Robust Outlier Arm Identification

Yinglun Zhu [1] **Sumeet Katariya** [2] **Robert Nowak** [1]

Abstract

We study the problem of Robust Outlier Arm Identification (ROAI), where the goal is to identify arms whose expected rewards deviate substantially from the majority, by adaptively sampling from their reward distributions. We compute the outlier threshold using the median and median absolute deviation of the expected rewards. This is a robust choice for the threshold compared to using the mean and standard deviation, since it can identify outlier arms even in the presence of extreme outlier values. Our setting is different from existing pure exploration problems where the threshold is pre-specified as a given value or rank. This is useful in applications where the goal is to identify the set of promising items but the cardinality of this set is unknown, such as finding promising drugs for a new disease or identifying items favored by a population. We propose two δ-PAC algorithms for ROAI, which includes the first UCB-style algorithm for outlier detection, and derive upper bounds on their sample complexity. We also prove a matching, up to logarithmic factors, worst case lower bound for the problem, indicating that our upper bounds are generally unimprovable. Experimental results show that our algorithms are both robust and about 5x sample efficient compared to state-of-the-art.

1. Introduction

Multi-armed bandits are commonly used to identify the optimal items (arms) among multiple candidates through adaptive queries (pure exploration setting (Jamieson & Nowak, 2014)). Every item is associated with an unknown probability distribution, and when a bandit algorithm selects (pulls) an item, it observes a value (reward) sampled from this distribution. Depending on its objective and the history of observed values, the bandit algorithm has to decide which item to sample at every time t, so as to identify the optimal items using as few samples as possible. Pure exploration bandit algorithms have been proposed for various objectives, such as identifying arms with the largest rewards (Jamieson et al., 2014; Jamieson & Nowak, 2014; Chen et al., 2016), identifying arms above a given threshold (Locatelli et al., 2016; Mukherjee et al., 2017; Xu et al., 2019) or clustering arms (Katariya et al., 2018; 2019).

In this paper, we study bandit algorithms for identifying outlier arms. Outlier arms are defined as those with expected rewards that are outliers relative to the overall set of expected rewards (e.g., arms with expected rewards that are several deviations above the mean/median of the overall set of expected rewards). The outlier detection problem has wide applications in scientific discovery (Grün et al., 2015; Chaudhary et al., 2015), fraud detection (Porwal & Mukund, 2018), medicine (Schiff et al., 2017), and public health (Hauskrecht et al., 2013). In contrast to passive outlier detection algorithms which identify outlier items using a pre-sampled dataset, bandit algorithms actively query items with the goal of identifying outliers using as few samples as possible. This is important because it can lead to early detection of fraud for example. Outlier arms subsume *good* arms with expected rewards substantially above the average, and most applications mentioned in good arm identification (Kano et al., 2019) apply to our setting.

As observed in Zhuang et al. (2017), bandit outlier detection cannot be reduced to best arm(s) identification in bandits because of the inherent double exploration dilemma - the threshold is unknown and any algorithm must balance exploring individual arms and exploring the outlier threshold. Zhuang et al. (2017) define the outlier threshold $\bar{\theta}$ using the k-sigma rule applied to the mean $\bar{\mu}$ and standard deviation $\bar{\sigma}$ of the expected rewards i.e., $\bar{\theta} = \bar{\mu} + k \cdot \bar{\sigma}$. However this threshold can fail to identify the correct outlier arms because the mean and standard deviation are themselves sensitive to outlier values (non-robust estimators). It can also miss outliers when the number of arms is small. In this paper, we define the threshold using the k-sigma rule applied to the *median* and the *median absolute deviation*, which are robust estimators with the highest possible breakdown point 0.5. This is the recommended practice in literature (Hampel, 1974; Huber, 2004; Swallow & Kianifard, 1996), and emphasized by Leys et al. (2013) in their aptly titled paper:

[1]University of Wisconsin-Madison [2]Amazon. Correspondence to: Yinglun Zhu <yinglun@cs.wisc.edu>.

Proceedings of the 37^{th} International Conference on Machine Learning, Online, PMLR 119, 2020.

"Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median". Similarly, Chung et al. (2008) conduct extensive experiments to compare the two methods and show that the median-based threshold identifies outliers that were missed by the mean-based threshold. We show through our theoretical and empirical results that this robust threshold not only identifies outliers more accurately, but it also requires fewer samples to do so than the mean-based threshold.

1.1. Contributions and Paper Organization

We make the following contributions. In Section 2 we formally define the Robust Outlier Arm Identification (ROAI) problem with justifications from Huber's ϵ-contamination model. In Section 3, we propose two algorithms for the ROAI problem, which includes the first UCB-style algorithm for outlier detection. We theoretically prove the correctness our algorithms and derive their sample complexity upper bounds in Section 4. A matching, up to logarithmic factors, worst case lower bound is provided in Section 5, indicating our upper bounds are generally tight. We further generalize our algorithms to settings with known contamination upper bound in Section 6. Experiments conducted in Section 7 show that our algorithms are *both* more robust and more sample efficient than previous state-of-the-art. We conclude our paper in Section 8 with open problems. All proofs are deferred to the Appendix due to lack of space.

1.2. Related Work

The pure exploration problem in the multi-armed bandit setting has a long history, starting from the work of (Bechhofer, 1958; Paulson et al., 1964). The aim of pure exploration is to identify an arm or arms with certain properties. For example, the best-arm identification problem involves correctly deciding which arm has the largest expected reward. The instance-dependent sample complexity bound on the best arm identification problem was analyzed/improved by (Even-Dar et al., 2002; 2006; Gabillon et al., 2012; Karnin et al., 2013; Jamieson et al., 2014; Jamieson & Nowak, 2014; Chen et al., 2016). The problem was also generalized to the setting of identifying the top-m arm (Kalyanakrishnan et al., 2012; Chen et al., 2017a;b); the thresholding bandit (Locatelli et al., 2016; Mukherjee et al., 2017; Xu et al., 2019) which identifies all arms with expected reward above a given threshold θ; and the good arm identification problem (Kano et al., 2019; Katz-Samuels & Jamieson, 2019), where for a given ϵ "good arms" have expected reward within ϵ of the largest. Lower bounds developed in the pure exploration setting (Mannor & Tsitsiklis, 2004; Chen & Li, 2015; Kaufmann et al., 2016; Garivier & Kaufmann, 2016; Simchowitz et al., 2017) shed light on the optimality of existing algorithms.

