Appendices

A Unbiased noisy partial feedback

A.1 Proposition 1

Proof. By Lemma 1 applied to $X_{i,t_1}, X_{i,t_2}, \ldots$ for an arm i for F full delayed feedback, we have w.p. $1 - {}^{\delta_f}/_n$:

$$\left| \frac{1}{F} \sum_{f=1}^{F} X_{i,t_f} - \mu_i \right| \leq C\left(\sigma_i, F, {}^{\delta_f}/_n\right). \tag{7}$$

For any a, $E[Y_{i,t_F+p}|X_{i,t_F} = a] = a$, and $E[Y_{i,t_F+p} - a|X_{i,t_F} = a] = 0$. Conditioned on $X_{i,t_F} = a$, $(Y_{i,t_F+p} - a)|(X_{i,t_F} = a)$ is sub-Gaussian by assumption.

Therefore, conditioned on $X_{i,t_F} = a$, by Lemma 1 applied to $\frac{1}{P}\sum_{p=1}^{P}(Y_{i,t_F+p}|X_{i,t_F} = a) - a$ for an arm i computed using P partial feedback for the F-th pull, we have w.p. $1 - {}^{\delta_p}/_n$:

$$\left| \frac{1}{P} \sum_{p=1}^{P} Y_{i,t_F+p} - a \right| \leq C\left(\sigma_i^{(p)}, P, {}^{\delta_p}/_n\right). \tag{8}$$

Given that the result does not depend on the value a, we have:

$$\left| \frac{1}{P} \sum_{p=1}^{P} \left(Y_{i,t_F+p}|X_{i,t_F}\right) - X_{i,t_F} \right| \leq C\left(\sigma_i^{(p)}, P, {}^{\delta_p}/_n\right). \tag{9}$$

From a union bound Eq. (7) and Eq. (9), we have w.p. $1 - {}^{\delta_f}/_n - {}^{\delta_p}/_n$:

$$\left| \frac{1}{F} \left[\sum_{f=1}^{F-1} X_{i,t_f} + \frac{1}{P} \sum_{p=1}^{P} Y_{i,t_F+p} \right] - \mu_i \right| \leq C\left(\sigma_i, F, {}^{\delta_f}/_n\right) + \frac{1}{F} C\left(\sigma_i^{(p)}, P, {}^{\delta_p}/_n\right). \tag{10}$$

Union bounding Eq.(10) over all arms, we have w.p. $1 - \delta_f - \delta_p$:

$$\left| \frac{1}{F} \left[\sum_{f=1}^{F-1} X_{i,t_f} + \frac{1}{P} \sum_{p=1}^{P} Y_{i,t_F+p} \right] - \mu_i \right| \leq C\left(\sigma_i, F, {}^{\delta_f}/_n\right) + \frac{1}{F} C\left(\sigma_i^{(p)}, P, {}^{\delta_p}/_n\right) \ \forall i \in [1, n] \tag{11}$$

finishing the proof. □

A.2 Theorem 1

At any given time $t \geq 1$, $F \in \mathbb{N}, P \in [1, D_F]$, we observe $F - 1$ full feedback, $X_{i,t_{1:F-1}}$ for an arbitrary arm $i \in [1, n]$. Accordingly, we have the following two cases to consider as per Algorithm 2.

- **Case (a):** $C(\sigma_i, F-1, {}^{\delta}/_n) < C(\sigma_i, F, {}^{\delta_f^*}/_n) + \frac{1}{F} C(\sigma_i^{(p)}, P, {}^{\delta_p^*}/_n)$

$$\widehat{\mu}_i = \frac{1}{F-1} \sum_{f=1}^{F-1} X_{i,t_f}$$
$$C_i = C\left(\sigma_i, F-1, {}^{\delta}/_n\right).$$

- **Case (b):** otherwise

$$\widehat{\mu}_i = \frac{1}{F} \left[\sum_{f=1}^{F-1} X_{i,t_f} + \frac{1}{P} \sum_{l=1}^{P} Y_{i,t_F+l} \right]$$
$$C_i = C(\sigma_i, F, {}^{\delta_f^*}/_n) + \frac{1}{F} C(\sigma_i^{(p)}, P, {}^{\delta_p^*}/_n).$$

Define $\mathcal{E}_i = \{\forall t \geq 1, |\widehat{\mu}_i - \mu_i| \leq C_i\}$ be the event that the lower and upper confidence bounds of arm i trap the true mean μ_i for all $t \geq 1$ where $\widehat{\mu}_i$ and C_i are chosen as described above at time t. Let S_t, A_t, R_t denote the set of surviving, accepted, and rejected arms at time t. We can then state and prove the following lemma.

