcal{P}(C) - E_C(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*, \sigma)\right),
\end{aligned}$$

where $E_C(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^*, \sigma) = \exp(\boldsymbol{f}_{i_t}^\intercal \boldsymbol{\theta}_a^* + {}^{\sigma^2}/_2) \cdot \frac{\Phi_{0,1}(C_{i_t,a}^{(1)})}{\Phi_{0,1}(C_{i_t,a}^{(2)})}$.

E ASlib Overview

ASlib (Bischl et al. 2016) is a curated collection of over 25 different algorithm selection problems, called scenarios, based on different algorithmic problem classes such as SAT, TSP, CSP. Each scenario is made up of several instances for which the performance of a set of algorithms has been evaluated using a certain cutoff to avoid unacceptably long algorithm runs. Table 2 gives an overview of the examined scenarios including relevant statistics.

F Software and Hyperparameter Settings

All experiments are based on Python 3 implementations. A complete list of used packages and the corresponding version number can be found both on Github and in the README of the accompanying source code appendix. Below we give a short (non-exhaustive) list of the most important software used for the corresponding approaches and the hyperparameter settings used for the evaluation.

F.1 Our Approaches

Software All presented approaches (LinUCB and Thompson variants) were implemented in Python by using scipy[3] and numpy[4].

[3] https://www.scipy.org/

[4] https://numpy.org/

Scenario	#I	#F	#A	C	T	MT	minT	maxT
ASP-POTASSCO	1294	138	11	600.00	0.20	0.20	0.14	0.28
BNSL-2016	1179	86	8	7200.00	0.28	0.18	0.12	0.59
CPMP-2015	527	22	4	3600.00	0.28	0.28	0.19	0.37
CSP-2010	2024	86	2	5000.00	0.20	0.20	0.14	0.25
CSP-MZN-2013	4642	155	11	1800.00	0.70	0.70	0.48	0.87
CSP-Minizinc-Time-2016	100	95	20	1200.00	0.50	0.52	0.28	0.82
GRAPHS-2015	5725	35	7	1e+08.00	0.07	0.04	0.03	0.27
MAXSAT-PMS-2016	601	37	19	1800.00	0.39	0.19	0.12	0.96
MAXSAT-WPMS-2016	630	37	18	1800.00	0.58	0.46	0.21	0.96
MAXSAT12-PMS	876	37	6	2100.00	0.41	0.43	0.23	0.57
MAXSAT15-PMS-INDU	601	37	29	1800.00	0.49	0.42	0.12	0.96
MIP-2016	218	143	5	7200.00	0.20	0.10	0.04	0.45
PROTEUS-2014	4021	198	22	3600.00	0.60	0.63	0.37	0.67
QBF-2011	1368	46	5	3600.00	0.55	0.51	0.42	0.72
QBF-2014	1254	46	14	900.00	0.56	0.56	0.37	0.71
QBF-2016	825	46	24	1800.00	0.36	0.31	0.20	0.61
SAT03-16_INDU	2000	483	10	5000.00	0.25	0.24	0.19	0.32
SAT11-HAND	296	115	15	5000.00	0.61	0.62	0.50	0.68
SAT11-INDU	300	115	18	5000.00	0.33	0.33	0.28	0.39
SAT11-RAND	600	115	9	5000.00	0.47	0.45	0.40	0.58
SAT12-ALL	1614	115	31	1200.00	0.54	0.53	0.25	0.74
SAT12-HAND	767	115	31	1200.00	0.67	0.64	0.52	0.93
SAT12-INDU	1167	115	31	1200.00	0.50	0.36	0.26	0.99
SAT12-RAND	1362	115	31	1200.00	0.73	0.87	0.27	0.98
SAT15-INDU	300	54	28	3600.00	0.24	0.21	0.13	0.74
TSP-LION2015	3106	122	4	3600.00	0.10	0.03	0.00	0.32

Table 2: Overview of ASlib scenarios including their number of instances (#I), number of features (#F), number of algorithms (#A), cutoff time (C), average (T), median (MT), minimum (minT) and maximum (maxT) relative number of timeouts per algorithms.

Hyperparameter Settings If not stated differently at the beginning of the corresponding experiment (e.g., sensitivity analysis), the following hyperparameters were used:

- Thompson variants
  - $\sigma = 1.0$
  - $\lambda = 0.5$
- LinUCB variants
  - $\lambda = 1.0$
  - $\alpha = 1.0$
  - $\sigma = 10.0$ (for the _rev variants)
  - $\tilde{\sigma}^2 = 0.25$ (for the rand_ variants)

The values of these parameters were chosen according to a hyperparameter sensitivity analysis (cf. Sec. G).

Caveat All Thompson variants rely on sampling from the multi-variate normal distribution, which we implemented using the 'np.random.multivariate_normal' method. Unfortunately, this method seems to have a bug, which is caused by the underlying BLAS implementation of the corresponding SVD, which is performed as part of the method. Changing to various versions of numpy and BLAS did not solve the problem for us. As a consequence, some of the repetitions of the experiments of some scenarios did not complete for some Thompson variants. Below you can find a table indicating how many repetitions are *missing* for the corresponding variant on the corresponding scenario. We reported the bug and hope to add the missing values to the experimental results once the bug is fixed. However, due to the very few amount of data points missing, we do not assume a relevant change for the final results.