In all of the above settings, the subset of arms of interest is determined by a user-defined parameter, e.g., m, θ, and ϵ. Outlier arm identification cannot be cast in these settings, since the cut-off cannot be a prespecified threshold or rank. The cut-off depends on the overall distribution of expected rewards, which is unknown in advance. In other words, outlier arm identification has an *instance-dependent identification target*. Bandit problems with instance-dependent identification targets have attracted some attention recently. One of the work (Katariya et al., 2019) studies the problem of identifying the largest gap in the ordering of the expected rewards, which provides a natural separation of the arms into two groups or clusters. Another line of work (Zhuang et al., 2017) focuses on identifying outlier arms with an outlier threshold adaptive to the bandit instance. Specifically, they use the threshold $\bar{\theta} = \bar{\mu} + k \cdot \bar{\sigma}$, with $\bar{\mu}$ and $\bar{\sigma}$ being the mean and standard deviation of distribution of expected rewards, respectively. The parameter k is usually chosen as 2 or 3 according to the famous three-sigma rule.

Our work focuses on robust and sample-efficient approaches to the outlier arm identification problem. We model our setting through Huber's ϵ-contamination model (Huber et al., 1964) and apply robust estimators with the highest possible breakdown point (Donoho & Huber, 1983; Rousseeuw & Hubert, 2011), i.e., median and median absolute deviation (MAD), in building the outlier threshold. Robust statistics were previously incorporated in the bandit setting (Altschuler et al., 2019), but they mainly deal with traditional settings, i.e., best arm identification, with each reward distribution being contaminated rather than identifying instance-adaptive subsets. Although our work could also be generalized to the setting with contaminated reward distribution by incorporating their techniques, we do not pursue this direction here.

2. Problem Setting and Notations

We consider the standard multi-armed bandit setting where there are n arms and the reward of each arm follows a 1-subgaussian distribution with mean y_i. The goal of the agent is to identify outlier arms whose expected rewards *substantially* deviate from the *majority*, in the fixed confidence and pure exploration setting. Without loss of generality, we assume $y_i \geq y_{i+1}$ and $n = 2m - 1$, so that the median arm is unambiguous.[1] We also only consider identifying outliers with high rewards; identifying outliers with low rewards is analogous. Let $y_{(m)} = \text{median}\{y_i\}$ denote the expected reward of the *median* arm, and let $\text{AD}_i = |y_i - y_{(m)}|$ represent the absolute deviation of arm i from the median. Let $\text{AD}_{(m)} = \text{median}\{|y_i - y_{(m)}|\}$ denote the Median Absolute Deviation (MAD) of expected reward. Note that $y_{(m)}$

[1] If $n = 2m$, we choose the median as m without loss of generality.

and $\mathrm{AD}_{(m)}$ serve as the first two robust moments of the means of the underlying bandit instance $\{y_i\}_{i=1}^n$. We define outlier arms to be arms whose mean is greater than the threshold θ given by

$$\theta = y_{(m)} + k \cdot \mathrm{AD}_{(m)}, \tag{1}$$

where $k \geq 1$ is a user-specified parameter. The goal of the agent is to *identify outlier arms using as few samples as possible*. Specifically, we are interested in designing adaptive algorithms that return the subset of outlier arms $S_o = \{i \in [n] : y_i > \theta\}$ (we assume $y_i \neq \theta$, $\forall\, i \in [n]$). We call this setting *Robust Outlier Arm Identification* (ROAI). For a given error probability $\delta \in (0, 1)$, we say an algorithm is δ-PAC if it correctly identifies S_o with probability at least $1 - \delta$ using a finite number of samples.

Our choice of the threshold is justified under Huber's ϵ-contamination model, where with probability $1-\epsilon$ the mean y_i is drawn from an unknown *meta* distribution P with mean μ and standard deviation σ, and with probability ϵ the mean y_i is drawn from a contamination distribution. Note that sample median and MAD enjoy the highest possible breakdown point 0.5 (Donoho & Huber, 1983; Rousseeuw & Hubert, 2011). Hence, our threshold in Eq. (1) (up to scaling of $\mathrm{AD}_{(m)}$) is a more robust estimator of the true threshold as compared to existing thresholds constructed using the sample mean and sample standard deviation (which have a breakdown point of 0) (Zhuang et al., 2017). Furthermore, for many common meta distributions including the normal and uniform distribution, Altschuler et al. (2019) prove tight non-asymptotic concentration results for the median and MAD constructed from contaminated samples.

Given our assumption of $y_i \geq y_{i+1}$, let the outlier set be $S_o = \{1, \ldots, n_1\}$ where n_1 is *unknown*. For a given set $\{z_i\}_{i=1}^n$, we use $z_{(k)}$ to denote the k-th largest value in $\{z_i\}$; particularly, we use $z_{(m)} := \mathrm{median}\{z_i\}$.

3. Algorithms

We formally introduce our algorithms in the section. We first provide a subroutine for constructing confidence intervals (CIs) of various quantities including the outlier threshold in Section 3.1; and then introduce our elimination- and LUCB-style algorithms in Section 3.2.

For any arm $i \in [n]$ and time t, we use $S_{i,t}$ and $N_{i,t}$ to denote the sum of rewards and number of pulls; and use $\hat{y}_{i,t} = S_{i,t}/N_{i,t}$ to denote the empirical mean reward. For any quantity $q \in \{y_i, y_{(m)}, \mathrm{AD}_i, \mathrm{AD}_{(m)}, \theta\}$, we use $L_{q,t}, U_{q,t}, \mathcal{I}_{q,t}$ to denote the lower bound, upper bound, and the CI respectively of q at time t.

3.1. Construction of Confidence Intervals (CIs)

The CI of individual arms i can easily be constructed using Hoeffding's inequality as $[L_{y_i,t}, U_{y_i,t}] = [\hat{y}_{i,t} - \beta_{N_{i,t}}, \hat{y}_{i,t} - \beta_{N_{i,t}}]$, where $\beta_s = \sqrt{\log(4ns^2/\delta)/2s}$.

