arXiv:2302.02526v2 [cs.LG] 4 Mar 2023

隐私与鲁棒赌博机

Yulian Wu [*†] Xingyu Zhou [*‡] Youming Tao [§] Di Wang [¶]

摘要

我们研究私有且稳健的多臂老虎机（MAB），其中智能体接收 Huber 污染的重尾奖励，同时需要确保差分隐私。我们首先给出其极小极大下界，刻画遗憾相对于隐私预算、污染程度和重尾性的信息论极限。随后，我们提出一种元算法，该算法基于一个私有稳健均值估计子程序 PRM，而该子程序本质上仅依赖奖励截断和拉普拉斯机制。针对两种不同的重尾设置，我们给出 PRM 的具体方案，从而实现近乎最优的遗憾。作为主要结果的推论，我们还首次给出了私有重尾多臂老虎机（即不存在污染时）的极小极大下界。此外，我们提出的两种基于截断的 PRM 实现了估计精度、隐私与鲁棒性之间的最优权衡。最后，我们通过实验研究验证了理论结果。

[*]贡献相同。
[†]阿卜杜拉国王科技大学，沙特阿拉伯图瓦勒。电子邮件： yulian.wu@kaust.edu.sa
[‡]韦恩州立大学，美国底特律。电子邮件： xingyu.zhou@wayne.edu[§]山东大学，中国青岛。电子邮件： ym.tao99@mail.sdu.edu.cn
[¶]阿卜杜拉国王科技大学，沙特阿拉伯图瓦勒。电子邮箱： di.wang@kaust.edu.sa

目录

1 引言 32 **相关工作 4**

3 预备知识 **4**
3.1 隐私与鲁棒多臂老虎机.3.2 私有鲁棒多臂老虎机的遗憾. 4

64 下界 65 我们的方法：一种元算法 76 上界 86.1 有限原始矩情形.96.2 有限中心矩情形.10**7 实验 13**7.1 实验设置.147.2 结果与讨论.14**8 结论 159 致谢 15A 实用引理 18B 第4 节的证明 19C 第6.1 节的证明 22D 第6.2 节的证明 27**

1 引言

多臂老虎机（MAB） [BF85] 问题为在老虎机反馈下进行不确定性序贯决策提供了一个基础框架，并已在医学[GPKW17], 金融 [HBD+11； SWJZ15],推荐系统 [CB13], 以及在线广告 [SBF17],等领域得到了广泛应用。以金融领域的投资组合选择为例。在每一决策轮次中 $t \in [T]$，学习智能体选择一个动作 $a_t \in [K]$（即，选择一组特定资产供用户使用 t），并获得一项奖励 r_t（例如，相应的收益），该收益是从与该投资组合选择相关联的未知概率分布中独立同分布抽取的。目标是学会最大化累计收益。

在实践中，将广受赞誉的多臂老虎机（MAB）框架应用于现实场景（例如上文的金融示例），需要同时处理鲁棒性和隐私问题。一方面，众所周知，金融数据通常具有重尾特征（而非次高斯特征） [Rac03； Hul12]。此外，金融领域收到的收益数据往往包含离群值 [AHMRV19]，这是数据污染所致。另一方面，金融领域对隐私的担忧日益加剧 [LMM20； CSW22； CMW22]。例如，即使对手无法直接访问数据集，仍然可以通过与定价平台交互并观察其决策，来重构其他客户的个人信息 [FLJLPR14]。

受此启发，一系列关于多臂老虎机的研究致力于设计能够应对重尾奖励 [BCL13], 对手污染 [LMP18; KPK19], 或二者兼顾 [BMM22]的鲁棒算法。近期另一系列研究则探讨了在不同信任模型下，如何通过差分隐私（DP）实现多臂老虎机的隐私保护，例如中心化差分隐私 [SS19; AB22], 本地差分隐私 [RZLS20; TWZW21] 以及分布式差分隐私 [TKMS21; CZ22]。此外，近期研究还揭示了均值估计问题中鲁棒性与隐私之间的紧密联系（例如，鲁棒性诱导隐私 [HKMN22] ，反之亦然[GH22]）。鉴于此，本文关注的一个基本问题是：

是　否有一种简单的算法可以同时应对多臂老虎机中的隐私性和鲁棒性　　?

我们的贡献。 我们通过证明一种简单的基于截断的算法能够在多臂老虎机中实现遗憾、隐私和鲁棒性之间近乎最优的权衡，对此给出了肯定的回答。其核心直觉是，奖励截断不仅有助于减少离群值（由于重尾和污染所致），还能够限制其敏感度，而这对于差分隐私是必要的。为使这一直觉更加严谨，我们采取了以下原则性方法。

（i）我们首先建立隐私与鲁棒多臂老虎机的极小极大遗憾下界，即同时具有隐私约束和 Huber 污染的重尾多臂老虎机的极小极大遗憾下界 [Hub64]（见第4节）。这刻画了遗憾相对于隐私预算、污染程度和重尾性的信息论极限。作为副产品，我们的结果还给出了私有重尾多臂老虎机（即无污染情形）的首个极小极大下界，从而解决了 Tao等人 提出的一个开放问题。 [TWZW21]。

(ii)为匹配下界，我们首先提出一种元算法（见第 5 节）。该算法基于分批逐次淘汰的思想，并依赖一个通用的私有稳健均值估计子程序，记为 PRM。随后，针对两种不同的（重尾）奖励分布设定（即原始矩或中心矩有限），我们为子程序 PRM 提出相应方案；这两种方案均仅需结合截断与拉普拉斯机制，即可同时保证鲁棒性和隐私。借助这些方案，我们的元算法能够获得几乎匹配的遗憾上界（见第 6 节）。实验研究也印证了我们的理论结果。

(iii) 在此过程中，我们还得到了若干具有独立研究价值的结果。特别是，我们提出的 PRM表明，在一维情形下，截断足以帮助私有且稳健的均值估计实现最优的高概率集中。此外，在不存在污染的情况下，我们的遗憾上界

不仅与有限原始矩的私有重尾多臂老虎机的最优上界相匹配，还首次给出了有限中心矩情形下的结果，从而对私有老虎机进行了全面研究。

受篇幅限制，技术引理和所有证明均收录于附录中。

2 相关工作

鲁棒多臂老虎机。 关于鲁棒老虎机的研究大致可分为两类。第一类工作主要关注总污染有界的设定，即观测奖励与真实奖励之间的累积差异被某个常数 [LMP18]所界定。第二类工作考虑 Huber α污染模型 [Hub64] （这也是本文关注的重点）或类似的 α比例模型。在这些情形下，每一轮的奖励都可能以概率 $\alpha \in [0, 1]$ [AGKS20； KPK19； MTCD21], 受到任意分布的污染， α或者至多有 [NT20]比例的奖励受到任意污染。该类现有工作主要关注轻尾设定，其中真实内点分布为高斯分布或次高斯分布，并使用鲁棒中位数或截尾均值估计量。近期的一项工作 [BMM22] 研究了内点分布仅具有有限方差的设定，并使用 Huber 估计量建立问题相关界。相比之下，我们从极小极大遗憾的角度出发，即研究问题无关界，同时也考虑隐私。

私有多臂老虎机。 除了上述关于奖励轻尾的私有多臂老虎机的结果外，Tao等人 [TWZW21] 研究了在差分隐私的中心模型和本地模型下，具有有限原始矩的私有重尾多臂老虎机。然而，这种情况下的最优极小极大遗憾仍然未知，而且如何为具有有限中心矩的重尾分布设计私有算法也尚不明确。本文作为主要结果的副产品，解决了这两个问题。

鲁棒与私有均值估计。 我们的工作也与鲁棒与私有均值估计有关，尤其是一维情形。在 Huber 模型下关于鲁棒性的研究中，Lai等人首先为高斯分布的中位数（由对称性可知也适用于均值）建立了一个高概率集中界。 [LRV16]。最近，Mukherjee 等人 [MTCD21]针对一般的次高斯内点分布，利用截尾均值估计量给出了高概率均值集中结果；而 Prasad等人 [PBR19] 则关注重尾情形。在隐私方面，Kamath 等人 [KSU20], 的工作与此密切相关，他们通过均值中位数法，为具有有限中心矩的私有重尾分布首次给出了高概率均值集中结果。值得注意的是，近期的研究在理解鲁棒性与隐私之间的密切关系方面取得了令人振奋的进展（例如，鲁棒性诱导隐私 [HKMN22]，反之亦然 [GH22]）。从这个角度看，我们的结果表明，对于一维均值估计问题，仅靠截断就足以同时实现鲁棒性与隐私。

3 预备知识

本节首先正式介绍我们的私有鲁棒多臂老虎机问题，然后阐述其遗憾概念。

3.1 私有与鲁棒多臂老虎机

如前所述，鲁棒性旨在同时应对奖励污染和可能存在的重尾内点分布。为此，我们首先介绍以下两类重尾奖励分布。

定义 3.1（有限 k 阶原始矩）。 如果 $\mathbb{R}$ 上的一个分布属于

$$\mathcal{P}_k = \left\{ P : \mathbb{E}_{X\sim P}\left[|X|^k \right] \leqslant 1 \right\}, \quad k \geq 2. \tag{1}$$

定义 3.2（有限的 k 阶中心矩）。 若 $\mathbb{R}$ 上的一个分布满足以下条件，则称其具有有限的 k 阶中心矩：

$$\mathcal{P}_k^c = \left\{ P : \mathbb{E}_{X\sim P}\left[|X-\mu|^k \right] \leqslant 1 \right\}, \quad k \geq 2. \tag{2}$$

其中 $\mu := \mathbb{E}_{X\sim P}[X] \in [-D, D]$，即 D 是其均值的有限范围，但其值可以任意大。

我们进一步考虑著名的 Huber 污染模型 [Hub64]，并将其应用于重尾多臂老虎机。

定义 3.3（带有 Huber 污染的重尾多臂老虎机）。 给定污染水平 $\alpha \in [0,1]$。对于每一轮 $t \in [T]$，动作 a_t的观测奖励[1] x_t 以概率 $1-\alpha$独立地从真实分布$P_{a_t} \in \mathcal{P}_k$（或 $P_{a_t} \in \mathcal{P}_k^c$）中抽样；否则，则从某个任意且未知的污染分布 $G_{a_t} \in \mathcal{G}$中抽样。

除了鲁棒性之外，我们还通过差分隐私（DP）的视角考虑多臂老虎机中的隐私保护。具体而言，我们考虑多臂老虎机标准的差分隐私中心模型（例如 [MT15]），在该模型中，学习智能体可以访问用户的原始数据（即奖励），并保证其输出（即动作序列）在概率上对两条相邻的奖励序列不可区分。由于存在污染，学习智能体在第 t 轮访问的奖励数据可能已经受到污染。更准确地说，令 $D_T = (x_1, \ldots, x_T) \in \mathbb{R}^T$表示学习过程中生成的奖励序列，并令 $\mathcal{M}(D_T) = (a_1, \ldots, a_T) \in [K]^T$表示学习算法 $\mathcal{M}$推荐的所有动作构成的序列。在此设定下，我们给出以下形式化定义。