Scenario	Approach	#Missing seeds
CSP-MZN-2013	bj_thompson	2
PROTEUS-2014	bj_thompson	1
PROTEUS-2014	thompson_rev	1
PROTEUS-2014	thompson	2
SAT03-16_INDU	bj_thompson_rev	2
SAT03-16_INDU	bj_thompson	1
SAT03-16_INDU	thompson_rev	1
SAT03-16_INDU	thompson	3
SAT12-RAND	bj_thompson_rev	1
TSP-LION2015	bj_thompson	1

F.2 Degroote Approach

Software We re-implemented the Degroote approach using scikit-learn[5] in Python. In particular the linear model is implemented using the LinearRegression estimator from scikit-learn.

Hyperparameter Settings

- Epsilon-Greedy linear regression
  - $\epsilon = 0.05$
- the hyperparameters of underlying models from scikit-learn were set according to their default values

The value of the hyperparameter ϵ is as suggested by the authors in (Degroote et al. 2018).

G Sensitivity Analysis

In this section we provide a sensitivity analysis of the most important hyperparameters of our presented approaches. To keep the amount of experiments to perform in a reasonable dimension, we limit ourselves to the most advanced variant of both LinUCB and Thompson we presented in this work. Moreover, we selected six scenarios from the ASlib covering a range of algorithmic problems, number of instances and features for this analysis. All figures described in the following display the average PAR10 score over ten seeds for different settings of the corresponding hyperparameter. The error bars indicate the corresponding standard deviation.

G.1 Thompson Variants

Figure 2 displays the average PAR10 score over ten seeds for different settings of σ on a selection of scenarios where $\lambda = 0.5$ is fixed. Overall, small values of σ tend to lead to better results indicating that sampling $\widetilde{\boldsymbol{\theta}}_a$, i.e., our belief about the weight vector according to the posterior distribution, should be based on a rather small variance and hence, not too much exploration. This is quite in line with our findings regarding the amount of exploration of the LinUCB variants.

Figure 3 displays the average PAR10 score over 10 seeds for different settings of λ on a selection of scenarios where $\sigma = 10$ is fixed. Overall a clear trend whether small or large values of λ lead to good results seems hard to detect indicating that the performance is rather robust with respect to the choice of λ.

G.2 LinUCB Variants

Figure 4 displays the average PAR10 score over ten seeds for different settings of σ on a selection of scenarios where $\lambda = 0.5$ and $\widetilde{\sigma}^2 = 0.25$ are fixed. In contrast to the Thompson variants previously discussed, where small values of σ tend to lead to better results, here, large values of σ tend to lead to better PAR10 scores indicating that the noise terms (defined in (5)) have a large standard deviation.

Figure 5 displays the average PAR10 score over ten seeds for different settings of α on a selection of scenarios where $\sigma = 1$ and $\widetilde{\sigma}^2 = 0.25$ are fixed. Overall, no clear trend can be observed whether small or large values of α lead to better results.

Figure 6 displays the average PAR10 score over ten seeds for different settings of $\widetilde{\sigma}^2$ on a selection of scenarios where $\alpha = 1$ and $\lambda = 0.5$ are fixed. Once again, overall no clear trend can be observed whether small or large values of $\widetilde{\sigma}$ lead to good results, due to the wide error bars.

H Detailed Performance Values

Table 3 shows the average PAR10 scores (averaged over 10 seeds) and the corresponding standard deviation of all discussed approach variants and all Degroote variants for reference. Once again, the best value for each scenario is printed in bold, whereas the second best is underlined. As elaborated on in the main paper, the Thompson variants achieve the best performance.

[5] https://scikit-learn.org/stable/

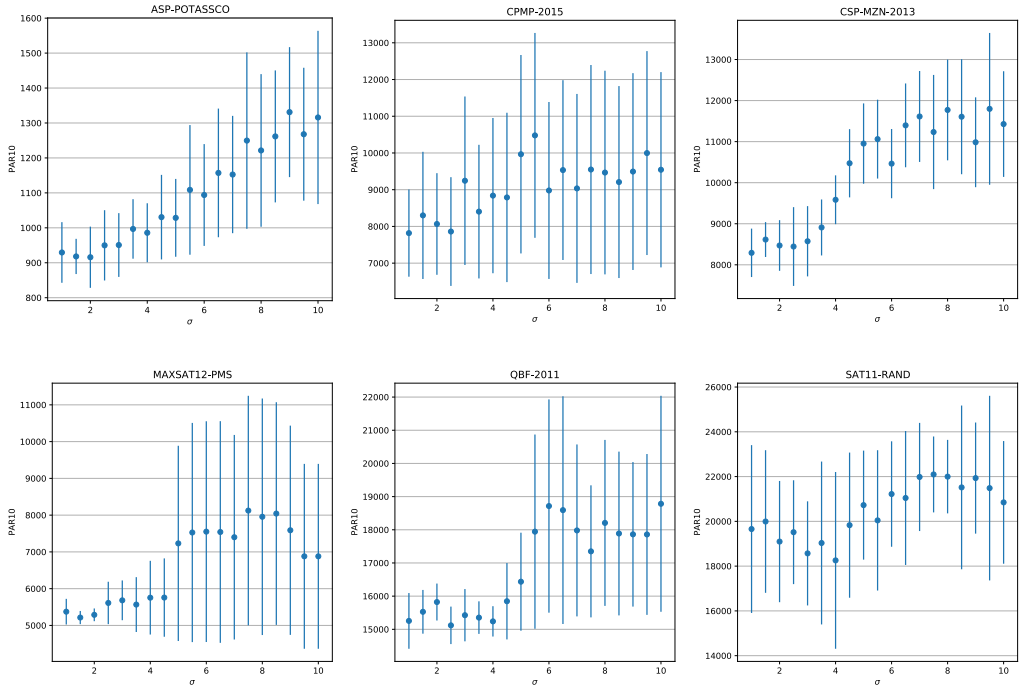


Figure 2: Sensitivity analysis for parameter σ of approach bj_thompson_rev.