The construction of CIs for the median ($\mathcal{I}_{y_{(m)},t}$), MAD ($\mathcal{I}_{\mathrm{AD}_{(m)},t}$), and the outlier threshold ($\mathcal{I}_{\theta,t}$), which are needed for ascertaining whether an arm is an outlier, is explained in Algorithm 1. On line 1, the CI $\mathcal{I}_{y_{(m)},t}$ is constructed using the CIs of all arms. This is necessary because the identity of the median arm may be unknown. If the median arm can be unambiguously determined, this CI reduces to the CI of the median-th arm. The CI $\mathcal{I}_{\mathrm{AD}_{(m)},t}$ is similarly constructed from $\mathcal{I}_{\mathrm{AD}_i,t}$. We set $\widehat{\mathrm{AD}}_{i,t}$ and $\hat{\theta}_t$ as the midpoint of their corresponding confidence intervals.

Algorithm 1 Construction of Confidence Intervals

Input: CIs of individual arms $\{\mathcal{I}_{y_i,t}\}_{i=1}^n$
Output: CIs $\mathcal{I}_{y_{(m)},t}, \mathcal{I}_{\mathrm{AD}_i,t}, \mathcal{I}_{\mathrm{AD}_{(m)},t}, \mathcal{I}_{\theta,t}$

1: $L_{y_{(m)},t} = \mathrm{median}\{L_{y_i,t}\}$
$U_{y_{(m)},t} = \mathrm{median}\{U_{y_i,t}\}$
$\mathcal{I}_{y_{(m)},t} = [L_{y_{(m)},t}, U_{y_{(m)},t}]$
2: **for** $i = 1, \ldots, n$ **do**
3: $L_{\mathrm{AD}_i,t} = \max\{L_{y_i,t} - U_{y_{(m)},t}, L_{y_{(m)},t} - U_{y_i,t}\}$
$U_{\mathrm{AD}_i,t} = \max\{U_{y_i,t} - L_{y_{(m)},t}, U_{y_{(m)},t} - L_{y_i,t}\}$
$\mathcal{I}_{\mathrm{AD}_i,t} \in [L_{\mathrm{AD}_i,t}, U_{\mathrm{AD}_i,t}]$
$\widehat{\mathrm{AD}}_{i,t} = (U_{\mathrm{AD}_i,t} + L_{\mathrm{AD}_i,t})/2$
4: **end for**
5: $L_{\mathrm{AD}_{(m)},t} = \mathrm{median}\{L_{\mathrm{AD}_i,t}\}$
$U_{\mathrm{AD}_{(m)},t} = \mathrm{median}\{U_{\mathrm{AD}_i,t}\}$
$\mathcal{I}_{\mathrm{AD}_{(m)},t} = [L_{\mathrm{AD}_{(m)},t}, U_{\mathrm{AD}_{(m)},t}]$
6: $L_{\theta,t} = L_{y_{(m)},t} + k \cdot L_{\mathrm{AD}_{(m)},t}$
$U_{\theta,t} = U_{y_{(m)},t} + k \cdot U_{\mathrm{AD}_{(m)},t}$
$\mathcal{I}_{\theta,t} = [L_{\theta,t}, U_{\theta,t}]$ and $\hat{\theta}_t = (U_{\theta,t} + L_{\theta,t})/2$

3.2. Algorithms

We introduce our elimination-style (Even-Dar et al., 2006) algorithm `ROAIElim` and LUCB-style (Kalyanakrishnan et al., 2012) algorithm `ROAILUCB` in this section. Any pure exploration bandit algorithm is specified through its sampling, stopping, and recommendation rule (Kaufmann et al., 2016). The stopping and recommendation rules are the same for both algorithms. We stop at the first time t such that $\{i \in [n] : \mathcal{I}_{y_i,t} \cap \mathcal{I}_{\theta,t} \neq \emptyset\} = \emptyset$, and upon stopping we output the empirical subset of outlier arms $\hat{S}_{o,t} = \{i \in [n] : \hat{y}_{i,t} > \hat{\theta}_t\}$. We present our two algorithms next.

ROAIElim: The pseudocode of ROAIElim is given in Algorithm 2. At round t, ROAIElim constructs three active sets for the median, the MAD, and the threshold. Each of these active sets contains arms whose CIs overlap with the respective CI. Since the threshold is constructed from the median and the MAD, any of these arms can contribute towards shrinking the CI of the threshold, and hence ROAIElim samples all arms in the union of these active sets.

Algorithm 2 ROAIElim

Input: Error tolerance ϵ, probability of failure δ, and outlier detection parameter k
Output: Subset of outlier arms $\hat{S}_{o,t}$

Initialize $A_{E,1} = A_{E,1}^{\text{median}} = A_{E,1}^{\text{MAD}} = A_{E,1}^{\theta} = [n]$
for $t = 1, 2, \dots$ **do**
  Sample arms in $A_{E,t}$ and update $\{\mathcal{I}_{i,t}\}_{i \in A_{E,t}}$
  Update CIs using Algorithm 1
  Set

$$A_{E,t+1}^{\text{median}} = \{i \in [n] : \mathcal{I}_{y_i,t} \cap \mathcal{I}_{y_{(m)},t} \neq \emptyset\} \cap A_{E,t}^{\text{median}}$$

$$A_{E,t+1}^{\text{MAD}} = \{i \in [n] : \mathcal{I}_{\text{AD}_i,t} \cap \mathcal{I}_{\text{AD}_{(m)},t} \neq \emptyset\} \cap A_{E,t}^{\text{MAD}}$$

$$A_{E,t+1}^{\theta} = \{i \in [n] : \mathcal{I}_{y_i,t} \cap \mathcal{I}_{\theta,t} \neq \emptyset\} \cap A_{E,t}^{\theta}$$

$$A_{E,t+1} = A_{E,t+1}^{\text{median}} \cup A_{E,t+1}^{\text{MAD}} \cup A_{E,t+1}^{\theta}$$

  If $A_{E,t+1}^{\theta} = \emptyset$, stop and return $\hat{S}_{o,t}$
end for

ROAILUCB: The pseudocode of ROAILUCB is presented in Algorithm 3. We use the notation $J_{\kappa_i,t}$ to denote κ_i arms with the largest empirical means $\{\hat{y}_{i,t}\}$, and $J_{\kappa_i,t}^{\text{AD}}$ to denote the κ_i arms with the largest empirical absolute deviations $\{\widehat{\text{AD}}_{i,t}\}$. Since we are mainly interested in shrinking confidence intervals around the median quantity, we set $\kappa_1 = m - 1$ and $\kappa_2 = m$.

Motivated by the LUCB algorithm (Kalyanakrishnan et al., 2012), ROAILUCB finds the 4 arms at the median boundary, 4 arms at the MAD boundary, and 2 arms at the threshold boundary, and samples arms in the union of these sets. Unlike ROAIElim, ROAILUCB samples at most 10 arms in each round.

