Supplementary Materials

1 The Empirical Results for the $20k$ Datasets of single cell RNA-seq and Netflix users

On the $20,000$ cell dataset, we note that `Med-dit` stopped within 140 distance evaluations per arm in each of the 1000 trials (with 80 distance evaluations per arm on average) and never returned the wrong answer. `RAND` needs around 700 distance evaluations per point to obtain 2% error rate.

On the $20,000$ user dataset, we note that `Med-dit` stopped within 600 distance evaluations per arm in each of the 1000 trials (with 500 distance evaluations per arm on average) and never returned the wrong answer. `RAND` needs around 4500 distance evaluations per point to obtain 2% error rate.

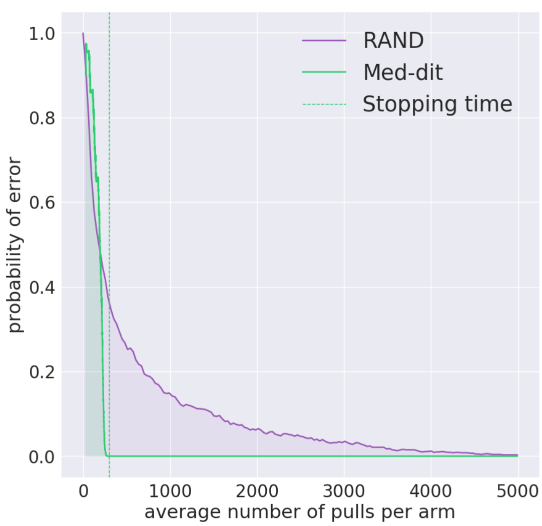

Figure 9: Small Netflix-prize dataset: We computed the true medoid of the data-set by brute force. The y-axis shows the probability that the estimated medoid does *not* correspond to the true medoid as a function of the number of pulls per arm. We note that `Med-dit` has a stopping condition while `RAND` does not. However we ignore the stopping condition for `Med-dit` here. `Med-dit` stops after 500 distance evaluations per point without failing in any of the 1000 trials, while `RAND` takes around 4500 distance evaluations to reach a 2% probability of error.

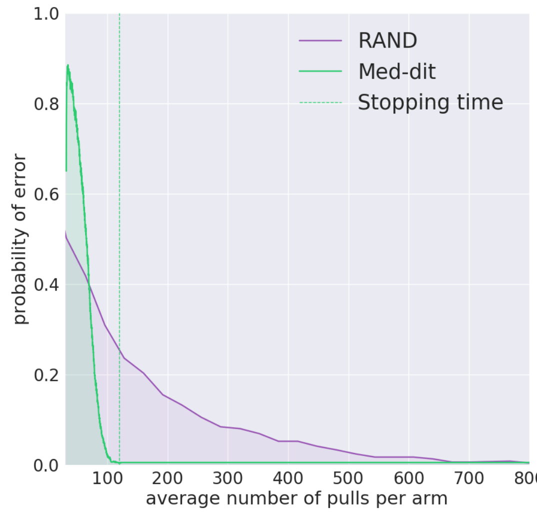

Figure 8: Small single cell dataset: We computed the true medoid of the dataset by brute force. The y-axis shows the probability that the estimated medoid does *not* correspond to the true medoid as a function of the number of pulls per arm. We note that `Med-dit` has a stopping condition while `RAND` does not. However we ignore the stopping condition for `Med-dit` here. `Med-dit` stops after 80 distance evaluations per point without failing in any of the 1000 trials, while `RAND` takes around 650 distance evaluations to reach a 2% probability of error.

2 The $O(n \log n)$ Distance Evaluations Under Gaussian Prior

We assume that the mean distances of each point μ_i are i.i.d. samples of $N(\gamma, 1)$. We note that this implies that $\Delta_i,\ 1 \le i \le n$ are n i.i.d. random variables. Let Δ be a random variable with the same law as Δ_i.

From the concentration of the minimum of n gaussians, we have that

$$\min_i \mu_i + \sqrt{2 \log n} \xrightarrow{p} \gamma.$$

This gives us that

$$\Delta - \sqrt{2 \log n} \xrightarrow{d} \mathcal{N}(0, 1).$$

We note that by Eq (2), we have that the expected number of distance evaluations M is of the order of

$$\mathbb{E}[M] \le n\mathbb{E}\left[\frac{\log n}{\Delta^2} \wedge n\right],$$

where the expection is taken with respect to the randomness of Δ.