定义 3.4（多臂老虎机的差分隐私）。 对于任意 $\varepsilon > 0$，若学习算法 $\mathcal{M}: \mathbb{R}^T \to [K]^T$ isε-DP 对所有仅在一个元素上不同的序列 $D_T, D_T' \in \mathbb{R}^T$ 以及所有事件 $E \subset [K]^T$ 均满足以下条件，则

$$\mathbb{P}\left[\mathcal{M}(D_T) \in E\right] \leqslant e^{\varepsilon} \cdot \mathbb{P}\left[\mathcal{M}\left(D_T'\right) \in E\right].$$

在本文中， 我们将利用广为人知的拉普拉斯机制来保证差分隐私。

定义 3.5（拉普 拉斯机制）。 给定函数 f: $\mathcal{X}^n \to \mathbb{R}^d$，拉普拉斯机制定义为

$$\mathcal{M}_L(D, f, \varepsilon) = f(D) + (Y_1, Y_2, \cdots, Y_d),$$

其中 Y_i 是独立同分布地从拉普拉斯分布 [2]Lap($\frac{\Delta_1(f)}{\varepsilon}$) 中抽取的，其中 $\Delta_1(f)$ 是函数 f 的 ℓ_1-敏感度，即， $\Delta_1(f) = \sup_{D\sim D'} ||f(D) - f(D')||_1$.那么，对于任意 $\varepsilon > 0$，拉普拉斯机制满足ε-DP。

在接下来的章节中，为简洁起见，我们将用隐私与鲁棒多臂老虎机来指代我们的研究设定，即带有 Huber 污染和隐私约束的重尾多臂老虎机。

[1]这里我们在污染情形下使用 x_t，以区别于标准奖励 r_t。[2]对于参数 λ，拉普拉斯分布的概率密度函数为 $\text{Lap}(\lambda)(x) = \frac{1}{2\lambda}\exp(-\frac{|x|}{\lambda})$。

3.2 私有鲁棒多臂老虎机的遗憾

在污染情形下，使用观测到的（受污染的）奖励计算标准遗憾 $\{x_t\}_{t\in[T]}$ 并无明确的定义 [NT20]。相反，相关文献关注的是干净遗憾，即根据期望的真实未污染奖励来衡量，并与事后最佳策略进行比较 [NT20; BMM22; CKMY22]。因此，令 μ_a 表示臂 $a \in [K]$ 的内点分布均值，且令 $\mu^* = \max^{a\in[K]} \mu_a$。我们还令 Π^ε 表示所有 ε-差分隐私多臂老虎机算法的集合，并令 $\mathcal{E}_{\alpha,\ k}$ 表示所有带有 Huber 污染的重尾多臂老虎机实例的集合。

定义 3.6 （干净遗憾）。 固定一个算法 $\pi \in \Pi^\varepsilon$ 和一个实例 $\nu \in \mathcal{E}_{\alpha,\ k}$。则，$\pi$ 在 ν 下的干净遗憾为

$$\mathcal{R}_T(\pi,\nu) := \mathbb{E}_{\pi,\nu}[T\mu^* - \sum_{t=1}^{T} \mu_{a_t}].$$

注意，此处的期望是针对由受污染的环境和 ε-差分隐私多臂老虎机算法所产生的随机性取的，而均值则是真实内点分布的均值。

为了刻画私有鲁棒多臂老虎机问题的内在难度，我们也关注其极小极大遗憾。

定义 3.7（极小　极大遗憾）。 我们的私有鲁棒多臂老虎机问题的极小极大遗憾定义为

$$\mathcal{R}^{\text{minimax}}_{\varepsilon,\alpha,k} := \inf_{\pi\in\Pi^\varepsilon} \sup_{\nu\in\mathcal{E}_{\alpha,k}} \mathbb{E}_{\pi,\nu}[T\mu^* - \sum_{t=1}^{T} \mu_{a_t}]. \tag{3}$$

4 下界

我们首先给出极小极大遗憾的如下下界，它刻画了隐私预算（通过 ε）、污染程度（通过 α）以及奖励的重尾性（通过 k）对遗憾的根本影响。

定理 4.1. 考虑一个私有且鲁棒的多臂老虎机问题，其中内点分布具有有限的 k 阶原始矩（或中心矩）（$k \geq 2$）。那么，其极小极大遗憾满足

$$\mathcal{R}^{minimax}_{\varepsilon,\alpha,k} = \Omega\left(\sqrt{KT} + \left(\frac{K}{\varepsilon}\right)^{1-\frac{1}{k}} T^{\frac{1}{k}} + T\alpha^{1-\frac{1}{k}}\right).$$

我们先来解读上述结果，它基本上取三个项中的最大值。第一项来自高斯奖励下的标准遗憾，第二项刻画了隐私和重尾奖励导致的额外遗憾代价，最后一项则表示污染和重尾奖励导致的额外遗憾代价。请注意，对于给定的 k，隐私和污染的影响是可分离的。将我们的下界与相关结果进行比较也很有意义，以下备注旨在进行这一比较。

注释4.2。首先，当 $k = \infty$ 且 $\alpha = 0$时，我们的下界恢复了私有多臂老虎机在次高斯奖励下的当前最优下界 [AB22]；其次，我们注意到，即使 $\alpha = 0$，关于私有重尾多臂老虎机的极小极大遗憾（即与问题无关的）下界仍无现有结果。事实上，近期的一项工作将其列为开放问题 [TWZW21]。因此，我们的下界不仅解决了

这一问题[3]，，还涵盖了污染。最后，在没有隐私保护的情况下，近期的一项工作 [BMM22] 为鲁棒多臂老虎机建立了与问题相关的遗憾下界，而我们关注的是与问题无关的下界。因此，其结果与我们的结果不可比较。

现在，尚需探究能否通过某些算法达到这一下界，这正是接下来两节的主要关注点。

5 我们的方法：AMeta-Algorithm

本节首先介绍一种用于私有稳健多臂老虎机的元算法。它不仅能以统一的方式处理原始矩或中心矩有界的内点分布，还突出了一个关键组件，即基于截断这一简单思路构建的私有稳健均值估计子程序。

从较高层次来看，我们的元算法可以视为著名的逐步臂消除 [EMMM06] 的批次版本，并结合私有稳健均值估计子程序 PRM （见算法 1）。具体而言，它将时间范围 T 划分为大小呈指数增长的批次，并通过 PRM 得到均值估计，逐步消除次优臂。更具体地说，该算法根据批次大小分为两个阶段。当批次大小小于阈值 $\mathcal{T}$时，它会随机推荐动作（第 5–7 行；稍后将对此作进一步说明）。否则，对于批次 τ中的每个活跃臂 a，它首先将 a 分配给一批 $B_\tau = 2^\tau$ 新用户，并观测可能受污染的奖励（第 8 行）。然后，它调用子程序 PRM，为每个活跃臂 a 计算私有稳健均值估计（第 12 行）。具体而言，它只使用最近一个批次内的奖励（即“遗忘”），并采用适当的奖励截断阈值 M_τ。最后，它采用经典的臂消除思路，并适当选择置信半径 β_τ，以高置信度移除次优臂（第 18–20 行）。

现在，我们通过说明算法的主要组成部分如何协同工作，进一步阐释算法设计背后的直觉。首先，设置第一阶段（即 $B_\tau \le \mathcal{T}$）的原因是，当样本量较小时， PRM的均值估计无法实现高概率集中。因此，在这一阶段不能采用臂消除，因为这可能会淘汰最优臂。请注意，除了我们选择的随机选择外，第一阶段也可以采用其他方法（见下文备注 5.1）。其次，在第二阶段，分批和遗忘的思路是以最少噪声实现隐私（从而获得更低遗憾）的关键。这是因为此时任何单次奖励反馈都只会影响一次估计计算。这与标准的臂消除方法（例如 [EMMM06]）形成鲜明对比：标准方法中的每个均值估计都基于截至目前的所有样本（因为没有采用分批），因此，单次奖励变化可能会影响 $O(T)$次均值估计[4]。第三，在 PRM 中，奖励截断这一简单思路被证明对鲁棒性和隐私都极为有用。一方面，截断有助于减小离群值（由重尾和污染导致）的影响；另一方面，截断还有助于限制敏感度，这是实现隐私所必需的。事实上，正如我们稍后将展示的，经过恰当调节的截断阈值能够使我们在遗憾、隐私和鲁棒性之间实现近乎最优的权衡。最后，与第一阶段不同，当批次大小大于 $\mathcal{T}$时，由于均值估计具有高概率集中性，我们现在可以有较高把握地消除次优臂（稍后将详细说明 PRM 的具体取值，以及 $\mathcal{T}$的选择）。

注 5.1。 第一阶段的算法选择可以灵活处理。例如，不必在整个批次中都执行随机选定的动作，也可以在每一

[3]在 [TWZW21], 中，作者考虑了一种略有不同的设定：重尾分布只有有限的 $(1+v)$阶矩，且满足 $v \in (0,1]$。不过，只需令 $k = 1+v$，我们的结果便可直接推广到这一设定。 [4]可以使用基于树的算法 [CSS11] 将其降至 $O(\log T)$，但它仍是次优的 [SS19]。

算法 1私有且稳健的臂消除

输入： 臂数 K，时间范围 T，隐私预算 ε，Huber参数 $\alpha \in (0,1]$, 错误概率 $\delta \in (0,1]$, 内点分布参数，即 k 以及可选的 D

初始化： $\tau = 0$，臂的活跃集 $\mathcal{S} = \{1, \cdots, K\}$。

对于 批次 $\tau = 1, 2, \cdots$ **执行**

  设置批次大小 $B_\tau = 2^\tau$

  if $B_\tau < T$ **然后**

    随机选择一个动作 $a \in [K]$

    执行动作 a B_τ 次

  else

    对于每个活跃臂 $a \in \mathcal{S}$ **执行**

      对于 i 从 1 到 B_τ **do**

        拉取臂 a , 观察受污染奖励 x_i^a

        如果拉取次数总数达到 T，**退出**

      结束循环

      设置截断阈值 M_τ

      设置附加参数 Φ

      计算估计值 $\widetilde{\mu}_a = \mathrm{PRM}(\{x_i^a\}_{i=1}^{B_\tau}, M_\tau, \Phi)$

    结束循环

    设置置信半径 β_τ

    Let $\widetilde{\mu}_{\max}$ =最大值$_{a \in \mathcal{S}}$ $\widetilde{\mu}_a$

    从 $\mathcal{S}$ 中移除所有臂 a，使得 $\widetilde{\mu}_{\max} - \widetilde{\mu}_a > 2\beta_\tau$