4. Analysis

In Section 4.1, we discuss correctness and sample complexity results of our algorithms. We compare the robustness and sample complexity of our algorithms with previous work in Section 4.2. The proofs can be found in the Appendix.

Algorithm 3 ROAILUCB

Input: Error tolerance ϵ, probability of failure δ, and outlier detection parameter k
Output: Subset of outlier arms $\hat{S}_{o,t}$

Initialize $A_{L,1} = [n]$
for $t = 1, 2, \dots$ **do**
  Sample arms in $A_{L,t}$ and update $\{\mathcal{I}_{y_i,t}\}_{i \in A_{L,t}}$
  Update CIs using Algorithm 1
  Set

$$A_{L,t+1}^{\text{median}} = \arg\min_{i \in J_{m-1,t}} \{L_{y_i,t}\} \cup \arg\min_{i \in J_{m,t}} \{L_{y_i,t}\} \cup \arg\max_{i \notin J_{m-1,t}} \{U_{y_i,t}\} \cup \arg\max_{i \notin J_{m,t}} \{U_{y_i,t}\}$$

$$A_{L,t+1}^{\text{MAD}} = \arg\min_{i \in J_{m-1,t}^{\text{AD}}} \{L_{\text{AD}_i,t}\} \cup \arg\min_{i \in J_{m,t}^{\text{AD}}} \{L_{\text{AD}_i,t}\} \cup \arg\max_{i \notin J_{m-1,t}^{\text{AD}}} \{U_{\text{AD}_i,t}\} \cup \arg\max_{i \notin J_{m,t}^{\text{AD}}} \{U_{\text{AD}_i,t}\}$$

$$A_{L,t+1}^{\theta} = \arg\min_{i \in \hat{S}_{o,t}} \{L_{y_i,t}\} \cup \arg\max_{i \in \hat{S}_{n,t}} \{U_{y_i,t}\} \cap \{i \in [n] : \mathcal{I}_{y_i,t} \cap \mathcal{I}_{\theta,t} \neq \emptyset\}$$

$$A_{L,t+1} = A_{L,t+1}^{\text{median}} \cup A_{L,t+1}^{\text{MAD}} \cup A_{L,t+1}^{\theta}$$

  If $A_{L,t+1}^{\theta} = \emptyset$, stop and return $\hat{S}_{o,t}$
end for

4.1. Correctness and Sample Complexity

Lemma 1 shows the correctness of CIs in Algorithm 1. We use it to prove the correctness of our algorithms in Theorem 1.

Lemma 1. *Suppose*

$$\mathbb{P}\left(\forall t \in \mathbb{N}, \forall i \in [n], y_i \in \mathcal{I}_{y_i,t}\right) \geq 1 - \delta.$$

Then the CIs returned by Algorithm 1 are valid with probability $1 - \delta$, i.e., for $q \in \{y_{(m)}, \{\text{AD}_i\}_{i=1}^n, \text{AD}_{(m)}, \theta\}$,

$$\mathbb{P}\left(\forall t \in \mathbb{N}, q \in \mathcal{I}_{q,t}\right) \geq 1 - \delta.$$

Theorem 1 (Correctness)**.** ROAIElim *and* ROAILUCB *are δ-PAC.*

In order to state our sample complexity bounds, we first introduce some new notations. Define

$$\Delta_i^{\theta} = |\theta - y_i|, \quad \Delta_*^{\theta} = \min_{i \in [n]} \{\Delta_i^{\theta}\},$$
$$\Delta_i^{\text{median}} = |y_{(m)} - y_i|, \quad \Delta_i^{\text{MAD}} = |\,\text{AD}_{(m)} - \text{AD}_i\,|,$$
$$\Delta_i^* = \max\{\Delta_*^{\theta}, \min\{\Delta_i^{\theta}, \Delta_i^{\text{median}}, \Delta_i^{\text{MAD}}\}\}. \tag{2}$$

Theorem 2 (Sample Complexity). *With probability at least $1-\delta$, the sample complexity of* ROAIElim *and* ROAILUCB *is upper bounded by*

$$Ck^2 \sum_{i=1}^{n} \frac{\log\left(nk/\delta\Delta_i^*\right)}{(\Delta_i^*)^2}, \tag{3}$$

where C is a universal constant.

The sample complexity is inversely proportional to Δ_i^* defined in Eq. (2). In order to interpret the sample complexity, we consider two cases. If there exists arms whose means are close to the threshold θ, i.e., Δ_*^θ is small, then in order to classify these arms correctly, we need to estimate θ and consequently the median and the MAD accurately. Hence the complexity of sampling an arm depends on its gaps from $y_{(m)}, \mathrm{AD}_{(m)}, \theta$. Conversely, if all the arm means are widely separated from the threshold, i.e., Δ_*^θ is large and there is a clear distinction between normal and outlier arms, then we do not need to estimate θ accurately, and the sample complexity is $O(n/(\Delta_*^\theta)^2)$.

We highlight that the proof of Theorem 2 is non-trivial and cannot be reduced to existing techniques. The existing works (Kalyanakrishnan et al., 2012; Katariya et al., 2018) deal with scenarios where the positions of the separating boundaries depend only on the arm means, and furthermore they are user-specified. This holds true only for the median in our case, it does not hold for the AD, MAD, and the threshold because their values do not depend on a single arm. The CIs of these estimators have varying degree of uncertainty and we quantify these in our Lemmas. The technical contributions may be of independent interest and we refer the reader to our proofs in the Appendix.

4.2. Comparison to Previous Work

We compare our setting and analysis to algorithms by Zhuang et al. (2017), which is the only work study outlier detection in the bandit setting.

To deal with the *unknown* μ and σ, (Zhuang et al., 2017) use the sample mean $\bar{\mu} = \sum_{i=1}^n y_i/n$ and sample standard deviation $\bar{\sigma} = \sqrt{\sum_{i=1}^n (y_i - \bar{\mu})^2/n}$ to approximate μ and σ, respectively, and define the outlier threshold to be $\bar{\theta} = \bar{\mu} + k \cdot \bar{\sigma}$. As discussed in Section 2, these estimators have a breakdown point 0 and are very sensitive to outliers; a single extreme outlier arm can ruin their threshold.

Algorithms developed in (Zhuang et al., 2017) also require the reward distribution of the arms to be strictly bounded; our analysis is general and works for any sub-gaussian distributions.