To show this, it is enough to show that,

$$\mathbb{E}\left[\frac{\log n}{\Delta^2} \wedge n\right] \le C \log n,$$

for some constant C.

To compute that for this prior, we divide the real line into three intervals, namely

- $\left(-\infty, \sqrt{\frac{\log n}{n}}\right]$,
- $\left(\sqrt{\frac{\log n}{n}}, c\sqrt{\log n}\right)$,
- $\left[c\sqrt{\log n}, \infty\right)$,

and compute the expectation on these three ranges. We note that for if $\Delta \in (-\infty, \sqrt{\frac{\log n}{n}}]$, while for $\Delta \in (\sqrt{\frac{\log n}{n}}, \infty]$, $\frac{\log n}{\Delta^2} \le n$. Thus we have that,

$$\mathbb{E}\left[\frac{\log n}{\Delta^2} \wedge n\right] \le \overbrace{\mathbb{E}\left[n\mathbb{I}\left(\Delta \le \sqrt{\frac{\log n}{n}}\right)\right]}^{\mathbf{I}} + \overbrace{\lim_{\delta,\epsilon \to 0} \mathbb{E}\left[\frac{\log n}{\Delta^2}\mathbb{I}\left(\sqrt{\frac{\log n}{n^{(1-\epsilon)}}} \le \Delta \le c\frac{\sqrt{\log n}}{n^{\delta}}\right)\right]}^{\mathbf{II}}$$
$$+ \underbrace{\mathbb{E}\left[\frac{\log n}{\Delta^2}\mathbb{I}\left(\Delta \ge c\sqrt{\log n}\right)\right]}_{\mathbf{III}},$$

where we use the Bounded Convergence Theorem to establish **II**.

We next show that all three terms in the above equation are of $O(\log n)$. We first show the easy cases of **I** and **III** and then proceed to **II**.

- To establish that **I** is $O(\log n)$, we start by defining, $q_1 = P[\Delta < \frac{\sqrt{\log n}}{n}]$

$$\mathbb{E}\left[n\mathbb{I}\left(\Delta \le \frac{\sqrt{\log n}}{n}\right)\right] = nq_1.$$

Further note that,

$$q_1 \leq \exp\left(-\frac{1}{2}\left(\sqrt{2\log n} - \sqrt{\frac{\log n}{n}}\right)^2\right),$$
$$= \left(\frac{1}{n}\right)^{(1-\frac{1}{\sqrt{2n}})^2}.$$

Thus

$$\frac{q_1 n}{\log n} \leq \frac{n^{1-(1-\frac{1}{\sqrt{2n}})^2}}{\log n},$$
$$\leq \exp((\sqrt{\frac{2}{n}}(1+o(1))\log n - \log\log n),$$
$$= o(1).$$

- To establish that **III** is $O(\log n)$, we note that,

$$\mathbb{E}\left[\frac{\log n}{\Delta^2}\mathbb{I}\left(\Delta \geq c\sqrt{\log n}\right)\right] \leq \frac{1}{c^2}P(\Delta \geq c\sqrt{\log n}),$$
$$\leq \frac{1}{c^2},$$
$$= \Theta(1).$$

- Finally to establish that **II** is $O(\log n)$, we note that,

$$\lim_{\delta,\epsilon\to 0}\mathbb{E}\left[\frac{\log n}{\Delta^2}\mathbb{I}\left(\sqrt{\frac{\log n}{n^{(1-\epsilon)}}} \leq \Delta \leq c\frac{\sqrt{\log n}}{n^\delta}\right)\right] \leq \lim_{\delta,\epsilon\to 0} n^{2-\epsilon}P\left(\Delta \leq c\frac{\sqrt{\log n}}{n^\delta}\right),$$
$$\leq \lim_{\delta,\epsilon\to 0} n^{1-\epsilon}\exp\left(-\frac{1}{2}\left(\sqrt{2\log n} - \sqrt{\log n}\frac{c}{n^\delta}\right)^2\right),$$
$$= \lim_{\delta,\epsilon\to 0} n^{1-\epsilon}\left(\frac{1}{n}\right)^{(1-\frac{c}{\sqrt{2}n^\delta})^2},$$
$$= \lim_{\delta,\epsilon\to 0} n^{1-\epsilon-(1-\frac{c}{\sqrt{2}n^\delta})^2},$$
$$= \lim_{\delta,\epsilon\to 0} n^{-\epsilon+\frac{\sqrt{2}c}{n^\delta}-\frac{c^2}{2n^{2\delta}}}.$$

Letting δ to go to 0 faster than ϵ, we see that,

$$\lim_{\delta,\epsilon\to 0}\mathbb{E}\left[\frac{\log n}{\Delta^2}\mathbb{I}\left(\sqrt{\frac{\log n}{n^{(1-\epsilon)}}} \leq \Delta \leq c\frac{\sqrt{\log n}}{n^\delta}\right)\right] \leq O(\log n).$$

This gives us that under this model, we have that under this model,

$$\mathbb{E}[M] \leq O(n\log n).$$

3 Proof of Theorem 1

Let M be the total number of distance evaluations when the algorithm stops. As defined above, let $T_i(t)$ be the number of distance evaluations of point i up to time t.

