A List 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	
$\mathbb{1}\{\cdot\}$	indicator function
$\mathbb{N}$	set of natural numbers (without 0), i.e., $\mathbb{N} = \{1, 2, 3, \dots\}$
$\mathbb{R}$	set of real numbers
D	observation domain (categorical or numerical)
$\mathcal{A} = [n]$	set of arms
n	number of arms
k	maximal possible subset size
B	budget for the learner
$\mathcal{Q}_{\leq k}$	all subsets of $\mathcal{A}$ of size $\leq k$: $\{Q \subseteq \mathcal{A} \mid 2 \leq \lvert Q\rvert \leq k\}$
$\mathcal{Q}_{\leq k}(i)$	all subsets in $\mathcal{Q}_{\leq k}$ which contain arm i: $\{Q \in \mathcal{Q}_{\leq k} \mid i \in Q\}$
$\mathcal{Q}_{=k}$	all subsets of $\mathcal{A}$ of size k : $\{Q \subseteq \mathcal{A} \mid \lvert Q\rvert = k\}$
$\mathcal{Q}_{=k}(i)$	all subsets of $\mathcal{A}$ of size k which contain arm i : $\{Q \in \mathcal{Q}_{=k} \mid i \in Q\}$
$\mathbf{o}_Q(t)$	observed feedback vector by querying Q for the t-th time
Modelling related	
s	relevant statistic for the decision making process
$s_{i\lvert Q}(t)$	statistics for arm $i \in Q$ derived by the observed feedback at the t-th usage of query set Q
$\mathbf{s}_Q(t)$	vector of statistics for all arms in the query set Q after its t-th usage: $(s_{i\lvert Q})_{i\in Q}(t)$
$S_{i\lvert Q}$	limit of the statistics for arm i in query set Q: $\lim_{t\to\infty} s_{i\lvert Q}(t)$
i^*	best arm or generalized Condorcet winner: $\forall Q \in \mathcal{Q}_{\leq k}$ with $i^* \in Q$ it holds that $S_{i^*\lvert Q} > S_{j\lvert Q}$ for any $j \in Q\backslash\{i^*\}$
$s_i^{\mathcal{B}}(t)$	Borda score of arm i at time t: $\sum_{Q\in\mathcal{Q}_{=k}(i)} s_{i\lvert Q}(t) / \lvert\mathcal{Q}_{=k}(i)\rvert$
$S_i^{\mathcal{B}}$	limit Borda score of arm i: $\lim_{t\to\infty} s_i^{\mathcal{B}}(t)$
$i^*_{\mathcal{B}}$	generalized Borda winner: $i^*_{\mathcal{B}} \in \arg\max_{i\in\mathcal{A}} S_i^{\mathcal{B}}$
$n_Q(t)$	number of times query set Q was used until time t
$\gamma_{i\lvert Q}(t)$	point-wise smallest non-increasing function bounding the difference $\lvert s_{i\lvert Q}(t) - S_{i\lvert Q}\rvert$ (rate of convergence)
$\bar{\gamma}_Q(t)$	maximal $\gamma_{i\lvert Q}(t)$ over all $i \in Q$
$\bar{\gamma}(t)$	maximal $\gamma_Q(t)$ over all $Q \in \mathcal{Q}_{\leq k}$
$\gamma_{i\lvert Q}^{-1}(\alpha)$	quasi-inverse of $\gamma_{i\lvert Q}$: $\min\{t \in \mathbb{N} \mid \gamma_{i\lvert Q}(t) \leq \alpha\}$
$\bar{\gamma}_Q^{-1}(t)$	minimal $\gamma_{i\lvert Q}(t)$ over all $i \in Q$
$\bar{\gamma}^{-1}(t)$	minimal $\gamma_Q(t)$ over all $Q \in \mathcal{Q}_{\leq k}$
$\hat{\gamma}_i(t)$	rate of convergence of the Borda score for arm i: $\frac{1}{\lvert\mathcal{Q}_{=k}(i)\rvert} \sum_{Q\in\mathcal{Q}_{=k}(i)} \gamma_{i\lvert Q}(t)$
$\hat{\gamma}_{i,j}^{\max}(t)$	$\max\{\hat{\gamma}_i(t), \hat{\gamma}_j(t)\}$.
$\Delta_{i\lvert Q}$	gap of the limit statistic of arm $i \in Q$ to the limit statistic of the generalized Condorcet winner: $\lvert S_{i^*\lvert Q} - S_{i\lvert Q}\rvert$ for any $Q \in \mathcal{Q}_{\leq k}(i) \cap \mathcal{Q}_{\leq k}(i^*)$
$S_{(l)\lvert Q}, \Delta_{(l)\lvert Q}$	l-th order statistic of $\{S_{i\lvert Q}\}_{i\in Q}$ for $l \in \{1, 2, \dots, \lvert Q\rvert\}$ and its gap $\Delta_{(l)\lvert Q} = S_{i^*\lvert Q} - S_{(l)\lvert Q}$
Algorithm related	
f	function from $[k]$ to $[k]$ specifying the nature of the arm elimination strategy
$R, R^{\mathbb{A}}$	number of rounds of the learning algorithm ($\mathbb{A}$)
$P_r, P_r^{\mathbb{A}}$	number of partitions of the learning algorithm ($\mathbb{A}$) in round r
$\mathbb{A}_{r,j}$	j-th partition in round r
$\mathbb{A}_r(i^*)$	the partition in round r containing i^* (emptyset otherwise)
b_r	budget used in round r for a partition
$z_{\mathbb{A}}$	sufficient budget for learning algorithm $\mathbb{A}$ to return i^* (or $i^*_{\mathcal{B}}$ if $\mathbb{A}$ is ROUNDROBIN)
ROUNDROBIN	the naïve algorithm introduced in Section C
CSE	the generic *combinatorial successive elimination* algorithm (Algorithm 1)
CSWS	the *combinatorial successive winner stays* algorithm resulting by using $f(x) = 1$ in CSE
CSR	the *combinatorial successive rejects* algorithm resulting by using $f(x) = x - 1$ in CSE
CSH	the *combinatorial successive halving* algorithm resulting by using $f(x) = \lceil x/2\rceil$ in CSE
SH	the *successive halving* algorithm for pure exploration settings in standard multi-armed bandits (cf. [25])
GBW	Generalized Borda winner
GCW	Generalized Condorcet winner

B Proofs for Section 3

In this section, we prove the general lower bounds on the necessary budget for identifying the generalized Condorcet winner (GCW), the generalized Borda winner (GBW) or the generalized Copeland winner (GCopeW). For this purpose, let us first fix some further notation. If Alg is a possibly probabilistic algorithm and $\mathbf{s}$ is fixed, we write $\mathrm{Alg}(\mathbf{s})$ for the output of Alg executed on the instance $\mathbf{s}$. We restrict ourselves only to algorithms whose output is solely determined by the sequence of observations it has received as well as the corresponding statistics. Moreover, for $Q \in \mathcal{Q}_{\leq k}$, we write $B_Q(\mathrm{Alg}, \mathbf{s}) \in \mathbb{N} \cup \{\infty\}$ for the number of times Alg queries Q when started on instance $\mathbf{s}$. Note that $\mathrm{Alg}(\mathbf{s})$ as well as $B_Q(\mathrm{Alg}, \mathbf{s})$ and $B(\mathrm{Alg}, \mathbf{s}) = \sum_{Q \in \mathcal{Q}_{\leq k}} B_Q(\mathrm{Alg}, \mathbf{s})$ are random variables, because they depend on the innate randomness of Alg.
Given $\mathbf{s}$, let us write $\mathrm{GCW}(\mathbf{s})$, $\mathrm{GBW}(\mathbf{s})$ and $\mathrm{GCopeW}(\mathbf{s})$ for the set of all GCWs, GBWs and GCopeWs of $\mathbf{s}$, respectively. In case $|\mathrm{GCW}(\mathbf{s})| = 1$, $|\mathrm{GBW}(\mathbf{s})| = 1$ resp. $|\mathrm{GCopeW}(\mathbf{s})| = 1$, with a slight abuse of notation, we may denote by $\mathrm{GCW}(\mathbf{s})$, $\mathrm{GBW}(\mathbf{s})$ resp. $\mathrm{GCopeW}(\mathbf{s})$ simply the only GCW, GBW resp. GCopeW of $\mathbf{s}$. Recall that the GCW, the GBWs and the GCopeWs of $\mathbf{s}$ only depend on the limits $\mathbf{S} = (S_{i|Q})_{Q \in \mathcal{Q}_{\leq k}, i \in Q}$ with $S_{i|Q} = \lim_{t \to \infty} s_{i|Q}(t)$.

Definition B.1. *Let* Alg *be a (possibly probabilistic) sequential algorithm.*

(i) Alg *solves* $\mathcal{P}_{\mathrm{GCW}}(\mathbf{S}, \boldsymbol{\gamma})$ *if* $\mathbb{P}(\mathrm{Alg}(\mathbf{s}) \in \mathrm{GCW}(\mathbf{s})) = 1$ *for any* $\mathbf{s}$ *in* $\mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$.

(ii) Alg *solves* $\mathcal{P}_{\mathrm{GBW}}(\mathbf{S}, \boldsymbol{\gamma})$ *if* $\mathbb{P}(\mathrm{Alg}(\mathbf{s}) \in \mathrm{GBW}(\mathbf{s})) = 1$ *for any* $\mathbf{s}$ *in* $\mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$.

(iii) Alg *solves* $\mathcal{P}_{\mathrm{GCopeW}}(\mathbf{S}, \boldsymbol{\gamma})$ *if* $\mathbb{P}(\mathrm{Alg}(\mathbf{s}) \in \mathrm{GCopeW}(\mathbf{s})) = 1$ *for any* $\mathbf{s}$ *in* $\mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$.

B.1 Proof of Theorem 3.1 (i): Lower Bound for GCW Identification

The proof of (i) in Theorem 3.1 is prepared with the next lemma.

Lemma B.2. *Let* Alg *be a deterministic solution to* $\mathcal{P}_{\mathrm{GCW}}(\mathbf{S}, \boldsymbol{\gamma})$ *and* $\mathbf{s}, \mathbf{s}' \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$.

(i) If $\mathrm{Alg}(\mathbf{s}) \neq \mathrm{Alg}(\mathbf{s}')$*, then*

$$\exists Q \in \mathcal{Q}_{\leq k}, i \in Q,\, t \in \{1, \ldots, \min\{B_Q(\mathrm{Alg}, \mathbf{s}), B_Q(\mathrm{Alg}, \mathbf{s}')\}\} \,:\, s_{i|Q}(t) \neq s'_{i|Q}(t).$$

(ii) If $\mathbf{s}$ *and* $\mathbf{s}'$ *coincide on* $\{t < B'\}$ *and on* $\widetilde{\mathcal{Q}} \subseteq \mathcal{Q}_{\leq k}$ *in the sense that*

$$\forall Q \in \mathcal{Q}_{\leq k}, \forall i \in Q, \forall t < B' \,:\, s_{i|Q}(t) = s'_{i|Q}(t) \tag{1}$$

and

$$\forall Q \in \widetilde{\mathcal{Q}}, \forall i \in Q, \forall t \in \mathbb{N} \,:\, s_{i|Q}(t) = s'_{i|Q}(t), \tag{2}$$

then $\mathrm{Alg}(\mathbf{s}) \neq \mathrm{Alg}(\mathbf{s}')$ *implies*

$$\exists Q \in \mathcal{Q}_{\leq k} \setminus \widetilde{\mathcal{Q}} \,:\, \min\{B_Q(\mathrm{Alg}, \mathbf{s}), B_Q(\mathrm{Alg}, \mathbf{s}')\} \geq B'.$$

Proof. (i) To prove the contraposition, suppose that

$$\forall Q \in \mathcal{Q}_{\leq k}, i \in Q,\, t \in \{1, \ldots, \min\{B_Q(\mathrm{Alg}, \mathbf{s}), B_Q(\mathrm{Alg}, \mathbf{s}')\}\} \,:\, s_{i|Q}(t) = s'_{i|Q}(t) \tag{3}$$

holds.
Claim 1: $B_Q(\mathrm{Alg}, \mathbf{s}) = B_Q(\mathrm{Alg}, \mathbf{s}')$ for any $\mathcal{Q} \in \mathcal{Q}_{\leq k}$.
Proof: Assume this was not the case. Let $Q \in \mathcal{Q}_{\leq k}$ be the first set, for which Alg exceeds its budget on one of $\mathbf{s}, \mathbf{s}'$ but does not reach it on the other instance, and suppose w.l.o.g. $B_Q(\mathrm{Alg}, \mathbf{s}) > B_Q(\mathrm{Alg}, \mathbf{s}')$. Since Alg has observed until this point exactly the same feedback on $\mathbf{s}$ as on $\mathbf{s}'$, this is a contradiction as Alg is deterministic. ■
Combining Claim 1 and (3) yields that Alg observes on $\mathbf{s}$ exactly the same feedback as on $\mathbf{s}'$ until its termination. Since Alg is deterministic, this implies $\mathrm{Alg}(\mathbf{s}) = \mathrm{Alg}(\mathbf{s}')$.

(ii) If $\mathrm{Alg}(\mathbf{s}) \neq \mathrm{Alg}(\mathbf{s}')$, then (i) together with (2) yields

$$\exists Q \in \mathcal{Q}_{\leq k} \setminus \widetilde{\mathcal{Q}}, i \in Q, t \leq \min\{B_Q(\mathrm{Alg}, \mathbf{s}), B_Q(\mathrm{Alg}, \mathbf{s}')\} \,:\, s_{i|Q}(t) \neq s'_{i|Q}(t),$$

and thus (1) implies

$$\exists Q \in \mathcal{Q}_{\leq k} \setminus \widetilde{\mathcal{Q}} \,:\, \min\{B_Q(\mathrm{Alg}, \mathbf{s}), B_Q(\mathrm{Alg}, \mathbf{s}')\} \geq B'.$$

□

Lemma B.2 is the main ingredient for the proof of Theorem 3.1, as we first analyze the lower bound for deterministic algorithms and then apply Yao's minimax principle [51] to infer the lower bound for any randomized algorithm.

Proof of Theorem 3.1 (i). We split the proof into two parts.

Part 1: The statement holds in case Alg **is a deterministic algorithm.**
Abbreviate $B' := \min_{Q\in\mathcal{Q}_{\leq k}} \min_{j\in Q} \gamma_{j|Q}^{-1}\left(\frac{S_{(1)|Q} - S_{(|Q|)|Q}}{2}\right)$. Fix a family $\{\pi_Q\}_{Q\in\mathcal{Q}_{\leq k}}$ of permutations $\pi_Q : Q \mapsto Q$ such that $S_{\pi_Q(1)|Q} = S_{(1)|Q}$ holds for any $Q \in \mathcal{Q}_{\leq k}(1)$, and define $\mathbf{s} = (s_{i|Q}(t))_{Q\in\mathcal{Q}_{\leq k}, i\in Q, t\in\mathbb{N}}$ via

$$s_{i|Q}(t) := \begin{cases} \frac{S_{(1)|Q}+S_{(|Q|)|Q}}{2}, & \text{if } t < B', \\ S_{\pi_Q(i)|Q}, & \text{if } t \geq B'. \end{cases}$$

Regarding our assumption on $\mathbf{S}$, $\mathrm{GCW}(\mathbf{s}) = 1$ holds by construction. For $t < B' \leq \gamma_{i|Q}^{-1}\left(\frac{S_{(1)|Q} - S_{(|Q|)|Q}}{2}\right)$, which implies $\gamma_{i|Q}(t) \geq \frac{S_{(1)|Q} - S_{(|Q|)|Q}}{2}$, we have due to $S_{(1)|Q} \geq S_{i|Q} \geq S_{(|Q|)|Q}$ the inequality

$$\begin{aligned}
&\left|s_{i|Q}(t) - \lim\nolimits_{t\to\infty} s_{i|Q}(t)\right| = \left|\frac{S_{(1)|Q} + S_{(|Q|)|Q}}{2} - S_{i|Q}\right| \\
&\leq \max\left\{S_{(1)|Q} - \frac{S_{(1)|Q} + S_{(|Q|)|Q}}{2}, \frac{S_{(1)|Q} + S_{(|Q|)|Q}}{2} - S_{(|Q|)|Q}\right\} \\
&= \frac{S_{(1)|Q} - S_{(|Q|)|Q}}{2} \\
&\leq \gamma_{i|Q}(t)
\end{aligned}$$

for any $i \in Q$. This shows $\mathbf{s} \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$.

For any $l \in \{2, \ldots, n\}$ define an instance $\mathbf{s}^l = (s^l_{i|Q}(t))_{Q\in\mathcal{Q}_{\leq k}, i\in Q, t\in\mathbb{N}}$ such that $s^l_{\cdot|Q}(\cdot) = s_{\cdot|Q}(\cdot)$ for any $Q \in \mathcal{Q}_{\leq k}$ with $l \notin Q$ and

$$s^l_{i|Q}(t) := \begin{cases} \frac{S_{(1)|Q}+S_{(|Q|)|Q}}{2}, & \text{if } t < B', \\ S_{(1)|Q}, & \text{if } t \geq B' \text{ and } i = l, \\ S_{l|Q}, & \text{if } t \geq B' \text{ and } i = \mathrm{argmax}_{j\in Q} S_{j|Q} \\ s_{i|Q}(t), & \text{else}, \end{cases}$$

for all $Q \in \mathcal{Q}_{\leq k}(l)$, $i \in Q$ and $t \in \mathbb{N}$. According to its definition, we have $\mathrm{GCW}(\mathbf{s}^l) = l$, and similarly as above one may check $\mathbf{s}^l \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$.

Since Alg solves $\mathcal{P}_{\mathrm{GCW}}(\mathbf{S}, \boldsymbol{\gamma})$, it satisfies $\mathrm{Alg}(\mathbf{s}) = 1 \neq 2 = \mathrm{Alg}(\mathbf{s}^2)$. Regarding that $\mathbf{s}$ and $\mathbf{s}^2$ coincide on $\{t < B'\}$ and on $\{Q \in \mathcal{Q}_{\leq k} \,|\, 1 \notin Q \text{ or } 2 \notin Q\}$ in the sense of (1) and (2), Lemma B.2 (ii) assures the existence of some $Q_1 \in \mathcal{Q}_{\leq k}$ with $1 \in Q_1$ and $i_1 := 2 \in Q_1$ such that $B_{Q_1}(\mathrm{Alg}, \mathbf{s}) \geq \min\{B_{Q_1}(\mathrm{Alg}, \mathbf{s}), B_{Q_1}(\mathrm{Alg}, \mathbf{s}^{i_1})\} \geq B'$. Let $F_1 := [n] \setminus Q_1$ and fix an arbitrary $i_2 \in F_1$. Then, $\mathrm{Alg}(\mathbf{s}) = 1 \neq i_2 = \mathrm{Alg}(\mathbf{s}^{i_2})$ and since $\mathbf{s}$ and $\mathbf{s}^{i_2}$ coincide on $\{t < B'\}$ and $\{Q \in \mathcal{Q}_{\leq k} \,|\, i_2 \notin Q\}$, Lemma B.2 (ii) yields the existence of some $Q_2 \in \mathcal{Q}_{\leq k}$ with $i_2 \in Q_2$ such that $B_{Q_2}(\mathrm{Alg}, \mathbf{s}) \geq \min\{B_{Q_2}(\mathrm{Alg}, \mathbf{s}), B_{Q_2}(\mathrm{Alg}, \mathbf{s}^{i_2})\} \geq B'$. From $i_2 \in F_1 = [n] \setminus Q_1$ and $i_2 \in Q_2$ we infer $Q_1 \neq Q_2$. With this, we define $F_2 := F_1 \setminus Q_2 = [n] \setminus (Q_1 \cup Q_2)$.