Lemma 2. *Assume $\mathcal{E}_i$ holds for an arbitrary arm $i \in S_t$ and $i \notin S_{t+1}$. Then, the following statements hold:*

- *$i \in A_{t+1}$ if $i \leq k$.*
- *$i \in R_{t+1}$ if $i > k$.*

Proof. By definition, $S_{t+1} \cup A_{t+1} \cup R_{t+1} = S_t$. Recursing over $t, t-1, ...0$, we note that $S_{t+1} \cup A_{t+1} \cup R_{t+1} = \{1, 2, \ldots, n\}$. Since the lemma assumes that arm $i \notin S_{t+1}$, either $i \in A_{t+1}$ or $i \in R_{t+1}$.

We will prove the first statement of the lemma by contradiction. For an arbitrary $i \leq k$, let us assume $i \in R_{t+1}$. This implies that $UCB_i < \max^{(k)}_{j \in S_t} LCB_j$. Since by assumptions on the lemma the lower and upper confidence bounds of any arm trap its true mean, we have $UCB_i \geq \mu_i$ and $\max^{(k)}_{j \in S_t} LCB_j \leq \mu_k$. Hence, we obtain $\mu_i < \mu_k$ which is a contradiction since $i \leq k$. The second statement holds true by symmetry. ☐

Since both Proposition 1 and Eq. (5) hold true w.p. at least $1 - \delta/n$ for all arms, we get that $\cap_{i=1}^n \mathcal{E}_i$ holds true w.p. at least $1-\delta$ (union bound) regardless of the set of $\{\widehat{\mu}_i\}_{i=1}^n$ and $\{C_i\}_{i=1}^n$ picked by the algorithm. Combining the union bound with Lemma 2, the algorithm outputs the top-k set w.p. at least $1 - \delta$ if it terminates.

B Biased noisy partial feedback

B.1 Proposition 2

Proof. By Lemma 1 applied to $X_{i,t_1}, X_{i,t_2}, \ldots$ for an arm i for F full delayed feedback, we have w.p. $1 - {}^{\delta_f}\!/n$:

$$\left| \frac{1}{F} \sum_{f=1}^{F} X_{i,t_f} - \mu_i \right| \leq C\left(\sigma_i, F, {}^{\delta_f}\!/n\right). \tag{12}$$

For any a, $E[Y_{i,t_F+p} | X_{i,t_F} = a] = a + b_i$, and $E[Y_{i,t_F+p} - a - b_i | X_{i,t_F} = a] = 0$. Conditioned on $X_{i,t_F} = a$, $(Y_{i,t_F+p} - a - b_i) | (X_{i,t_F} = a)$ is sub-Gaussian by assumption. Therefore, conditioned on $X_{i,t_F} = a$, by Lemma 1 applied to $\frac{1}{P} \sum_{p=1}^{P} (Y_{i,t_F+p} - b_i) - X_{i,t_F}$ for the (incomplete) F-th pull of an arm i with P partial feedback, we have w.p. $1 - {}^{\delta_p}\!/n$:

$$\left| \frac{1}{P} \sum_{p=1}^{P} (Y_{i,t_F+p} - b_i) - X_{i,t_F} \right| \leq C\left(\sigma_i^{(p)}, P, {}^{\delta_p}\!/n\right). \tag{13}$$

Now, consider the $F-1$ random variables for all $f \in [1, F-1]$:

$$\frac{\sum_{p=1}^{D_f - 1} Y_{i,t_f+p}}{D_f - 1} - X_{i,f}. \tag{14}$$

The random variables in (14) are all sub-Gaussian with mean b_i and scale parameter $\sigma_i^{(p)}$. Hence, applying LIL on these random variables conditioning on b_i, we have w.p. $1 - {}^{\delta_b}\!/n$:

$$\left| \frac{1}{F-1} \sum_{f=1}^{F-1} \left(\frac{\sum_{p=1}^{D_f - 1} Y_{i,t_f+p}}{D_f - 1} - X_{i,D_f - 1} \right) - b_i \right| \leq C\left(\sigma_i^{(p)}, F-1, {}^{\delta_b}\!/n\right). \tag{15}$$

From a union bound of Eq. (12) and Eq. (13), we have w.p. $1 - {}^{\delta_f}\!/n - {}^{\delta_p}\!/n$:

$$\left| \frac{1}{F} \left[\sum_{f=1}^{F-1} X_{i,t_f} + \frac{1}{P} \sum_{p=1}^{P} (Y_{i,t_F+p} - b_i) \right] - \mu_i \right| \leq C\left(\sigma_i, F, {}^{\delta_f}\!/n\right) + \frac{1}{F} C\left(\sigma_i^{(p)}, P, {}^{\delta_p}\!/n\right). \tag{16}$$