We first assume that $[\hat{\mu}_i(t) - C_i(t), \hat{\mu}_i(t) - C_i(t)]$ are true $\left(1 - \frac{2}{n^3}\right)$-confidence interval (Recall that $\delta = \frac{2}{n^3}$) and show the result. Then we prove this statement.

Let i^* be the true medoid. We note that if we choose to update arm $i \neq i^*$ at time t, then we have

$$\hat{\mu}_i(t) - C_i(t) \leq \hat{\mu}_{i^*}(t) - C_{i^*}(t).$$

For this to occur, at least one of the following three events must occur:

$$\begin{aligned}
\mathcal{E}_1 &= \left\{\hat{\mu}_{i^*}(t) \geq \mu_{i^*}(t) + C_{i^*}(t)\right\}, \\
\mathcal{E}_2 &= \left\{\hat{\mu}_i(t) \leq \mu_i(t) - C_i(t)\right\}, \\
\mathcal{E}_3 &= \left\{\Delta_i = \mu_i - \mu_{i^*} \leq 2C_i(t)\right\}.
\end{aligned}$$

To see this, note that if none of $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ occur, we have

$$\hat{\mu}_i(t) - C_i(t) \overset{(a)}{>} \mu_i - 2C_i(t) \overset{(b)}{>} \mu_1 \overset{(c)}{>} \hat{\mu}_1 - C_1(t),$$

where (a), (b), and (c) follow because $\mathcal{E}_2$, $\mathcal{E}_3$, and $\mathcal{E}_1$ do not hold respectively.

We note that as we compute $\left(1 - \frac{2}{n^3}\right)$-confidence intervals at most n times for each point. Thus we have at most n^2 computations of $\left(1 - \frac{2}{n^3}\right)$-confidence intervals in total.

Thus $\mathcal{E}_1$ and $\mathcal{E}_2$ do not occur during any iteration with probability $\left(1 - \frac{2}{n}\right)$, because

$$\text{w.p. } (1 - \frac{2}{n}) : |\mu_i - \hat{\mu}_i(t)| \leq C_i(t),\ \forall\, i \in [n],\ \forall\, t. \tag{3}$$

This also implies that with probability $1 - \Theta\left(\frac{1}{n}\right)$ the algorithm does not stop unless the event $\mathcal{E}_3$, a deterministic condition stops occurring.

Let ζ_i be the iteration of the algorithm when it evaluates a distance to point i for the last time. From the previous discussion, we have that the algorithm stops evaluating distances to points i when the following holds.

$$\begin{aligned}
C_i(\zeta_i) \leq \frac{\Delta_i}{2} &\implies \frac{\Delta_i}{2} \geq \sqrt{\frac{2\sigma^2 \log n^3}{T_i(\zeta_i)}} \text{ or } C_i(\zeta_i) = 0, \\
&\implies T_i(\zeta_i) \geq \frac{24\sigma^2}{\Delta_i^2} \log n \text{ or } T_i(\zeta_i) \geq 2n.
\end{aligned}$$

Thus with probability $(1 - o(1))$, the algorithm returns i^* as the medoid with at most M distance evaluations, where

$$M \leq \sum_{i \in [n]} T_i(\zeta_i) \leq \sum_{i \in [n]} \left(\frac{24\sigma^2}{\Delta_i^2} \log n \wedge 2n\right).$$

To complete the proof, we next show that $[\hat{\mu}_i(t) - C_i(t), \hat{\mu}_i(t) - C_i(t)]$ are true $\left(1 - \frac{2}{n^3}\right)$-confidence intervals of μ_i.

We observe that if point i is picked by `Med-dit` less than n times at time t, then $T_i(t)$ is equal to the number of times the point is picked. Further $C_i(t)$ is the true $(1 - \delta)$-confidence interval from Eq (1).