In order to represent the performance of our approaches in a more detailed way than it is the case in Table 3, we have plotted in Figure 7 the averaged cumulative PAR10 regret curves (regret wrt. the oracle) of the best Thompson, the best LinUCB and the Degroote approach along with their standard deviation. Here, the cumulative regret up to time T of an approach s is defined as $\sum_{t=1}^{T} l(i_t, s(h_t, i_t)) - \sum_{t=1}^{T} l(i_t, s^*(h_t, i_t))$, where s^* is the oracle and l as in (3) with $\mathcal{P}(C) = 10C$. It is not difficult to see that LinUCB cannot compete with the other approaches in many cases and also features a much larger standard deviation than the others. However, in several cases such as Figures 7o, 7r, or 7z, the differences become much more subtle. Comparing the Thompson variant with the Degroote approach, we see that the former is at least competitive with the latter on almost all scenarios, while being even better on some scenarios (e.g. Fig. 7b, 7k, 7m, 7o). Of course, there are also a few scenarios where the Degroote approach performs better (e.g. Fig. 7n).

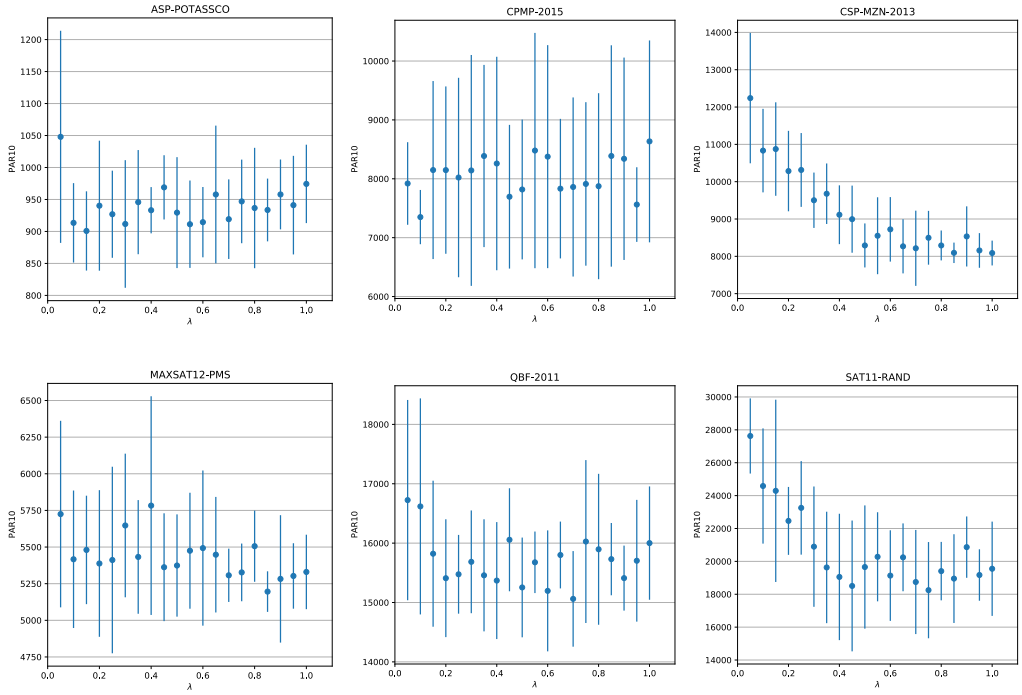


Figure 3: Sensitivity analysis for parameter λ of approach bj_thompson_rev.