Inductively, whenever $F_l \neq \emptyset$, we may select an element $i_{l+1} \in F_l$ and infer from Lemma B.2 (ii), due to $\mathrm{Alg}(\mathbf{s}) = 1 \neq i_{l+1} = \mathrm{Alg}(\mathbf{s}^{i_{l+1}})$ and the similarity of $\mathbf{s}$ and $\mathbf{s}^{i_{l+1}}$ on $\{t < B'\}$ and $\{Q \in \mathcal{Q}_{\leq k} | i_{l+1} \notin Q\}$, the existence of a set $Q_{l+1} \in \mathcal{Q}_{\leq k}$ with $i_{l+1} \in Q_{l+1}$ such that $B_{Q_{l+1}}(\mathrm{Alg}, \mathbf{s}) \geq B'$, and define $F_{l+1} := F_l \setminus Q_{l+1}$. Then, $i_{l+1} \in F_l = [n] \setminus (Q_1 \cup \cdots \cup Q_l)$

and $i_{l+1} \in Q_{l+1}$ assure $Q_{l+1} \notin \{Q_1, \dots, Q_l\}$. This procedure terminates at the smallest l' such that $F_{l'} = \emptyset$, and $Q_1, \dots, Q_{l'}$ are distinct. Regarding that $|F_{l+1}| - |F_l| \le |Q_l| \le k$ for all $l \in \{1, \dots, l'-1\}$, we have $l' \ge \lceil \frac{n}{k} \rceil$. Consequently,

$$B(\mathrm{Alg}, \mathbf{s}) \ge \sum_{l=1}^{l'} B_{Q_l}(\mathrm{Alg}, \mathbf{s}) \ge \left\lceil \frac{n}{k} \right\rceil B'$$

holds, which shows the claim for deterministic algorithms with regard to the definition of B'.

Part 2: The statement holds for arbitrary Alg**.**
Let $\mathfrak{A}$ be the set of all deterministic algorithms[3] and $\mathbf{s}$ be the instance from the first part. Write $\delta_{\mathbf{s}}$ for the probability distribution on $\{\mathbf{s}\}$, which assigns $\mathbf{s}$ probability one, i.e., the Dirac measure on $\mathbf{s}$. Note that for any randomized algorithm Alg there exists a probability distribution P on $\mathfrak{A}$ such that $\mathrm{Alg} \sim P$. By applying Yao's minimax principle [51] and using part one we conclude

$$\begin{aligned} \mathbb{E}[B(\mathrm{Alg}, \mathbf{s})] = \mathbb{E}_{\mathrm{Alg}' \sim P}[B(\mathrm{Alg}', \mathbf{s})] &\ge \inf\nolimits_{\mathrm{Alg} \in \mathfrak{A}} \mathbb{E}_{\mathbf{s}' \sim \delta_{\mathbf{s}}}[B(\mathrm{Alg}, \mathbf{s}')] \\ &= \inf\nolimits_{\mathrm{Alg} \in \mathfrak{A}} B(\mathrm{Alg}, \mathbf{s}) \ge \left\lceil \frac{n}{k} \right\rceil B', \end{aligned}$$

where B' is as in part one. □

Remark B.3. *(i) The above proof reveals even a stronger version of Theorem 3.1 (i). Indeed, in the proof we explicitly construct n distinct instances $\mathbf{s}^1 := \mathbf{s}, \dots, \mathbf{s}^n \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$ with $\mathrm{GCW}(\mathbf{s}^l) = l$ for all $l \in [n]$, and in fact show: Any (possibly random) algorithm* Alg*, which is able to correctly identify the best arm for any $\mathbf{s}' \in \{\mathbf{s}_1, \dots, \mathbf{s}_n\}$ (i.e.,* Alg *does not necessarily have to solve $\mathcal{P}_{\mathrm{GCW}}(\mathbf{S}, \boldsymbol{\gamma})$) fulfills*

$$\mathbb{E}\left[B(\mathrm{Alg}, \mathbf{s})\right] \ge \left\lceil \frac{n}{k} \right\rceil \min_{Q \in \mathcal{Q}_{\le k}} \min_{j \in Q} \gamma_{j|Q}^{-1} \left(\frac{S_{(1)|Q} - S_{(|Q|)|Q}}{2} \right).$$

(ii) Condition (iii) in the definition of $\mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$ assures that the term $S_{(1)|Q}$ resp. $S_{(|Q|)|Q}$ in our lower bound from Theorem 3.1 coincides with $S'_{(1)|Q}$ resp. $S'_{(|Q|)|Q}$, when $S'_{i|Q} := \lim_{t \to \infty} s_{i|Q}(t)$ for $\mathbf{s} \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$.

B.2 Proof of Theorem 3.1 (ii): Lower Bound for GBW Identification

Recall that $\mathrm{GBW}(\mathbf{s})$ is the set of elements $i \in [n]$, for which the limits $S_{i|Q} = \lim_{t \to \infty} s_{i|Q}(t)$ have the highest Borda score

$$S_i^{\mathcal{B}} = \frac{\sum_{Q \in \mathcal{Q}_{=k}(i)} S_{i|Q}}{|\mathcal{Q}_{=k}(i)|} = \frac{\sum_{Q \in \mathcal{Q}_{=k}(i)} S_{i|Q}}{\binom{n-1}{k-1}}.$$

We call $\mathbf{S} = (S_{i|Q})_{Q \in \mathcal{Q}_{\le k}, i \in Q}$ **homogeneous** if $(S_{(1)|Q}, \dots, S_{(|Q|)|Q})$ does not depend on Q. Thus, if $\mathbf{S}$ is homogeneous, we may simply write $S_{(l)}$ for $S_{(l)|Q}$ for any $Q \in \mathcal{Q}_{=k}$.

The next two lemmata serves as a preparation for the proof of (ii) and (iii) in Theorem 3.1.

Lemma B.4. *For any $\mathcal{W} \subseteq \mathcal{Q}_{=k}$ we have $\sum_{j=1}^n |\mathcal{Q}_{=k}(j) \cap \mathcal{W}| = k|\mathcal{W}|$.*

Proof of Lemma B.4. Let $\mathcal{W} \subseteq \mathcal{Q}_{=k}$ be fixed. For any $Q = \{i_1, \dots, i_k\} \in \mathcal{Q}_{=k} \cap \mathcal{W}$ we have that $Q \in \mathcal{Q}_{=k}(i_l) \cap \mathcal{W}$ for any $l \in [k]$, whereas $Q \notin \mathcal{Q}_{=k}(j) \cap \mathcal{W}$ for any $j \in [n] \setminus \{i_1, \dots, i_k\}$. Hence,

$$\sum\nolimits_{j=1}^n |\mathcal{Q}_{=k}(j) \cap \mathcal{W}| = k \left| \bigcup\nolimits_{j=1}^n (\mathcal{Q}_{=k}(j) \cap \mathcal{W}) \right| = k \left| \left(\bigcup\nolimits_{j=1}^n \mathcal{Q}_{=k}(j) \right) \cap \mathcal{W} \right| = k|\mathcal{W}|.$$

□

Lemma B.5. *For any $\mathcal{W}' \subseteq \mathcal{Q}_{=k}$ and $\mathcal{W} := \mathcal{Q}_{=k} \setminus \mathcal{W}'$ with $|\mathcal{W}'| < \frac{(1-1/n)k}{k+n-2} \binom{n}{k}$ there exists $j \in [n] \setminus \{1\}$ with $|\mathcal{Q}_{=k}(j) \cap \mathcal{W}| > |\mathcal{Q}_{=k}(1) \cap \mathcal{W}'|$.*

[3] At any time $t \in \mathbb{N}$, a deterministic algorithm $\mathrm{Alg} \in \mathfrak{A}$ may either make a query $Q \in \mathcal{Q}_{\le k}$ or terminate with a decision $\mathrm{X} \in \{1, \dots, n\}$. Thus, $\mathfrak{A}$ is a countable set.

Proof of Lemma B.5. For $j \in [n] \setminus \{1\}$ abbreviate $a_j := |\mathcal{Q}_{=k}(j) \cap \mathcal{W}| - |\mathcal{Q}_{=k}(1) \cap \mathcal{W}'|$. Due to

$$\begin{aligned} |\mathcal{W}| &= \binom{n}{k} - |\mathcal{W}'| \\ &> \binom{n}{k} - \left(1 - \frac{1}{n}\right) \frac{k}{k+n-2} \binom{n}{k} \\ &= \binom{n}{k} - \frac{k\binom{n}{k} - \frac{k}{n}\binom{n}{k}}{k+n-2} \\ &= \frac{1}{k+n-2}\left(\binom{n-1}{k-1} + (n-2)\binom{n}{k}\right) \end{aligned}$$

we have

$$k|\mathcal{W}| - \binom{n-1}{k-1} - (n-2)\left(\binom{n}{k} - |\mathcal{W}|\right) > 0.$$

By using Lemma B.4 and the fact that $(\mathcal{W} \cap \mathcal{Q}_{=k}(1)) \cup (\mathcal{W}' \cap \mathcal{Q}_{=k}(1)) = \mathcal{Q}_{=k}(1)$ is a disjoint union, we obtain

$$\begin{aligned} &\sum\nolimits_{j \neq 1} a_j = \sum\nolimits_{j \neq 1} |\mathcal{Q}_{=k}(j) \cap \mathcal{W}| - (n-1)|\mathcal{Q}_{=k}(1) \cap \mathcal{W}'| \\ &= \sum\nolimits_{j \in [n]} |\mathcal{Q}_{=k}(j) \cap \mathcal{W}| - |\mathcal{Q}_{=k}(1) \cap \mathcal{W}| - |\mathcal{Q}_{=k}(1) \cap \mathcal{W}'| - (n-2)|\mathcal{Q}_{=k}(1) \cap \mathcal{W}'| \\ &= k|\mathcal{W}| - |\mathcal{Q}_{=k}(1)| - (n-2)|\mathcal{Q}_{=k}(1) \cap \mathcal{W}'| \\ &\geq k|\mathcal{W}| - \binom{n-1}{k-1} - (n-2)|\mathcal{W}'| \\ &= k|\mathcal{W}| - \binom{n-1}{k-1} - (n-2)\left(\binom{n}{k} - |\mathcal{W}|\right) > 0. \end{aligned}$$

Consequently, there exists $j \in [n] \setminus \{1\}$ with $a_j > 0$. □

Proof of Theorem 3.1 (ii). Similarly as in the proof of Theorem 3.1 (i), we proceed in two steps.

Part 1: The statement holds in case Alg **is deterministic.**
Abbreviate $B' := \overline{\gamma}^{-1}\left(\frac{S_{(1)} - S_{(|Q|)}}{2}\right)$ and fix a family of permutations $(\pi_Q)_{Q \in \mathcal{Q}_{\leq k}}$ with $S_{(1)|Q} = S_{\pi_Q(1)|Q}$ for all $Q \in \mathcal{Q}_{\leq k}(1)$. Exactly as in the proof of Theorem 3.1 (i), we define $\mathbf{s} = (s_{i|Q}(t))_{Q \in \mathcal{Q}_{\leq k}, i \in Q, t \in \mathbb{N}}$ via

$$s_{i|Q}(t) := \begin{cases} \frac{S_{(1)|Q} + S_{(|Q|)|Q}}{2}, & \text{if } t < B' \\ S_{\pi_Q(i)|Q}, & \text{if } t \geq B'. \end{cases}$$

In the proof of Theorem 3.1 (i) we have already verified $\mathbf{s} \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$. For any $j \in \{2, \ldots, m\}$ and $Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{Q}_{=k}(j)$ we have $S_{1|Q} > S_{j|Q}$, and using that $|\mathcal{Q}_{=k}(i') \setminus \mathcal{Q}_{=k}(j')|$ is the same for every distinct $i', j' \in [n]$ we thus have

$$\begin{aligned} \sum\nolimits_{Q \in \mathcal{Q}_{=k}(1)} S_{1|Q} &= \sum\nolimits_{Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{Q}_{=k}(j)} S_{1|Q} + S_{(1)} \cdot |\mathcal{Q}_{=k}(1) \setminus \mathcal{Q}_{=k}(j)| \\ &> \sum\nolimits_{Q \in \mathcal{Q}_{=k}(j) \cap \mathcal{Q}_{=k}(1)} S_{j|Q} + S_{(1)} \cdot |\mathcal{Q}_{=k}(j) \setminus \mathcal{Q}_{=k}(1)| \\ &> \sum\nolimits_{Q \in \mathcal{Q}_{=k}(j)} S_{j|Q}. \end{aligned}$$

As $|\mathcal{Q}_{=k}(1)| = |\mathcal{Q}_{=k}(j)|$, this shows $\mathrm{GBW}(\mathbf{s}) = 1$.

In the following, we will show that

$$\mathcal{W}' := \{Q \in \mathcal{Q}_{=k} \,:\, \text{Alg started on } \mathbf{s} \text{ queries } Q \text{ at least } B' \text{ times}\}$$

contains at least $\frac{(1-1/n)k}{k+n-2}\binom{n}{k}$ elements. For this, let us assume on the contrary $|\mathcal{W}'| < \frac{(1-1/n)k}{k+n-2}\binom{n}{k}$ and write $\mathcal{W} := \mathcal{Q}_{=k} \setminus \mathcal{W}'$. Lemma B.5 allows us to fix a $j \in [n] \setminus \{1\}$ with $|\mathcal{Q}_{=k}(j) \cap \mathcal{W}| >$

$|\mathcal{Q}_{=k}(1) \cap \mathcal{W}'|$. Now, define $\mathbf{s}' = (s'_{i|Q}(t))_{Q \in \mathcal{Q}_{\leq k}, i \in Q, t \in \mathbb{N}}$ via $s'_{\cdot|Q}(\cdot) = s_{\cdot|Q}(\cdot)$ for any $Q \in (\mathcal{Q}_{\leq k} \setminus (\mathcal{Q}_{=k}(1) \cup \mathcal{Q}_{=k}(j))) \cup \mathcal{W}'$ and[4]

$$s'_{i|Q}(t) := \begin{cases} s_{i|Q}(t), & \text{if } t < B' \text{ or } \{1, j\} \not\subseteq Q, \\ S_{(1)}, & \text{if } i = j \in Q \text{ and } t \geq B', \\ S_{(|Q|)}, & \text{if } i = 1 \in Q \text{ and } t \geq B', \\ S_{1|Q}, & \text{if } t \geq B', i = \arg\min_{l' \in Q} S_{l'|Q} \text{ and } 1 \in Q \not\ni j, \\ S_{j|Q}, & \text{if } t \geq B', i = \mathrm{argmax}_{l' \in Q} S_{l'|Q} \text{ and } j \in Q \not\ni 1, \\ S_{i|Q}, & \text{otherwise,} \end{cases}$$

for $Q \in (\mathcal{Q}_{=k}(1) \cup \mathcal{Q}_{=k}(j)) \cap \mathcal{W}$. Similarly as for $\mathbf{s}$, we see $\mathbf{s}' \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$. The corresponding limit values $S'_{i|Q} = \lim_{t \to \infty} s'_{i|Q}(t)$ fulfill

$$\forall Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{W} : S'_{1|Q} = S_{(|Q|)} \quad \text{and} \quad \forall Q \in \mathcal{Q}_{=k}(j) \cap \mathcal{W} : S'_{j|Q} = S_{(1)},$$

and trivially also $S_{(|Q|)} \leq S'_{i|Q} \leq S_{(1)}$ for any $Q \in \mathcal{Q}_{=k}, i \in Q$. Therefore, by choice of j, the corresponding Borda scores $(S')_i^{\mathcal{B}}$ for $\mathbf{s}'$ fulfill

$$\begin{aligned} \binom{n-1}{k-1}(S')_1^{\mathcal{B}} &= \sum\nolimits_{Q \in \mathcal{Q}_{=k}(1)} S'_{1|Q} = \sum\nolimits_{Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{W}'} S_{(1)} + \sum\nolimits_{Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{W}} S_{(|Q|)} \\ &= |\mathcal{Q}_{=k}(1) \cap \mathcal{W}'| \cdot S_{(1)} + |\mathcal{Q}_{=k}(1) \cap \mathcal{W}| \cdot S_{(|Q|)} \\ &< |\mathcal{Q}_{=k}(j) \cap \mathcal{W}| \cdot S_{(1)} + |\mathcal{Q}_{=k}(j) \cap \mathcal{W}'| \cdot S_{(|Q|)} \\ &\leq \sum\nolimits_{Q \in \mathcal{Q}_{=k}(j)} S'_{j|Q} = \binom{n-1}{k-1}(S')_j^{\mathcal{B}}, \end{aligned}$$

where we have used that $|\mathcal{Q}_{=k}(1) \cap \mathcal{W}'| + |\mathcal{Q}_{=k}(1) \cap \mathcal{W}| = |\mathcal{Q}_{=k}(1)| = |\mathcal{Q}_{=k}(j) \cap \mathcal{W}'| + |\mathcal{Q}_{=k}(j) \cap \mathcal{W}|$. This show $1 \notin \mathrm{GBW}(\mathbf{s}')$. But since $s_{\cdot|\cdot}(\cdot) = s'_{\cdot|\cdot}(\cdot)$ holds on $\{t < B'\}$ as well as on $\mathcal{W}'$, Alg observes for $\mathbf{s}$ until termination exactly the same feedback as for $\mathbf{s}'$. Consequently, it outputs for both instances the same decision. Since $\mathrm{GBW}(\mathbf{s}) = 1 \notin \mathrm{GBW}(\mathbf{s}')$, it makes on at least one of the instances a mistake, which contradicts the correctness of Alg.
Thus, $|\mathcal{W}'| \geq \frac{(1-1/n)k}{k+n-2}\binom{n}{k}$ has to hold and we conclude

$$B(\mathrm{Alg}, \mathbf{s}) \geq \sum\nolimits_{Q \in \mathcal{W}'} B_Q(\mathrm{Alg}, \mathbf{s}) \geq |\mathcal{W}'| \cdot B' \geq \left(1 - \frac{1}{n}\right) \frac{k}{k+n-2} \binom{n}{k} B'.$$