However, if point i is picked for the n-th time at iteration t (line 7 of Algorithm 1) then the empirical mean is computed by evaluating all $n - 1$ distances (many distances again). Hence $T_i(t) = 2(n - 1)$. As we know the mean distance of point i exactly, $C_i(t) = 0$ is still the true confidence interval.

We remark that if a point i is picked to be updated the $(n+1)$-th time, as $C_i(t) = 0$, we have that $\forall j \neq i$,

$$\hat{\mu}_i(t) + C_i(t) = \hat{\mu}_i(t) - C_i(t) < \hat{\mu}_j(t) - C_j(t).$$

This gives us that the stopping criterion is satisfied and point i is declared as the medoid.

4 Extensions to the Theoretical Results

In this section we discuss two extensions to the theoretical results presented in the main article: 1. we can actually relax the sub-Gaussian condition to the condition that the random variables have finite variances while still having the same $O(n \log n)$ sample complexity; 2. there exists a medoid algorithm with a $O(n \log \log n)$ distance evaluations, which improves the sample complexity of `Med-dit` by a $\log n$ factor and is essentially almost-lienar.

In short, the sub-Gaussian condition can be weakened the condition that the random variables of sampling with replacement from $\mathcal{D}_i$ have finite variances. This is by using some refined estimator to estimate the average distance rather than using the empirical mean [Bubeck et al., 2013].

The $O(n \log \log n)$ sample complexity can be achieved by another best-arm algorithm called `exponential-gap` [Karnin et al., 2013]. This algorithm is $\log n$ faster than `UCB-medoid`. But this improvement is at the sacrifice of a much larger constant factor.

4.1 Weakening the Sub-Gaussian Assumption

Careful readers may notice that in order to have the $O(n \log n)$ sample complexity, `UCB-medoid` relies on a concentration bound where the tail probability decays exponentially. In the main article, this is achieved by assuming that for each point x_i, the random variable of sampling with replacement from $\mathcal{D}_i$ is σ-sub-Gaussian. As a result, we can have the sub-Gaussian tail bound that for any point x_i at time t, with probability at least $1-\delta$, the empirical mean $\hat{\mu}_i$ satisfies

$$|\mu_i - \hat{\mu}_i| \leq \sqrt{\frac{2\sigma^2 \log \frac{2}{\delta}}{T_i(t)}}.$$

In fact, as pointed out by Bubeck et al. [2013], to achieve the $O(n \log n)$ sample complexity, all we need is a performance guarantee like the one shown above for the empirical mean. To be more precise, we need the following property:

Assumption 1. *[Bubeck et al., 2013] Let $\epsilon \in (0,1]$ be a positive parameter and let c, v be positive constants. Let $X_1, \cdots, X_T$ be i.i.d. random variables with finite mean μ. Suppose that for all $\delta \in (0,1)$, there exists an estimator $\hat{\mu} = \hat{\mu}(T, \delta)$ such that, with probability at least $1-\delta$,*

$$|\mu - \hat{\mu}| \leq v^{\frac{1}{1+\epsilon}} \left(\frac{c \log \frac{2}{\delta}}{T} \right)^{\frac{\epsilon}{1+\epsilon}}.$$

Remark 2. *If the distribution of X_j satisfies σ-sub-Gaussian condition, then Assumption 1 is satisfied for $\epsilon = 1$, $c = 2$, and variance factor $v = \sigma^2$.*

However, Assumption 1 can be satisfied with conditions much weaker than the sub-Gaussian condition. One way is by substituing the empirical mean estimator by some refined mean estimator that gives the exponential tail bound. Specifically, as suggested by Bubeck et al. [2013], we can use Catoni's M estimator [Catoni et al., 2012].

Catoni's M estimator is defined as follows: let $\psi : \mathbb{R} \to \mathbb{R}$ be a continuous strictly increasing function satisfying

$$-\log(1 - x + \frac{x^2}{2}) \leq \psi(x) \leq \log(1 + x + \frac{x^2}{2}).$$

Let $\delta \in (0,1)$ be such that $T > 2\log(\frac{1}{\delta})$ and introduce

$$\alpha_\delta = \sqrt{\frac{2 \log \frac{1}{\delta}}{T(\sigma^2 + \frac{2\sigma^2 \log \frac{1}{\delta}}{T - 2\log \frac{1}{\delta}})}}.$$

If $X_1, \cdots, X_T$ are i.i.d. random variables, the Catoni's estimator is defined as the unique value $\hat{\mu}_C = \hat{\mu}_C(T, \delta)$ such that

$$\sum_{i=1}^{n} \psi(\alpha_\delta(X_i - \hat{\mu}_C)) = 0.$$

Catoni [Catoni et al., 2012] proves that if $T \geq 4 \log \frac{1}{\delta}$ and the X_j have mean μ and variance at most σ^2, then with probability at least $1 - \delta$,

$$|\hat{\mu}_C - \mu| \leq 2\sqrt{\frac{\sigma^2 \log \frac{2}{\delta}}{T}}. \tag{4}$$

The corresponding modification to `Med-dit` is as follows.