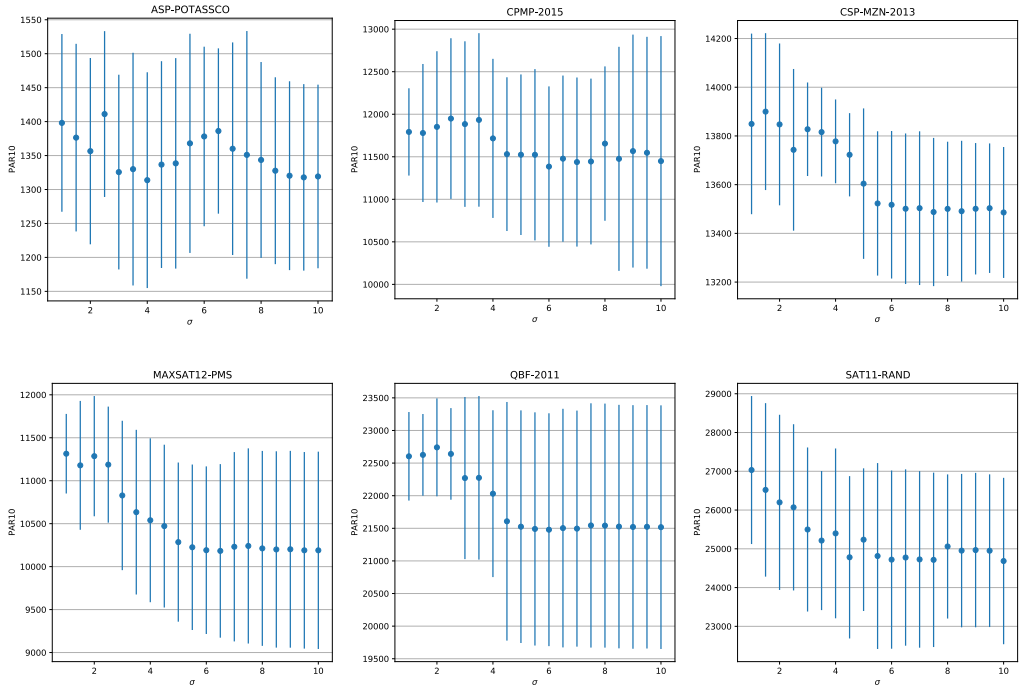


Figure 4: Sensitivity analysis for parameter σ of approach rand_bclinucb_rev.

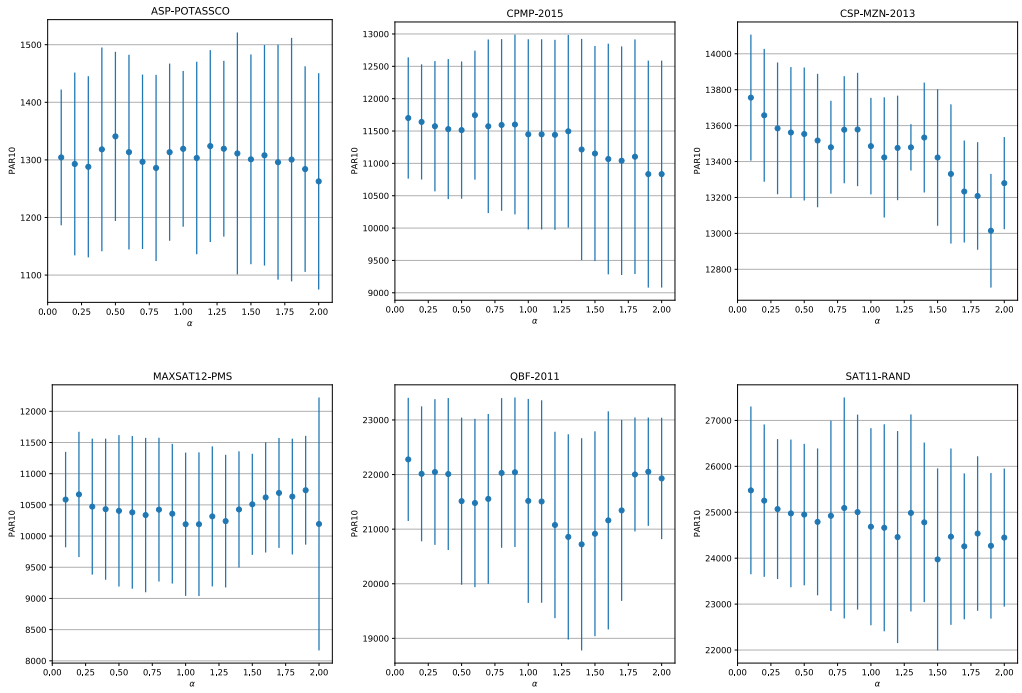


Figure 5: Sensitivity analysis for parameter α of approach rand_bclinucb_rev.

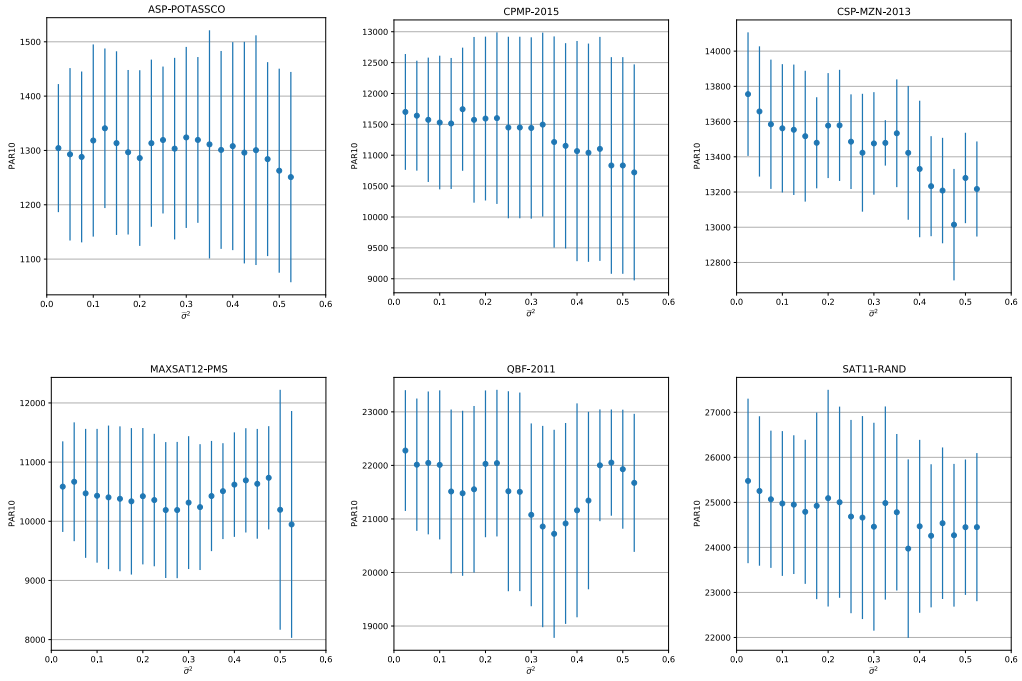


Figure 6: Sensitivity analysis for parameter $\widetilde{\sigma}^2$ of approach rand_bclinucb_rev.