  结束条件判断

22: </style}结束循环

轮中执行随机选定的动作。此外，还可以根据 PRM 得出的均值估计，选择贪心策略或概率贪心策略；这两种方式也都只使用各臂在上一批次中收集到的奖励。这些选择的理论保证都相同，不过其中一些有助于提升经验性能。

接下来 我们给出以下评论，将我们的元算法置于现有文献的背景下。

评论 5.2（与现有文献的比较）。 对于私有多臂老虎机（无污染），最先进的方法同样基于分批和遗忘的思路 [SS19；CZ22] 以实现最优遗憾。对于鲁棒多臂老虎机（无隐私），现有研究采用了不同的稳健均值估计方法。例如，Niss 和 Tewari [NT20] 以及 Mukherjee 等人 [MTCD21]针对次高斯内点分布使用截尾均值估计量，而 Basu 等人 [BMM22] 则采用 Huber 估计量处理仅具有有界方差的内点分布。我们首次通过一种简单的截断型估计量，同时研究隐私和鲁棒性，进而揭示多臂老虎机中隐私与鲁棒性之间的密切关系。这也补充了近期关于揭示（高维）统计学中这两者联系的研究进展 [HKMN22；GH22]。

6 上界

本节将为元算法的两种具体实例推导遗憾上界：一种对应有限原始矩情形，另一种对应有限中心矩情形。特别地，所得结果与我们的下界至多相差一个对数因子，表明它们具有近似最优性。

有限原始矩情形下的算法 2 PRM

1: 输入： 一组数据 $\{x_i\}_{i=1}^n$、截断参数 M、附加参数 $\Phi = \{\varepsilon\}$2: **对于** $i = 1, 2, \ldots\ldots,$ n **执行**3: 截断数据 $\bar{x}_i = x_i \cdot \mathbb{1}_{\{|x_i| \le M\}}$4: **循环结束** 5: 返回隐私估计 $\widetilde{\mu} = \frac{\sum_{i=1}^n \bar{x}_i}{n}$ + 拉普拉斯分布($\frac{2M}{n\varepsilon}$)

6.1 有限原始矩情形

本节将重点讨论内点分布具有有限 k阶原始矩的隐私与鲁棒多臂老虎机，如定义3.1所述。具体而言，我们首先介绍此情形下 PRM 的选择（见算法2），并证明其集中性质，这对我们实现元算法起着关键作用。

算法 PRM2 中的机制就是一种基于截断的拉普拉斯机制。也就是说，它首先使用阈值 M 截断所有收到的数据（第 3 行）。随后，在经验均值上添加拉普拉斯噪声以保护隐私（第 5 行）。我们再次强调，此处的截断既有助于鲁棒性（通过移除离群值），也有助于隐私保护（通过限制经验均值的敏感度）。

与多臂老虎机算法的标准设计一样，关键在于利用均值估计量的集中性。为此，我们首先给出以下高概率集中结果，该结果针对算法 PRM 2所返回的均值估计。

定理 6.1（均值估计的集中性）。给定一组 *Huber* 污染数据 $\{x_i\}_{i=1}^n$，其中内点分布满足定义 *3.1*，且均值为 μ，令 $\widetilde{\mu}$ 为算法 *2*得到的均值估计。那么，对于任意隐私预算 $\varepsilon > 0$ 以及 $\delta \in (0, 1)$，以下结论成立：

未污染情形。对于 $\alpha = 0$，有

$$|\widetilde{\mu} - \mu| = O\left(\sqrt{\frac{\log(1/\delta)}{n}} + \frac{M\log(1/\delta)}{n\varepsilon} + \frac{1}{M^{k-1}}\right),$$

以至少 $1-\delta$ 的概率。因此，选择截断阈值 $M = \Theta\left(\left(\frac{n\varepsilon}{\log(1/\delta)}\right)^{\frac{1}{k}}\right)$，可得

$$|\widetilde{\mu} - \mu| = O\left(\sqrt{\frac{\log(1/\delta)}{n}} + \left(\frac{\log(1/\delta)}{n\varepsilon}\right)^{1-\frac{1}{k}}\right).$$

污染情形。对于 $\alpha \in (0, 1]$ 且 $n = \Omega\left(\frac{\log(1/\delta)}{\alpha}\right)$，以至少$1-\delta$的概率，有以下结论

$$|\widetilde{\mu} - \mu| = O\left(\sqrt{\frac{\log(1/\delta)}{n}} + \frac{M\log(1/\delta)}{n\varepsilon} + \frac{1}{M^{k-1}} + \alpha M\right).$$

因此，选择截断阈值 $M = \Theta\left(\min\left\{\left(\frac{n\varepsilon}{\log(1/\delta)}\right)^{\frac{1}{k}}, \alpha^{-\frac{1}{k}}\right\}\right)$，可得 $|\widetilde{\mu} - \mu| \le \beta$，其中

$$\beta = O\left(\sqrt{\frac{\log(1/\delta)}{n}} + \left(\frac{\log(1/\delta)}{n\varepsilon}\right)^{1-\frac{1}{k}} + \alpha^{1-\frac{1}{k}}\right).$$

根据上述结果，我们可以作出几点评论。首先，对于未污染情形，我们的集中性结果由标准的次高斯项，以及一个由隐私和重尾数据带来的新项构成。它可以转化为样本复杂度界；也就是说，为了保证 $|\widetilde{\mu}-\mu| \le \eta$ 对任意 $\eta \in (0,1)$ 成立，所需样本量为 $n \ge O(\frac{\log(1/\delta)}{\eta^2} + \log(1/\delta)\,\varepsilon\eta^{\frac{k}{k-1}})$，这是最优的，因为它与私有重尾均值估计的下界相匹配（参见 Hopkins 等人的定理 7.2 [HKM22]）。其次，对于污染情形，还会出现一个额外的偏差项 $O(\alpha^{1-1/k})$，而且已知该项在信息论意义下也是最优的 [Dia18]。因此，通过截断，算法2 所实现的准确性、隐私与鲁棒性之间的权衡是最优的，这也表明它有望集成到我们的元算法中。

现在，基于集中性结果，我们可以相应地为元算法设置其他缺失参数。具体而言，我们有以下定理，给出了具体实例及其性能保证。

定理 6.2（性能保证）。考虑一个私有鲁棒多臂老虎机，其内点分布满足定义 *3.1* 以及 $\alpha \in (0,1]$。令算法 *1* 采用算法 *2* 以及 M_τ，β_τ 由定理 *6.1* 给出，其中将 n 替换为 B_τ。令 $\mathcal{T} = \Omega(\frac{\log(1/\delta)}{\alpha})$ 且 $\delta = 1/T$。则算法 *1* 是 ε-*DP*，其遗憾上界为

$$O\left(\sqrt{KT\log T} + \left(\frac{K\log T}{\varepsilon}\right)^{\frac{k-1}{k}} T^{\frac{1}{k}} + T\alpha^{1-\frac{1}{k}} + \frac{K\log T}{\alpha}\right).$$

上述定理给出了隐私与鲁棒多臂老虎机中第一个可实现的遗憾保证。前三项与定理 4.1 中的下界相差至多一个 $\log T$ 因子。最后一个加性项主要是由于均值集中结果仅在样本量大于 $\mathcal{T} = \Omega(\frac{\log(1/\delta)}{\alpha})$ 时成立。因此，每个次优臂至少必须被选择 $\Omega(\frac{\log(1/\delta)}{\alpha})$ 次。然而，当 T 足够大且 α 为常数时，最后一项会被其他项所支配。

注 6.3. 对于 $\alpha = 0$ 的情形，利用定理 6.1 中的未污染集中界并进行相同的分析，我们可以得到遗憾上界 $O(\sqrt{KT}\log T + (\frac{K}{\log_\varepsilon T})^{\frac{k-1}{k}} T^{\frac{1}{k}})$，这也在 $\log T$ 因子的范围内匹配下界。

6.2 有限中心矩情形

上一节中有限原始矩情形的设定可能并不完全令人满意，因为它实质上假设各臂的均值被限制在一个较小的范围内（因此也限制了次优间隙）。因此，本节转而讨论内点分布具有有限 k 阶中心矩的隐私与鲁棒多臂老虎机问题，如定义 3.2 所述。为此，我们首先需要一种新的 PRM，因为像算法 2 中那样仅在零点附近截断，现在行不通。

我们提出的新 PRM 见算法 3，它包含两个步骤。其思路很简单：第一步旨在获得均值的粗略估计，这是必要的，因为现在均值可能远离零点。然后，第二步围绕初始估计进行截断，得到最终结果。具体而言，第一步中，我们首先在范围 $[-D, D]$ 内构造区间；根据定义 3.2，真实均值假定包含在该范围内。随后，我们通过拉普拉斯机制计算私有直方图。将经验质量最大的区间的左端点作为初始估计。接着，在第二步中，只需围绕初始估计进行截断，并再次添加拉普拉斯噪声以实现隐私保护。

备注 6.4。值得注意的是，先前关于一维重尾情形下稳健均值估计的研究中也采用了类似的两步估计思路 [PBR19； KSU20； LBY22]。然而，

算法 3 PRM （有限中心矩情形）

1: 输入： 一组数据 $\{x_i\}_{i=1}^{2n}$、截断参数 M、附加参数 $\Phi = \{\varepsilon, D, r\}$,$r \in \mathbb{R}$。

2: // 第一步：初始估计3: $B_j = [j, j+r)$
$j \in \mathcal{J} = \{-D, -D+r, \ldots, D-r\}$4: 使用第一折数据计
算私有直方图:

$$\widetilde{p}_j = \frac{\sum_{i=1}^{n} \mathbb{1}_{\{X_i \in B_j\}}}{n} + \mathrm{Lap}\left(\frac{2}{n\varepsilon}\right)$$

5: 获取初始估计 $J = \mathrm{argmax}_{j \in \mathcal{J}} \widetilde{p}_j$6: // 第二步：最终估
计
7: 使用第二折数据获取最终估计量: $\widetilde{\mu} = J + \frac{1}{n}\sum_{i=n+1}^{2n}(X_i - J)\mathbb{1}_{\{|X_i - J| \le M\}}$ + 拉普拉斯分布 $\left(\frac{2M}{n\varepsilon}\right)$

我们的算法设计和分析存在若干差异。具体而言， [PBR19]考虑的是 Huber 模型下不受隐私约束的均值估计，而我们进一步施加了差分隐私要求。因此，与 [PBR19], 相比，我们两步中的估计量形式各不相同，尽管它们具有相同的高层次直觉。另一方面， [KSU20]考虑的是差分隐私下的均值估计，并未像我们这样考虑 Huber 污染。此外，我们的第二个估计基于截断，而他们的方法采用均值中位数法。事实上，正如后文将展示的（见备注6.7），当我们的结果退化为未污染情形时，其性能优于 [KSU20]中的结果。最后， [LBY22][5] 同时考虑了 Huber 污染和局部差分隐私，并建立了相应的均方误差（MSE）结果。相比之下，我们考虑的是中央差分隐私，目标是建立高概率尾部集中界。为此，与 [LBY22]中的方法相比，我们采用了不同的截断方法（即在第 7 行使用指示函数）。