Since $1 - \frac{1}{n} \geq 1/2$ and $k \leq n+2$ hold by assumption, we have in particular

$$B(\mathrm{Alg}, \mathbf{s}) \geq \frac{k}{4n}\binom{n}{k}\overline{\gamma}^{-1}\left(\frac{S_{(1)} - S_{(|Q|)}}{2}\right) = \frac{1}{4}\binom{n-1}{k-1}\overline{\gamma}^{-1}\left(\frac{S_{(1)} - S_{(|Q|)}}{2}\right) \in \Omega\left(\binom{n-1}{k-1}\right).$$

Part 2: The statement holds for arbitrary Alg.
Similarly as for the proof of (i) in Theorem 3.1, the proof follows by means of Yao's minimax principle. □

Remark B.6. *(i) To compare the bounds for* RoundRobin *in Theorem C.1 with the lower bound from Theorem 3.1 (ii) suppose in the following* $\mathbf{S}$ *to be homogeneous with* $S_{(1)} > S_{(2)}$ *and let* $\boldsymbol{\gamma}$ *be homogeneous in the sense that* $\gamma_{i|Q}(t) = \gamma(t)$ *for all* $Q \in \mathcal{Q}_{=k}, i \in Q, t \in \mathbb{N}$ *for some* $\gamma : \mathbb{N} \to [0, \infty)$. *Moreover, let* $\mathbf{s}$ *be the instance from the proof of Theorem 3.1 (ii), and denote by* $\mathbf{S}$ *the family of limits* $S_{i|Q} = \lim_{t \to \infty} s_{i|Q}(t)$, $Q \in \mathcal{Q}_{\leq k}, i \in Q$. *Let us write* $S^{\mathcal{B}}_{(1)}, \ldots, S^{\mathcal{B}}_{(n)}$ *for the order statistics of* $\{S^{\mathcal{B}}_i\}_{i \in [n]}$, *i.e.,* $S^{\mathcal{B}}_{(1)} \geq \cdots \geq S^{\mathcal{B}}_{(n)}$. *Then,* RoundRobin *returns a GBW of* $\mathbf{s} \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$ *if it is executed with a budget* B *at least*

$$z_{\mathrm{RR}} = \binom{n}{k} B_1 \quad \textit{with} \quad B_1 := \overline{\gamma}^{-1}\left(\frac{S^{\mathcal{B}}_{(1)} - S^{\mathcal{B}}_{(2)}}{2}\right).$$

[4] That is, for constructing $\mathbf{s}'$, we proceed for $Q \in \mathcal{W}$ as follows: If $\{1, j\} \subseteq Q$, we exchange $S_{1|Q}$ with $S_{j|Q}$. If $1 \in Q \not\ni j$, we exchange $S_{1|Q}$ with $S_{(Q)|Q}$. And if $j \in Q \not\ni 1$, we exchange $S_{j|Q}$ with $S_{(1)|Q}$.

In comparison to this, the lower bound just shown reveals that any (possibly deterministic) solution to $\mathcal{P}_{\mathrm{GBW}}(\mathbf{S},\boldsymbol{\gamma})$ fulfills

$$\mathbb{E}[B(\mathrm{Alg},\mathbf{s})] \geq \left(1-\frac{1}{n}\right)\frac{k}{k+n-2}\binom{n}{k}B_2 \quad \textit{with} \quad B_2 := \overline{\gamma}^{-1}\left(\frac{S_{(1)}-S_{(|Q|)}}{2}\right).$$

Consequently, the optimality-gap between the upper and lower bound is of the order

$$B_1^{-1}B_2\left(1-\frac{1}{n}\right)\frac{k}{k+n-2}.$$

(ii) In the proof of Theorem 3.1 (ii), where we showed that $|\mathcal{W}'| \geq \frac{(1-1/n)k}{k+n-2}\binom{n}{k}$ leads to a contradiction, we have constructed an instance $\mathbf{s}' \in \mathfrak{S}(\mathbf{S},\boldsymbol{\gamma})$ with $\mathrm{GBW}(\mathbf{s}) = 1 \notin \mathrm{GBW}(\mathbf{s}')$ such that Alg observes on $\mathbf{s}$ the same feedback as on $\mathbf{s}'$. To finish the proof, we have only used that Alg is correct for $\mathbf{s}$ and for $\mathbf{s}'$, but we did not require correctness of Alg on any instance $\mathbf{s}'' \in \mathfrak{S}(\mathbf{S},\boldsymbol{\gamma}) \setminus \{\mathbf{s},\mathbf{s}'\}$. The construction of $\mathbf{s}'$ therein dependeds on the behaviour of Alg only by means of the choices of $\mathcal{W}$ and j in the proof, i.e., we have the dependence $\mathbf{s}' = \mathbf{s}'(\mathcal{W},j)$. Recall that for constructing $\mathbf{s}'$ we used that $|\mathcal{W}| = |\mathcal{Q}_{=k}| - |\mathcal{W}'| \geq \binom{n}{k} - \frac{(1-1/n)k}{k+n-2}\binom{n}{k}$, so that for $j \in [n]\setminus\{1\}$, the set

$$\left\{\mathbf{s}'(\mathcal{W},j)\,\middle|\,\mathcal{W}\subseteq\mathcal{Q}_{=k} \textit{ with } |\mathcal{W}| \geq \binom{n}{k} - \frac{(1-1/n)k}{k+n-2}\binom{n}{k} \textit{ and } j\in[n]\setminus\{1\}\right\}$$

of possible choices for $\mathbf{s}'$ has at most

$$N := (n-1)\sum_{l=\lceil\binom{n}{k}-\frac{(1-1/n)k}{k+n-2}\binom{n}{k}\rceil}^{\binom{n}{k}}\binom{\binom{n}{k}}{l}$$

elements, say $\mathbf{s}'_1,\dots,\mathbf{s}'_N$. Thus, the formulation of the theorem may be strengthened in the following way:
If $\mathbf{S}$ is homogeneous and $\boldsymbol{\gamma}$ fixed, then there exist $N+1$ instances $\mathbf{s},\mathbf{s}'_1,\dots,\mathbf{s}'_N$ with the following property: Whenever a (possibly probabilistic) sequential testing algorithm Alg correctly identifies the GBW for any of these $N+1$ instances, then

$$\mathbb{E}\left[B(\mathcal{A},\mathbf{s})\right] \geq \left(1-\frac{1}{n}\right)\frac{k}{k+n-2}\binom{n}{k}\overline{\gamma}^{-1}\left(\frac{S_{(1)}-S_{(|Q|)}}{2}\right).$$

B.3 Proof of Theorem 3.1 (iii): Lower Bound for GCopeW Identification

Recall that $\mathrm{GCopeW}(\mathbf{s})$ is the set of elements $i \in [n]$, for which the limits $S_{i|Q} = \lim_{t\to\infty} s_{i|Q}(t)$ have the highest Copeland score

$$S_i^{\mathcal{C}} = \frac{\sum_{Q\in\mathcal{Q}_{=k}(i)} \mathbb{1}\{S_{i|Q} = S_{(1)|Q}\}}{|\mathcal{Q}_{=k}(i)|} = \frac{\sum_{Q\in\mathcal{Q}_{=k}(i)} \mathbb{1}\{S_{i|Q} = S_{(1)|Q}\}}{\binom{n-1}{k-1}}.$$

Proof of Theorem 3.1.(iii). Similarly as in the proofs (i) and (ii) Theorem 3.1, we proceed in two steps.

Part 1: The statement holds in case Alg is deterministic.
Abbreviate $B' := \min_{Q\in\mathcal{Q}_{\leq k}}\min_{i\in Q}\ \gamma_{i|Q}^{-1}\left(\frac{S_{(1)|Q}-S_{(|Q|)|Q}}{2}\right)$ and fix a family of permutations $(\pi_Q)_{Q\in\mathcal{Q}_{\leq k}}$ with $S_{(1)|Q} = S_{\pi_Q(1)|Q}$ for all $Q \in \mathcal{Q}_{\leq k}(1)$. Exactly as in the proofs of the lower bounds for GCW and GBW identification, we define $\mathbf{s} = (s_{i|Q}(t))_{Q\in\mathcal{Q}_{\leq k}, i\in Q, t\in\mathbb{N}}$ via

$$s_{i|Q}(t) := \begin{cases} \frac{S_{(1)|Q}+S_{(|Q|)|Q}}{2}, & \text{if } t < B' \\ S_{\pi_Q(i)|Q}, & \text{if } t \geq B'. \end{cases}$$

In the proof of the lower bound of GCW identification we have already verified $\mathbf{s} \in \mathfrak{S}(\mathbf{S},\boldsymbol{\gamma})$. For any $j \in \{2,\dots,m\}$ and $Q \in \mathcal{Q}_{=k}(1)\cap\mathcal{Q}_{=k}(j)$ we have $S_{1|Q} > S_{j|Q}$, and using that $|\mathcal{Q}_{=k}(i')\setminus\mathcal{Q}_{=k}(j')|$

is the same for every distinct $i', j' \in [n]$ we thus have

$$
\begin{aligned}
&\sum_{Q \in \mathcal{Q}_{=k}(1)} \mathbb{1}\{S_{1|Q} = S_{(1)|Q}\} \\
&\qquad = \sum_{Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{Q}_{=k}(j)} \mathbb{1}\{S_{1|Q} = S_{(1)|Q}\} + \sum_{Q \in \mathcal{Q}_{=k}(1) \setminus \mathcal{Q}_{=k}(j)} \mathbb{1}\{S_{1|Q} = S_{(1)|Q}\} \\
&\qquad = \sum_{Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{Q}_{=k}(j)} \mathbb{1}\{S_{1|Q} = S_{(1)|Q}\} + |\mathcal{Q}_{=k}(1) \setminus \mathcal{Q}_{=k}(j)| \\
&\qquad > \sum_{Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{Q}_{=k}(j)} \mathbb{1}\{S_{j|Q} = S_{(1)|Q}\} + |\mathcal{Q}_{=k}(j) \setminus \mathcal{Q}_{=k}(1)| \\
&\qquad \geq \sum_{Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{Q}_{=k}(j)} \mathbb{1}\{S_{j|Q} = S_{(1)|Q}\} + \sum_{Q \in \mathcal{Q}_{=k}(j) \setminus \mathcal{Q}_{=k}(1)} \mathbb{1}\{S_{j|Q} = S_{(1)|Q}\} \\
&\qquad = \sum_{Q \in \mathcal{Q}_{=k}(j)} \mathbb{1}\{S_{j|Q} = S_{(1)|Q}\}.
\end{aligned}
$$

As $|\mathcal{Q}_{=k}(1)| = |\mathcal{Q}_{=k}(j)|$, this shows $\mathrm{GCopeW}(\mathbf{s}) = 1$.
Similarly as in the proof of (ii), we will show indirectly that

$$
\mathcal{W}' \coloneqq \{Q \in \mathcal{Q}_{=k} \,:\, \text{Alg started on } \mathbf{s} \text{ queries } Q \text{ at least } B' \text{ times}\}
$$

contains at least $\frac{(1-1/n)k}{k+n-2}\binom{n}{k}$ elements. For this purpose, let us assume on the contrary $|\mathcal{W}'| < \frac{(1-1/n)k}{k+n-2}\binom{n}{k}$ and write $\mathcal{W} \coloneqq \mathcal{Q}_{=k} \setminus \mathcal{W}'$. Lemma B.5 allows us to fix a $j \in [n] \setminus \{1\}$ with $|\mathcal{Q}_{=k}(j) \cap \mathcal{W}| > |\mathcal{Q}_{=k}(1) \cap \mathcal{W}'|$. Now, define $\mathbf{s}' = (s'_{i|Q}(t))_{Q \in \mathcal{Q}_{\leq k}, i \in Q, t \in \mathbb{N}}$ analogously as in the proof of (ii), i.e., via $s'_{\cdot|Q}(\cdot) = s_{\cdot|Q}(\cdot)$ for any $Q \in (\mathcal{Q}_{\leq k} \setminus (\mathcal{Q}_{=k}(1) \cup \mathcal{Q}_{=k}(j))) \cup \mathcal{W}'$ and

$$
s'_{i|Q}(t) \coloneqq \begin{cases} s_{i|Q}(t), & \text{if } t < B' \text{ or } \{1, j\} \not\subseteq Q, \\ S_{(1)|Q}, & \text{if } i = j \in Q \text{ and } t \geq B', \\ S_{(|Q|)|Q}, & \text{if } i = 1 \in Q \text{ and } t \geq B', \\ S_{1|Q}, & \text{if } t \geq B', i = \arg\min_{l' \in Q} S_{l'|Q} \text{ and } 1 \in Q \not\ni j, \\ S_{j|Q}, & \text{if } t \geq B', i = \mathrm{argmax}_{l' \in Q} S_{l'|Q} \text{ and } j \in Q \not\ni 1, \\ S_{i|Q}, & \text{otherwise,} \end{cases}
$$

for $Q \in (\mathcal{Q}_{=k}(1) \cup \mathcal{Q}_{=k}(j)) \cap \mathcal{W}$. Similarly as for $\mathbf{s}$, we see $\mathbf{s}' \in \mathfrak{S}(\mathbf{S}, \boldsymbol{\gamma})$. The corresponding limit values $S'_{i|Q} = \lim_{t \to \infty} s'_{i|Q}(t)$ fulfill

$$
\forall Q \in \mathcal{Q}_{=k}(1) \cap \mathcal{W} : S'_{1|Q} = S_{(|Q|)|Q} \quad \text{and} \quad \forall Q \in \mathcal{Q}_{=k}(j) \cap \mathcal{W} : S'_{j|Q} = S_{(1)|Q},
$$

and trivially also $S_{(|Q|)|Q} \leq S'_{i|Q} \leq S_{(1)|Q}$ for any $Q \in \mathcal{Q}_{=k}, i \in Q$. Therefore, by choice of j, the corresponding Copeland scores $(S')_i^{\mathcal{C}}$ for $\mathbf{s}'$ fulfill

$$\begin{aligned}
\binom{n-1}{k-1}(S')_1^{\mathcal{C}} &= \sum\nolimits_{Q\in\mathcal{Q}_{=k}(1)} \mathbb{1}\{S'_{1|Q} = S'_{(1)|Q}\} \\
&= \sum_{Q\in\mathcal{Q}_{=k}(1)\cap\mathcal{W}'} \mathbb{1}\{S'_{1|Q} = S'_{(1)|Q}\} + \sum_{Q\in\mathcal{Q}_{=k}(1)\cap\mathcal{W}} \mathbb{1}\{S'_{1|Q} = S'_{(1)|Q}\} \\
&= \sum_{Q\in\mathcal{Q}_{=k}(1)\cap\mathcal{W}'} \mathbb{1}\{S_{1|Q} = S_{(1)|Q}\} + \sum_{Q\in\mathcal{Q}_{=k}(1)\cap\mathcal{W}} \mathbb{1}\{S_{(|Q|)|Q} = S_{(1)|Q}\} \\
&= |\mathcal{Q}_{=k}(1)\cap\mathcal{W}'| \\
&< |\mathcal{Q}_{=k}(j)\cap\mathcal{W}| \\
&= \sum_{\mathcal{Q}_{=k}(j)\cap\mathcal{W}} \mathbb{1}\{S_{(1)|Q} = S_{(1)|Q}\} \\
&= \sum_{\mathcal{Q}_{=k}(j)\cap\mathcal{W}} \mathbb{1}\{S'_{j|Q} = S'_{(1)|Q}\} \\
&\leq \sum_{\mathcal{Q}_{=k}(j)\cap\mathcal{W}} \mathbb{1}\{S'_{j|Q} = S'_{(1)|Q}\} + \sum_{\mathcal{Q}_{=k}(j)\cap\mathcal{W}'} \mathbb{1}\{S'_{j|Q} = S'_{(1)|Q}\} \\
&= \sum_{Q\in\mathcal{Q}_{=k}(j)} \mathbb{1}\{S'_{j|Q} = S'_{(1)|Q}\} \\
&= \binom{n-1}{k-1}(S')_j^{\mathcal{C}},
\end{aligned}$$

where we used that $S_{(1)|Q} = S'_{(1)|Q}$. This shows $1 \notin \mathrm{GCopeW}(\mathbf{s}')$. But since $s_{\cdot|\cdot}(\cdot) = s'_{\cdot|\cdot}(\cdot)$ holds on $\{t < B'\}$ as well as on $\mathcal{W}'$, Alg observes for $\mathbf{s}$ until termination exactly the same feedback as for $\mathbf{s}'$. Consequently, it outputs for both instances the same decision. Since $\mathrm{GCopeW}(\mathbf{s}) = 1 \notin \mathrm{GCopeW}(\mathbf{s}')$, it makes on at least one of the instances a mistake, which contradicts the correctness of Alg.

Thus, $|\mathcal{W}'| \geq \frac{(1-1/n)k}{k+n-2}\binom{n}{k}$ has to hold and we conclude

$$B(\mathrm{Alg},\mathbf{s}) \geq \sum\nolimits_{Q\in\mathcal{W}'} B_Q(\mathrm{Alg},\mathbf{s}) \geq |\mathcal{W}'|\cdot B' \geq \left(1-\frac{1}{n}\right)\frac{k}{k+n-2}\binom{n}{k}B'.$$

Since $1-\frac{1}{n} \geq 1/2$ and $k \leq n+2$ hold by assumption, we have in particular

$$\begin{aligned}
B(\mathrm{Alg},\mathbf{s}) &\geq \frac{k}{4n}\binom{n}{k}\min_{Q\in\mathcal{Q}_{\leq k}}\min_{i\in Q}\gamma_{i|Q}^{-1}\left(\frac{S_{(1)|Q}-S_{(|Q|)|Q}}{2}\right) \\
&= \frac{1}{4}\binom{n-1}{k-1}\min_{Q\in\mathcal{Q}_{\leq k}}\min_{i\in Q}\gamma_{i|Q}^{-1}\left(\frac{S_{(1)|Q}-S_{(|Q|)|Q}}{2}\right) \in \Omega\left(\binom{n-1}{k-1}\right).
\end{aligned}$$

Part 2: The statement holds for arbitrary Alg**.**
Similarly as for the proofs of the lower bound of (i) and (ii) of this theorem, the proof follows by means of Yao's minimax principle. □

C Generalized Borda Winner Identification

Let ROUNDROBIN be the algorithm, which enumerates all possible subsets of the fixed subset size k, chooses each subset in a round-robin fashion and returns the arm with the highest empirical Borda score $s_i^{\mathcal{B}}$ after the available budget is exhausted. It is a straightforward baseline method, which we analyze theoretically in terms of the sufficient and necessary budget to return a generalized Borda winner (GBW) $i^*_{\mathcal{B}}$. For this purpose, let $\hat{\gamma}_i(t) = \frac{1}{|\mathcal{Q}_{=k}(i)|}\sum_{Q\in\mathcal{Q}_{=k}(i)}\gamma_{i|Q}(t)$ and $\hat{\gamma}_{i,j}^{\max}(t) = \max\{\hat{\gamma}_i(t), \hat{\gamma}_j(t)\}$.