Scenario \ Approach	bj_thompson_rev	bj_thompson	thompson_rev	thompson	bclinucb_rev	bclinucb	blinducb_rev
ASP-POTASSCO	929.45 ± 86.72	949.38 ± 62.38	**902.64** ± 78.43	916.28 ± 72.73	1264.98 ± 303.58	1361.64 ± 239.21	1471.79 ± 43.06
BNSL-2016	9656.86 ± 318.43	9638.04 ± 378.05	9467.01 ± 252.52	**9361.41** ± 262.75	31224.30 ± 12227.55	35665.49 ± 8813.31	30500.86 ± 852.64
CPMP-2015	7818.47 ± 1187.55	8241.01 ± 1164.85	8158.72 ± 1268.83	8499.38 ± 2059.18	10441.68 ± 2449.07	11437.13 ± 2265.92	11667.65 ± 369.54
CSP-2010	8138.63 ± 820.90	8295.76 ± 699.43	7892.67 ± 692.83	8103.96 ± 887.50	8296.32 ± 2170.35	8837.75 ± 2487.60	10348.84 ± 197.67
CSP-MZN-2013	8291.83 ± 589.63	8207.06 ± 532.70	8171.21 ± 594.49	8472.17 ± 760.25	12665.29 ± 647.82	15060.69 ± 1033.48	14646.81 ± 64.27
CSP-Minizinc-Time-2016	**4741.99** ± 505.13	4811.54 ± 409.79	4759.50 ± 306.03	4942.91 ± 326.89	6284.19 ± 1377.37	7515.73 ± 831.41	7138.65 ± 283.28
GRAPHS-2015	4.11e+07 ± 3.71e+06	4.13e+07 ± 4.39e+06	4.21e+07 ± 3.37e+06	3.97e+07 ± 4.76e+06	1.53e+08 ± 1.15e+08	1.53e+08 ± 1.15e+08	1.11e+08 ± 1.64e+07
MAXSAT-PMS-2016	2774.15 ± 218.67	2853.44 ± 210.21	2808.51 ± 218.55	**2763.58** ± 134.14	12391.75 ± 6441.49	16674.43 ± 298.37	15087.51 ± 399.80
MAXSAT-WPMS-2016	6548.69 ± 183.77	6304.15 ± 166.98	6592.87 ± 210.25	6527.34 ± 213.44	7036.99 ± 3325.59	16479.67 ± 482.00	15948.34 ± 203.76
MAXSAT12-PMS	5373.99 ± 348.92	5347.39 ± 291.87	5408.40 ± 482.42	5324.88 ± 208.49	8927.94 ± 3294.10	11787.82 ± 385.81	11918.42 ± 279.90
MAXSAT15-PMS-INDU	3040.87 ± 196.70	3046.05 ± 128.34	**3032.08** ± 90.71	3080.24 ± 130.48	14579.90 ± 814.65	16529.05 ± 274.72	15482.08 ± 299.25
MIP-2016	**7961.45** ± 765.53	8081.57 ± 845.74	8746.73 ± 1159.36	8776.59 ± 823.11	19319.18 ± 12407.96	23090.18 ± 10445.65	20305.81 ± 1787.80
PROTEUS-2014	14223.14 ± 766.02	**13484.34** ± 541.83	14115.69 ± 768.16	13550.56 ± 426.67	20267.46 ± 3744.35	23480.31 ± 416.14	23330.42 ± 206.84
QBF-2011	15253.81 ± 839.93	15708.25 ± 784.81	15178.86 ± 904.72	14902.82 ± 834.91	20051.24 ± 3346.02	23794.53 ± 2734.83	24467.88 ± 266.89
QBF-2014	**3552.30** ± 192.75	3629.40 ± 220.68	3679.96 ± 256.03	3599.79 ± 193.80	4305.61 ± 623.06	5145.37 ± 760.32	5669.88 ± 58.89
QBF-2016	**4770.36** ± 595.50	5082.59 ± 718.71	5045.16 ± 848.59	4937.47 ± 710.69	5765.73 ± 1157.23	7969.34 ± 1830.92	7631.75 ± 323.55
SAT03-16_INDU	12128.48 ± 477.79	**11980.15** ± 193.67	12154.46 ± 221.01	12225.57 ± 501.47	14973.15 ± 1628.24	14800.11 ± 2043.18	13682.55 ± 263.59
SAT11-HAND	30436.08 ± 1196.65	30484.08 ± 1379.35	30085.51 ± 764.32	29547.31 ± 871.64	32028.44 ± 132.51	32086.20 ± 1315.24	30670.46 ± 525.38
SAT11-INDU	17083.58 ± 490.45	17540.58 ± 530.82	17028.84 ± 479.15	17360.28 ± 1064.28	18029.73 ± 1914.17	17944.22 ± 1514.41	17488.94 ± 420.92
SAT11-RAND	19656.39 ± 3747.94	18061.78 ± 2770.70	19061.88 ± 2522.11	**16535.77** ± 2649.93	22349.94 ± 2672.21	23399.06 ± 2254.92	32256.56 ± 333.93
SAT12-ALL	5110.38 ± 221.26	**4720.22** ± 432.14	5132.48 ± 395.74	4812.94 ± 387.46	6433.41 ± 991.59	7657.90 ± 673.88	8548.46 ± 145.06
SAT12-HAND	7707.60 ± 219.43	7443.01 ± 180.51	7509.02 ± 199.39	**7309.74** ± 138.02	8086.51 ± 1043.74	9345.40 ± 669.01	8395.15 ± 289.83
SAT12-INDU	6227.99 ± 714.29	4511.68 ± 76.33	4945.79 ± 228.37	**4428.49** ± 142.91	9741.33 ± 3131.30	11083.35 ± 274.76	10586.81 ± 417.60
SAT12-RAND	5489.81 ± 649.90	**4008.79** ± 206.59	4523.33 ± 170.56	4157.19 ± 277.10	4541.08 ± 1857.92	11024.73 ± 212.35	11118.21 ± 91.24
SAT15-INDU	7885.50 ± 583.23	**7700.27** ± 310.65	7856.08 ± 522.84	7879.61 ± 546.81	9501.14 ± 3536.88	23644.43 ± 2846.82	9691.58 ± 716.04
SAT18-EXP	24942.85 ± 847.88	25201.41 ± 681.42	**24906.56** ± 540.36	25015.28 ± 1031.38	25310.89 ± 1669.61	33234.71 ± 5109.02	26762.26 ± 760.54
TSP-LION2015	**1160.50** ± 396.73	1226.11 ± 309.42	1411.06 ± 329.16	1213.43 ± 370.00	8324.41 ± 4816.93	8436.53 ± 4575.03	4810.21 ± 212.49
avgrank	3.074074	3.074074	2.925926	**2.555556**	8.0	11.518519	10.740741