与之前一样，我们首先介绍新 PRM的集中性质，该性质将在我们的元算法的具体实例化中体现。具体而言，我们先给出以下一般定理，然后陈述两个更为详细的推论。

定理 *6.5*（均值估计的集中性）。给定一组 *Huber* 污染数据 $\{x_i\}_{i=1}^{2n}$，其中内点分布满足定义 *3.2*，且均值为 μ，设 $\widetilde{\mu}$ 为算法 *3* 给出的均值估计。对于任意 $\alpha \in (0, \alpha_{\max})$、 $\varepsilon \in (0, 1]$ 以及 $\delta \in (0, 1)$，存在一些常数 $\mathcal{T}(\alpha, \varepsilon, \delta)$、 r、 M和 $D \ge 2r$，使得对所有 $n \ge \mathcal{T}(\alpha, \varepsilon, \delta)$，其概率至少为 $1 - \delta$

$$|\widetilde{\mu} - \mu| \le O\left(\sqrt{\frac{\log(1/\delta)}{n}} + \frac{M\log(1/\delta)}{n\varepsilon} + \frac{1}{M^{k-1}} + \alpha M\right),$$

其中 $\alpha_{\max} < 1$ 为崩溃点。

上述定理与原始矩情形（定理 6.1）遵循相同的模式。主要区别在于阈值 $\mathcal{T}(\alpha, \varepsilon, \delta)$ 和崩溃点 $\alpha_{\max}$，以下结果对它们进行了总结。

[5]具体而言，我们参考 LB [Y22]的第一个 arXiv 版本。

推论 6.6(均值集中，$\alpha = 0$)。设定与定理 *6.5* 相同的假设。对于任意$\varepsilon \in (0,1]$, 设置 $r = 10^{1/k}$ 且 $M = \Theta\left(\frac{n\varepsilon}{\log(1/\delta)}\right)^{1/k}$，则对于所有 $n \geq \Omega(\log(D/\delta)/\varepsilon)$ 以及 $D \geq 2r$，对于任意 $\delta \in (0,\ 1)$，以至少 $1-\delta$ 的概率，$|\tilde{\mu} - \mu| \leq \beta$ 其中

$$\beta = O\left(\sqrt{\frac{\log(1/\delta)}{n}} + \left(\frac{\log(1/\delta)}{n\varepsilon}\right)^{1-\frac{1}{k}}\right).$$

换言之，取样本数 n 使得

$$n \geq O\left(\frac{\log(1/\delta)}{\eta^2} + \frac{\log(1/\delta)}{\varepsilon\eta^{\frac{k}{k-1}}} + \frac{\log(D/\delta)}{\varepsilon}\right),$$

我们以至少 $1-\delta$ 的概率有 $|\tilde{\mu} - \mu| \leq \eta$。

注 6.7. 上述引理严格改进了 [KSU20 定理 3 中的结果。5][6]特别是，该结果使用均值中位数法，并在第三项中达到 $\log D \cdot \log(1/\delta)\,\varepsilon$。相比之下，我们的第三项是加性的，而非乘性的。事实上，我们的集中性是最优的，与一维情形的下界相匹配（参见 [HKM22 定理 7。2]）。

推论 *6.8*（均值集中，$\alpha > 0$）。设定理 *6.5* 中的相同假设成立。对于任意 $\varepsilon \in (0,1]$ 且 $\alpha \in (0,\ 0.133)$, 令 $r = \iota^{1/k}$，其中 $\iota = \frac{1-\alpha}{0.249-\alpha}$ 且 $M = \Theta(\min\{\left(\frac{n\varepsilon}{\log(1/\delta)}\right)^{1/k}, (\alpha)^{-1/k}\})$。那么，存在常数 c_1，使得对所有 n，满足 $n \geq \mathcal{T} = \Omega(\max\{\frac{\iota}{\log(1/\delta)\,\varepsilon}, \frac{c_1}{\log(D/\delta)\,\varepsilon}, \frac{\log(1/\delta)}{\alpha^2}\})$ 且 $D \geq 2r$ 时，对于任意 $\delta \in (0,\ 1)$，以至少 $1-\delta$ 的概率，$|\tilde{\mu} - \mu| \leq \beta$ 满足

$$\beta = O\left(\sqrt{\frac{\log(1/\delta)}{n}} + \left(\frac{\log(1/\delta)}{n\varepsilon}\right)^{1-\frac{1}{k}} + \alpha^{1-\frac{1}{k}}\right).$$

注释 6.9。上述集中性与定理6.1中的形式相同。具体而言，当样本量 n较大时，它具有最优集中性（对于较小的 α）。现在，$\mathcal{T}$关于 n 的阈值取决于 α、ε。值得注意的是，即使对于没有隐私保护的次高斯内点分布，现有的集中性结果也存在一个阈值 $\mathcal{T} = \frac{1}{\log(1/\delta)\,\alpha^2}$（参见 Mukherjee 等人 [MTCD21]中的引理 4.1）。

现在，我们需要利用上述两个集中性结果，为元算法设计具体的实例化方案，并证明其性能保证。

我们的第一个实例化方案适用于未污染情形，即 $\alpha = 0$。因此，此时的鲁棒性仅针对重尾奖励，同时仍能保证隐私。

定理 6.10（性能保证，$\alpha = 0$）。考虑一个私有鲁棒多臂老虎机，其内点分布满足定义 *3.1* 以及 $\alpha = 0$。设算法 *1* 采用算法 *3*进行实例化，并且 r、M_τ、β_τ 由推论 *6.6* 给出，其中 n 替换为 B_τ。令 $\mathcal{T} = \Omega(\frac{\log(D/\delta)}{\varepsilon})$，且 $\delta = 1/T$。则算法 *1* 满足ε-差分隐私，其遗憾上界为

$$O\left(\sqrt{KT\log T} + \left(\frac{K\log T}{\varepsilon}\right)^{\frac{k-1}{k}} T^{\frac{1}{k}} + \gamma\right),$$

其中 $\gamma := O\left(\frac{KD\log(\frac{DT}{\varepsilon})}{\varepsilon}\right)$。

[6]我们还注意到，Kamath 等人 [KSU20]的主要关注点并非实现最优估计。

注释6.11。 据我们所知，这是首个针对满足有限中心矩假设的隐私重尾老虎机问题的结果。当前最先进的结果仅关注较简单的情形，即有限原始矩假设 [TWZW21]。

最后，借助推论 6.8，我们得到了元算法的第二种实现，用于处理污染情形。

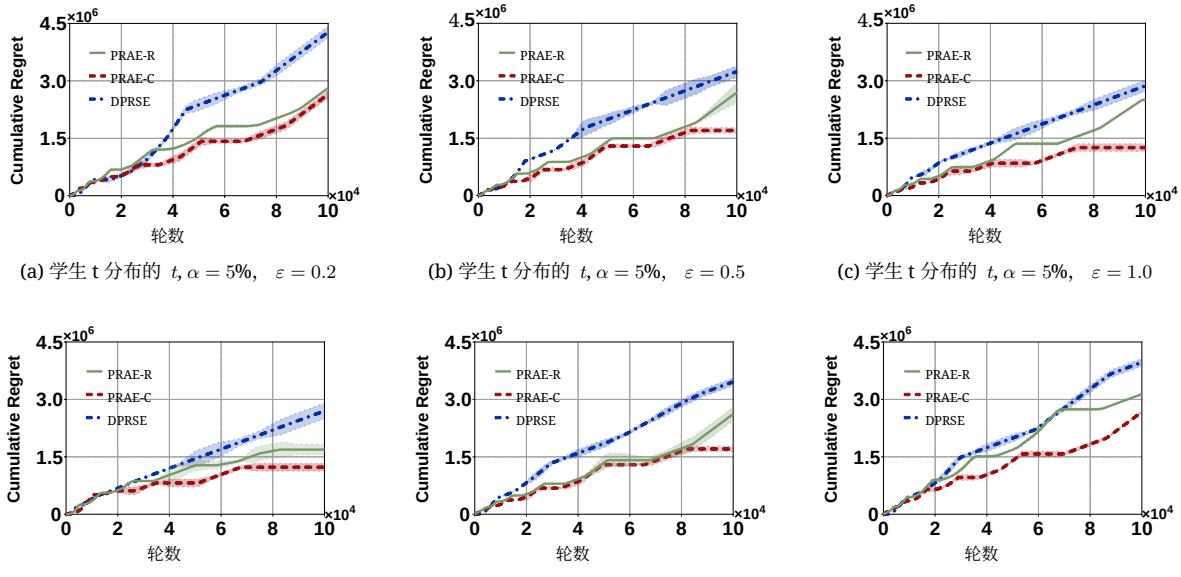
定理 6.12 （性能保证， $\alpha > 0$）。 考虑一个私有鲁棒多臂老虎机，其内点分布满足定义 *3.1* 和 $\alpha \in (0,\ 0。133)$。令算法 *1* 采用算法 *3* 实例化，并令 r， $\mathcal{T}$， M_τ， β_τ 由推论 *6.8* 给出，其中 n 替换为 B_τ。设 $\delta = 1/T$，则算法 *1* 是 ε-差分隐私，其遗憾上界为

$$O\left(\sqrt{KT\log T} + \left(\frac{K\log T}{\varepsilon}\right)^{\frac{k-1}{k}} T^{\frac{1}{k}} + T\alpha^{1-\frac{1}{k}} + \widehat{\gamma}\right),$$

其中 $\widehat{\gamma}\colon = O\left(\underbrace{DK_{\text{对数}}}_{\alpha^2}\ \underbrace{T}_{+}\ \underbrace{\iota DK_{\text{对数}}}_{\varepsilon}\ \underbrace{T}_{+}\ \underbrace{DK_{\text{对数}}(}_{\varepsilon}\ \underbrace{DT}_{)}\right)$ 且 $\iota = \frac{1-\alpha}{0.249-\alpha}$。

上述上界与我们的下界也相差至多 $O(\widehat{\gamma})$，当 T 足够大且 α为常数时， D这一项会被其他项主导。

7 实验



(a) 学生 t 分布的 $t, \alpha = 5\%,\ \varepsilon = 0.2$ (b) 学生 t 分布的 $t, \alpha = 5\%,\ \varepsilon = 0.5$ (c) 学生 t 分布的 $t, \alpha = 5\%,\ \varepsilon = 1.0$

(d) 帕累托分布， $\alpha = 2\%,\ \varepsilon = 0.5$ (e) 帕累托分布， $\alpha = 5\%,\ \varepsilon = 0.5$ (f) 帕累托分布， $\alpha = 10\%,\ \varepsilon = 0.5$