Theorem C.1. ROUNDROBIN *returns $i^*_{\mathcal{B}}$ if it is executed with a budget $B \geq z_{\mathrm{RR}}$, where*

$$z_{\mathrm{RR}} := \binom{n}{k} \max_{\rho \in \mathcal{A}, \rho \neq i^*_{\mathcal{B}}} \left(\hat{\gamma}^{\max}_{i^*_{\mathcal{B}},\rho}\right)^{-1} \left(\frac{S^{\mathcal{B}}_{i^*_{\mathcal{B}}} - S^{\mathcal{B}}_{\rho}}{2}\right).$$

The latter bound is tight in a worst-case scenario, as the following result shows (cf. Sec. D.1 for the proofs).

Theorem C.2. *For any asymptotical Borda scores $S^{\mathcal{B}}_1, \ldots, S^{\mathcal{B}}_n$, there exists a corresponding instance* $\mathbf{s}$ *such that if $B < z_{\mathrm{RR}}$ then* ROUNDROBIN *will not return $i^*_{\mathcal{B}}$.*

Thus, ROUNDROBIN is already nearly-optimal (up to a factor $\mathcal{O}(n/k)$) with respect to worst-case scenarios due to Theorem 3.1 (see Rem. B.6 for a more detailed discussion.).

D Proofs of Section 4

In this section we provide the detailed proofs of Section 4. We assume throughout that $\frac{B}{\binom{n}{k}}$ is a natural number, i.e., the budget is a multiple of $\binom{n}{k.}$

D.1 Proof of Theorems C.1 and C.2

Proof of Theorem C.1. After relabeling the arms in round r we may assume w.l.o.g. $i^*_{\mathcal{B}} = 1$. We will prove the theorem by contradiction and therefore assume

$$\begin{aligned}
&\rho = \mathrm{argmax}_{i \in \mathcal{A}}\, s^{\mathcal{B}}_i\left(\frac{B}{\binom{n}{k}}\right) \neq 1 \\
&\Rightarrow s^{\mathcal{B}}_1\left(\frac{B}{\binom{n}{k}}\right) < \max_{j=2,\ldots n} s^{\mathcal{B}}_j\left(\frac{B}{\binom{n}{k}}\right) = s^{\mathcal{B}}_\rho\left(\frac{B}{\binom{n}{k}}\right) \\
&\Rightarrow S^{\mathcal{B}}_1 - S^{\mathcal{B}}_\rho < s^B_\rho\left(\frac{B}{\binom{n}{k}}\right) - S^{\mathcal{B}}_\rho + S^{\mathcal{B}}_1 - s^{\mathcal{B}}_1\left(\frac{B}{\binom{n}{k}}\right) \\
&\quad = \frac{1}{|\mathcal{Q}_{=k}(\rho)|}\sum_{Q \in \mathcal{Q}_{=k}(\rho)}\left(s_{\rho|Q}\left(\frac{B}{\binom{n}{k}}\right) - S_{\rho|Q}\right) + \frac{1}{|\mathcal{Q}_{=k}(1)|}\sum_{Q \in \mathcal{Q}_{=k}(1)}\left(S_{1|Q} - s_{1|Q}\left(\frac{B}{\binom{n}{k}}\right)\right) \\
&\Rightarrow S^{\mathcal{B}}_1 - S^{\mathcal{B}}_\rho < \hat{\gamma}_\rho\left(\frac{B}{\binom{n}{k}}\right) + \hat{\gamma}_1\left(\frac{B}{\binom{n}{k}}\right) \\
&\Rightarrow S^{\mathcal{B}}_1 - S^{\mathcal{B}}_\rho < 2 \cdot \hat{\gamma}^{\max}_{1,\rho}\left(\frac{B}{\binom{n}{k}}\right),
\end{aligned}$$

where $\hat{\gamma}_i(t) = \frac{1}{|\mathcal{Q}_{=k}(i)|}\sum_{Q \in \mathcal{Q}_{=k}(i)} \gamma_{i|Q}(t)$ and $\hat{\gamma}^{\max}_{i,j}(t) = \max\{\hat{\gamma}_i(t), \hat{\gamma}_j(t)\}$. With this, however, we can derive

$$\Rightarrow\ z_{RR} = \left(\hat{\gamma}^{\max}_{1,\rho}\right)^{-1}\left(\frac{S^{\mathcal{B}}_1 - S^{\mathcal{B}}_\rho}{2}\right)\binom{n}{k} \geq B,$$

which contradicts the assumption we make on the budget B. Thus, it holds that the returned arm is $\rho = 1$. ☐

Proof of Theorem C.2. Let $\beta(t)$ be an arbitrary, monotonically decreasing function of t with $\lim_{t\to\infty}\beta(t) = 0$. We define for all $j \in \mathcal{A}$ with $j \neq i^*_{\mathcal{B}}$ the empirical Borda scores to be $s^{\mathcal{B}}_j(t) = S^{\mathcal{B}}_j + \beta(t)$ and $s^{\mathcal{B}}_{i^*_{\mathcal{B}}}(t) = S^{\mathcal{B}}_{i^*_{\mathcal{B}}} - \beta(t)$, where $(S^{\mathcal{B}}_i)_{i\in[n]}$ are arbitrary real values such that $S^{\mathcal{B}}_{i^*_{\mathcal{B}}}$ is the unique maximum for some $i^*_{\mathcal{B}} \in [n]$. We can again assume after relabeling all arms that w.l.o.g. that $i^*_{\mathcal{B}} = 1$ and $\mathrm{argmax}_{j=2,\dots,n} S^{\mathcal{B}}_j = 2$. Note that $\hat{\gamma}_i(t) = \beta(t)$ for all $i \in \mathcal{A}$. In light of these considerations, ROUNDROBIN returns 1 as the best arm if and only if

$$
\begin{aligned}
s_1^{\mathcal{B}}\left(\frac{B}{\binom{n}{k}}\right) > \max_{j=2,\dots,n} s_j^{\mathcal{B}}\left(\frac{B}{\binom{n}{k}}\right) \quad & \Leftrightarrow S_1^{\mathcal{B}} - \hat{\gamma}_1\left(\frac{B}{\binom{n}{k}}\right) > \max_{j=2,\dots,n} S_j^{\mathcal{B}} + \hat{\gamma}_j\left(\frac{B}{\binom{n}{k}}\right) \\
& \Leftrightarrow S_1^{\mathcal{B}} - \hat{\gamma}_1\left(\frac{B}{\binom{n}{k}}\right) > S_2^{\mathcal{B}} + \hat{\gamma}_2\left(\frac{B}{\binom{n}{k}}\right) \\
& \Leftrightarrow \hat{\gamma}_1\left(\frac{B}{\binom{n}{k}}\right) + \hat{\gamma}_2\left(\frac{B}{\binom{n}{k}}\right) < S_1^{\mathcal{B}} - S_2^{\mathcal{B}} \\
& \Leftrightarrow 2 \cdot \hat{\gamma}_{1,2}^{\max}\left(\frac{B}{\binom{n}{k}}\right) < S_1^{\mathcal{B}} - S_2^{\mathcal{B}} \\
& \Leftrightarrow B \geq \binom{n}{k}\left(\hat{\gamma}_{1,2}^{\max}\right)^{-1}\left(\frac{S_1^{\mathcal{B}} - S_2^{\mathcal{B}}}{2}\right).
\end{aligned}
$$

Thus, the necessary budget is z_{RR} in this case concluding the claim. □

D.2 Proofs of Theorem 4.1 and 4.2

Proof of Theorem 4.1. For the sake of convenience, let us abbreviate $[R] := \{1, \dots, R\}$ and $\mathbb{A}_{rj} := \mathbb{A}_{r,j}$ in the following. By possibly relabeling the arms and query sets queried by the algorithm, we can assume w.l.o.g. $i^* = 1$ and $\mathbb{A}_r(1) = \mathbb{A}_{r1}$ for all $r \in [R]$ in the following. In particular, we have $S_{1|\mathbb{A}_{r1}} = S_{(1)|\mathbb{A}_{r1}}$ for all $r \in [R]$. We prove the correctness of the algorithm indirectly. Thus, we start by assuming that the best arm is not contained in the last partition (i.e., the remaining active arm):

$$
\begin{aligned}
& \mathbb{A}_{R+1} \neq \{1\} \\
\Leftrightarrow & \exists r \in [R] : 1 \notin \mathbb{A}_{r+1} \land 1 \in \mathbb{A}_r \\
\Rightarrow & \exists r \in [R] : \sum_{i \in \mathbb{A}_{r1}} \mathbb{1}\{s_{i|\mathbb{A}_{r1}}(b_r) \geq s_{1|\mathbb{A}_{r1}}(b_r)\} > f(|\mathbb{A}_{r1}|) \\
\Rightarrow & \exists r \in [R] : \sum_{i \in \mathbb{A}_{r1}} \mathbb{1}\{S_{1|\mathbb{A}_{r1}} - S_{i|\mathbb{A}_{r1}} \leq S_{1|\mathbb{A}_{r1}} - s_{1|\mathbb{A}_{r1}}(b_r) - S_{i|\mathbb{A}_{r1}} + s_{i|\mathbb{A}_{r1}}(b_r)\} > f(|\mathbb{A}_{r1}|) \\
\Rightarrow & \exists r \in [R] : \sum_{i \in \mathbb{A}_{r1}} \mathbb{1}\{S_{1|\mathbb{A}_{r1}} - S_{i|\mathbb{A}_{r1}} \leq |S_{1|\mathbb{A}_{r1}} - s_{1|\mathbb{A}_{r1}}(b_r)| + |S_{i|\mathbb{A}_{r1}} - s_{i|\mathbb{A}_{r1}}(b_r)|\} > f(|\mathbb{A}_{r1}|) \\
\Rightarrow & \exists r \in [R] : \sum_{i \in \mathbb{A}_{r1}} \mathbb{1}\{S_{1|\mathbb{A}_{r1}} - S_{i|\mathbb{A}_{r1}} \leq 2\bar{\gamma}_{\mathbb{A}_{r1}}(b_r)\} > f(|\mathbb{A}_{r1}|) \\
\Rightarrow & \exists r \in [R] : S_{1|\mathbb{A}_{r1}} - S_{(f(\mathbb{A}_{r1})+1)|\mathbb{A}_{r1}} \leq 2\bar{\gamma}_{\mathbb{A}_{r1}}(b_r) \\
\Rightarrow & \exists r \in [R] : \left\lfloor \frac{B}{P_r R} \right\rfloor = b_r \leq \bar{\gamma}_{\mathbb{A}_{r1}}^{-1}\left(\frac{S_{1|\mathbb{A}_{r1}} - S_{(f(\mathbb{A}_{r1})+1)|\mathbb{A}_{r1}}}{2}\right) \\
\Rightarrow & \exists r \in [R] : B \leq P_r R \left\lceil \bar{\gamma}_{\mathbb{A}_{r1}}^{-1}\left(\frac{S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}}{2}\right)\right\rceil \\
\Rightarrow & B \leq R \max\nolimits_{r \in [R]} P_r \left\lceil \bar{\gamma}_{\mathbb{A}_{r1}}^{-1}\left(\frac{S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}}{2}\right)\right\rceil = z\left(f, R, \{P_r\}_{1 \leq r \leq R}\right),
\end{aligned}
$$

which contradicts the assumption we make on the budget B. Thus, it holds that the remaining active arm in round $R+1$ is $i^* = 1$. □

Remark D.1. *Using the definition of $\bar{\gamma}_Q(t)$ and $\bar{\gamma}(t)$ we can derive the following more coarser bounds on the sufficient budget:*

$$
\begin{aligned}
z_1\left(f, R, \{P_r\}_{1 \leq r \leq R}\right) &= R \max\nolimits_{r \in [R]} P_r \left\lceil \bar{\gamma}^{-1}\left(\frac{S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}}{2}\right)\right\rceil, \\
z_2\left(f, R, \{P_r\}_{1 \leq r \leq R}\right) &= R\left(\max\nolimits_{r \in [R]} P_r\right) \max\nolimits_{Q \in \mathcal{Q}_{\leq k}} \left\lceil \bar{\gamma}^{-1}\left(\frac{S_{1|Q} - S_{(f(|Q|)+1)|Q}}{2}\right)\right\rceil.
\end{aligned}
$$

Proof of Theorem 4.2. After relabeling, we may suppose w.l.o.g. $i^* = 1$. Let $\beta : \mathbb{N} \to (0, \infty)$ be an arbitrary strictly decreasing function with $\beta(t) \to 0$ as $t \to \infty$ and

$$\left\{ \frac{S_{1|Q} - S_{j|Q}}{2} : Q \in \mathcal{Q}_{\leq k}, j \in Q \right\} \subseteq \beta(\mathbb{N}).$$

Then, β is invertible on $\beta(\mathbb{N})$ and its inverse function $\beta^{-1} : \beta(\mathbb{N}) \to \mathbb{N}$ trivially fulfills $\beta^{-1}(\alpha) = \min\{t \in \mathbb{N} : \beta(t) \leq \alpha\}$ for all $\alpha \in \beta(\mathbb{N})$. Define for any $Q \in \mathcal{Q}_{\leq k}$ and $i \in Q$ the family of statistics by means of

$$s_{i|Q}(t) := \begin{cases} S_{i|Q} - \beta(t), & \text{if } i = \mathrm{argmax}_{j \in Q} S_{j|Q}, \\ S_{i|Q} + \beta(t), & \text{otherwise,} \end{cases}$$

and note that $\overline{\gamma}_Q(t) = \beta(t)$ for all $Q \in \mathcal{Q}_{\leq k}$ and $t \in \mathbb{N}$. Writing $b_r = \left\lfloor \frac{B}{RP_r} \right\rfloor$ we obtain due to the choice of β that

$$\begin{aligned}
& B < R \max_{r \in [R]} P_r \bar{\gamma}_{\mathbb{A}_{r1}}^{-1} \left(\frac{S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}}{2} \right) \\
\Rightarrow & \exists r \in [R] : B < RP_r \min \left\{ t \in \mathbb{N} : \bar{\gamma}_{\mathbb{A}_{r1}}(t) \leq \frac{S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}}{2} \right\} \\
\Rightarrow & \exists r \in [R] : b_r < \min \left\{ t \in \mathbb{N} : \beta(t) \leq \frac{S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}}{2} \right\} = \beta^{-1} \left(\frac{S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}}{2} \right) \\
\Rightarrow & \exists r \in [R] : 2\beta(b_r) > S_{1|\mathbb{A}_{r1}} - S_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}} = s_{1|\mathbb{A}_{r1}}(b_r) + \beta(b_r) - \left(s_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}(b_r) - \beta(b_r) \right) \\
\Rightarrow & \exists r \in [R] : s_{1|\mathbb{A}_{r1}}(b_r) < s_{(f(|\mathbb{A}_{r1}|)+1)|\mathbb{A}_{r1}}(b_r) \\
\Rightarrow & \exists r \in [R] : 1 \notin \mathbb{A}_{r+1} \\
\Rightarrow & 1 \notin \mathbb{A}_{R+1}.
\end{aligned}$$

This shows that $z(f, R, \{P_r\}_{1 \leq r \leq R})$ is the necessary budget for returning the best arm i^* in this scenario. □

D.3 Proof of Corollary 4.3

For sake of convenience, we provide the entire pseudo-code of CSWS in Algorithm 3, which results by using $f(x) = x - 1$ as well as P_r^{CSWS} and R^{CSWS} as defined in Section 4.1 in Algorithm 1.

Proof of Corollary 4.3 (CSWS case). Suppose $B > 0$ to be arbitrary but fixed. First, note that there are at most $\lceil \log_k(n) \rceil$ rounds within the first while-loop and at most 1 in the second, so that we have at most $\lceil \log_k(n) \rceil + 1$ many rounds in total. The total number of partitions in round $r \in \{1, \ldots, \lceil \log_k(n) \rceil + 1\}$ is at most $\left\lceil \frac{n}{k^r} \right\rceil$. Abbreviating $R := R^{\mathrm{CSWS}}$ and $P_r := P_r^{\mathrm{CSWS}}$ for the moment, the budget allocated to a partition in round r is by definition $b_r = \lfloor \frac{B}{RP_r} \rfloor = \left\lfloor \frac{B}{\lceil \frac{n}{k^r} \rceil \cdot (\lceil \log_k(n) \rceil + 1)} \right\rfloor$. Hence, the total budget used by CSWS is

$$\sum_{r=1}^{\lceil \log_k(n) \rceil + 1} \#\{\text{partitions in round r}\} \cdot b_r = \sum_{r=1}^{\lceil \log_k(n) \rceil + 1} \left\lceil \frac{n}{k^r} \right\rceil \left\lfloor \frac{B}{\lceil \frac{n}{k^r} \rceil \cdot (\lceil \log_k(n) \rceil + 1)} \right\rfloor \leq B.$$

Thus, the stated correctness of CSWS follows directly from Theorem 4.1.

□

For sake of convenience, we provide the entire pseudo-code of CSR in Algorithm 4, which results by using $f(x) = 1$ as well as P_r^{CSR} and R^{CSR} as defined in Section 4.1 in Algorithm 1.