Scenario \ Approach	blinducb	rand_bclinucb_rev	rand_bclinucb	rand_blinducb_rev	rand_blinducb	degroote_EpsilonGreedy_LinearRegression
ASP-POTASSCO	1204.20 ± 161.11	1319.22 ± 135.21	1232.59 ± 64.63	1280.76 ± 24.99	1198.79 ± 135.63	1047.13 ± 46.50
BNSL-2016	17369.40 ± 1553.72	32178.72 ± 6600.21	23650.56 ± 2177.42	23563.48 ± 907.10	18531.72 ± 2259.16	12510.26 ± 1291.03
CPMP-2015	11518.55 ± 1316.62	11449.01 ± 1469.61	10664.02 ± 1014.65	10439.34 ± 308.03	11350.81 ± 1338.76	**6991.97** ± 501.36
CSP-2010	9563.85 ± 1821.74	9533.88 ± 1360.51	9378.99 ± 270.41	9847.48 ± 316.70	8796.21 ± 1534.82	**7593.13** ± 208.94
CSP-MZN-2013	12113.14 ± 593.67	13485.81 ± 268.79	12516.57 ± 90.27	13551.04 ± 95.32	11595.72 ± 914.74	**8034.62** ± 113.78
CSP-Minizinc-Time-2016	5944.38 ± 481.06	7043.06 ± 440.10	6104.17 ± 404.23	6291.28 ± 431.44	5353.62 ± 364.20	5258.70 ± 406.91
GRAPHS-2015	6.91e+07 ± 4.15e+07	1.74e+08 ± 8.54e+07	1.23e+08 ± 4.48e+06	7.85e+07 ± 1.71e+07	6.72e+07 ± 4.00e+07	**3.45e+07** ± 1.42e+06
MAXSAT-PMS-2016	5785.25 ± 1369.27	15107.67 ± 238.80	7320.34 ± 877.35	11887.90 ± 403.61	5405.64 ± 495.98	3279.54 ± 133.00
MAXSAT-WPMS-2016	12940.70 ± 727.41	15326.14 ± 285.53	11268.89 ± 443.14	13492.59 ± 174.10	12518.20 ± 585.24	**6287.21** ± 541.69
MAXSAT12-PMS	9829.22 ± 2477.59	10189.82 ± 1148.46	7761.18 ± 632.24	9661.09 ± 254.48	9580.82 ± 2684.49	**5308.11** ± 129.30
MAXSAT15-PMS-INDU	12020.48 ± 1801.39	15342.87 ± 313.35	9842.87 ± 419.03	12936.66 ± 495.89	10416.87 ± 930.70	3867.70 ± 255.98
MIP-2016	10280.18 ± 646.54	21472.48 ± 10159.74	21224.37 ± 3076.88	16238.36 ± 1625.48	10102.14 ± 783.52	10644.68 ± 3405.18
PROTEUS-2014	23496.49 ± 957.96	22713.85 ± 325.06	21945.81 ± 518.38	22380.67 ± 173.49	23508.44 ± 1081.76	15622.29 ± 784.60
QBF-2011	24105.19 ± 2408.37	21517.52 ± 1867.77	19664.68 ± 805.56	21819.49 ± 613.82	24190.56 ± 1452.98	**13912.24** ± 356.69
QBF-2014	4951.40 ± 630.45	5414.49 ± 240.11	4976.71 ± 128.30	5213.18 ± 81.48	4836.72 ± 617.95	4116.15 ± 116.27
QBF-2016	6080.61 ± 828.17	8530.38 ± 426.05	6019.49 ± 291.23	6811.82 ± 209.17	6234.42 ± 943.66	5346.29 ± 210.05
SAT03-16_INDU	12836.59 ± 385.99	14204.97 ± 1429.35	13505.17 ± 405.65	13421.35 ± 251.79	12671.99 ± 458.24	12754.50 ± 200.55
SAT11-HAND	33755.48 ± 470.48	30710.59 ± 932.73	31157.03 ± 843.87	30519.78 ± 633.95	33151.03 ± 689.91	**29544.70** ± 952.78
SAT11-INDU	17407.75 ± 887.53	17785.96 ± 942.15	17900.53 ± 797.85	17404.10 ± 671.54	17792.97 ± 645.24	**17018.24** ± 647.90
SAT11-RAND	32812.38 ± 484.62	24684.78 ± 2146.11	22754.43 ± 401.94	28780.03 ± 577.18	32382.10 ± 453.97	21008.77 ± 530.22
SAT12-ALL	8235.03 ± 393.14	7388.34 ± 313.26	7096.01 ± 138.67	7417.48 ± 76.56	8231.51 ± 319.87	5650.32 ± 214.36
SAT12-HAND	9944.53 ± 702.20	8626.88 ± 250.21	8214.43 ± 185.22	8164.12 ± 126.91	9560.68 ± 629.12	7634.24 ± 267.89
SAT12-INDU	11709.07 ± 47.37	10733.23 ± 400.92	8703.58 ± 987.22	9159.07 ± 827.64	11568.96 ± 45.53	4755.52 ± 206.95
SAT12-RAND	8270.41 ± 249.25	10699.76 ± 153.19	9015.20 ± 189.65	10432.59 ± 97.18	8109.35 ± 140.40	5023.73 ± 174.68
SAT15-INDU	11203.29 ± 3070.49	12188.32 ± 3435.99	10349.12 ± 897.36	9599.63 ± 676.35	10656.72 ± 3172.70	8220.22 ± 525.13
SAT18-EXP	29573.37 ± 2210.54	27153.72 ± 924.36	26709.51 ± 679.87	26361.79 ± 1018.39	30237.60 ± 1497.64	25272.35 ± 881.19
TSP-LION2015	9991.79 ± 294.01	9168.26 ± 2612.01	6255.37 ± 194.34	4588.53 ± 191.98	9422.52 ± 655.46	1634.79 ± 112.29
avgrank	9.481481	10.592593	8.074074	8.518519	8.777778	3.666667