图 1：PRAE-R、PRAE-C 和 DPRSE 的累积遗憾比较。**上：**奖励由 Student's t 分布生成，固定 Huber参数 $\alpha = 10\%$，隐私预算 $\varepsilon \in \{0.2, 0.5, 1.0\}$ 变化。**下：**奖励由帕累托分布生成，固定隐私预算 $\varepsilon = 1.0$ ，Huber参数 $\alpha \in \{2\%,\ 5\%,\ 10\%\}$。

本节将通过实证评估我们的隐私鲁棒臂消除算法的实际性能。当子程序 PRM 对于有限原始矩情形为算法2，对于有限中心矩情形为算法3 时，这两种算法分别简称为 PRAE-R 和 PRAE-C。我们将它们与 Tao 等人的 DPRSE算法进行比较 [TWZW21], ，后者实现了差分隐私重尾多臂老虎机的最优遗憾界。

7.1 实验设置

我们考虑有 $K = 5$ 个臂的情况，并且每个臂的均值都在 $[0, 100]$ 范围内。具体而言，我们令各臂的均值线性递减，即对于每个臂 $a \in [k]$，令 $\mu_a = 100 - \frac{100}{K-1}(a-1)$。对于真实内点奖励生成，我们考虑以下两种重尾分布：

- *Pareto*分布： 对于臂 a 的每次拉取，我们生成一个从分布 $\mu_a + \eta - 2.5$ 中采样的奖励，其中 $\eta \sim \frac{sx_m^s}{x^{s+1}} \mathbb{1}_{\{x \geq x_m\}}$ 对于 $x \in \mathbb{R}$，并设置形状参数 $s = 2.5$ 和尺度参数 $x_m = 1.5$。

- *Student's t*-分布： 对于臂 a 的每次拉取，我们生成一个从分布$\mu_a + \eta$中采样的奖励，其中 $\eta \sim \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$。这里我们将自由度设为 $\nu = 2.5$。在这两种情况下，随机奖励的二阶中心矩均为 5。上述两种分布的主要区别在于，Student's t-分布是对称的，而 Pareto 分布是单侧的。为了生成受污染的奖励，我们考虑均值为零、标准差为 50 的高斯分布。

对于 Student 的 t-分布情形，我们将在不同的隐私预算$\varepsilon \in \{0.2, 0.5, 1\}$下固定 Huber参数 $\alpha = 5\%$，而对于 Parote 分布情形，我们固定 $\varepsilon = 0.5$ 并改变污染水平$\alpha = \{2\%, 5\%, 10\%\}$。对于每次实验，我们重复 30 次，并设置总轮数 $T = 10^5$（因此我们设定 $\delta = 10^{-5}$）。我们将报告相对于轮数的平均累积遗憾 $\mathcal{R}_T$。

7.2 结果与讨论

我们在图1中展示了我们的结果。对于所有情形，PRAE-R 和 PRAE-C 都实现了更小的累积遗憾，因此其预期性能优于 DPRSE。更具体地说，从图 1(d)-1(f)中可以看到，当 Huber参数 α 增大时，DPRSE 的遗憾会发散并增大，而 PRAE-R 和 PRAE-C 仅受到有限影响。这是因为 DPRSE 采用了更激进的截断阈值，从而纳入了更多离群值。相比之下，PRAE-R 和 PRAE-C 中的截断阈值经过精心设计，因此能够针对受污染的奖励提供鲁棒性。此外，通过观察误差条，我们发现三种方法在对称和单侧重尾分布下都很稳定（例如，见图 1(b)和 1(e)）。因此，我们可以得出结论：我们的 PRAE-R 和 PRAE-C 方法优于基线方法 DPRSE。

对于 PRAE-R 和 PRAE-C，我们还可以观察到，当 ε 变小或 α 变大时，两种分布类型下的遗憾都会增加，这是因为当$1/\varepsilon$ 和 α 足够大时，遗憾界与它们成正比。此外，与 PRAE-R 相比，我们可以看到 PRAE-C 在所有实验中的遗憾都更低。这是因为 PRAE-C 的 PRM 子程序利用了每个臂均值的先验信息（即范围 D），从而为算法1提供了更精细的性能界。总的来说，以上结果均印证了我们的理论。

8 结论

本文研究了 Huber 污染模型及差分隐私约束下，奖励具有重尾特性的私有稳健多臂老虎机问题。我们提出了一种元算法，其基础是一个私有稳健均值估计子程序 PRM。针对两种不同的重尾设置，我们分别给出了 PRM 的具体方案，两者都仅依赖于截断和拉普拉斯机制。此外，我们还为这些算法建立了遗憾上界，它们与我们推导出的极小极大下界几乎匹配。我们还进行了实验，以支持我们的理论分析。

9 致谢

YW 和 DW 的部分研究经费由阿卜杜拉国王科技大学的 BAS/1/1689-01-01、URF/1/4663-01-01 和 FCC/1/1976-49-01 项目资助。XZ 的部分研究经费由 NSF CNS-2153220 资助。我们感谢 Mengchu Li 的深刻讨论，并感谢他指出了 [LBY22] 的第一个 arXiv 版本。

参考文献

[AB22] A. Azize 和 D. Basu。“当隐私遇上部分信息：差分隐私赌博机的精细分析”。载于：*arXiv* 预印本 *arXiv:2209.02570*（2022）。[AGKS20] P. Awasthi、S. Gollapudi、K. Kollias 和 A. Sadhwani。“对抗性污染下的在线学习”。载于：（2020）。[AHMRV19] J. Adams、D. Hayunga、S. Mansi、D. Reeb 和 V. Verardi。“识别和处理金融领域的离群值”。载于：*Financial Management* 48.2（2019），第 345–384 页。[BCL13] S. Bubeck、N. Cesa-Bianchi 和 G. Lugosi。“重尾多臂老虎机”。载于：*IEEE* 信息论汇刊59.11（2013），第 7711–7717 页。[BF85] D. A. Berry 和 B. Fristedt。“多臂老虎机问题：实验的序贯分配（统计与应用概率专著）”。载于：伦敦：*Chapman and Hall* 5.71-87（1985），第 7–7 页。[BMM22] D. Basu、O.-A. Maillard 和 T. Mathieu。“自然污染的多臂老虎机：遗憾下界与稳健乐观算法”。载于：*arXiv* 预印本 *arXiv:2203.03186*（2022）。[CB13] S. Caron 和 S. Bhagat。“混合赌博机：改进社交网络冷启动推荐的方法”。载于：第 *7* 届社交网络挖掘与分析研讨会论文集。2013，第 1–9 页。[CGR18] M. Chen、C. Gao 和 Z. Ren。“Huber 污染模型下的稳健协方差与散布矩阵估计”。载于：统计年刊 46.5（2018），第 1932–1960 页。[CKMY22] S. Chen、F. Koehler、A. Moitra 和 M. Yau。“在线无分布稳健性：Huber 污染下的回归与上下文赌博机”。载于：*2021* 年 *IEEE* 第 *62* 届计算机科学基础年会（*FOCS*）。IEEE。2022，第 684–695 页。[CMW22] X. Chen、S. Miao 和 Y. Wang。“非参数需求模型下个性化定价中的差分隐私”。载于：*Operations Research*（2022）。[CSS11] T.-H. H. Chan、E. Shi 和 D. Song。“统计数据的私有持续发布”。载于：*ACM* 信息与系统安全汇刊（*TISSEC*）14.3（2011），第 1–24 页。

[CSW22] X. Chen、D. Simchi-Levi 和 Y. Wang。“结合需求学习的隐私保护动态个性化定价”。载于：《管理科学》68.7（2022），第 4878–4898 页。[CZ22] S. R. Chowdhury 和 X. Zhou。“多臂老虎机中的分布式差分隐私”。载于：arXiv 预印本 arXiv:2206.05772（2022）。[Dia18] I. Diakonikolas。“算法高维稳健统计”。载于：网页 http://www. iliasdiakonikolas. org/simons-tutorial-robust. html（2018）。[DMNS06] C. Dwork、F. McSherry、K. Nissim 和 A. Smith。“在私有数据分析中根据敏感度校准噪声”。载于：密码学理论会议。Springer。2006，第 265–284 页。[DR+14] C. Dwork、A. Roth 等。“差分隐私的算法基础”。载于：Found. Trends Theor. Comput. Sci. 9.3-4（2014），第 211–407 页。[EMMM06] E. Even-Dar、S. Mannor、Y. Mansour 和 S. Mahadevan。“多臂老虎机与强化学习问题中的动作消除和停止条件”。载于：机器学习研究期刊 7.6（2006）。[FLJLPR14] M. Fredrikson、E. Lantz、S. Jha、S. Lin、D. Page 和 T. Ristenpart。“药物基因组学中的隐私：个性化华法林剂量给药的 {端到端} 案例研究”。载于：第 23 届 USENIX 安全研讨会（USENIX Security 14）。2014，第 17–32 页。[GH22] K. Georgiev 和 S. B. Hopkins。“隐私诱导稳健性：信息-计算间隙与稀疏均值估计”。载于：arXiv 预印本 arXiv:2211.00724（2022）。[GPKW17] B. Guti´errez、L. Peter、T. Klein 和 C. Wachinger。“利用多臂老虎机从大型医疗数据中智能选择训练集”。载于：国际医学图像计算与计算机辅助干预会议。Springer。2017，第 38–45 页。[HBD+11] M. Hoffman、E. Brochu、N. De Freitas 等。“贝叶斯优化的投资组合配置”。载于：不确定性人工智能会议。Citeseer。2011，第 327–336 页。[HKM22] S. B. Hopkins、G. Kamath 和 M. Majid。“通过平方和指数机制实现纯差分隐私下的高效均值估计”。载于：第 54 届 ACM SIGACT 计算理论研讨会论文集。2022，第 1406–1417 页。[HKMN22] S. B. Hopkins、G. Kamath、M. Majid 和 S. Narayanan。“统计估计中的稳健性蕴含隐私”。载于：arXiv 预印本 arXiv:2212.05015（2022）。[Hub64] P. J. Huber。“位置参数的稳健估计”。载于：Ann. Math. Statist. 35.4（1964），第 73–101 页。[Hul12] J. Hull。《金融机构风险管理》，+ 网站。第 733 卷。John Wiley & Sons，2012。[KPK19] S. Kapoor、K. K. Patel 和 P. Kar。“容错多臂老虎机学习”。载于：机器学习 108.4（2019），第 687–715 页。[KSU20] G. Kamath、V. Singhal 和 J. Ullman。“重尾分布的私有均值估计”。载于：学习理论会议。PMLR。2020，第 2204–2235 页。[KV17] V. Karwa 和 S. Vadhan。“有限样本差分隐私置信区间”。载于：arXiv 预印本 arXiv:1711.03908（2017）。[LBY22] M. Li、T. B. Berrett 和 Y. Yu。“论稳健性与局部差分隐私”。载于：arXiv 预印本 arXiv:2201.00751（2022）。