Proof of Corollary 4.3 (CSR case). Suppose $B > 0$ to be arbitrary but fixed. First, note that there are at most $\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \right\rceil$ rounds within the first while-loop and at most $k - 1$ in the second, so that we have at most $\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \right\rceil + k - 1$ many rounds in total. The total number of partitions in round

Algorithm 3 Combinatorial Successive Winner Stays (CSWS)

Input: set of arms $[n]$, subset size $k \leq n$, sampling budget B

Initialization: For each $r \in \{1, \ldots, \lceil \log_k(n) \rceil + 1\}$ let $b_r := \left\lfloor \frac{B}{\lceil \frac{n}{k^r} \rceil \cdot (\lceil \log_k(n) \rceil + 1)} \right\rfloor$, $\mathbb{A} \leftarrow [n]$, $r \leftarrow 1$

while $|\mathbb{A}_r| \geq k$ **do**
  $J = \lceil \frac{n}{k^r} \rceil$
  $\mathbb{A}_{r1}, \mathbb{A}_{r2}, \ldots, \mathbb{A}_{r,J} \leftarrow Partition(\mathbb{A}_r, k)$
  if $|\mathbb{A}_{r,J}| < k$ **then**
    $\mathcal{R} \leftarrow \mathbb{A}_{r,J}$, $J \leftarrow J - 1$
  else
    $\mathcal{R} \leftarrow \emptyset$
  end if
  $\mathbb{A}_{r+1} \leftarrow \emptyset$
  for $j \in [J]$ **do**
    Play the set $\mathbb{A}_{r,j}$ for b_r times
    For all $i \in \mathbb{A}_{r,j}$, update $s_{i|\mathbb{A}_{r,j}}(b_r)$
    Let $w \in \text{argmax}_i \; s_{i|\mathbb{A}_{r,j}}(b_r)$
    $\mathbb{A}_{r+1} \leftarrow \mathbb{A}_{r+1} \cup \{w\}$
  end for
  $\mathbb{A}_{r+1} \leftarrow \mathbb{A}_{r+1} \cup \mathcal{R}$
  $r \leftarrow r + 1$
end while
$\mathbb{A}_{r+1} \leftarrow \emptyset$
while $|\mathbb{A}_r| > 1$ **do**
  Play the set $\mathbb{A}_r$ for b_r times
  For all $i \in \mathbb{A}_r$, update $s_{i|\mathbb{A}_r}(b_r)$
  Let $w \in \text{argmax}_i \; s_{i|\mathbb{A}_r}(b_r)$
  $\mathbb{A}_{r+1} \leftarrow \mathbb{A}_{r+1} \cup \{w\}$
  $r \leftarrow r + 1$
end while

Output: The remaining item in $\mathbb{A}_r$

$r \in \{1, \ldots, \left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \right\rceil + k - 1\}$ is at most $\left\lceil \frac{n(1-\frac{1}{k})^{r-1}}{k} \right\rceil$. The budget allocated to a partition in round r (i.e., b_r) is by definition given by

$$b_r = \lfloor B/(R^{\text{CSR}} P_r^{\text{CSR}}) \rfloor = \left\lfloor \frac{B}{\left\lceil \frac{n(1-\frac{1}{k})^{r-1}}{k} \right\rceil \left(\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \right\rceil + k - 1 \right)} \right\rfloor .$$

Consequently, the total budget used by CSR is

$$\begin{aligned} &\sum_{r=1}^{\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \right\rceil + k - 1} \#\{\text{partitions in round r}\} \cdot b_r \\ &= \sum_{r=1}^{\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \right\rceil + k - 1} \left\lceil \frac{n(1-\frac{1}{k})^{r-1}}{k} \right\rceil \left\lfloor \frac{B}{\left\lceil \frac{n(1-\frac{1}{k})^{r-1}}{k} \right\rceil \left(\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \right\rceil + k - 1 \right)} \right\rfloor \\ &\leq B. \end{aligned}$$

Therefore, the statement follows from Theorem 4.1.

□

For sake of convenience, we provide the entire pseudo-code of CSH in Algorithm 5, which results by using $f(x) = \lceil x/2 \rceil$ as well as P_r^{CSH} and R^{CSH} as defined in Section 4.1 in Algorithm 1.

Proof of Corollary 4.3 (CSH case). Suppose $B > 0$ to be arbitrary but fixed. First, note that there are at most $\lceil \log_2(n) \rceil$ rounds within the first while-loop and at most $\lceil \log_2(k) \rceil$ in the second, so that

Algorithm 4 Combinatorial Successive Reject (CSR)

Input: set of arms $[n]$, subset size $k \leq n$, sampling budget B

Initialization: For each $r \in \{0, \ldots, \lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \rceil\}$ let $b_r := \left\lfloor \frac{B}{\left\lceil \frac{n(1-\frac{1}{k})^{r-1}}{k} \right\rceil \left(\left\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \right\rceil + k - 1 \right)} \right\rfloor$,

$\mathbb{A} \leftarrow [n]$, $r \leftarrow 1$

while $|\mathbb{A}_r| \geq k$ **do**
    $J = \lceil \frac{n(1-\frac{1}{k})^{r-1}}{k} \rceil$
    $\mathbb{A}_{r1}, \mathbb{A}_{r2}, \ldots, \mathbb{A}_{r,J} \leftarrow Partition(\mathbb{A}_r, k)$
    if $|\mathbb{A}_{r,J}| < k$ **then**
        $\mathcal{R} \leftarrow \mathbb{A}_{r,J}$, $J \leftarrow J - 1$
    else
        $\mathcal{R} \leftarrow \emptyset$
    end if
    $\mathbb{A}_{r+1} \leftarrow \mathbb{A}_r$
    for $j \in [J]$ **do**
        Play the set $\mathbb{A}_{r,j}$ for b_r times
        For all $i \in \mathbb{A}_{r,j}$, update $s_{i|\mathbb{A}_{r,j}}(b_r)$
        Let $w \in \arg\min_i \; s_{i|\mathbb{A}_{r,j}}(b_r)$
        $\mathbb{A}_{r+1} = \mathbb{A}_{r+1} \backslash \{w\}$
    end for
    $\mathbb{A}_{r+1} \leftarrow \mathbb{A}_{r+1} \cup \mathcal{R}$
    $r \leftarrow r + 1$
end while
$\mathbb{A}_{r+1} \leftarrow \mathbb{A}_r$
while $|\mathbb{A}_r| > 1$ **do**
    Play the set $\mathbb{A}_r$ for b_r times
    For all $i \in \mathbb{A}_r$, update $s_{i|\mathbb{A}_r}(b_r)$
    Let $w \in \arg\min_i \; s_{i|\mathbb{A}_r}(b_r)$
    $\mathbb{A}_{r+1} = \mathbb{A}_{r+1} \backslash \{w\}$
    $r \leftarrow r + 1$
end while

Output: The remaining item in $\mathbb{A}_r$

we have at most $\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil$ many rounds in total. The total number of partitions in round $r = 1, \ldots, \lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil$ is at most $\left\lceil \frac{n}{2^{r-1}k} \right\rceil$. The budget allocated to a partition in round r is

$$b_r = \lfloor B/(R^{\mathrm{CSH}} P_r^{\mathrm{CSH}}) \rfloor = \left\lfloor \frac{B}{\left\lceil \frac{n}{2^{r-1}k} \right\rceil \cdot (\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil)} \right\rfloor.$$

In particular, the total budget used by CSH is

$$\begin{aligned} &\sum_{r=1}^{\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil} \#\{\text{partitions in round r}\} \cdot b_r \\ &= \sum_{r=1}^{\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil} \left\lceil \frac{n}{2^{r-1}k} \right\rceil \cdot \left\lfloor \frac{Bk}{\lceil \frac{n}{2^{r-1}} \rceil (\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil)} \right\rfloor \\ &\leq B. \end{aligned}$$

Once again, Theorem 4.1 allows us to conclude the proof. □

E Proofs of Section 5

E.1 Stochastic Numerical Feedback: Proof of Corollary 5.1

A rich class of statistics can be obtained by applying a linear functional $U(F) = \int r(x) \mathrm{d}F(x)$, where F is a cumulative distribution function and $r : \mathbb{R} \to \mathbb{R}$ some measurable function, on the empirical distribution function [49], i.e., for any $x \in \mathbb{R}$ and any multiset of (reward) observations O

Algorithm 5 Combinatorial Successive Halving (CSH)

Input: set of arms $[n]$, subset size $k \leq n$, sampling budget B

Initialization: For each $r \in \{0, \dots, \lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil\}$ let $b_r := \left\lfloor \frac{Bk}{\lceil \frac{n}{2^{r-1}} \rceil (\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil)} \right\rfloor$,

$\mathbb{A} \leftarrow [n]$, $r \leftarrow 1$

while $|\mathbb{A}_r| \geq k$ **do**
  $J = \lceil \frac{n}{2^{r-1}k} \rceil$
  $\mathbb{A}_{r1}, \mathbb{A}_{r2}, \dots, \mathbb{A}_{r,J} \leftarrow Partition(\mathbb{A}_r, k)$
  if $|\mathbb{A}_{r,j}| < k$ **then**
    $\mathcal{R} \leftarrow \mathbb{A}_{r,j}$, $J \leftarrow J - 1$
  else
    $\mathcal{R} \leftarrow \emptyset$
  end if
  for $j \in [J]$ **do**
    Play the set $\mathbb{A}_{r,j}$ for b_r times
    For all $i \in \mathbb{A}_{r,j}$, update $s_{i|\mathbb{A}_{r,j}}(b_r)$
    Define $\bar{s} \leftarrow$ Median$(\{s_{i|\mathbb{A}_{r,j}}(b_r)\}_{i \in \mathbb{A}_{r,j}})$
    $\mathbb{A}_{r+1} \leftarrow \{i \in \mathbb{A}_{r,j} | s_{i|\mathbb{A}_{r,j}}(b_r) \leq \bar{s}\}$
  end for
  $\mathbb{A}_{r+1} \leftarrow \mathbb{A}_{r+1} \cup \mathcal{R}$
  $r \leftarrow r + 1$
end while
$\mathbb{A}_r \leftarrow \mathbb{A}_r \cup \{k - |\mathbb{A}_r|$ random elements from $[n] \backslash \mathbb{A}_r\}$
while $|\mathbb{A}_r| > 1$ **do**
  Play the set $\mathbb{A}_r$ for b_r times
  For all $i \in \mathbb{A}_r$, update $s_{i|\mathbb{A}_r}(b_r)$
  Define $\bar{s} \leftarrow$ Median$(\{s_{i|\mathbb{A}_r}(b_r)\}_{i \in \mathbb{A}_r})$
  $\mathbb{A}_{r+1} \leftarrow \{i \in \mathbb{A}_r | s_{i|\mathbb{A}_r}(b_r) \leq \bar{s}\}$
  $r \leftarrow r + 1$
end while

Output: The remaining item in $\mathbb{A}_r$

$$\tilde{s}(O, x) = \tfrac{1}{|O|} \textstyle\sum_{o \in O} \mathbb{1}\{x \leq o\}.$$

This leads to the statistics

$$s_{i|Q}(t) = U(\tilde{s}(o_{i|Q}(1), \dots, o_{i|Q}(t), \cdot)) = \sum_{s=1}^{t} \frac{r(o_{i|Q}(s))}{t},$$

which converge to $S_{i|Q} = \mathbb{E}_{X \sim \nu_{i|Q}}[r(X)]$ by the law of large numbers, provided these expected values exist. In this section we show the following result which generalizes Corollary 5.1 for statistics of the above kind.

Corollary E.1. *Let f, R and $\{P_r\}_{r \in [R]}$ be as in Theorem 4.1 and suppose that $r(o_{i|Q}(t))$ are σ-sub-Gaussian and such that their means $S_{i|Q} := \mathbb{E}_{X \sim \nu_{i|Q}}[r(X)]$ satisfy (A2). Then, there is a function*

$$C(\delta, \varepsilon, k, R, \sigma) \in \mathcal{O}\left(\sigma^2 \varepsilon^{-2} \ln\left(\nicefrac{kR}{\delta} \ln\left(\nicefrac{kR\sigma}{\varepsilon\delta}\right)\right)\right)$$

with the following property: If i^* is the GCW and $\sup_{Q \in \mathcal{Q}_{\leq k}(i^*)} \Delta_{(f(|Q|)+1)|Q} \leq \varepsilon$, then Algorithm 1 used with a budget B larger than $C(\delta, \varepsilon, k, R, \sigma) \cdot R \max_{r \in [R]} P_r$ returns i^* with probability at least $1 - \delta$.

Note that we immediately obtain the proof for Corollary 5.1 as a special case of Corollary E.1 by using the the identity function $r(x) = x$.

The following two lemmata serve as a preparation for the proof of Corollary E.1. The proof of Lemma E.2 is an adaptation of the proof of Lemma 3 in [24].

Lemma E.2. *Let $X_1, X_2, \dots \sim \mathcal{X}$ be iid real-valued random variables and $r : \mathbb{R} \to \mathbb{R}$ such that $r(\mathcal{X})$ is σ^2-sub-Gaussian. For any $\epsilon \in (0, 1)$ and $\delta \in (0, \log(1 + \epsilon)/e)$ one has with probability at* *least* $1 - \frac{(2+\epsilon)}{\epsilon}\left(\frac{\delta}{\log(1+\epsilon)}\right)^{(1+\epsilon)}$ *for any* $t \geq 1$

$$\sum_{i=1}^{t} r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)] \leq (1+\sqrt{\epsilon})\sqrt{2\sigma^2(1+\epsilon)t\log\left(\frac{\log((1+\epsilon)t)}{\delta}\right)}.$$

Moreover, the same concentration inequality holds for $-\left(\sum_{i=1}^{t} r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)]\right)$ *as well.*

Proof. We denote in the following $\psi(x) = \sqrt{2\sigma^2 x \log\left(\frac{\log(x)}{\delta}\right)}$ and $R_t = \sum_{i=1}^{t} r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)]$ and define a sequence of integers (u_k) as $u_0 = 1$ and $u_{k+1} = \lceil (1+\epsilon)u_k \rceil$. The maximal *Azuma-Hoeffding Inequality* states that for any martingale difference sequence $S_1, S_2, \ldots$ with each element being σ^2-sub-Gaussian, it holds that for any $\alpha > 0$, $n \geq 1$:

$$\mathbb{P}\left(\max_{i\in[n]} S_i - S_0 \geq \alpha\right) \leq \exp\left(-\frac{\alpha^2}{2\sum_{j=1}^{n}\sigma_j^2}\right).$$

In the following let $\mathcal{F}_0 = \{\emptyset, \Omega\}$ be the trivial σ-algebra and for $k \in \{1, \ldots, n\}$ let $\mathcal{F}_k = \sigma(X_1, \ldots, X_k)$ be the σ-algebra generated by the observations $X_1, \ldots, X_k$. Then

$$\begin{aligned}\mathbb{E}[R_{t+1}|\mathcal{F}_t] &= \mathbb{E}[r(X_{t+1}) - \mathbb{E}_{X\sim\mathcal{X}}[r(X)] + R_t|\mathcal{F}_t] \\ &= \mathbb{E}[r(X_{t+1})|\mathcal{F}_t] - \mathbb{E}_{X\sim\mathcal{X}}[r(X)] + \mathbb{E}[R_t|\mathcal{F}_t] \\ &= R_t\end{aligned}$$

which shows the martingale property of R_t. Note, that $R_0 = 0$ and $R_{t+1} - R_t = r(X_{t+1}) - \mathbb{E}_{X\sim\mathcal{X}}[r(X)]$, which is according to the assumption σ^2-sub-Gaussian and has zero mean, for any $t \in \mathbb{N}$. Thus, we can apply the maximal Azuma-Hoeffding inequality for $R_1, R_2, \ldots, R_t$.

Step 1.
In the first step of the proof we derive a bound for the probability of a lower bound of R_{u_k} for $k \geq 1$. For this we use the union bound, the maximal Azuma-Hoeffding inequality, the fact that $u_k \geq (1+\epsilon)^k$, a sum-integral comparison and some simple transformations and obtain

$$\begin{aligned}&\mathbb{P}\left(\exists k \geq 1 \,:\, R_{u_k} \geq \sqrt{1+\epsilon}\psi(u_k)\right) \\ &\leq \sum_{k=1}^{\infty} \mathbb{P}\left(R_{u_k} \geq \sqrt{1+\epsilon}\psi(u_k)\right) \\ &\leq \sum_{k=1}^{\infty} \exp\left(-\frac{(1+\epsilon)\psi(u_k)^2}{2u_k\sigma^2}\right) \\ &= \sum_{k=1}^{\infty} \exp\left(-(1+\epsilon)\log\left(\frac{\log(u_k)}{\delta}\right)\right) \\ &\leq \sum_{k=1}^{\infty} \exp\left(-(1+\epsilon)\log\left(\frac{\log((1+\epsilon)^k)}{\delta}\right)\right) \\ &= \sum_{k=1}^{\infty}\left(\frac{\delta}{k\log((1+\epsilon))}\right)^{(1+\epsilon)} \\ &= \left(\frac{2\delta}{\log((1+\epsilon))}\right)^{(1+\epsilon)} \sum_{k=1}^{\infty}\left(\frac{1}{k}\right)^{(1+\epsilon)} \\ &= \left(\frac{\delta}{\log((1+\epsilon))}\right)^{(1+\epsilon)}\left(1 + \sum_{k=2}^{\infty}\left(\frac{1}{k}\right)^{(1+\epsilon)}\right) \\ &\leq \left(\frac{\delta}{\log((1+\epsilon))}\right)^{(1+\epsilon)}\left(1 + \int_{k=1}^{\infty}\left(\frac{1}{k}\right)^{(1+\epsilon)}\right)\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{\delta}{\log((1+\epsilon))}\right)^{(1+\epsilon)} \left(1 + \left[-\frac{1}{\epsilon}\left(\frac{1}{k}\right)^{\epsilon}\right]_1^{\infty}\right) \\
&= \left(\frac{\delta}{\log((1+\epsilon))}\right)^{(1+\epsilon)} \left(1 + \frac{1}{\epsilon}\right).
\end{aligned}
$$

Step 2.
Next, we bound the probability that the difference between some R_s and R_t exceeds a lower bound for some $s = u_k$, $k \in \mathbb{N}$ and $s \leq t \leq u_{k+1}$. Note that $R_t - R_{u_k}$ and R_{t-u_k} have the same distribution, such that we obtain

$$
\begin{aligned}
&\mathbb{P}\left(\exists t \in \{u_k+1, \dots, u_{k+1}-1\} \;:\; R_t - R_{u_k} \geq \sqrt{\epsilon}\psi(u_{k+1})\right) \\
&\quad = \mathbb{P}\left(\exists t \in [u_{k+1} - u_k - 1] \;:\; R_t \geq \sqrt{\epsilon}\psi(u_{k+1})\right) \\
&\quad \leq \exp\left(-\frac{\epsilon\psi(u_{k+1})^2}{2\sigma^2(u_{k+1} - u_k - 1)}\right) \\
&\quad = \exp\left(-\frac{\epsilon u_{k+1}}{u_{k+1} - u_k - 1}\log\left(\frac{\log(u_{k+1})}{\delta}\right)\right) \\
&\quad \leq \exp\left(-\frac{\epsilon u_{k+1}}{(1+\epsilon)u_k + 1 - u_k - 1}\log\left(\frac{\log(u_{k+1})}{\delta}\right)\right) \\
&\quad = \exp\left(-\frac{u_{k+1}}{u_k}\log\left(\frac{\log(u_{k+1})}{\delta}\right)\right) \\
&\quad \leq \exp\left(-(1+\epsilon)\log\left(\frac{\log(u_{k+1})}{\delta}\right)\right) \\
&\quad \leq \left(\frac{\delta}{(k+1)\log(1+\epsilon)}\right)^{1+\epsilon},
\end{aligned}
$$

where we used once again the maximal Azuma-Hoeffding inequality and that $u_{k+1} \geq (1+\epsilon)u_k$ as well as that $\frac{u_{k+1}}{u_k} \geq 1 + \epsilon$. For all possible $k \in \mathbb{N}$ we get with the union bound and a similar sum-integral comparison as above

$$
\begin{aligned}
&\mathbb{P}\left(\exists k \in \mathbb{N},\, \exists t \in \{u_k+1, \dots, u_{k+1}-1\} \;:\; R_t - R_{u_k} \geq \sqrt{\epsilon}\psi(u_{k+1})\right) \\
&\leq \sum_{k=1}^{\infty}\left(\frac{\delta}{(k+1)\log(1+\epsilon)}\right)^{1+\epsilon} \\
&= \sum_{k=2}^{\infty}\left(\frac{\delta}{k\log(1+\epsilon)}\right)^{1+\epsilon} \\
&\leq \int_{k=1}^{\infty}\left(\frac{\delta}{k\log(1+\epsilon)}\right)^{1+\epsilon} \\
&= \left(\frac{\delta}{\log(1+\epsilon)}\right)^{1+\epsilon}\frac{1}{\epsilon}
\end{aligned}
$$

Step 3.
Finally, by combining Step 1 and 2 we can infer that for any $k \geq 0$ and $t \in \{u_k+1, \dots, u_{k+1}-1\}$ it holds

$$
\begin{aligned}
R_t &= R_t - R_{u_k} + R_{u_k} \\
&\leq \sqrt{\epsilon}\psi(u_{k+1}) + \sqrt{1+\epsilon}\psi(u_k) \\
&\leq \sqrt{\epsilon}\psi((1+\epsilon)t) + \sqrt{1+\epsilon}\psi(t) \\
&\leq (1+\sqrt{\epsilon})\psi((1+\epsilon)t),
\end{aligned}
$$

with probability at least $1 - \frac{2+\epsilon}{\epsilon}\left(\frac{\delta}{\log(1+\epsilon)}\right)^{1+\epsilon}$ leading to the first claim of the lemma.