Table 3: Average PAR10 scores (averaged over 10 seeds) and the corresponding standard deviation of all discussed approach variants and the Degroote approach.

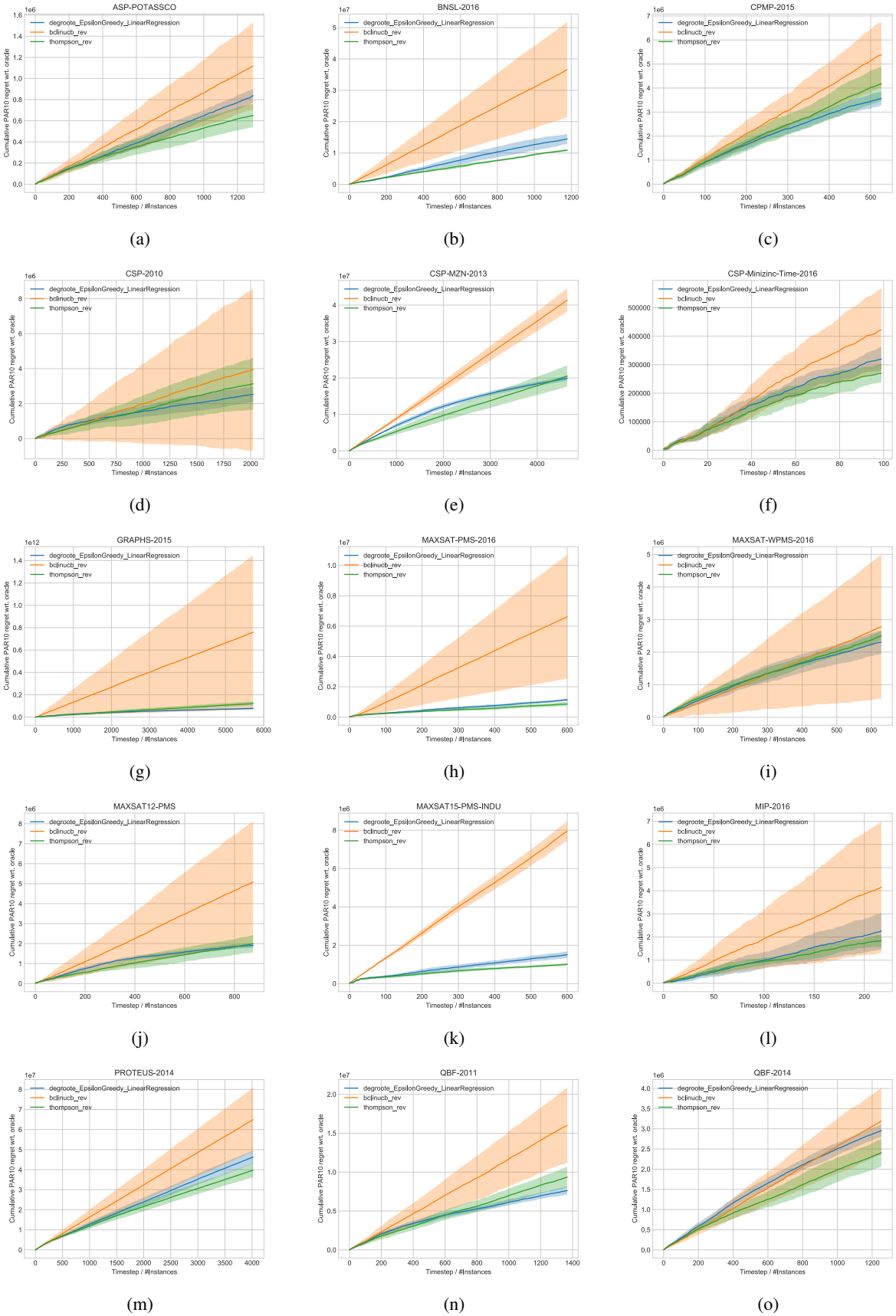