[LMM20] Y. M. Lei、S. Miao 和 R. Momot。《隐私保护的个性化收益管理》。载于：巴黎高等商学院研究论文第 MOSI-2020-1391 号（2020）。[LMP18] T. Lykouris、V. Mirrokni 和 R. Paes Leme。《对抗性污染下的稳健随机赌博机》。载于：第 50 届 ACM SIGACT年度计算理论研讨会论文集。2018 年，第 114–122 页。[LRV16] K. A. Lai、A. B. Rao 和 S. Vempala。《均值与协方差的无先验估计》。载于：2016 年 IEEE 第 57 届年度计算机科学基础研讨会（FOCS）。IEEE。2016 年，第 665–674 页。[LS20] T. Lattimore 和 C. Szepesv´ari。《赌博机算法》。剑桥大学出版社，2020 年。[MT15] N. Mishra 和 A. Thakurta。《（近乎）最优的差分隐私随机多臂赌博机》。载于：第 31 届人工智能不确定性会议（UAI）论文集。2015 年，第 592–601 页。[MTCD21] A. Mukherjee、A. Tajer、P.-Y. Chen 和 P. Das。《奖励污染下随机赌博机中的基于均值的最佳臂识别》。载于：神经信息处理系统进展 34（2021），第 9651–9662 页。[NA21] J. P. Near 和 C. Abuah。《差分隐私编程》。载于：URL：https://uvm（2021）。[NT20] L. Niss 和 A. Tewari。《眼见未必为实：对 ε-污染具有鲁棒性的 UCB赌博机算法》。载于：人工智能不确定性会议。PMLR。2020 年，第 450–459 页。[PBR19] A.Prasad、S.Balakrishnan 和 P. Ravikumar。《稳健均值估计的统一方法》。载于：arXiv 预印本 arXiv:1907.00927（2019）。[Rac03] S. T. Rachev。《金融重尾分布手册：金融手册，第 1 卷》。Elsevier，2003 年。[RZLS20] W. Ren、X. Zhou、J. Liu 和 N. B. Shroff。《具有局部差分隐私的多臂老虎机》。载于：arXiv 预印本 arXiv:2007.03121（2020）。[SBF17] E. M. Schwartz、E. T. Bradlow 和 P. S. Fader。《通过多臂赌博机实验利用展示广告获取客户》。载于：《营销科学》36.4（2017），第 500–522 页。[SS19] T. Sajed 和 O. Sheffet。《基于最优隐私停止规则的最优隐私随机多臂赌博机算法》。载于：第 36 届国际机器学习大会（ICML）论文集。2019 年，第 5579–5588 页。[SWJZ15] W. Shen、J. Wang、Y.-G. Jiang 和 H. Zha。《基于正交赌博机学习的投资组合选择》。载于：第二十四届国际人工智能联合会议。2015 年。[TKMS21] J. Tenenbaum、H. Kaplan、Y. Mansour 和 U. Stemmer。《洗牌模型下的差分隐私多臂赌博机》。载于：神经信息处理系统进展 34（2021）。[TWZW21] Y. Tao、Y. Wu、P. Zhao 和 D. Wang。《（局部）差分隐私重尾多臂赌博机的最优速率》。载于：arXiv 预印本 arXiv:2106.02575（2021）。[Ver18] R. Vershynin。《高维概率：数据科学中的应用导论》。第 47 卷。剑桥大学出版社，2018 年。

实用引理

引理 A.1 (后处理 [DR+14])。 设 $\mathcal{M}: \mathcal{X} \to \mathcal{Y}$ 是一个随机化算法，且满足 (ε, δ)- 差分隐私。设 $f: \mathcal{Y} \to \mathcal{Z}$ 是任意随机映射。那么 $f \circ \mathcal{M}: \mathcal{X} \to \mathcal{Z}$ 满足(ε, δ)-差分隐私。

引理 *A.2*（组合定理 [*DR*+14]）。设 $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_h$ 是一系列随机算法，其中 $\mathcal{M}_1: \mathcal{X}^n \to \mathcal{Y}_1, \mathcal{M}_2: \mathcal{Y}_1 \times \mathcal{X}^n \to \mathcal{Y}_2, \ldots, \mathcal{M}_h: \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_{h-1} \times \mathcal{X}^n \to \mathcal{Y}_h$。假设对于每个 $i \in [h]$ 和 $y_1 \in \mathcal{Y}_1, y_2 \in \mathcal{Y}_2, \ldots, y_h \in \mathcal{Y}_h$，都有 $\mathcal{M}_i(y_1, \ldots, y_{i-1}, \cdot): \mathcal{X}^n \to \mathcal{Y}_i$ 是 ε_i-差分隐私。那么，依次运行算法 $\mathcal{M}_i$ 的算法 $\mathcal{M}: \mathcal{X}^n \to \mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \mathcal{Y}_h$ 对于$\varepsilon = \sum_{i=1}^{h} \varepsilon_i$ 是 ε-差分隐私。

引理 A.3（并行组合 [NA21]）。假设有 n ε个差分隐私机制$\{\mathcal{M}_i\}_{i=1}^n$ 以及 n 个不相交数据集，记为 $\{D_i\}_{i=1}^n$。那么，对于在相应的 $\mathcal{M}_i$ 上分别应用每个 D_i的算法，它是 ε-差分隐私。

引理 A.4（马尔可夫不等式）。 若 $Y \in \mathbb{R}$ 是一个随机变量，且 $a > 0$，则有

$$\mathbb{P}(|Y| \geqslant a) \leqslant \frac{\mathbb{E}\left(|Y|^k\right)}{a^k}$$

引理 A.5 （切比雪夫不等式）。 对于一个实值随机变量 $Y \in \mathbb{R}$、 $a > 0$ 和 $k \in \mathbb{N}$，有

$$\mathbb{P}(|Y - \mathbb{E}Y| \geqslant a) = \mathbb{P}\left(|Y - \mathbb{E}Y|^k \geqslant a^k\right) \leqslant \frac{\mathbb{E}\left(|Y - \mathbb{E}Y|^k\right)}{a^k}$$

引理 A.6（拉普拉斯随机变量的尾界 [DMNS06]）。 若 $X \sim$ 拉普拉斯分布(b)，则

$$\mathbb{P}(|X| \geq t \cdot b) = \exp(-t).$$

引理 A.7 （霍夫丁不等式）。 设 Z_1、…、Z_n 为独立有界随机变量，且对所有 i，$Z_i \in [a, b]$，其中 $-\infty < a < b < \infty$。则

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}\left(Z_i - \mathbb{E}\left[Z_i\right]\right)\right| \geqslant t\right) \leqslant 2\exp\left(-\frac{2nt^2}{(b-a)^2}\right)$$

引理 A.8 （H¨older 不等式）。 设 X、Y 是定义在 $\mathbb{R}$ 上的随机变量，并设 $k > 1$。则，

$$\mathbb{E}[|XY|] \leqslant \left(\mathbb{E}\left[|X|^k\right]\right)^{\frac{1}{k}} \left(\mathbb{E}\left[|Y|^{\frac{k}{k-1}}\right]\right)^{\frac{k-1}{k}}$$

引理 A.9 (伯恩斯坦不等式 [版本18])。 设 $X_1, \cdots X_n$ 是 n 相互独立的零均值随机变量。假设 $|X_i| \leqslant M$ 且 $\mathbb{E}[X_i^2] \leq s$ 对所有 i均成立。那么对于任意 $t > 0$，有

$$\mathbb{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| \geqslant t\right\} \leqslant 2\exp\left(-\frac{\frac{1}{2}t^2 n}{s + \frac{1}{3}Mt}\right)$$

B 第4节证明

引理 B.1(多臂老虎机中 ε-差分隐私 [AB22] 的 KL 散度上界)。 若 π 是一个满足ε-差分隐私的机制，则对于两个实例 $\nu_1 = (r_a : \ a \in [K])$ 和 $\nu_2 = (r'_a : \ a \in [K])$ 有

$$KL\left(\mathbb{P}^T_{\pi,\nu_1} \| \mathbb{P}^T_{\pi,\nu_2}\right) \leq 6\varepsilon \mathbb{E}_{\pi,\nu_1}\left[\sum_{t=1}^{T} TV(r_{a_t} \| r'_{a_t})\right]$$

其中，$TV(r_a \| r'_a)$ 是 r_a 与 r'_a之间的全变差距离。

引理 B.2 (定理 5.1，见 [CGR18])。 设 R_1和 R_2是定义在 $\mathcal{X}$上的两个分布。若对于某个 $\alpha \in [0,1]$, 有 $TV(R_1, R_1) = \frac{\alpha}{1-\alpha}$，则存在定义在同一概率空间上的两个分布 G_1和G_2，使得

$$(1-\alpha)R_1 + \alpha G_1 = (1-\alpha)R_2 + \alpha G_2.$$

定理 4.1 的证明。设 Π 为所有策略的集合， Π^ε 为所有 ε-差分隐私策略的集合。我们将对应于由均值为 $\mu \in \mathbb{R}^K$、方差相同为 σ_k^2 的 K-高斯分布构成的集合的环境记为 k，其中 σ_k 的取值由此确定，从而使分布的 k阶原始矩有界于 1，即 $\mathcal{E}^K_{\mathcal{N}}(\sigma_k) \triangleq \left\{\left(\mathcal{N}\left(\mu_i, \sigma_k^2\right)\right)_{i=1}^K : \ \mu = (\mu_1, \ldots, \mu_K) \in \mathbb{R}^K\right\}$。由于 $\Pi^\varepsilon \subset \Pi$，我们可以得到

$$\mathcal{R}_T^{\text{minimax}}(\pi, \nu) \geqslant \inf_{\pi \in \Pi} \sup_{\nu \in \mathcal{E}^K_{\mathcal{N}}(\sigma_k)} \text{Reg}_T(\pi, \nu) \geqslant \Omega(\sqrt{KT})$$

其中，最后一个不等式由 Lattimore 和 Szepesv´ari [LS20]中的定理 15.2 得出。

情形 1：未污染情形。根据极小极大遗憾的定义，我们知道$\mathcal{R}_\varepsilon^{\text{minimax}}{}_{,\alpha} \geq \mathcal{R}_{\varepsilon\,0}^{\text{minimax}}{}_{,}$. 因此，我们首先推导无污染的私有老虎机下界。

考虑两种 个环境。在第一个环境 ν_1中，最优臂（记为 a_1）服从

$$r_{a_1} = \begin{cases} 1/\gamma & \text{with probability of } \frac{1}{2}\gamma^k \\ 0 & \text{with probability of } 1 - \frac{1}{2}\gamma^k \end{cases}$$

其中 $\gamma \in (0,1]$。我们可以验证 $\mathbb{E}[r_{a_1}] = \frac{1}{2}\gamma^{k-1}$和 $\mathbb{E}[r_{a_1}^k] = \frac{1}{2} \leq 1$。任何其他次优臂 $a \neq a_1$在ν_1 中都服从相同的奖励分布

$$r_a = \begin{cases} 1/\gamma & \text{with probability of } \frac{3}{10}\gamma^k \\ 0 & \text{with probability of } 1 - \frac{3}{10}\gamma^k \end{cases}$$