Step 4.

Note that $\tilde{R}_t = t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)] - \sum_{i=1}^t r(X_i)$ is a martingale difference sequence with $\tilde{R}_{t+1} - \tilde{R}_t = -R_t + R_{t+1} = \mathbb{E}_{X\sim\mathcal{X}}[r(X)] - r(X_{t+1})$, which is according to the assumption σ^2-sub-Gaussian and has zero mean, for any $t \in \mathbb{N}$. Thus, repeating Step 1–3 for $\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_t$ shows the second claim of the lemma. □

Lemma E.3. *Let $X_1, X_2, \dots \sim \mathcal{X}$ be iid real-valued random variables and $r : \mathbb{R} \to \mathbb{R}$ such that $r(\mathcal{X})$ is σ^2-sub-Gaussian. For any $\gamma \in (0,1)$ we have*

$$\mathbb{P}\left(\exists t \in \mathbb{N} : \left|\sum_{i=1}^t r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)]\right| > (1+\sqrt{1/2})\sqrt{3\sigma^2 t \ln\left(\frac{10^{2/3}\ln(3t/2)}{\gamma^{2/3}\ln(3/2)}\right)}\right) \le \gamma.$$

Proof. Let $\gamma \in (0,1)$ be fixed and $\varepsilon := 1/2$. Then, $\gamma' := \left(\frac{\gamma}{10}\right)^{2/3}\ln(3/2)$ fulfills

$$\frac{2+\varepsilon}{\varepsilon}\left(\frac{\gamma'}{\ln(1+\varepsilon)}\right)^{1+\varepsilon} = 5\left((\gamma/10)^{2/3}\right)^{3/2} = \gamma/2$$

and moreover $\gamma' < (1/10)^{2/3}\ln(3/2) < e^{-1}\ln(3/2)$. Consequently, Lemma E.2 yields with

$$\tilde{c}_\gamma(t) := (1+\sqrt{\varepsilon})\sqrt{2\sigma^2(1+\varepsilon)t\ln\left(\frac{\ln((1+\varepsilon)t)}{\gamma'}\right)} = (1+\sqrt{1/2})\sqrt{3\sigma^2 t\ln\left(\frac{10^{2/3}\ln(3t/2)}{\gamma^{2/3}\ln(3/2)}\right)}$$

that

$$\mathbb{P}\left(\exists t \in \mathbb{N} : \sum_{i=1}^t r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)] > \tilde{c}_\gamma(t)\right) \le \gamma/2.$$

as well as

$$\mathbb{P}\left(\exists t \in \mathbb{N} : -\left(\sum_{i=1}^t r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)]\right) > \tilde{c}_\gamma(t)\right) \le \gamma/2.$$

Thus, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\exists t \in \mathbb{N} : \left|\sum_{i=1}^t r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)]\right| > \tilde{c}_\gamma(t)\right) \\
&\quad\le \mathbb{P}\left(\exists t \in \mathbb{N} : \sum_{i=1}^t r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)] > \tilde{c}_\gamma(t)\right) \\
&\quad\quad + \mathbb{P}\left(\exists t \in \mathbb{N} : -\left(\sum_{i=1}^t r(X_i) - t \cdot \mathbb{E}_{X\sim\mathcal{X}}[r(X)]\right) > \tilde{c}_\gamma(t)\right) \\
&\quad\le \gamma/2 + \gamma/2 = \gamma.
\end{aligned}$$

□

We are now ready to prove Corollary E.1.

Proof of Corollary E.1. Recall the definition of $\tilde{c}_\gamma(t)$ from the proof of Lemma E.3 and let

$$c_\gamma(t) := \frac{2}{t}\tilde{c}_\gamma(t) = 2(1+\sqrt{1/2})\sqrt{\frac{3\sigma^2}{t}\ln\left(\frac{10^{2/3}\ln(3t/2)}{\gamma^{2/3}\ln(3/2)}\right)}$$

for any $\gamma \in (0,1)$, $t \in \mathbb{N}$. For any fixed γ, $c_\gamma : \mathbb{N} \to (0,\infty), t \mapsto c_\gamma(t)$ is strictly monotonically decreasing with $\lim_{t\to\infty} c_\gamma(t) = 0$. Contraposition of (1) in [24] states

$$t > \frac{1}{c}\ln\left(\frac{2\ln((1+\varepsilon)/(c\omega))}{\omega}\right) \Rightarrow c > \frac{1}{t}\ln\left(\frac{\ln((1+\varepsilon)t)}{\omega}\right) \qquad \forall t \ge 1, \varepsilon \in (0,1), c > 0, \omega \le 1.$$

For any $\alpha > 0$ and $\gamma \in (0,1)$, using this with $\omega = \frac{\gamma^{2/3}\ln(3/2)}{10^{2/3}}, c = \frac{\alpha^2}{12(1+\sqrt{1/2})^2\sigma^2}$ and $\varepsilon = 1/2$ reveals

$$\begin{aligned} c_\gamma^{-1}(\alpha) &= \min\{t \in \mathbb{N} : c_\gamma(t) \le \alpha\} \\ &= \min\left\{t \in \mathbb{N} : \frac{1}{t}\ln\left(\frac{10^{2/3}\ln(3t/2)}{\gamma^{2/3}\ln(3/2)}\right) \le \frac{\alpha^2}{12(1+\sqrt{1/2})^2\sigma^2}\right\}. \end{aligned}$$

Thus, we have $c \ge \frac{1}{t}\ln\left(\frac{\ln((1+\varepsilon)t)}{\omega}\right)$ and we know, that this statement is true if $t \ge \frac{1}{c}\ln\left(\frac{2\ln((1+\varepsilon)/(c\omega))}{\omega}\right)$. In particular also for the smallest such t, for which holds $t \le \left\lceil \frac{1}{c}\ln\left(\frac{2\ln((1+\varepsilon)/(c\omega))}{\omega}\right)\right\rceil + 1$. It follows

$$c_\gamma^{-1}(\alpha) \le \left\lceil \frac{12(1+\sqrt{1/2})^2\sigma^2}{\alpha^2}\ln\left(\frac{2\cdot 10^{2/3}}{\gamma^{2/3}\ln(3/2)}\ln\left(\frac{18\cdot 10^{2/3}(1+\sqrt{1/2})^2\sigma^2}{\gamma^{2/3}\ln(3/2)\,\alpha^2}\right)\right)\right\rceil + 1,$$

which is of the order $\mathcal{O}(\sigma^2\alpha^{-2}\ln\ln(\alpha^{-1}\sigma)\ln\gamma^{-1})$.

Now, suppose $\max_{Q\in\mathcal{Q}_{\le k}(i^*)} \Delta_{(f(|Q|)+1)|Q} \le \varepsilon$ and that Algorithm 1 is started with a budget B larger than

$$c_{\delta/(kR)}^{-1}(\varepsilon/2)\cdot R\max\nolimits_{r\in[R]} P_r.$$

Recall that $\gamma_{i|Q}(t) = |s_{i|Q}(t) - S_{i|Q}|$, $s_{i|Q}(t) = \frac{1}{t}\sum_{s=1}^t r(o_{i|Q}(s))$ and $S_{i|Q} = \mathbb{E}_{X\sim\nu_{i|Q}}[r(X)]$. With this, we obtain for any possible sequence of partitions $(E_r)_{r\in[R]} \in (\mathcal{Q}_{\le k})^R$ with $\mathbb{P}(\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]) > 0$ that

$$\begin{aligned} &\mathbb{P}\left(\exists t\in\mathbb{N}, r\in[R], i\in E_r : \gamma_{i|E_r}(t) \ge c_{\delta/(kR)}(t)\middle|\,\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]\right) \\ &\le \sum_{r\in[R]}\sum_{i\in E_r}\mathbb{P}\left(\exists t\in\mathbb{N} : \gamma_{i|E_r}(t) \ge c_{\delta/(kR)}(t)\middle|\,\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]\right) \\ &= \sum_{r\in[R]}\sum_{i\in E_r}\mathbb{P}\left(\exists t\in\mathbb{N} : \Big|\frac{1}{t}\sum_{t'=1}^t r(o_{i|E_r}(t')) - \mathbb{E}_{X\sim\nu_{i|E_r}}[r(X)]\Big| \ge c_{\delta/(kR)}(t)\,\middle|\,\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]\right) \\ &= \sum_{r\in[R]}\sum_{i\in E_r}\mathbb{P}\left(\exists t\in\mathbb{N} : \frac{1}{t}\Big|\sum_{t'=1}^t r(o_{i|E_r}(t')) - t\cdot\mathbb{E}_{X\sim\nu_{i|E_r}}[r(X)]\Big| \ge c_{\delta/(kR)}(t)\,\middle|\,\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]\right) \\ &= \sum_{r\in[R]}\sum_{i\in E_r}\mathbb{P}\left(\exists t\in\mathbb{N} : \Big|\sum_{t'=1}^t r(o_{i|E_r}(t')) - t\cdot\mathbb{E}_{X\sim\nu_{i|E_r}}[r(X)]\Big| \ge \tilde{c}_{\delta/(kR)}(t)\,\middle|\,\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]\right) \\ &\le \sum_{r\in[R]}\sum_{i\in E_r}\frac{\delta}{kR} \le \delta, \end{aligned}$$

where we used Lemma E.3 for the second last inequality. Using the law of total probability for all possible sequences of partitions $(E_r)_{r\in[R]}$, we see that the event

$$\mathcal{E} := \left\{\exists t\in\mathbb{N}, r\in[R], i\in\mathbb{A}_r(i^*) : \gamma_{i|\mathbb{A}_r(i^*)}(t) \ge c_{\delta/(kR)}(t)\right\}$$

occurs with probability

$$\begin{aligned} &\mathbb{P}(\mathcal{E}) \\ &= \sum_{(E_r)_{r\in[R]}}\mathbb{P}\left(\exists t\in\mathbb{N}, r\in[R], i\in E_r : \gamma_{i|E_r}(t) \ge c_{\delta/(kR)}(t)\,\middle|\,\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]\right) \\ &\quad\times\ \mathbb{P}(\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]) \\ &\le \delta\sum\nolimits_{(E_r)_{r\in[R]}}\mathbb{P}(\mathbb{A}_r(i^*) = E_r\,\forall r\in[R]) = \delta. \end{aligned}$$

On $\mathcal{E}^c$ we have $\overline{\gamma}_{\mathbb{A}_r(i^*)}(t) < c_{\delta/(kR)}(t)$ for all $t\in\mathbb{N}, r\in[R]$ and thus in particular $\overline{\gamma}_{\mathbb{A}_r(i^*)}^{-1}(\alpha) \ge c_{\delta/(kR)}^{-1}(\alpha)$ for any $\alpha\in(0,\infty)$. Since $\max_{Q\in\mathcal{Q}_{\le k}(i^*)}\Delta_{(f(|Q|)+1)|Q} \le \varepsilon$, Theorem 4.1 thus lets us conclude

$$
\begin{aligned}
\mathbb{P}\left(\text{Alg. 1 returns } i^*\right) &\geq \mathbb{P}\left(B > R\max_{r\in[R]} P_r \overline{\gamma}^{-1}_{\mathbb{A}_r(i^*)}\left(\frac{\Delta_{(f(|\mathbb{A}_r(i^*)|)+1)|\mathbb{A}_r(i^*)}}{2}\right)\right)\\
&\geq \mathbb{P}\left(\left\{B > R\max_{r\in[R]} P_r \overline{\gamma}^{-1}_{\mathbb{A}_r(i^*)}\left(\varepsilon/2\right)\right\}\cap \mathcal{E}^c\right)\\
&\geq \mathbb{P}\left(\left\{B > R\max_{r\in[R]} P_r c^{-1}_{\delta/(kR)}\left(\varepsilon/2\right)\right\}\cap \mathcal{E}^c\right)\\
&= \mathbb{P}(\mathcal{E}^c) \geq 1-\delta,
\end{aligned}
$$

where the equality holds due to the assumption on B. Consequently, we can conclude the proof by defining

$$
\begin{aligned}
C(\delta,\varepsilon,k,R) &\coloneqq c^{-1}_{\delta/(kR)}(\varepsilon/2)\\
&\leq \left\lceil \frac{48(1+\sqrt{1/2})^2\sigma^2}{\varepsilon^2}\ln\left(\frac{2(10kR)^{2/3}}{\delta^{2/3}\ln(3/2)}\ln\left(\frac{72\cdot(10kR)^{2/3}(1+\sqrt{1/2})^2\sigma^2}{\delta^{2/3}\varepsilon^2\ln(3/2)}\right)\right)\right\rceil + 1\\
&\in \mathcal{O}\left(\frac{\sigma^2}{\varepsilon^2}\ln\left(\frac{kR}{\delta}\ln\left(\frac{kR\sigma}{\varepsilon\delta}\right)\right)\right).
\end{aligned}
$$

□

E.2 Stochastic Preference Winner Feedback: Proof of Corollary 5.2

The following two lemmata serve as a preparation for the proof of Corollary 5.2. But first let us introduce the $(k-1)$-simplex

$$
\mathcal{S}_k = \left\{(p_i)_{i\in[k]} \in [0,1]^k \ : \ \sum\nolimits_{i=1}^k p_i = 1 \wedge \forall i : p_i \geq 0\right\}.
$$

Lemma E.4 (Dvoretzky-Kiefer-Wolfowitz inequality for categorical random variables)**.** *Let $\{X_t\}_{t\in\mathbb{N}}$ be a sequence of iid random variables $X_t \sim \mathrm{Cat}(\mathbf{p})$ for some $\mathbf{p} \in \mathcal{S}_k$. For $t \in \mathbb{N}$ let $\hat{\mathbf{p}}^t$ be the corresponding empirical distribution after the t observations $X_1,\ldots,X_t$, i.e., $\hat{p}^t_i = \frac{1}{t}\sum_{s=1}^t \mathbf{1}_{\{X_s=i\}}$ for all $i\in[k]$. Then, we have for any $\varepsilon > 0$ and $t\in\mathbb{N}$ the estimate*

$$
\mathbb{P}\left(\left|\left|\hat{\mathbf{p}}^t - \mathbf{p}\right|\right|_\infty > \varepsilon\right) \leq 4e^{-t\varepsilon^2/2}.
$$

Proof. Confer [19, 35] as well as Theorem 11.6 in [29]. Moreover, note that the cumulative distribution functions F resp. $\hat{F}^t$ of $X_1 \sim \mathrm{Cat}(\mathbf{p})$ resp. $\hat{\mathbf{p}}^t$ fulfill $p_j = F(j) - F(j-1)$ and $\hat{p}^t_j = \hat{F}^t(j) - \hat{F}^t(j-1)$ and thus

$$
|\hat{p}^t_j - p_j| \leq |\hat{F}^t(j) - F(j)| + |\hat{F}^t(j-1) - F(j-1)|.
$$

for each $j \in [k]$. □

Lemma E.5. *For every $\beta \in [1, e/2]$, $c_1, c_2 > 0$ the number*

$$
x \coloneqq \frac{\beta}{c_1}\left(\ln\left(\frac{c_2 e}{c_1^\beta}\right) + \ln\ln\left(\frac{c_2}{c_1^\beta}\right)\right)
$$

fulfills $c_1 x \geq \ln(c_2 x^\beta)$.

