Rebuttal Letter for Anonymous Reviewer 3

"Pure Exploration of Multi-Armed Bandits with Heavy-Tailed Payoffs" (Paper ID: 332)

First of all, we (the authors) would like to express our sincere gratitude to the anonymous reviewer for the comments on our paper entitled "*Pure Exploration of Multi-Armed Bandits with Heavy-Tailed Payoffs*". We would like to address the comments on the practical applications of our proposed algorithms as below, where important discussions are highlighted by the blue color for reading convenience.

Comments: "*I am not very convinced by the practicalness of the proposed two algorithms, not sure how it is useful in real life, in the sense that you might want to play with some real data to demonstrate the performance that I would expect in the next version of this draft.*"

Response. We thank the reviewer for the constructive comments. The proposed two algorithms (i.e., one is for fixed confidence, and the other is for fixed budget) in our paper are for the task of pure exploration of Multi-Armed Bandits (MAB) with heavy-tailed payoffs, with the goal to identify the optimal arm among a given decision-arm set. There have been many investigations illustrating that pure exploration of MAB can be applied into practical applications, e.g., identification of active users in social networks [1] and crowdsourcing [2]. In general, pure exploration of MAB is a sub-task of high-level applications, because the output of optimal arm at the end of exploration from bandit algorithms should be exploited in a main task of sequential decisions, e.g., crowdsourcing. In practice, our proposed two algorithms are applicable for scenarios for optimal learning with heavy-tailed payoffs.

To demonstrate the practicalness of our algorithms, we encounter two significant challenges. One is to demonstrate pure exploration of MAB with heavy-tailed payoffs as a sub-task in practical scenarios, and the other is to determine parameters of heavy-tailed distributions as prior information for our algorithms, i.e., the parameters of p, B, C in Table 2 of our paper. We tackle the first challenge by one of financial applications in the market of cryptocurrency, and successfully solve the second challenge by taking advantage of statistical analyses with the whole collected real data.

To clearly demonstrate the practicalness of our algorithms, we present three steps as below.

1) A Financial Application in Cryptocurrency.

It has been pointed out that financial data show the inherent characteristic of heavy tails [3-5], because the probability of events with a large deviation is high in financial markets. It is also worth mentioning that many other practical scenarios are related to heavy-tailed distributions, e.g., bioinformatics [6] and networks [7]. Due to the availability of financial data on the Internet, we choose a financial application of identifying the most profitable cryptocurrency over a period of time in a given pool of digital currencies.

We get hourly price data of cryptocurrency via the Internet[1], which include the open price, the closed price, the highest price and the lowest price of each hour. From historical financial data in digital currency, we observe that high fluctuations of price of cryptocurrency reflect a significant characteristic of heavy

[1] https://www.cryptocompare.com/

TABLE I: Ten selected cryptocurrencies in experiments.

full name	symbol	market value in April 2018 (unit: billion US dollar)
Bitcoin	BTC	155
Ethereum Classic	ETC	66
Ripple	XRP	32
Bitcoin Cash	BCH	23
EOS	EOS	15
Litecoin	LTC	8
Cardano	ADA	8
Stellar	XLM	7
IOTA	IOT	5
NEO	NEO	5

TABLE II: Statistical property of ten selected cryptocurrencies in experiments with hourly returns from Feb. 3rd, 2018 to Apr. 27th, 2018. KS-test1 denotes Kolmogrov-Smirnov (KS) test with a null hypothesis that real data follow a Gaussian distribution. KS-test2 denotes KS test with a null hypothesis that real data follow a *Student's t distribution*.

symbol	empirical statistics (mean, variance, skewness, kurtosis)	KS-test1 (statistic, $\bar{p}$-value)	KS-test2 (statistic, $\bar{p}$-value)
BTC	$(0.36 \times 10^{-3}, 0.54 \times 10^{-3}, 0.38, 3.02)$	$(0.08, 0.05 \times 10^{-1})$	$(0.05, 0.20)$
ETC	$(0.29 \times 10^{-3}, 1.03 \times 10^{-3}, 0.45, 2.30)$	$(0.07, 0.02)$	$(0.03, 0.89)$
XRP	$(0.33 \times 10^{-3}, 0.94 \times 10^{-3}, 0.86, 4.30)$	$(0.09, 0.04 \times 10^{-2})$	$(0.03, 0.61)$
BCH	$(0.78 \times 10^{-3}, 0.92 \times 10^{-3}, 1.27, 10.46)$	$(0.08, 0.01 \times 10^{-1})$	$(0.03, 0.64)$
EOS	$(\mathbf{1.56 \times 10^{-3}}, 1.18 \times 10^{-3}, 0.62, 2.45)$	$(0.09, 0.02 \times 10^{-2})$	$(0.03, 0.88)$
LTC	$(0.68 \times 10^{-3}, 0.86 \times 10^{-3}, 0.87, 3.64)$	$(0.10, 0.02 \times 10^{-2})$	$(0.04, 0.49)$
ADA	$(0.02 \times 10^{-3}, 1.22 \times 10^{-3}, 0.59, 4.31)$	$(0.07, 0.03)$	$(0.02, 0.99)$
XLM	$(0.62 \times 10^{-3}, 0.12 \times 10^{-3}, 0.32, 3.00)$	$(0.07, 0.02)$	$(0.03, 0.80)$
IOT	$(0.68 \times 10^{-3}, 0.11 \times 10^{-3}, 0.25, 1.77)$	$(0.07, 0.02)$	$(0.04, 0.57)$
NEO	$(-0.31 \times 10^{-3}, 1.26 \times 10^{-3}, 1.16, 7.52)$	$(0.10, 0.02 \times 10^{-2})$	$(0.04, 0.53)$