我们可以验证 $\mathbb{E}[r_a] = \frac{3}{10}\gamma^{k-1}$和 $\mathbb{E}[r_{a_1}^k] = \frac{3}{10} \leq 1$。于是，最优臂与次优臂之间的均值差为 $\Delta = \frac{1}{5}\gamma^{k-1}$。

对于算法 π和实例 ν_1，我们记 $i = \arg\min_{a \in \{2,\cdots,K\}} \mathbb{E}_{\pi,\ \nu_1}[N_a(T)]$. 因此，$\mathbb{E}_{\pi,\ \nu_1}[N_i(T)] \leqslant TK-1$。

现在，考虑另一个实例 ν_2，其中 r_{a_1}，$\cdots$，r_{a_k}与 ν_1中的相同，但第 i 个臂除外，使得

$$r'_i = \begin{cases} 1/\gamma & \text{with probability of } \frac{7}{10}\gamma^k \\ 0 & \text{with probability of } 1 - \frac{7}{10}\gamma^k \end{cases}$$

我们可以验证 $\mathbb{E}[r'_i] = \frac{7}{10}\gamma^{k-1}$ 和 $\mathbb{E}[(r'_i)^k] = \frac{7}{10} \leq 1$。那么在 ν_2 中，臂 i 是最优的。

根据经典遗憾分解，我们得到

$$\mathcal{R}_T(\pi,\nu_1)=(T-\mathbb{E}_{\pi,\nu_1}[N_1(T)])\Delta\geq\mathbb{P}^T_{\pi,\nu_1}\left[N_1(T)\leq\frac{T}{2}\right]\frac{T\Delta}{2}.$$

$$\mathcal{R}_T(\pi,\nu_2)=\Delta\mathbb{E}_{\pi,\nu_2}[N_1(T)]+\sum_{a\notin\{1,i\}}2\Delta\mathbb{E}_{\pi,\nu_2}[N_a(T)]\geq\mathbb{P}^T_{\pi,\nu_2}\left[N_1(T)\geq\frac{T}{2}\right]\frac{T\Delta}{2}.$$

应用 Bretagnolle–Huber 不等式（[LS20],定理 14.2），我们有

$$\begin{aligned}\mathcal{R}_T(\pi,\nu_1)+\mathcal{R}_T(\pi,\nu_2)&\geq\frac{T\Delta}{2}\left(\mathbb{P}^T_{\pi,\nu_1}\left[N_1(T)\leq\frac{T}{2}\right]+\mathbb{P}^T_{\pi,\nu_2}\left[N_1(T)\geq\frac{T}{2}\right]\right).\\&\geq\frac{T\Delta}{4}\exp\left(-\mathrm{KL}\left(\mathbb{P}^T_{\pi,\nu_1}\|\mathbb{P}^T_{\pi,\nu_2}\right)\right)\end{aligned}$$

根据引理B.1，我们可以得到边缘分布之间 KL 散度的上界。

$$\begin{aligned}\mathrm{KL}\left(\mathbb{P}^T_{\pi,\nu_1}\|\mathbb{P}^T_{\pi,\nu_2}\right)&\leq 6\varepsilon\mathbb{E}_{\pi,\nu_1}\left[\sum_{t=1}^{T}\mathrm{TV}(r_{a_t}\|r'_{a_t})\right]\\&\leq 6\varepsilon\mathbb{E}_{\pi,\nu_1}[N_i(T)]\mathrm{TV}(r_i\|r'_i)\end{aligned}$$

因为 ν_1 和 ν_i 仅在臂 i上有所不同。

因此，

$$\begin{aligned}\mathcal{R}_T(\pi,\nu_1)+\mathcal{R}_T(\pi,\nu_2)&\geq\frac{T\Delta}{4}\exp\left(-6\varepsilon\mathbb{E}_{\pi,\nu_1}[N_i(T)]\cdot\frac{2}{5}\gamma^k\right)\\&\geq\frac{T\gamma^{k-1}}{20}\exp\left(-\frac{12\cdot\varepsilon T\gamma^k}{5(K-1)}\right).\end{aligned}$$

取 $\gamma=\left(\frac{K-1}{T\varepsilon}\right)^{\frac{1}{k}}$，我们得到结果

$$\mathcal{R}_T(\pi,\nu_1)\geq\Omega\left(\left(\frac{K}{\varepsilon}\right)^{\frac{k-1}{k}}T^{\frac{1}{k}}\right).$$

情形 2：污染情形。对于 $\alpha\neq 0$ 和 $\alpha\in(0,1]$, 我们仍假设上述 ν_1 和 ν_2中的臂的真实分布相同。在第一个环境 ν_1中，最优臂（记为 a_1）服从

$$r_{a_1}=\begin{cases}1/\gamma & \text{with probability of } \frac{1}{2}\gamma^k\\ 0 & \text{with probability of } 1-\frac{1}{2}\gamma^k\end{cases}$$

其中 $\gamma\in(0,1]$。我们可以验证 $\mathbb{E}[r_{a_1}]=\frac{1}{2}\gamma^{k-1}$和 $\mathbb{E}[r^k_{a_1}]=\frac{1}{2}\leq 1$。任何其他次优臂 $a\neq a_1$在ν_1 中都遵循相同的奖励分布

$$r_a=\begin{cases}1/\gamma & \text{with probability of } \frac{3}{10}\gamma^k\\ 0 & \text{with probability of } 1-\frac{3}{10}\gamma^k\end{cases}$$

我们可以验证 $\mathbb{E}[r_a]=\frac{3}{10}\gamma^{k-1}$和 $\mathbb{E}[r^k_{a_1}]=\frac{3}{10}\leq 1$。那么，最优臂与次优臂之间的均值差为 $\Delta=\frac{1}{5}\gamma^{k-1}$。

并且，我们将 ν_1 的污染版本记为 $\tilde{\nu}_1$。对于算法 π 和实例 $\tilde{\nu}_1$，我们记$i = \arg\min_{a\in\{2,\cdots,K\}} \mathbb{E}_{\pi,\tilde{\nu}_1}[N_a(T)]$. 因此，$\mathbb{E}_{\pi,\tilde{\nu}_1}[N_i(T)] \leqslant \frac{T}{K-1}$。

现在，考虑另一个实例 ν_2，其中 r_{a_1}，$\cdots$，r_{a_k} 与 ν_1 中的相同，但第 i 个臂除外，即

$$r'_i = \begin{cases} 1/\gamma & \text{with probability of } \frac{7}{10}\gamma^k \\ 0 & \text{with probability of } 1 - \frac{7}{10}\gamma^k \end{cases}$$

我们可以验证 $\mathbb{E}[r'_i] = \frac{7}{10}\gamma^{k-1}$ 和 $\mathbb{E}[(r'_i)^k] = \frac{7}{10} \leq 1$。那么在 ν_2 中，臂 i 是最优的 .

此外，我们将 ν_2 的污染版本记为 $\tilde{\nu}_2$。取 $\gamma = \alpha^{\frac{1}{k}} \in (0, 1]$, 因为对于任意 $a \in [K]$, TV（$r_a \| r'_a$）$\leq \frac{2}{5}\gamma^k = \frac{2}{5}\alpha \leq \frac{\alpha}{1-\alpha}$, 根据引理 B.2，对于任意臂 $a \in [K]$，存在分布G_a 和 G'_a 使得

$$(1-\alpha)r_a + \alpha G_a = (1-\alpha)r'_a + \alpha G'_a.$$

我们考虑 $\tilde{\nu}_1 = \{x_a = (1-\alpha)r_a + \alpha G_a : \ a \in [K]\}$ 和 $\tilde{\nu}_2 = \{x'_a = (1-\alpha)r'_a + \alpha G'_a : \ a \in [K]\}$。现在根据经典遗憾分解，我们得到

$$\mathcal{R}_T(\pi, \tilde{\nu}_1) = (T - \mathbb{E}_{\pi,\tilde{\nu}_1}[N_1(T)])\Delta \geq \mathbb{P}^T_{\pi,\tilde{\nu}_1}\left[N_1(T) \leq \frac{T}{2}\right]\frac{T\Delta}{2}.$$

$$\mathcal{R}_T(\pi, \tilde{\nu}_2) = \Delta\mathbb{E}_{\pi,\tilde{\nu}_2}[N_1(T)] + \sum_{a\notin\{1,i\}} 2\Delta\mathbb{E}_{\pi,\tilde{\nu}_2}[N_a(T)] \geq \mathbb{P}^T_{\pi,\tilde{\nu}_2}\left[N_1(T) \geq \frac{T}{2}\right]\frac{T\Delta}{2}.$$

应用 Bretagnolle–Huber 不等式（[LS20],定理 14.2），我们得到

$$\begin{aligned}\mathcal{R}_T(\pi, \tilde{\nu}_1) + \mathcal{R}_T(\pi, \tilde{\nu}_2) &\geq \frac{T\Delta}{2}\left(\mathbb{P}^T_{\pi,\tilde{\nu}_1}\left[N_1(T) \leq \frac{T}{2}\right] + \mathbb{P}^T_{\pi,\tilde{\nu}_2}\left[N_1(T) \geq \frac{T}{2}\right]\right). \\ &\geq \frac{T\Delta}{4}\exp\left(-\mathrm{KL}\left(\mathbb{P}^T_{\pi,\tilde{\nu}_1} \| \mathbb{P}^T_{\pi,\tilde{\nu}_2}\right)\right)\end{aligned}$$

基于 引理 B.1，我们可以得到边缘分布之间 KL 散度的上界。

$$\mathrm{KL}\left(\mathbb{P}^T_{\pi,\tilde{\nu}_1} \| \mathbb{P}^T_{\pi,\tilde{\nu}_2}\right) \leq 6\varepsilon\mathbb{E}_{\pi,\tilde{\nu}_1}\left[\sum_{t=1}^{T}\mathrm{TV}(x_{a_t} \| x'_{a_t})\right]$$

由于，$\mathrm{TV}(x_a \| x'_a) = 0$ 对于 $\forall a \in [K]$、$\Delta = \frac{1}{5}\gamma^{k-1}$ 和 $\gamma = \alpha^{\frac{1}{k}}$。-我们得到

$$\mathcal{R}_T(\pi, \tilde{\nu}_1) \geq \Omega(T\alpha^{1-\frac{1}{k}}).$$

综合高斯情形、情形 1 和情形 2，可得

$$\mathcal{R}_T = \Omega\left(\max\left\{\sqrt{KT}, \left(\frac{K}{\varepsilon}\right)^{1-\frac{1}{k}} T^{\frac{1}{k}}, T\alpha^{1-\frac{1}{k}}\right\}\right).$$