Proof. This is Lemma 18 in [20]. □

Proof of Corollary 5.2. For $t\in\mathbb{N}$ and $\gamma\in(0,1)$ define

$$
c_\gamma(t) \coloneqq \sqrt{\frac{4\ln(2\pi^2t^2/(3\gamma))}{t}}
$$

and note that, for any fixed γ, the function $c_\gamma : \mathbb{N} \to (0,\infty), t \mapsto c_\gamma(t)$ is strictly monotonically decreasing with $\lim_{t\to\infty} c_\gamma(t) = 0$. For any $\alpha > 0$, $\gamma \in (0,1)$, we obtain via Lemma E.5 with the choices $\beta = 1$, $c_1 = \frac{\alpha^2}{8}$ and $c_2 = \sqrt{2/(3\gamma)}\pi$ the estimate

$$\begin{aligned} c_\gamma^{-1}(\alpha) &= \min\left\{t \in \mathbb{N} : 4\ln(2\pi^2 t^2/(3\gamma)) \le t\alpha^2\right\} \\ &= \min\left\{t \in \mathbb{N} : \ln\left(\sqrt{2/(3\gamma)}\pi t\right) \le \frac{\alpha^2}{8} t\right\} \\ &\le \left\lceil \frac{8}{\alpha^2}\left(\ln\left(\frac{8\sqrt{2/(3\gamma)}\pi e}{\alpha^2}\right) + \ln\ln\left(\frac{8\sqrt{2/(3\gamma)}\pi}{\alpha^2}\right)\right)\right\rceil + 1. \end{aligned}$$

Now, suppose $\max_{Q\in\mathcal{Q}_{\le k}(i^*)} \Delta_{(f(|Q|)+1)|Q} \le \varepsilon$ and that Algorithm 1 is started with a budget B larger than

$$c_{\delta/R}^{-1}(\varepsilon/2) \cdot R \max_{r\in[R]} P_r.$$

Recall that in this preference-based setting we use as the statistic the empirical mean of the (winner) observations we obtained for arm i after querying Q (with $i \in Q$) for t many times. In particular, we set

$$s_{i|Q}(t) = \frac{w_{i|Q}(t)}{t} = \frac{1}{t}\sum_{t'=1}^{t} o_{i|Q}(t'),$$

where $o_{i|Q}(t') = 1$ if arm i is the preferred (or winning) arm among the arms in Q, if Q is queried for the t'-th time, and 0 otherwise. Thus, $w_{i|Q}(t)$ is the total number of times arm i has won in the query set Q after t queries. Moreover, $\gamma_{i|Q}(t) = |s_{i|Q}(t) - S_{i|Q}|$, where $S_{i|Q} = p_{i|Q}$ and $o_{i|Q}(t') \sim \mathrm{Cat}(\mathbf{p_Q})$. With this, we obtain for any $t \in \mathbb{N}$ and any possible sequence of partitions $(E_r)_{r\in[R]} \in (\mathcal{Q}_{\le k})^R$ with $\mathbb{P}(\mathbb{A}_r(i^*) = E_r\, \forall r \in [R]) > 0$ that

$$\begin{aligned} &\mathbb{P}\left(\exists r \in [R] : \gamma_{i|E_r}(t) \ge c_{\delta/R}(t) \,\middle|\, \mathbb{A}_r(i^*) = E_r\, \forall r \in [R]\right) \\ &\le \sum_{r\in[R]} \mathbb{P}\left(\gamma_{i|E_r}(t) \ge c_{\delta/R}(t) \,\middle|\, \mathbb{A}_r(i^*) = E_r\, \forall r \in [R]\right) \\ &= \sum_{r\in[R]} \mathbb{P}\left(\max_{i\in E_r} \Big|\frac{1}{t}\sum_{t'=1}^{t} o_{i|E_r}(t') - S_{i|E_r}\Big| > \sqrt{\frac{4\ln(2\pi^2t^2/(3\gamma))}{t}} \,\middle|\, \mathbb{A}_r(i^*) = E_r\, \forall r \in [R]\right) \\ &\le \frac{6\delta}{\pi^2 t^2}, \end{aligned}$$

where we used Lemma E.4 in the last inequality. Using the law of total probability for all possible sequences of partitions $(E_r)_{r\in[R]}$, we see that the event

$$\mathcal{E} := \left\{\exists t \in \mathbb{N}, r \in [R] : \overline{\gamma}_{\mathbb{A}_r(i^*)}(t) \ge c_{\delta/R}(t)\right\}$$

occurs with probability

$$\begin{aligned} \mathbb{P}(\mathcal{E}) \le &\sum_{t\in\mathbb{N}} \sum_{(E_r)_{r\in[R]}} \mathbb{P}\left(\exists r \in [R] : \overline{\gamma}_{E_r}(t) \ge c_{\delta/R}(t) \,\middle|\, \mathbb{A}_r(i^*) = E_r\, \forall r \in [R]\right) \\ &\times\ \mathbb{P}(\mathbb{A}_r(i^*) = E_r\, \forall r \in [R]) \\ \le &\sum_{t\in\mathbb{N}} \frac{6\delta}{\pi^2t^2} \sum_{(E_r)_{r\in[R]}} \mathbb{P}(\mathbb{A}_r(i^*) = E_r\, \forall r \in [R]) \le \delta. \end{aligned}$$

On $\mathcal{E}^c$ we have $\overline{\gamma}_{\mathbb{A}_r(i^*)}(t) < c_{\delta/R}(t)$ for all $t \in \mathbb{N}, r \in [R]$ and thus in particular $\overline{\gamma}^{-1}_{\mathbb{A}_r(i^*)}(\alpha) \ge c^{-1}_{\delta/R}(\alpha)$ for any $\alpha \in (0,\infty)$. Since $\max_{Q\in\mathcal{Q}_{\le k}(i^*)} \Delta_{(f(|Q|)+1)|Q} \le \varepsilon$, Theorem 4.1 thus lets us conclude

$$\begin{aligned} \mathbb{P}\left(\text{Alg. 1 returns } i^*\right) &\ge \mathbb{P}\left(B > R\max_{r\in[R]} P_r \overline{\gamma}^{-1}_{\mathbb{A}_r(i^*)}\left(\frac{\Delta_{(f(|\mathbb{A}_r(i^*)|)+1)|\mathbb{A}_r(i^*)}}{2}\right)\right) \\ &\ge \mathbb{P}\left(\left\{B > R\max_{r\in[R]} P_r \overline{\gamma}^{-1}_{\mathbb{A}_r(i^*)}(\varepsilon/2)\right\} \cap \mathcal{E}^c\right) \\ &\ge \mathbb{P}\left(\left\{B > R\max_{r\in[R]} P_r c^{-1}_{\delta/R}(\varepsilon/2)\right\} \cap \mathcal{E}^c\right) \\ &= \mathbb{P}(\mathcal{E}^c) \ge 1-\delta, \end{aligned}$$

where the equality holds due to the assumption on B. Consequently, the statement holds with

$$
\begin{aligned}
C(\delta, \varepsilon, k, R) &\coloneqq c_{\delta/R}^{-1}(\varepsilon/2) \\
&\leq \left\lceil \frac{32}{\varepsilon^2} \left(\ln \left(\frac{32\sqrt{2R/(3\delta)}\pi e}{\varepsilon^2} \right) + \ln \ln \left(\frac{32\sqrt{2R/(3\delta)}\pi}{\varepsilon^2} \right) \right) \right\rceil + 1 \\
&\in \mathcal{O}\left(\frac{1}{\varepsilon^2} \ln \left(\frac{R}{\delta \varepsilon^4} \right) \right).
\end{aligned}
$$

□

F Comparisons of the Algorithms

In the following we summarize the theoretical results obtained for our proposed algorithms in a concise way. First of all, we give an overview of the individual key quantities of each algorithm in Table 2, where we assume w.l.o.g. that $\binom{n}{k}$ is a divisor of B in ROUNDROBIN to make the assignments of R, P_r and $f(s)$ for ROUNDROBIN well-defined. The maximal number of different query sets is derived in Section F.1.

Table 2: Comparison of the maximal number of rounds, the maximal number of partitions per round, the amount of retained arms from each partition and the maximal number of query sets for ROUNDROBIN and our proposed algorithms CSWS, CSR and CSH.

Alg.	R	P_r	$f(x)$	max #query_sets
ROUNDROBIN	1	$\binom{n}{k}$	$x \mapsto x$	$\binom{n}{k}$
CSWS	$\lceil \log_k(n) \rceil + 1$	$\lceil \frac{n}{k^r} \rceil$	$x \mapsto 1$	$R^{CSWS} + n \cdot \left(\frac{1-1/k^{\lceil \log_k(n) \rceil + 1}}{k-1} \right)$
CSR	$\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n} \right) \right\rceil + k - 1$	$\left\lceil \frac{n(1-\frac{1}{k})^{(r-1)}}{k} \right\rceil$	$x \mapsto x - 1$	$R^{CSR} + n \left(1 - \left(1 - \frac{1}{k}\right)^{\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n}\right) \rceil + k - 1} \right)$
CSH	$\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil$	$\lceil \frac{n}{2^{r-1}k} \rceil$	$x \mapsto \lceil \frac{x}{2} \rceil$	$R^{CSH} + \frac{2n}{k} \left(1 - 1/2^{\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil} \right)$

Using Remark D.1 we can derive the following sufficient budgets of the algorithms summarized in the following table, where $\pi(Q) \in Q$ be the $\lfloor \frac{|Q|}{2} \rfloor + 1$- th best arm with respect to $(S_{i|Q})_{i \in Q}$.

Table 3: Comparison of the sufficient budget for ROUNDROBIN and our proposed algorithms CSWS, CSR and CSH.

Algorithm	Sufficient budget
ROUNDROBIN	$\binom{n}{k} \max_{i \in \mathcal{A}, i \neq i^*_{\mathcal{B}}} \left(\hat{\gamma}^{\max}_{i^*_{\mathcal{B}}, i} \right)^{-1} \left(\frac{S^{\mathcal{B}}_{i^*_{\mathcal{B}}} - S^{\mathcal{B}}_i}{2} \right)$
CSWS	$\lceil \frac{n}{k} \rceil \left(\lceil \log_k(n) \rceil + 1 \right) \cdot \max_{Q \in \mathcal{Q}_{\leq k} : i^* \in Q} \max_{i \in Q \setminus \{i^*\}} \left\lceil \bar{\gamma}^{-1} \left(\frac{S_{i^*\|Q} - S_{i\|Q}}{2} \right) \right\rceil$
CSR	$\lceil \frac{n}{k} \rceil \left(\left\lceil \log_{1-\frac{1}{k}} \left(\frac{1}{n} \right) \right\rceil + k - 1 \right) \cdot \max_{Q \in \mathcal{Q}_{\leq k} : i^* \in Q} \min_{i \in Q \setminus \{i^*\}} \left\lceil \bar{\gamma}^{-1} \left(\frac{S_{i^*\|Q} - S_{i\|Q}}{2} \right) \right\rceil$
CSH	$\lceil \frac{n}{k} \rceil \left(\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil \right) \cdot \max_{Q \in \mathcal{Q}_{\leq k} : i^* \in Q} \left\lceil \bar{\gamma}^{-1} \left(\frac{S_{i^*\|Q} - S_{\pi(Q)\|Q}}{2} \right) \right\rceil$

In Section F.2 we compare these quantities for the special case, in which the gaps $\Delta_{i|Q} = S_{i^*|Q} - S_{i|Q}$ are all equal to some $\Delta > 0$, while in Section F.3 we derive the sufficient budgets resulting from Corollaries 5.1 and 5.2 for the reward setting and preference-based setting, respectively, to return the best arm with high probability in the stochastic setting. Note that if $\gamma_{i|Q}(t) = \gamma(t)$ and $S_{(2)|Q} = \cdots = S_{(|Q|)|Q}$ are fulfilled for all $Q \in \mathcal{Q}_{\leq k}, i \in Q$ and $t \in \mathbb{N}$, then the lower bound in Theorem 3.1 (i) matches the above upper bound for CSWS up to a factor $C = \lceil \log_k(n) \rceil + 1$.

F.1 Maximal Number of Different Query Sets

The maximal number of required query sets for each algorithm is $\sum_{r=1}^{R} P_r$. Note that this is a geometric series and thus the partial sum can easily be computed for each of our proposed algorithms.

CSWS By using the specified valued of R and P_r for CSWS, we obtain that the number of different query set is at most

$$\begin{aligned}
\sum_{r=1}^{R^{CSWS}} P_r^{CSWS} &= \sum_{r=1}^{\lceil \log_k(n) \rceil + 1} \left\lceil \frac{n}{k^r} \right\rceil \\
&\leq \lceil \log_k(n) \rceil + 1 + \sum_{r=1}^{\lceil \log_k(n) \rceil + 1} \frac{n}{k^r} \\
&= \lceil \log_k(n) \rceil + 1 + n \cdot \left(\sum_{r=0}^{\lceil \log_k(n) \rceil + 1} \left(\frac{1}{k}\right)^r - 1 \right) \\
&= \lceil \log_k(n) \rceil + 1 + n \cdot \left(\frac{1 - 1/k^{\lceil \log_k(n) \rceil + 2}}{1 - 1/k} - 1 \right) \\
&= \lceil \log_k(n) \rceil + 1 + n \cdot \left(\frac{1 - 1/k^{\lceil \log_k(n) \rceil + 1}}{k - 1} \right),
\end{aligned}$$

where we used for the inequality that $\lceil x \rceil \leq x + 1$ for any $x \in \mathbb{R}$.

CSR For CSR we get as an upper bound on the number of different query sets:

$$\begin{aligned}
\sum_{r=1}^{R^{CSR}} P_r^{CSR} &= \sum_{r=1}^{\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \rceil + k - 1} \left\lceil \frac{n\left(1 - \frac{1}{k}\right)^{r-1}}{k} \right\rceil \\
&\leq \left\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \right\rceil + k - 1 + \sum_{r=1}^{\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \rceil + k - 1} \frac{n\left(1 - \frac{1}{k}\right)^{r-1}}{k} \\
&= \left\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \right\rceil + k - 1 + \frac{n}{k} \sum_{r=0}^{\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \rceil + k - 2} \left(1 - \frac{1}{k}\right)^r \\
&= \left\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \right\rceil + k - 1 + \frac{n}{k} \frac{\left(1 - \left(1 - \frac{1}{k}\right)^{\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \rceil + k - 1}\right)}{\left(1 - \left(1 - \frac{1}{k}\right)\right)} \\
&= \left\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \right\rceil + k - 1 + n\left(1 - \left(1 - \frac{1}{k}\right)^{\lceil \log_{1-\frac{1}{k}}\left(\frac{1}{n}\right) \rceil + k - 1}\right).
\end{aligned}$$

CSH Similarly, we can obtain for CSH the following maximum number of different query sets:

$$\begin{aligned}
\sum_{r=1}^{R^{CSH}} P_r^{CSH} &= \sum_{r=1}^{\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil} \left\lceil \frac{n}{2^{r-1}k} \right\rceil \\
&\leq \lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil + \sum_{r=1}^{\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil} \frac{n}{2^{r-1}k} \\
&= \lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil + \frac{n}{k} \sum_{r=0}^{\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil - 1} \left(\frac{1}{2}\right)^r \\
&= \lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil + \frac{2n}{k}\left(1 - 1/2^{\lceil \log_2(n) \rceil + \lceil \log_2(k) \rceil}\right).
\end{aligned}$$

These upper bounds on the maximum number of different query sets are summarized in Table 2. Note that ROUNDROBIN by design queries the possible query sets of $\mathcal{Q}_{=k}$ in a round-robin fashion, so that the number of different query sets is indeed $|\mathcal{Q}_{=k}| = \binom{n}{k}$.

F.2 Comparison of Sufficient Budgets

In order to compare the derived sufficient budgets of the different algorithms (see Table 3), we consider in the following the setting where the generalized Condorcet winner coincides with the generalized Borda winner. In addition we assume that the limit statistic $S_{i|Q}$ for each arm $i \in \mathcal{A}$ has always the same difference to the limit of the optimal arm $S_{i^*|Q}$ if $i^* \in Q$. More precisely, for each arms $i \in \mathcal{A}$ and each query set $Q \in \mathcal{Q}_{\leq k}$ we have $\Delta_{i|Q} = \Delta$ for some fixed $\Delta > 0$. In this way, the γ-dependent term present in the sufficient budget for each algorithm is simply $\lceil \bar{\gamma}^{-1}\left(\frac{\Delta}{2}\right) \rceil$. As a consequence, we can neglect this term as it has no influence on the differences in the desired budgets for the various algorithms and the remaining term based on the product of the number of rounds, i.e. R, and the number of partitions in round 1, i.e. P_1, is driving the (rough) sufficient budget bounds (see Table 2). However, the number of partitions in round 1 is the same for all algorithms, so that we can neglect this term as well. With a slight abuse of denotation, we refer to this remainder term simply as the sufficient budget in the following. With these considerations, it is easy to see that ROUNDROBIN requires the highest sufficient budget even for moderate sizes of n if k is sufficiently lower than n. To get an impression how the sufficient budget behaves for the more sophisticated algorithms based on the successive elimination strategy, we plot these in Figure 2 as curves depending on the number of arms n for different subset sizes k. Note, that in contrast to CSWS and CSH, the sufficient budget of CSR is higher for bigger subset sizes k, since only a smaller proportion of all arms is discarded after each round. In the case $k = 2$ the number of rounds are all the same, so that consequently the sufficient budget is the same for all three algorithms.

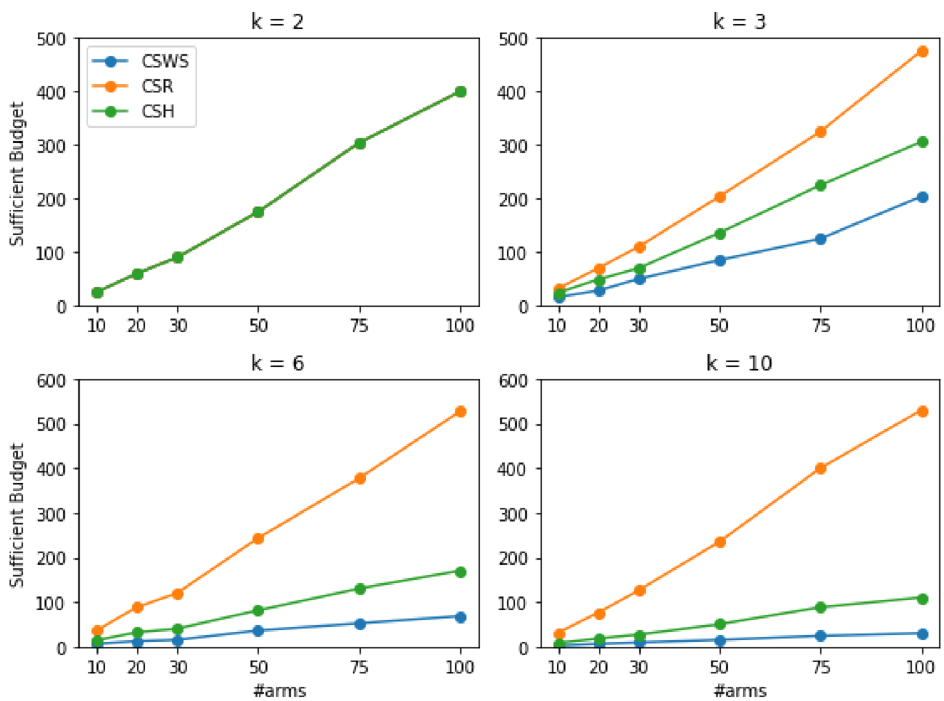

Figure 2: Comparison of required budget for our proposed algorithms for different values of the number of arms n and the subset size k.

F.3 Applications to Stochastic Settings

In Table 4 the sufficient budgets for our proposed algorithms in the stochastic setting with reward feedback and preference-based feedback are listed. Note, that these results are simply derived by applying Corollary 5.1 and resp. Corollary 5.2 with the specific instantiations of R and P_r for our algorithms (see Tables 2 and 3).

Table 4: Comparison of the sufficient budgets for our proposed algorithms CSWS, CSR and CSH in the reward and preference-based setting.