tails, which is pretty ideal for the task of pure exploration of MAB with heavy-tailed payoffs. For experiments, we choose top ten cryptocurrencies in terms of market value, with basic information shown in Table I.

The experimental design of identification for the most profitable cryptocurrency among the cryptocurrency pool in Table I is motivated by the practical scenario that an investor would like to invest a fixed budget of money in a cryptocurrency and get return as much as possible. Before the investor puts a large amount of real money into a cryptocurrency, a good solution is to test a small amount of (virtual) money

TABLE III: Estimated parameters for ten cryptocurrencies in experiments.

symbol	degree of freedom of *Student's t distribution*	(p, B, C) in our paper
BTC	3.50	
ETC	3.81	
XRP	**2.53**	
BCH	3.00	
EOS	2.90	
LTC	2.75	(2,1.577×10^{-3},1.575×10^{-3})
ADA	3.55	
XLM	3.81	
IOT	4.66	
NEO	3.13	

for investments over a certain time period and then find out which cryptocurrency is the ideal one, which exactly corresponds to the task of pure exploration of MAB with heavy-tailed payoffs in our paper. Note that pure exploration of cryptocurrency is a sub-task of the high-level task of sequential investments in the cryptocurrency market.

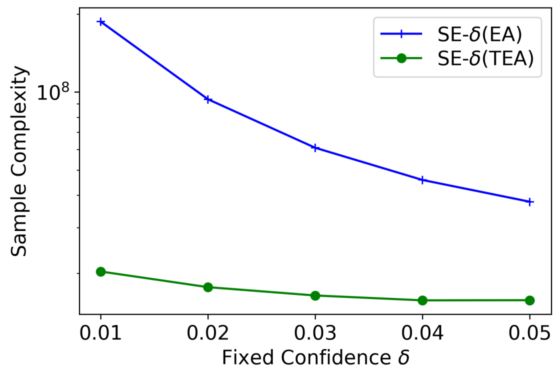
2) Statistical Analyses of Real Data from Cryptocurrency.

For experiments, we get hourly price data of the ten selected cryptocurrencies, and show the statistics of real data in Table II. In the table, we conduct a statistical analysis in hindsight with hourly returns of cryptocurrency from February 3rd, 2018 to April 27th, 2018. From the table, we find that the optimal option in hindsight is the cryptocurrency of EOS in terms of the maximal empirical mean of hourly payoffs. Note that an hourly payoff of 1.56×10^{-3} for EOS is significant, because the daily payoff in average is about 3.7%. Besides, we conduct Kolmogrov-Smirnov (KS) test by fitting real data of a cryptocurrency to a distribution. In particular, via KS test, we know that the null hypothesis of real data following a Gaussian distribution is rejected, because $\bar{p}$-value is less than a significant level of 0.05. We observe that real data of cryptocurrency are likely to follow a *Student's t distribution* via KS test, as shown in Table II.

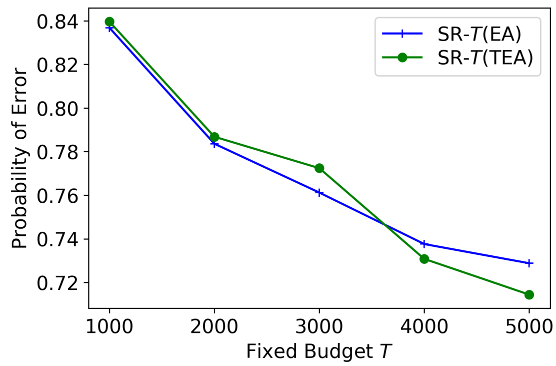
3) Pure Exploration of Cryptocurrency in Table I via Our Algorithms.