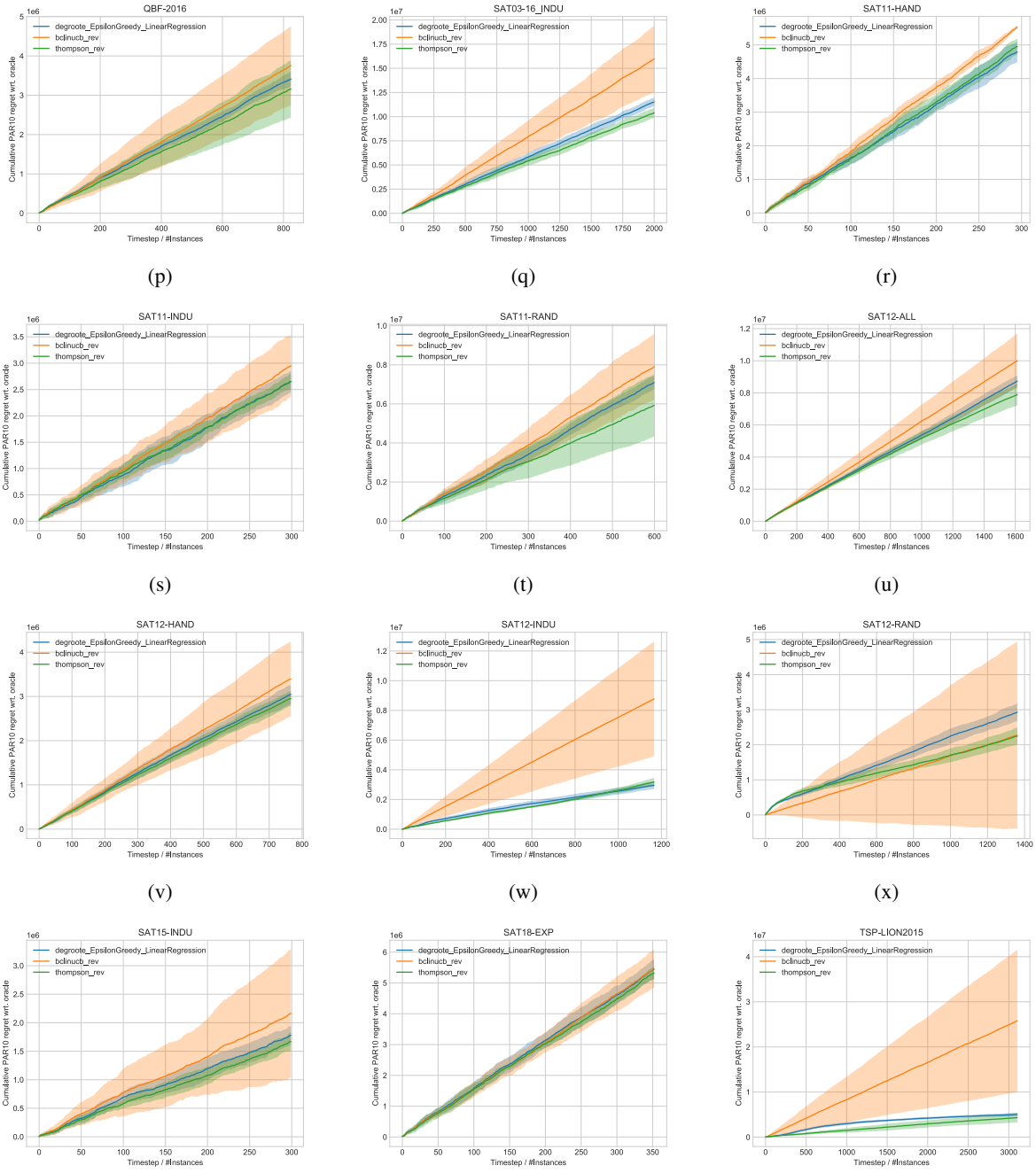


Figure 7: Cumulative PAR10 regret wrt. oracle.



(p) (q) (r)

(s) (t) (u)

(v) (w) (x)

(y) (z) (aa)

Figure 7: (Cont.) Cumulative PAR10 regret wrt. oracle.