Finally, although a direct comparison of sample complexities is not possible due to different definitions of outlier thresholds, we empirically see that our algorithms require fewer samples to achieve the same error rate.

5. Lower bound

In this section, we study lower bound on the expected number of samples needed to identify outlier arms by any δ-PAC algorithm, where the outlier threshold is defined by Eq. (1).

Our lower bound is instance-dependent. Recall that our upper bound scales like $\tilde{O}(\sum_{i\in[n]} 1/(\Delta_i^*)^2)$ where Δ_i^* is given by Eq. (2). The problem is easy when Δ_i^* is large, and the upper bound could potentially be large when Δ_i^θ is small. In this section we argue that this is unavoidable. We show that if Δ_i^θ is small enough, there exists a lower bound that matches the upper bound up to logarithmic factors. This indicates that our sample complexity upper bounds are generally *unimprovable*.

We apply the change of measure technique (Kaufmann et al., 2016), which give a lower bound in terms of the KL-divergence. To connect the KL-divergence to the Euclidean distance in our upper bound, we assume that the reward distribution of each arm is $\mathcal{N}(y_i, 1)$.[2] We use D_{y_i} to denote the distribution $\mathcal{N}(y_i, 1)$ as it is fully characterized by its mean y_i.

For a bandit instance $D_y = (D_{y_1}, \ldots, D_{y_n})$, assume without loss of generality that $y_i \geq y_{i+1}$ and that each arm is unambiguously identifiable as an outlier or normal arm, i.e., $y_i \neq \theta, \forall\, i \in [n]$. We use $\mathbb{E}_y(\cdot)$ to represent the expectation with respect to the bandit instance D_y and randomness in the algorithm. We develop lower bounds for the following subset of bandit instances.

Definition 1. *Let $\mathcal{M}_{n,\rho} = \{D_y = (D_{y_1}, \ldots, D_{y_n}) : y_i \neq \theta\}$ be a subset of bandit instances with θ defined in Eq. (1) and $k \geq 2$, and satisfying the following two conditions.*

1. *There exists a unique median $y_{(m)}$ and a unique MAD $\mathrm{AD}_{(m)}$, with*

$$\eta := 1/2 \cdot \min_{i\in\{m,m-1\}} \left\{ y_{(i)} - y_{(i+1)}, \mathrm{AD}_{(i)} - \mathrm{AD}_{(i+1)} \right\}.$$

2. *There exists a constant $\rho < \eta$ such that at least two arms l_1 and l_2 such that $\rho/2 < \theta - y_{l_i} < \rho$, and at least two arms u_1 and u_2 such that $\rho/2 < y_{u_i} - \theta < \rho$; furthermore, there exists no arm with mean in $[\theta - \rho/2, \theta + \rho/2]$.*

It is easy to see that $\mathcal{M}_{n,\rho} \neq \emptyset$ for reasonably large n. The conditions in Definition 1 are essentially to make sure that slightly changing the median $y_{(m)}$ or the MAD $\mathrm{AD}_{(m)}$ will incur a change in the set of outlier arms. Then, for

[2] The lower bound could be generalized to other distributions, as discussed in (Kaufmann et al., 2016).

any δ-PAC algorithm to correctly identify the subset of outlier arms, it is necessary to accurately identify the outlier threshold, which eventually leads to a matching sample complexity lower bound. We state our lower bound for the subset of bandit instances $\mathcal{M}_{n,\rho}$ next.

Theorem 3. *Suppose bandit instance $D_y \in \mathcal{M}_{n,\rho}$. Then for $\delta \leq 0.15$, any δ-PAC outlier arm identification algorithm $\mathcal{A}$ with outlier threshold constructed as in Eq. (1) and an almost surely finite stopping time τ, we have that*

$$\mathbb{E}_y[\tau] \geq \sum_{i\in[n]} \frac{1}{5\left(\Delta_i^*\right)^2} \log\left(\frac{1}{2.4\delta}\right).$$

In general for bandit instances outside $\mathcal{M}_{n,\rho}$ but with non-empty subset of outlier arms, the outlier identification problem is at least as hard as the top-n_1 arm identification problem where n_1 is the number of outlier arms *given* by an oracle. Thus, any lower bound for top-n_1 arm identification, e.g., Theorem 4 in (Kaufmann et al., 2016), applies as a general lower bound for the outlier arm identification problem.

6. Heuristic to Reduce Sample Complexity

The sample complexity of our algorithms is inversely proportional to $(\Delta_i^*)^2$ (see Eq. (2)), which could be as small as $(\min\{\Delta_i^\theta, \Delta_i^{\text{median}}, \Delta_i^{\text{MAD}}\})^2$ if Δ_*^θ is small. As n increases, there can be many arms with small Δ_i^{median} or Δ_i^{MAD} and the sample complexity can be high as a result. In general, we cannot circumvent this cost if the outlier threshold is constructed as in Eq. (1).

However, it might not be necessary to always construct outlier threshold using all n arms, and one heuristic approach is to construct threshold only from a subset of arms. Suppose we know, from an oracle, an upper bound $c < 0.5$ on the fraction of arms drawn from the contaminated distribution, we could then randomly draw a subset $\Omega \subseteq [n]$ of arms with cardinality $|\Omega| \geq 2\lfloor nc \rfloor + 1$. The cardinality requirement makes sure the fraction of contamination within the subset Ω is smaller than 0.5 so that the median and MAD are not arbitrarily destroyed by outliers; but of course the threshold constructed crucially depends on the selection of Ω. Although the outlier set computed from this modified threshold could differ from the outlier set computed from $[n]$, we could potentially enjoy a smaller sample complexity. We next state an upper bound on the sample complexity in this setting.[3] Empirical examinations of the performance are summarized in Section 7.2.

Corollary 1. *Suppose we run Algorithm 3 with $y_{(m)}$, $\text{AD}_{(m)}$ and θ constructed using arms in $\Omega \subseteq [n]$. Then, with probability at least $1-\delta$, the sample complexity is upper bounded by*

$$Ck^2 \sum_{i\in\Omega} \frac{\log\left(nk/(\delta\Delta_i^*)\right)}{(\Delta_i^*)^2} + C\sum_{i\notin\Omega} \frac{\log\left(n/(\delta\Delta_i^\theta)\right)}{(\Delta_i^\theta)^2},$$

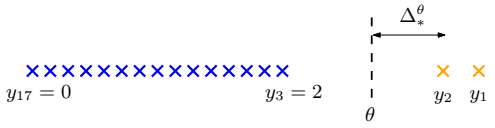
where $\Delta_i^* = \max\{\Delta_*^\theta, \min\{\Delta_i^\theta, \Delta_i^{\text{median}}, \Delta_i^{\text{MAD}}\}\}$ and C is a universal constant.