With the above statistical analyses, we can fit real data of cryptocurrency to a *Student's t distribution*, and obtain distributional parameters shown in Table III. From the table, we know that the smallest degree of freedom is 2.53 from XRP. Based on the property of *Student's t distribution*, we can choose $p = 2$ and estimate the parameters of B and C shown in the table. The estimated parameter of p reflects that financial data contain a finite variance, which is reasonable.

Similar to the experimental setting of our paper, i.e., running multiple epochs of experiments with each epoch containing ten independent experiments, we show the results on pure exploration of top ten



(a) fixed confidence



(b) fixed budget

Fig. 1: Pure exploration of cryptocurrency shown in Table I with the optimal option being **EOS**.

cryptocurrencies in Figure 1. Note that, due to limitation of data points in the setting of fixed confidence, we generate payoffs from *Student's t distributions* fitting to real data. But in the setting of fixed budget, we adopt exactly real financial data. We observe that, for fixed confidence, sample complexity decreases with increasing value of δ; and for fixed budget setting, probability of error converges to zero with increasing value of T. It is worth mentioning that, TEA algorithm outperforms EA algorithm in fixed-confidence setting when the value of δ is small, which is consistent to the findings in synthetic data. Besides, TEA algorithm is comparable to EA algorithm in fixed-budget setting because the truncated parameter in Algorithm 2 of our paper only has budget information and does not increase with the number of samples.

Further Discussions.

To demonstrate the practicalness of our proposed algorithms, we have successfully tackled two challenges. To the best of our knowledge, this is the first discussion on practical applications of pure exploration of MAB with heavy tails. In the above demonstrations, we have adopted the assumption that real data from February 3rd, 2018 to April 27th, 2018 are weakly stationary, which is a common and acceptable assumption in the analysis of financial data [8]. In fact, we still have the following considerations for the experiments on real data of cryptocurrency.

- We might develop a high-level system for sequential investments, which adopts pure exploration of MAB as a sub-task, to demonstrate cumulative returns in practical applications. The development of the investment system is a complicated task, which should be an independent effort of our work.
- In our proposed algorithms, the parameters of p, B, C have been estimated via the whole collected real data in hindsight. An alternative solution is to develop automatic learning for the parameters with samples increasing. The automatic learning of pure exploration of MAB is challenging, because the control of learning errors in distributional parameters still needs great efforts.
- For pure exploration of cryptocurrency in our experiments, the observed information is assumed to be partial because only the return information of the chosen cryptocurrency is available for each

round in experiments. In financial markets, full information (e.g., prices of all cryptocurrencies available for each hour) is available for an investor if we assume that the investment money from the investor does not affect the market. In practice, both settings (i.e., the partial information and the full information) are imperfect. In applications, we may consider develop bandit algorithms for pure exploration of MAB with semi-feedbacks, which is a trade-off between the partial information and the full information.

Overall, the above discussions have completely shown the practicalness of our proposed algorithms. Our work submitted to UAI 2018 is self-contained, because we have developed theoretical guarantees of algorithms, and demonstrated consistent experimental results via synthetic and real-world datasets. Finally, we would like to thank again the reviewer for pointing out the constructive comments, and we have added the experiment with real data into our paper.

References

[1] Jun, Kwang-Sung, et al. "Top Arm Identification in Multi-Armed Bandits with Batch Arm Pulls." AISTATS (2016).

[2] Zhou, Yuan, Xi Chen, and Jian Li. "Optimal PAC multiple arm identification with applications to crowdsourcing." ICML (2014).

[3] Müller, Ulrich A., Michel M. Dacorogna, and Oliver V. Pictet. "Heavy tails in high-frequency financial data." A practical guide to heavy tails: Statistical techniques and applications (1998).

[4] Panahi, Hanieh. "Model selection test for the heavy-tailed distributions under censored samples with application in financial data." International Journal of Financial Studies (2016).

[5] Oden, John, Kevin Hurt, and Susan Gentry. "A new class of heavy-tailed distribution and the stock market returns in Germany." Research in Business and Management (2017).

[6] Nguyen, Nha, et al. "Heavy-Tailed Noise Suppression and Derivative Wavelet Scalogram for Detecting DNA Copy Number Aberrations." IEEE/ACM Transactions on Computational Biology and Bioinformatics (2017).

[7] Markakis, Mihalis G., Eytan Modiano, and John N. Tsitsiklis. "Max-Weight scheduling in queueing networks with heavy-tailed traffic." IEEE/ACM Transactions on Networking (2014).

[8] Wang, Peijie. Financial Econometrics. Routledge (2005).