1. For the initialization step, sample each point $4 \log \frac{1}{\delta}$ times to meet the condition for the concentration bound of the Catoni's M estimator.

2. For each arm i, if $T_i(t) < n$, maintain the $1 - \delta$ confidence interval $[\hat{\mu}_{C,i} - C_i(t), \hat{\mu}_{C,i} + C_i(t)]$, where $\hat{\mu}_{C,i}$ is the Catoni's estimator of μ_i, and

$$C_i(t) = 2\sqrt{\frac{\sigma^2 \log \frac{2}{\delta}}{T_i(t)}}.$$

Proposition 1. *For $i \in [n]$, let $\Delta_i = \mu_i - \mu^*$. If we pick $\delta = \frac{1}{n^3}$ in the above algorithm, then with probability $1 - o(1)$, it returns the true medoid with the with number of distance evaluations M such that,*

$$M \leq 12n \log n + \sum_{i \in [n]} \left(\frac{48\sigma^2}{\Delta_i^2} \log n \wedge 2n \right)$$

Proof. Let $\delta = \frac{2}{n^3}$. The initialization step takes an extra $12n \log n$ distance computations. Following the same proof as Theorem 1, we can show that the modified algorithm returns the true medoid with probability at least $1 - \Theta(\frac{1}{n})$, and apart from the initialization, the total number of distance computations can be upper bounded by

$$\sum_{i \in [n]} \left(\frac{48\sigma^2}{\Delta_i^2} \log n \wedge 2n \right).$$

So the total number of distance computations can be upper bounded by

$$M \leq 12n \log n + \sum_{i \in [n]} \left(\frac{48\sigma^2}{\Delta_i^2} \log n \wedge 2n \right).$$

□

Remark 3. *By using the Catoni's estimator, instead of the sub-Gaussian assumption, we only require a much weaker assumption that the distance evaluations have finite variance. Yet, we achieve the same order of the distance evaluation complexity.*

4.2 On the $O(n \log \log n)$ Algorithm

The best-arm algrithm `exponential-gap` [Karnin et al., 2013] can be directly applied on the medoid problem, which takes $O(\sum_{i \neq i^*} \Delta_i^{-2} \log \log \Delta_i^{-2})$ distance evaluations, essentially $O(n \log \log n)$ if Δ_i are constants. It is an variation of the familiy of action elimination algorithm for the best-arm problem. A typical action elimination algorithm proceeds as follows: Maintaining a set Ω_k for $k = 1, 2, \cdots$, initialized as $\Omega_1 = [n]$. Then it proceeds in epoches by sampling the arms in Ω_k a predetermined number of times r_k, and maintains arms according to the rule:

$$\Omega_{k+1} = \{i \in \Omega_k : \hat{\mu}_a + C_a(t) < \hat{\mu}_i - C_i(t)\},$$

where $a \in \Omega_k$ is a reference arm, e.g. the arm with the smallest $\hat{\mu}_i + C_i(t)$. Then the algorithm terminates when Ω_k contains only one element.

The above vanilla version of the action elimination algorithm takes $O(n \log n)$ distance evaluations, same as `Med-dit`. The improvement by `exponential-gap` is by observing that the suboptimal $\log n$ factor is due to the large deviations of $|\hat{\mu}_a - \mu_a|$ with $a = \arg\min_{i \in \Omega_k} \hat{\mu}_i$. Instead, `exponential-gap` use a subroutine `median elimination` [Even-Dar et al., 2006] to determine an alternative reference arm a with smaller deviations and allows for the removal of the $\log n$ term, where `median elimination` takes $O(\frac{n}{\epsilon^2} \log \frac{1}{\delta})$ distance evaluations to return a ϵ-optimal arm. However, this will introduce a prohibitively large constant due to the use of `median elimination`. Regarding the technical details, we note both paper [Karnin et al., 2013, Even-Dar et al., 2006] assume the boundedness of the random variables for their proof, which is only used to have the hoeffding concentration bound. Therefore, with our sub-Gaussian assumption, the proof will follow symbol by symbol, line by line.