[3] See Appendix F for details of the algorithm.

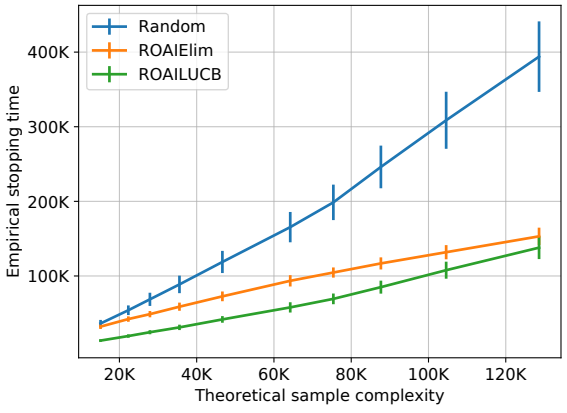
7. Experiments

We conduct three experiments. In Section 7.1, we verify the tightness of our sample complexity upper bounds in Section 4.1. In Section 7.2, we compare our setting to the non-robust version proposed by Zhuang et al. (2017) and empirically confirm the robustness of our thresholds as discussed in Section 4.2. Finally, in Section 7.3, we compare the anytime performance of our algorithms with baselines on a synthetic and a real-world dataset. For ease of comparison, we make the fraction of contamination deterministic rather than random as in the original Huber's contamination model. All our results are averaged over 500 runs. Error bar in Fig. 2, Fig. 4 and Fig. 5 are rescaled by $2/\sqrt{500}$. Our code is publicly available (Zhu et al., 2020).

7.1. Sample Complexity



(a)



(b)

Figure 1. (a) Configuration of the arm means, we vary Δ_*^θ to change hardness (b) Theoretical upper bound vs empirical stopping time, the linear relationship shows that our upper bounds are correct up to constants.

In order to test that the hardness predicted by our upper

bound scales correctly, we first plot the empirical stopping time of each algorithm against the theoretical sample complexity (Theorem 2 with $C = 10$). We choose the arm configuration in Fig. 1(a) containing 15 normal arms (in blue) with fixed means equally distributed from 0 to 2, an outlier threshold $\theta \approx 2.837$, and 2 outlier arms (in orange) above the outlier threshold. The distance between the outlier arms is fixed at 0.2. We decrease Δ_*^θ from 0.6 to 0.2, and this changes the theoretical sample complexity. Note that the threshold does not change. The reward of each arm is normally distributed with standard deviation 0.5. In Fig. 1(b), we plot the empirical stopping time of our algorithms against the theoretical sample complexity, and we see a linear relationship between the two, which suggests that our sample complexity in Theorem 2 is correct up to constants. Fig. 1(b) also shows that our adaptive algorithms always outperform random sampling, and the gains increase with the hardness of the problem.

7.2. Setting Comparison

In this section, we compare the robustness of our outlier threshold and the sample complexity upper bound of our algorithms to the threshold and algorithms considered by Zhuang et al. (2017). We introduce the nomenclature of the algorithms next. Round Robin (RR) and Weighted Round Robin (WRR) are algorithms proposed by Zhuang et al. (2017) which use a non-robust outlier threshold. We denote by ROAI-$\lambda n\epsilon$ the algorithm suggested in Section 6 that constructs the outlier threshold from a subset Ω of arms with $|\Omega| = \max\{\lambda\lfloor n\epsilon\rfloor + 1, 15\}$. For each run of this experiment, we generate the means of normal arms from $\mathcal{N}(0.3, 0.075^2)$ (clipped to the three-sigma range), and the means of outlier arms from $\texttt{Unif}(x, 1)$. We draw 105 arms in total. We multiply MAD with $1/(\Phi^{-1}(3/4)) \approx 1.4826$ to make it consistent for the true scale of normal distribution (Leys et al., 2013).

We first study robustness. In Fig. 2, we generate outlier arms from $\texttt{Unif}(0.7, 1)$ and vary the fraction ϵ of contaminated arms from 0 to 0.2, and compare the robustness of the proposed outlier threshold from different algorithms. We measure the robustness as deviation of the proposed threshold from the true threshold. The true threshold is chosen according to the three-sigma rule. It is clear that our outlier thresholds are much more robust to contamination.

We next compare the upper bounds on the sample complexity of different algorithms. We generate 10 outlier arms from $\texttt{Unif}(x, 1)$ with x varying from 0.6 to 0.9. In Fig. 3, we plot the median sample complexity upper bounds of each algorithm in log scale, ignoring universal constants. We notice that under these contamination settings, our sample complexity upper bounds are orders of magnitude smaller than the ones proposed in Zhuang et al. (2017). From Fig. 2 and Fig. 3, we also see the trade-off between robustness and sample complexity for our generalized algorithms suggested in Section 6.

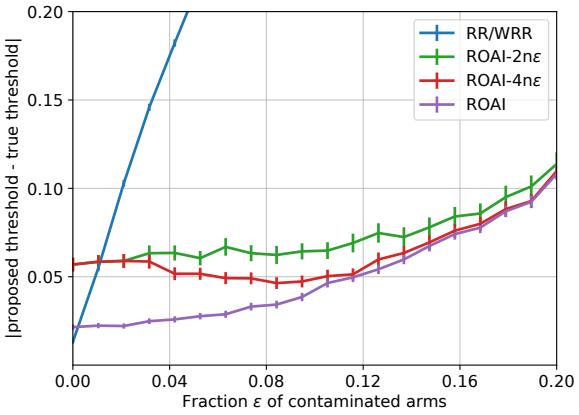


Figure 2. Deviation of the proposed outlier threshold from the true threshold as a function of the contamination level ϵ. This shows that our threshold is robust to contamination.

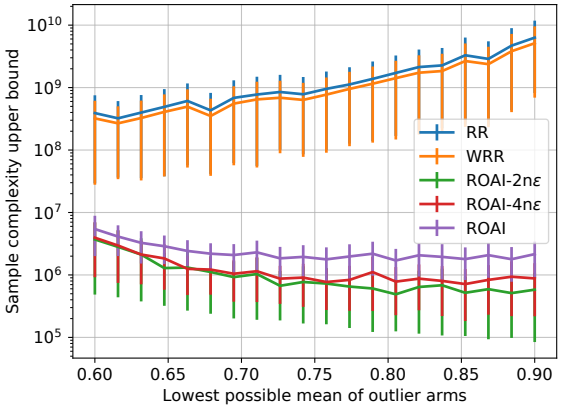


Figure 3. Sample complexity upper bounds as a function of the lowest possible mean of outlier arms, our upper bounds are smaller.