Algorithm 4 RacingBiasedPF (arm parameters $\{i, \sigma_i, \sigma_i^{(p)}\}_{i=1}^n$, top k, confidence δ)

1: Initialize global time step $t = 0$, surviving $S = \{i\}_{i=1}^n$, accepted $A = \{\}$, rejected $R = \{\}$.
2: Initialize per-arm full delayed feedback counter $F_i = 0$, empirical means $\hat{\mu}_i = 0$, confidence bounds $LCB_i = -\infty$, $UCB_i = \infty$ for all $i \in S$.
3: **while** S is not empty **do**
4:     **while** True **do**
5:         Increment $t \leftarrow t + 1$.
6:         Collect partial feedback $Y_{a,t}$.
7:         Update $\widehat{\mu}^{(p)}$ using $Y_{a,t}$ as per Proposition 2.
8:         Increment $P \leftarrow P + 1$.
9:         Set $C^{(partial)} \leftarrow C\left(\sigma_a, F_a + 1, \delta_f^*/n\right) + \frac{1}{F_a+1}\left[C\left(\sigma_a^{(p)}, P, \delta_p^*/n\right) + C\left(\sigma_a^{(p)}, F_a, \delta_b^*/n\right)\right]$
10:         Choose FOrP $\leftarrow \arg\min\left(C(\sigma_a, F_a, \delta/n), C^{(partial)}\right)$.
11:         Update $C_a \leftarrow C(\sigma_a, F_a, \delta/n)$ if FOrP $= F$ else $C^{(partial)}$.
12:         Update $\widehat{\mu}_a \leftarrow \widehat{\mu}^{(f)}$ if FOrP $= F$ else $\frac{F_a \widehat{\mu}^{(f)} + \widehat{\mu}^{(p)}}{F_a + 1}$.
13:         Update LCB_a, UCB_a.
14:         $A, R, S \leftarrow$ UpdateArmSets$(A, R, S, k, \{LCB_i, UCB_i)\}_{i \in S})$.
15:         **if** $P = D_{a,t_a}$ or $a \notin S$ **then**
16:             Break   ▷ Pull on termination/elimination
17:         **end if**
18:     **end while**
19:     Pull arm a where $a \leftarrow \arg\min_{a \in S} F_a$.
20:     Initialize start $t_a \leftarrow t$, partial feedback counter $P = 0$, partial mean $\widehat{\mu}^{(p)} = 0$, full mean $\widehat{\mu}^{(f)} \leftarrow \widehat{\mu}_i$.
21: **end while**
22: **return** A

From a union bound of Eq. (15) and Eq. (16), we have w.p. $1 - \delta_f/n - \delta_p/n - \delta_b/n$:

$$\left| \frac{1}{F}\left[\sum_{f=1}^{F-1} X_{i,t_f} + \frac{1}{P}\sum_{p=1}^{P}\left(Y_{i,t_F+p} - \frac{1}{F-1}\sum_{f=1}^{F-1}\left(\frac{\sum_{p=1}^{D_f-1} Y_{i,t_f+p}}{D_f - 1} - X_{i,D_f-1}\right)\right)\right] - \mu_i \right|$$
$$\leq C\left(\sigma_i, F, \delta_f/n\right) + \frac{1}{F}C\left(\sigma_i^{(p)}, P, \delta_p/n\right) + \frac{1}{F}C\left(\sigma_i^{(p)}, F-1, \delta_b/n\right). \tag{17}$$

Finally, union bounding Eq. (17) over all arms, we have w.p. $1 - \delta_f - \delta_p - \delta_b$:

$$\left| \frac{1}{F}\left[\sum_{f=1}^{F-1} X_{i,t_f} + \frac{1}{P}\sum_{p=1}^{P}\left(Y_{i,t_F+p} - \frac{1}{F-1}\sum_{f=1}^{F-1}\left(\frac{\sum_{p=1}^{D_f-1} Y_{i,t_f+p}}{D_f - 1} - X_{i,D_f-1}\right)\right)\right] - \mu_i \right|$$
$$\leq C\left(\sigma_i, F, \delta_f/n\right) + \frac{1}{F}C\left(\sigma_i^{(p)}, P, \delta_p/n\right) + \frac{1}{F}C\left(\sigma_i^{(p)}, F-1, \delta_b/n\right) \; \forall i \in [1, n] \tag{18}$$

finishing the proof. □

B.2 Algorithm

We provide the pseudocode for the racing procedures with biased partial feedback in Algorithm 4. As discussed previously, the algorithm is similar to Algorithm 2 with key differences in the mean and confidence bound estimators in Line 7 and Line 9 respectively.