Alg.	Budget in reward setting
CSWS	$\frac{1}{\epsilon^2}\ln\left(\frac{k(\lceil\log_k(n)\rceil+1)}{\delta}\ln\left(\frac{k(\lceil\log_k(n)\rceil+1)}{\epsilon\delta}\right)\right)\cdot(\lceil\log_k(n)\rceil+1)\left\lceil\frac{n}{k}\right\rceil$
CSR	$\frac{1}{\epsilon^2}\ln\left(\frac{k\left(\left\lceil\log_{1-\frac{1}{k}}\left(\frac{1}{n}\right)\right\rceil+k-1\right)}{\delta}\ln\left(\frac{k\left(\left\lceil\log_{1-\frac{1}{k}}\left(\frac{1}{n}\right)\right\rceil+k-1\right)}{\epsilon\delta}\right)\right)\cdot\left(\left\lceil\log_{1-\frac{1}{k}}\left(\frac{1}{n}\right)\right\rceil+k-1\right)\left\lceil\frac{n}{k}\right\rceil$
CSH	$\frac{1}{\epsilon^2}\ln\left(\frac{k(\lceil\log_2(n)\rceil+\lceil\log_2(k)\rceil)}{\delta}\ln\left(\frac{k(\lceil\log_2(n)\rceil+\lceil\log_2(k)\rceil)}{\epsilon\delta}\right)\right)\cdot(\lceil\log_2(n)\rceil+\lceil\log_2(k)\rceil)\left\lceil\frac{n}{k}\right\rceil$
Alg.	**Budget in preference-based setting**
CSWS	$\frac{1}{\epsilon^2}\ln\left(\frac{\lceil\log_k(n)\rceil+1}{\delta\epsilon^4}\right)\cdot(\lceil\log_k(n)\rceil+1)\left\lceil\frac{n}{k}\right\rceil$
CSR	$\frac{1}{\epsilon^2}\ln\left(\frac{\left\lceil\log_{1-\frac{1}{k}}\left(\frac{1}{n}\right)\right\rceil+k-1}{\delta\epsilon^4}\right)\cdot\left(\left\lceil\log_{1-\frac{1}{k}}\left(\frac{1}{n}\right)\right\rceil+k-1\right)\left\lceil\frac{n}{k}\right\rceil$
CSH	$\frac{1}{\epsilon^2}\ln\left(\frac{\lceil\log_2(n)\rceil+\lceil\log_2(k)\rceil}{\delta\epsilon^4}\right)\cdot(\lceil\log_2(n)\rceil+\lceil\log_2(k)\rceil)\left\lceil\frac{n}{k}\right\rceil$

G Further Experiments

In the following, we present some further experiments comparing our proposed algorithms with each other on synthetic data including a detailed description of the data generation and the experiment setting.

G.1 Synthetic Data

For each $Q \in \mathcal{Q}_{\leq k}$ with $Q = \{i_1, \ldots, i_{|Q|}\}$ we consider the case where the observation vector $\mathbf{o}_Q$ is a random sample from a multivariate Gaussian distribution with mean $\mu_Q = (\mu_{i_1|Q}, \ldots, \mu_{i_{|Q|}|Q})^\top$ and a diagonal covariance matrix $\mathrm{diag}(\sigma_{i_1|Q}, \ldots, \sigma_{i_{|Q|}|Q})$. Here, $\mu_{i_j|Q}$ are values in $[0, 1]$ for $i_j \neq i^*$ and $\sigma_{i|Q}$ in $[0.05, 0.2]$ (all randomly sampled). For any Q with $i^* \in Q$ we set $\mu_{i^*|Q} = \max_{j \in Q, j \neq i^*} \mu_{j|Q} + \varepsilon$ for some $\varepsilon > 0$, which ensures (A2) to hold for the expected values. In our experiments we always use a value of $\varepsilon = 0.1$. In the following we vary the values of $n \in \{50, 100\}$, $k \in \{2, 4, 6, 8, 10\}$ and $B \in \{50, 100, 200, 300, 500\}$.

We consider a reward setting and use the empirical mean as the statistic (see Section 5). We do not force the generalized Borda winner to be the same as the generalized Condorcet winner, but they naturally coincidence in most of the runs by sampling the observation vector as defined above.

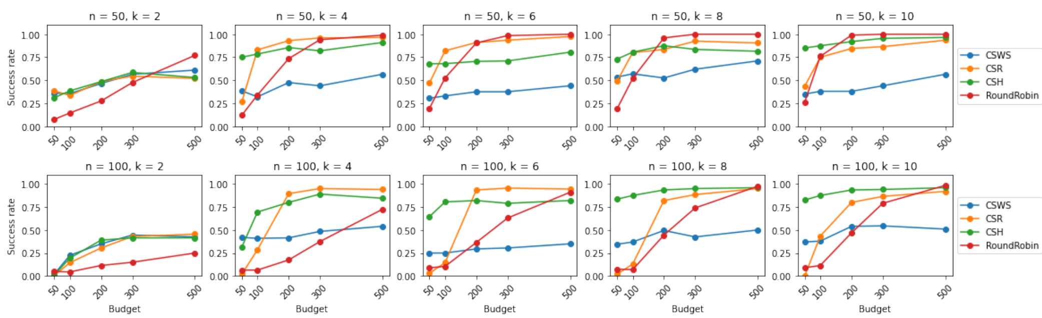

Figure 3: Success rates of our proposed algorithms for varying n, k and budget B in the reward setting.

The success rates of our proposed algorithms for identifying i^* given a budget B are shown in Figure 3. It is visible, that in particular for the challenging scenario, where the budget B and the subset size k are small and the number of arms n is large, both CSH and CSR perform well. Especially CSH has overall a solid performance.

Reward setting. In contrast to the experiments with reward feedback shown in the main paper, we try in the following experiments to force the generalized Borda winner to be different from the generalized Condorcet winner. For this purpose, we fix one random arm $i^*_{\mathcal{B}} \in [n] \backslash \{i^*\}$ as the prospective generalized Borda winner and set its expected value to $\mu_{i^*_{\mathcal{B}}|Q} = \max_{j \in Q, j \neq i^*_{\mathcal{B}}} \mu_{j|Q} + 2\epsilon$ for any $Q \in \mathcal{Q}_{\leq k}$ with $i^*_{\mathcal{B}} \in Q$ and $i^* \notin Q$. Thus, $i^*_{\mathcal{B}}$ is likely the generalized Borda winner and is different from the generalized Condorcet winner. Since our goal is to find the generalized Condorcet winner i^*, ROUNDROBIN will probably fail most of the times in finding i^*. This is due to the fact that ROUNDROBIN focuses on identifying $i^*_{\mathcal{B}}$, i.e., the the generalized Borda winner, which, however, does not coincidence with the generalized Condorcet winner i^*.

This suspicion is confirmed by the results of the experiments shown in Figure 4 illustrating the empirical success rates for finding the generalized Condorcet winner in the setting described above. Except for some cases where the subset size k is relatively large in comparison to the total number of arms, such that the generalized Condorcet winner is already contained in most of the seen subsets and hence is automatically also the generalized Borda winner, ROUNDROBIN performs poorly in finding the generalized Condorcet winner and is always outperformed by the algorithms based on the combinatorial successive elimination strategy in Section 4.1.

Preference-based setting with different GCW and GBW. In the preference-based setting we ignore the explicit numerical values of the observation vector and only use the information which

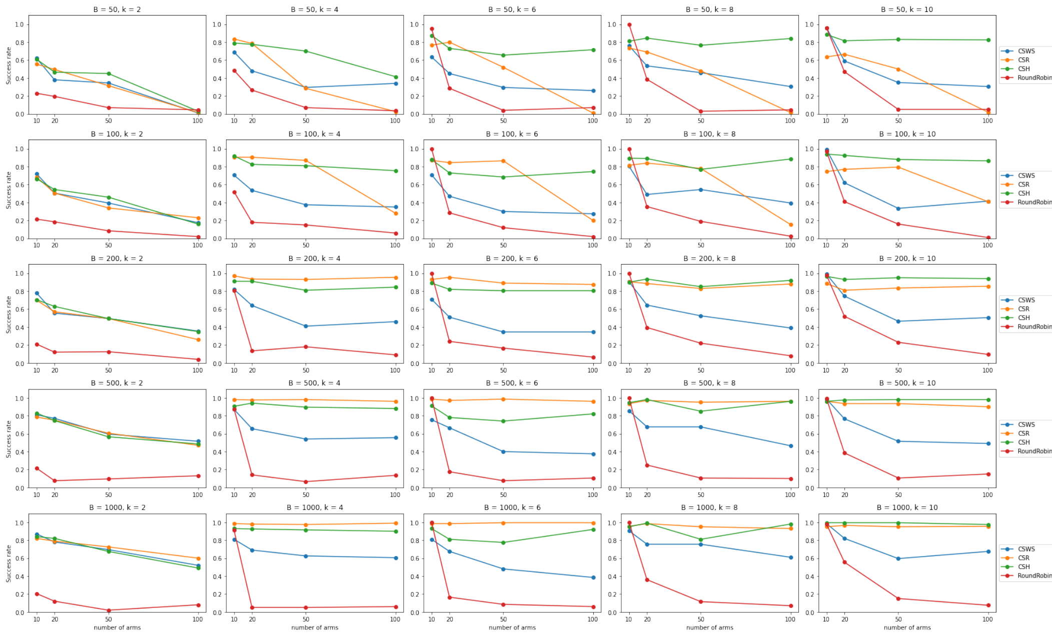

Figure 4: Success rates of our proposed algorithms for varying n, k and budget B in the reward setting with different generalized Condorcet winner and generalized Borda winner.

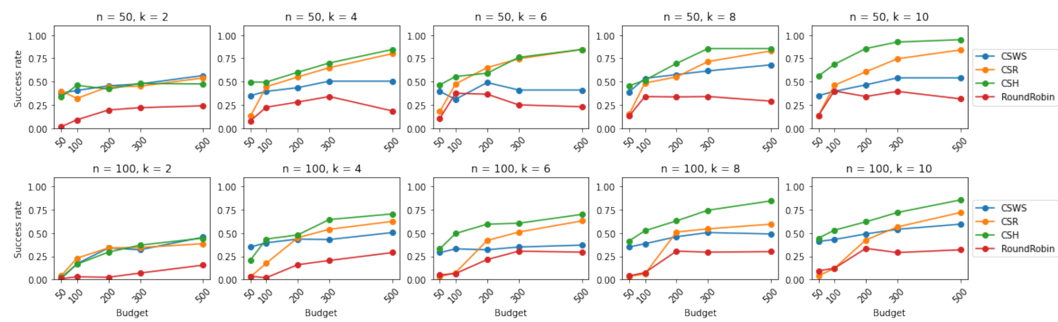

Figure 5: Success rates of our proposed algorithms for varying n, k and budget B in the preference-based setting with different generalized Condorcet winner and generalized Borda winner.

arm was (not) the winner, i.e., which had (not) the highest observation value in the query set used , formally $s_{i_j|Q}(t) = \frac{1}{t}\sum_{s=1}^{t} \mathbb{1}\{o_{i_j|Q}(s) = \max_{i=i_1,\dots i_{|Q|}} o_{i|Q}(s)\}$. Additionally, we fix one arm $i^*_{\mathcal{B}} \in [n]\backslash\{i^*\}$ and set $\mu_{i^*_{\mathcal{B}}|Q} = \max_{j\in Q, j\neq i^*_{\mathcal{B}}} \mu_{j|Q} + 2\varepsilon$ for any Q with $i^*_{\mathcal{B}} \in Q$ and $i^* \notin Q$. In this way, $i^*_{\mathcal{B}}$ is the generalized Borda winner and different from i^*.

The success rates of our proposed algorithms for identifying i^* in this setting are shown in Figure 5. As expected our methods outperform RoundRobin in all scenarios.

Preference-based setting. We now investigate the case, in which we do not force the generalized Borda winner and the generalized Condorcet winner to be different, thus they will naturally coincidence in most of the cases. This is achieved by considering the problem configuration as in the reward setting specified in Section 6, and ignoring the explicit numerical values (as in the preference-setting above).

The resulting success rates for finding the generalized Condorcet winner illustrated in Figure 6 are similar to the results in the reward setting for matching generalized Condorcet winner and generalized

Borda winner. This means that, in particular, when the budget is small, the number of arms is large and the subset size is small, the algorithms following the combinatorial successive elimination strategy outperform ROUNDROBIN. Note that this setting is arguably the most relevant setup for practical applications. Moreover, Figure 6 illustrates the natural effect one would expect for the number of arms n on success rates, namely that success rates decrease with a larger number of arms.

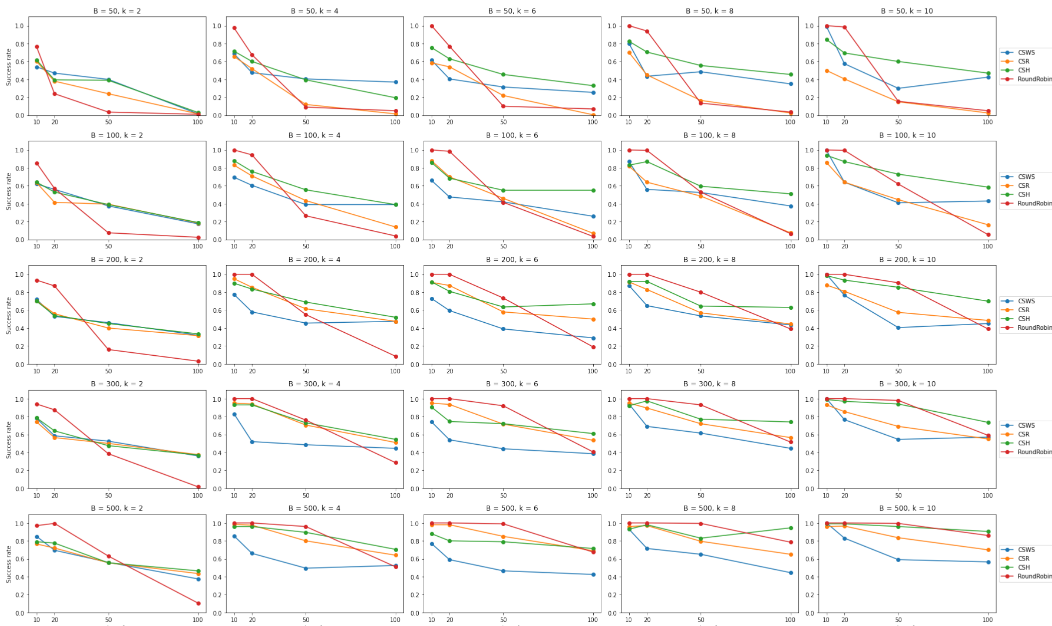

Figure 6: Success rates of our proposed algorithms for varying n, k and budget B in the preference-based setting with (mostly) matching generalized Condorcet winner and generalized Borda winner.

G.2 Statistics beyond the Arithmetic Mean

We consider in the following the reward setting, where each observation is random sampled from the following distribution

$$\boldsymbol{o}_Q(t) \sim \mathcal{N}\left(\left(\begin{array}{c} \mu_{1|Q} \\ \vdots \\ \mu_{|Q||Q} \end{array}\right), \left(\begin{array}{c} \sigma_{1|Q} \\ \vdots \\ \sigma_{|Q||Q} \end{array}\right)\right)$$

for $\mu_{i|Q}$ is sampled randomly from $[0,1]$ and $\sigma_{i|Q}$ from $[0.05, 0.2]$ for each arm $i \in Q$.

Median An alternative to the arithmetic mean would be to measure the quality of the arms by the median of the seen observations. In particular, when the observations are prone to outliers, the median provides a more robust statistic: $s_{i|Q}(t) = \textsc{Median}(o_{i|Q}(1), \dots, o_{i|Q}(t))$ for each arm $i \in Q$. The results for this setting are illustrated in Figure 7.

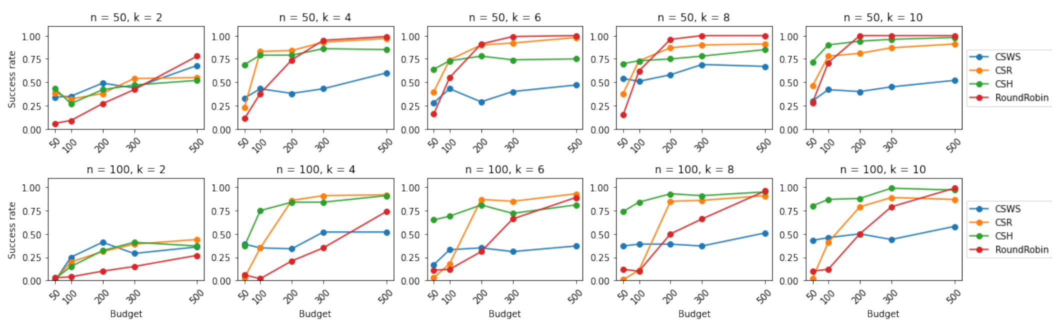

Figure 7: Success rates of our proposed algorithms for varying n, k and budget B in the rewards setting with (mostly) matching generalized Condorcet winner and generalized Borda winner and using the median as the statistic.

Power-Mean Another possibility is to use the so called power-mean, which is a compromise between the maximum and the arithmetic mean for a (multi)set of observations. Since the arithmetic mean is known to underestimate the true quality of an arm, while the maximum overestimates it, the power mean is often a good compromise, as it lies between the two. It is defined by $s_{i|Q}(t) = \left(\frac{1}{t}\sum_{t'=1}^{t} o_{i|Q}(t')^q\right)^{1/q}$ for each arm $i \in Q$ and a fixed $q \in \mathbb{N}$. We use in the following $q = 2$. The results for this setting are illustrated in Figure 8.

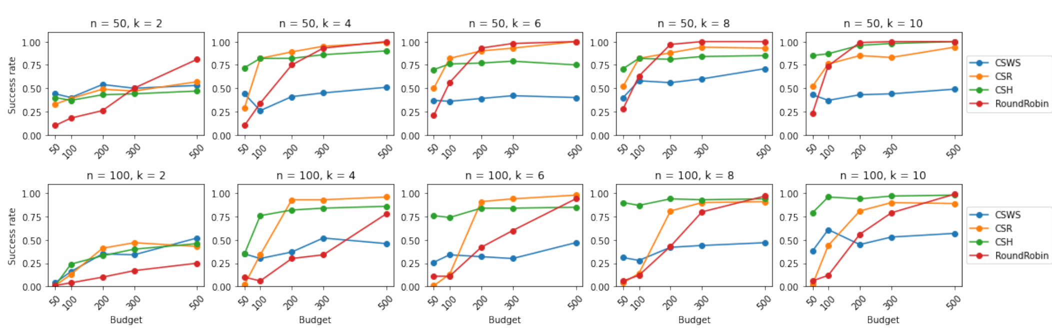

Figure 8: Success rates of our proposed algorithms for varying n, k and budget B in the rewards setting with (mostly) matching generalized Condorcet winner and generalized Borda winner and using the power mean as the statistic.