7.3. Anytime Performance

In this section, we examine the anytime empirical error rate of ROAILUCB, ROAIElim, random sampling and RR/WRR (Zhuang et al., 2017). Similar to Section 7.2, we generate 100 normal arm means from $\mathcal{N}(0.3, 0.075^2)$ and 5 outlier means from $\texttt{Unif}(0.8, 1)$. We draw rewards of each arm from a Bernoulli distribution with respect to its mean. We use Bernoulli distributions here as algorithms in Zhuang et al. (2017) only apply to arms with a strictly bounded

distribution. In order to simulate a run, we randomly draw means according to these two distributions and then draw rewards from these arms with fixed means till the end of the run. Under this setting, both our threshold (median-MAD) and the threshold in Zhuang et al. (2017) (mean-standard deviation) will lie in $[0.525, 0.8]$ with high probability. We filter out instances where the outlier sets (with respect to both thresholds) do not match the ground truth. The averaged minimum gap $\min\{|y_i - \theta|\}$ is 0.062 according to our threshold, and 0.063 according to theirs. In Fig. 4, we plot the fraction of times any algorithm fails to identify the correct set of outlier arms. We notice that `ROAILUCB` requires about 5x fewer samples than RR/WRR for the same error rate. Notice that RR is essentially random sampling with their threshold, and hence we use our threshold in the algorithm labeled Random. The empirical performance of RR/WRR is worse than Random.

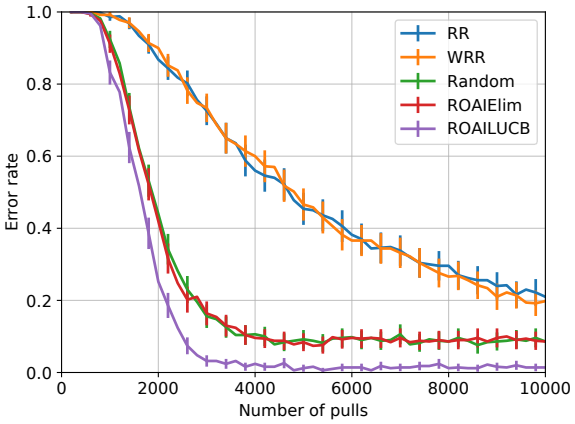


Figure 4. Fraction of times the outlier set is misidentified on synthetic data.

We also compare the performance of all algorithms on the real-world Wine Quality dataset (Sathe & Aggarwal, 2016), which is widely used to compare outlier detection algorithms. This dataset contains 129 wines, each having 13 features. 10 of these wines are labeled as outliers in the dataset. To obtain a 1d representation of each wine, we projected data points on the first principal component and then rescaled them to $[0, 1]$. We deleted 6 values closest to the threshold in this 1d representation so that the outlier set is the same according to both definitions. The 123 means thus obtained are plotted in Fig. 5(a) with the top-5 outliers in orange. We simulate each arm as a Bernoulli distribution. As in the previous experiment, `ROAILUCB` greatly outperform other algorithms, and RR/WRR is worse than random sampling.

The fact that `ROAIElim` performs similar to random sampling in terms of the anytime error rate is not new (Jamieson et al., 2014), elimination-style algorithms are very conservative initially. Fig. 1(b) does show that `ROAIElim` outperforms random sampling in terms of the empirical stopping time.

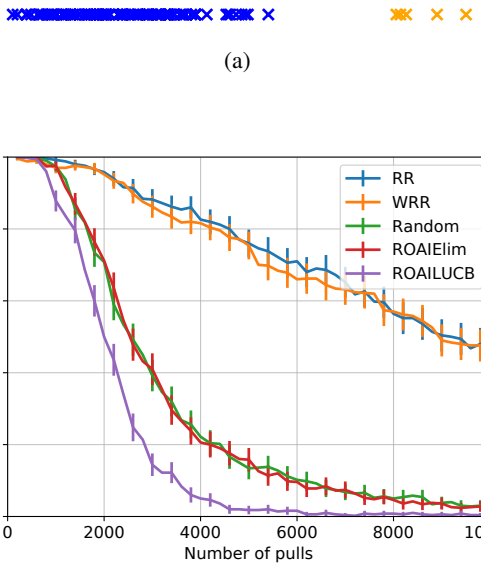


Figure 5. (a) 1d means obtained from the Wine Quality dataset (b) Fraction of times the outlier set is misidentified on this dataset.

8. Conclusion

This paper studies robust outlier arm identification problem, a pure exploration problem with instance-adaptive identification target in the multi-armed bandit setting. We propose two algorithms `ROAIElim` and `ROAILUCB`, and theoretically derive their correctness and sample complexity upper bounds. We also provide a matching, up to log factors, worst case lower bound, indicating our upper bounds are generally tight. We conduct experiments to show our algorithms are both robust and about 5x sample efficient compared to state-of-the-art.

We leave open several questions. First, the sample complexity of our algorithms is large when Δ_*^θ is small. We propose a heuristic to partially address this issue if an upper bound on the contamination ϵ is known in Section 6. Another potential approach is to add an error tolerance to allow arms close the threshold being misclassified, but that adds another user-specific parameter. We also leave open the problem of obtaining a tight instance dependent lower bound. Our current lower bound, even though instance-dependent, works only in the worst case, and we reduce the problem to top-n_1 arm identification in the general case.

Acknowledgements

Yinglun Zhu would like to thank Ardhendu Tripathy for helpful discussions, and thank Tuan Dinh for help with experiments in the early stage of this project.

References

Altschuler, J., Brunel, V.-E., and Malek, A. Best arm identification for contaminated bandits. *Journal of Machine Learning Research*, 20(91):1–39, 2019.

Bechhofer, R. E. A sequential multiple-decision procedure for selecting the best one of several normal populations with a common unknown variance, and its use with various experimental designs. *Biometrics*, 14(3):408–429, 1958.

Chaudhary, P., Naganathan, A. N., and Gromiha, M. M. Folding race: a robust method for predicting changes in protein folding rates upon point mutations. *Bioinformatics*, 31(13):2091–2097, 2015.

Chen, J., Chen, X., Zhang, Q., and Zhou, Y. Adaptive multiple-arm identification. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 722–730. JMLR. org, 2017a.

Chen, L. and Li, J. On the optimal sample complexity for best arm identification. *arXiv preprint arXiv:1511.03774*, 2015.

Chen, L., Li, J., and Qiao, M. Towards instance optimal bounds for best arm identification. *arXiv preprint arXiv:1608.06031*, 2016.

Chen, L., Li, J., and Qiao, M. Nearly instance optimal sample complexity bounds for top-k arm selection. *arXiv preprint arXiv:1702.03605*, 2017b.

Chung, N., Zhang, X. D., Kreamer, A., Locco, L., Kuan, P.-F., Bartz, S., Linsley, P. S., Ferrer, M., and Strulovici, B. Median absolute deviation to improve hit selection for genome-scale rnai screens. *Journal of biomolecular screening*, 13(2):149–158, 2008.

Donoho, D. L. and Huber, P. J. The notion of breakdown point. *A festschrift for Erich L. Lehmann*, 157184, 1983.

Even-Dar, E., Mannor, S., and Mansour, Y. Pac bounds for multi-armed bandit and markov decision processes. In *International Conference on Computational Learning Theory*, pp. 255–270. Springer, 2002.

Even-Dar, E., Mannor, S., and Mansour, Y. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of machine learning research*, 7(Jun):1079–1105, 2006.

Gabillon, V., Ghavamzadeh, M., and Lazaric, A. Best arm identification: A unified approach to fixed budget and fixed confidence. In *Advances in Neural Information Processing Systems*, pp. 3212–3220, 2012.

Garivier, A. and Kaufmann, E. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pp. 998–1027, 2016.

Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and Van Oudenaarden, A. Single-cell messenger rna sequencing reveals rare intestinal cell types. *Nature*, 525(7568):251–255, 2015.

Hampel, F. R. The influence curve and its role in robust estimation. *Journal of the american statistical association*, 69(346):383–393, 1974.

Hauskrecht, M., Batal, I., Valko, M., Visweswaran, S., Cooper, G. F., and Clermont, G. Outlier detection for patient monitoring and alerting. *Journal of biomedical informatics*, 46(1):47–55, 2013.

Huber, P. J. *Robust statistics*, volume 523. John Wiley & Sons, 2004.

Huber, P. J. et al. Robust estimation of a location parameter. *The annals of mathematical statistics*, 35(1):73–101, 1964.

Jamieson, K. and Nowak, R. Best-arm identification algorithms for multi-armed bandits in the fixed confidence setting. In *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6. IEEE, 2014.

Jamieson, K., Malloy, M., Nowak, R., and Bubeck, S. lilucb: An optimal exploration algorithm for multi-armed bandits. In *Conference on Learning Theory*, pp. 423–439, 2014.

Kalyanakrishnan, S., Tewari, A., Auer, P., and Stone, P. Pac subset selection in stochastic multi-armed bandits. 2012.

Kano, H., Honda, J., Sakamaki, K., Matsuura, K., Nakamura, A., and Sugiyama, M. Good arm identification via bandit feedback. *Machine Learning*, 108(5):721–745, 2019.

Karnin, Z., Koren, T., and Somekh, O. Almost optimal exploration in multi-armed bandits. In *International Conference on Machine Learning*, pp. 1238–1246, 2013.

Katariya, S., Jain, L., Sengupta, N., Evans, J., and Nowak, R. Adaptive sampling for coarse ranking. In *International Conference on Artificial Intelligence and Statistics*, pp. 1839–1848, 2018.

Katariya, S., Tripathy, A., and Nowak, R. Maxgap bandit: Adaptive algorithms for approximate ranking. *arXiv preprint arXiv:1906.00547*, 2019.

Katz-Samuels, J. and Jamieson, K. The true sample complexity of identifying good arms. *arXiv preprint arXiv:1906.06594*, 2019.

Kaufmann, E., Cappé, O., and Garivier, A. On the complexity of best-arm identification in multi-armed bandit models. *The Journal of Machine Learning Research*, 17 (1):1–42, 2016.

Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4):764–766, 2013.

Locatelli, A., Gutzeit, M., and Carpentier, A. An optimal algorithm for the thresholding bandit problem. *arXiv preprint arXiv:1605.08671*, 2016.

Mannor, S. and Tsitsiklis, J. N. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, 5(Jun):623–648, 2004.

Mukherjee, S., Naveen, K. P., Sudarsanam, N., and Ravindran, B. Thresholding bandits with augmented ucb. *arXiv preprint arXiv:1704.02281*, 2017.

Paulson, E. et al. A sequential procedure for selecting the population with the largest mean from k normal populations. *The Annals of Mathematical Statistics*, 35(1): 174–180, 1964.

Porwal, U. and Mukund, S. Credit card fraud detection in e-commerce: An outlier detection approach. *arXiv preprint arXiv:1811.02196*, 2018.

Rousseeuw, P. J. and Hubert, M. Robust statistics for outlier detection. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1):73–79, 2011.

Sathe, S. and Aggarwal, C. Lodes: Local density meets spectral outlier detection. In *Proceedings of the 2016 SIAM international conference on data mining*, pp. 171–179. SIAM, 2016.

Schiff, G. D., Volk, L. A., Volodarskaya, M., Williams, D. H., Walsh, L., Myers, S. G., Bates, D. W., and Rozenblum, R. Screening for medication errors using an outlier detection system. *Journal of the American Medical Informatics Association*, 24(2):281–287, 2017.

Simchowitz, M., Jamieson, K., and Recht, B. The simulator: Understanding adaptive sampling in the moderate-confidence regime. *arXiv preprint arXiv:1702.05186*, 2017.

Swallow, W. H. and Kianifard, F. Using robust scale estimates in detecting multiple outliers in linear regression. *Biometrics*, pp. 545–556, 1996.

Xu, Y., Chen, X., Singh, A., and Dubrawski, A. Thresholding bandit problem with both duels and pulls. *arXiv preprint arXiv:1910.06368*, 2019.

Zhu, Y., Katariya, S., and Nowak, R. Code for icml2020 paper robust outlier arm identification. 2020. URL https://github.com/yinglunz/ROAI_ICML2020.

Zhuang, H., Wang, C., and Wang, Y. Identifying outlier arms in multi-armed bandit. In *Advances in Neural Information Processing Systems*, pp. 5210–5219, 2017.