□

C 第6.1节的证明

定理 6.1 的证明。 我们用 P_k 表示奖励的有限原始矩分布，并用 $P_{\alpha,k}$ 表示在 α-Huber 污染下的 P_k。令 $\widehat{\mu} = \frac{1}{n} \sum_{\substack{i\in[n] \\ X_i\sim P_{\alpha,k}}} X_i \mathbb{1}_{(|X_i|\le M)}$ 且 $\mu = \mathbb{E}_{\overline{X_i}\sim P_k}[X_i]$。

$$
\begin{aligned}
|\tilde{\mu} - \mu| &\le \left|\mathrm{Lap}\left(\frac{2M}{n\varepsilon}\right)\right| + |\widehat{\mu} - \mu| \\
&\le \left|\mathrm{Lap}\left(\frac{2M}{n\varepsilon}\right)\right| + \left|\frac{1}{n}\sum_{\substack{i\in[n] \\ X_i\sim P_{\alpha,k}}} X_i\mathbb{1}_{(|X_i|\le M)} - \mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}]\right| + |\mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}] - \mu| \\
&= \left|\mathrm{Lap}\left(\frac{2M}{n\varepsilon}\right)\right| + \left|\frac{1}{n}\sum_{\substack{i\in[n] \\ X_i\sim P_{\alpha,k}}} X_i\mathbb{1}_{(|X_i|\le M)} - \mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}]\right| + |\mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|> M)}]| \\
&\overset{(a)}{\le} \frac{2M\log(2/\delta)}{n\varepsilon} + \left|\frac{1}{n}\sum_{\substack{i\in[n] \\ X_i\sim P_{\alpha,k}}} X_i\mathbb{1}_{(|X_i|\le M)} - \mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}]\right| + \mathbb{E}_{X_i\sim P_k}[|X_i|\mathbb{1}_{(|X_i|> M)}] \quad \text{w.p.} \quad 1-\frac{\delta}{2} \\
&\overset{(b)}{\le} \frac{2M\log(2/\delta)}{n\varepsilon} + \left|\frac{1}{n}\sum_{\substack{i\in[n] \\ X_i\sim P_{\alpha,k}}} X_i\mathbb{1}_{(|X_i|\le M)} - \mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}]\right| + (\mathbb{E}_{X_i\sim P_k}[|X_i|^k])^{\frac{1}{k}}(\mathbb{P}_{X_i\sim P_k}(|X_i|>M))^{\frac{k-1}{k}} \\
&\overset{(c)}{\le} \frac{2M\log(2/\delta)}{n\varepsilon} + \left|\frac{1}{n}\sum_{\substack{i\in[n] \\ X_i\sim P_{\alpha,k}}} X_i\mathbb{1}_{(|X_i|\le M)} - \mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}]\right| + \frac{1}{M^{k-1}}
\end{aligned}
$$

其中，不等式 (a) 由引理 A.6 得出，(b) 由引理 A.8 中的霍尔德不等式得出，而(c)由引理 A.4 中的马尔可夫不等式得出。

现在我们关注 $\left|\frac{1}{n}\sum_{\substack{i\in[n] \\ X_i\sim P_{\alpha,k}}} X_i\mathbb{1}_{(|X_i|\le M)} - \mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}]\right|$ 的上界。令 N_G 为

设 n 为按照 G 分布的样本中索引的集合， N_{P_k} 为按照 P_k分布的 n 个样本中的索引集合。那么

情形 1：未污染情形（$\alpha = 0$）现在，剩下唯一要做的就是给出上界

$$
\left|\frac{1}{n}\sum_{\substack{i\in[n] \\ X_i\sim P_k}} X_i\mathbb{1}_{(|X_i|\le M)} - \mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\le M)}]\right|.
$$

对于 $X_i \sim P_k$ $Y_i = X_i \mathbb{1}_{(|X_i| \le M}$ $|Y_i| \le M$ $Y_i = \mathbb{E}[Y_i^2] - \mathbb{E}[Y_i]^2 \le \mathbb{E}[Y_i^2] \le$，令），则和 $\mathbf{Var}()()$。然 $\mathbb{E}_{X_i \sim P_k}[X_i^2] \le 1$ $9,$ $1-\delta/2$
后，根据引理 A.中的伯恩斯坦不等式，以概率

$$\left| \frac{1}{n} \sum_{\substack{i \in [n] \\ X_i \sim P_k}} X_i \mathbb{1}_{(|X_i| \le M)} - \mathbb{E}_{X_i \sim P_k}[X_i \mathbb{1}_{(|X_i| \le M)}] \right| \le \sqrt{\frac{2\log(4/\delta)}{n}} + \frac{4M\log(4/\delta)}{3n}. \tag{4}$$

于是，以至少 $1-\delta$ 的概率，

$$|\tilde{\mu} - \mu| \le \sqrt{\frac{2\log(4/\delta)}{n}} + \frac{4M\log(4/\delta)}{3n} + \frac{2M\log(2/\delta)}{n\varepsilon} + \frac{1}{M^{k-1}}.$$

对于 $\varepsilon > 0$ 和 $\delta \in$ （0，1），以至少 $1-\delta$ 的概率，

$$|\tilde{\mu} - \mu| \le \sqrt{\frac{2\log(4/\delta)}{n}} + \frac{4M\log(4/\delta)}{n\varepsilon} + \frac{1}{M^{k-1}}.$$

取截断阈值 $M = \left(4 \quad {}^{n\varepsilon}{}_{\log(4/\delta)}\right)^{k}$，我们有 _

$$|\tilde{\mu} - \mu| \le \sqrt{\frac{2\log(4/\delta)}{n}} + 2\left(\frac{4\log(4/\delta)}{n\varepsilon}\right)^{1-\frac{1}{k}}.$$

情形 2：污染情形（$\alpha \in (0, \frac{1}{2}]$） _

$$\begin{aligned}
&\left| \frac{1}{n} \sum_{\substack{i \in [n] \\ X_i \sim P_{\alpha,k}}} X_i \mathbb{1}_{(|X_i| \le M)} - \mathbb{E}_{X_i \sim P_k}[X_i \mathbb{1}_{(|X_i| \le M)}] \right| \\
=& \left| \frac{1}{n} \sum_{i \in N_G} X_i \mathbb{1}_{(|X_i| \le M)} + \frac{1}{n} \sum_{i \in N_{P_k}} X_i \mathbb{1}_{(|X_i| \le M)} - \mathbb{E}_{X_i \sim P_k}[X_i \mathbb{1}_{(|X_i| \le M)}] \right| \\
\le& \underbrace{\left| \frac{1}{n} \sum_{i \in N_G} X_i \mathbb{1}_{(|X_i| \le M)} \right|}_{T_1} + \underbrace{\left| \frac{1}{n} \sum_{i \in N_{P_k}} X_i \mathbb{1}_{(|X_i| \le M)} - \mathbb{E}_{X_i \sim P_k}[X_i \mathbb{1}_{(|X_i| \le M)}] \right|}_{T_2}.
\end{aligned}$$

为了控制 T_1，我们可以将其写为

$$\begin{aligned}
T_1 &= \left| \frac{1}{n} \sum_{i \in N_G} X_i \mathbb{1}_{(|X_i| \le M)} \right| \\
&\le \frac{1}{n} \sum_{i \in N_G} |X_i| \, \mathbb{1}_{(|X_i| \le M)} \\
&\le \frac{|N_G|}{n} M.
\end{aligned}$$

那么 $\frac{|N_G|}{n}$ 可以视为伯努利分布的均值估计 $Ber(\alpha)$。根据引理 A.9中的伯恩斯坦不等式，我们以概率 $1-\delta/4$,

$$\left|\frac{|N_G|}{n}-\alpha\right| \leq \sqrt{\frac{2\alpha(1-\alpha)\log(8/\delta)}{n}}+\frac{2\log(8/\delta)}{3n}.$$

因此,

$$T_1 \leq \left(\alpha+\sqrt{\frac{2\alpha\log(8/\delta)}{n}}+\frac{2\log(8/\delta)}{3n}\right)M. \quad \text{with probability} 1-\delta/4$$

当 $n \geq \log(8/\delta)\,\alpha$ 时，我们有

$$T_1 \leq 4\alpha M.$$

为了对 T_2 进行界定，我们有

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i\in N_{P_k}} X_i\mathbb{1}_{(|X_i|\leq M)}-\mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\leq M)}]\right| \\
=&\left|\frac{1}{n}\sum_{\substack{i\in N_G\cup N_{P_k}\\ X_i\sim P_k}} X_i\mathbb{1}_{(|X_i|\leq M)}-\frac{1}{n}\sum_{\substack{i\in N_G\\ X_i\sim P_k}} X_i\mathbb{1}_{(|X_i|\leq M)}-\mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\leq M)}]\right| \\
\leq&\left|\frac{1}{n}\sum_{\substack{i\in N_G\\ X_i\sim P_k}} X_i\mathbb{1}_{(|X_i|\leq M)}\right|+\left|\frac{1}{n}\sum_{\substack{i\in [n]\\ X_i\sim P_k}} X_i\mathbb{1}_{(|X_i|\leq M)}-\mathbb{E}_{X_i\sim P_k}[X_i\mathbb{1}_{(|X_i|\leq M)}]\right| \\
\leq&4\alpha M+\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4M\log(16/\delta)}{3n} \quad \text{w.p.} \quad 1-\delta/4
\end{aligned}$$

其中，最后一个不等式基于对 T_1 的类似分析以及不等式（4）.

综合以上各项，以至少 $1-\delta$ 的概率，我们有,

$$|\tilde{\mu}-\mu| \leq \sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4M\log(16/\delta)}{3n}+8\alpha M+\frac{2M\log(2/\delta)}{n\varepsilon}+\frac{1}{M^{k-1}}.$$

因此，对于 $\varepsilon>0$，我们有

$$|\tilde{\mu}-\mu| \leq \sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4M\log(16/\delta)}{n\varepsilon}+\frac{1}{M^{k-1}}+8\alpha M.$$

取 $M=\min\left\{\left(\frac{n\varepsilon}{4\log(16/\delta)}\right)^{\frac{1}{k}},\ (8\alpha)^{-\frac{1}{k}}\right\}$，可得

$$|\tilde{\mu}-\mu| \leq \sqrt{\frac{2\log(16/\delta)}{n}}+2\left(\frac{4\log(16/\delta)}{n\varepsilon}\right)^{1-\frac{1}{k}}+2(8\alpha)^{1-\frac{1}{k}}.$$

□

定理 6.2 的证明。令 τ_0 为满足条件 $B_\tau < \log(16|\mathcal{S}|\tau^2/\delta)\,\alpha$ 的最大轮次。

对于所有轮次 $\tau \le \tau_0$，批次大小小于 2^{τ_0}。由于批次大小会翻倍，直到轮次 τ_0，每个臂 $a \in [K]$ 的拉取次数都小于 $2 \cdot 2^{\tau_0} \le 2^{\log(16|\mathcal{S}|\tau_0^2/\delta)}\,\alpha$。因此，每个 $a \in [K]$ 的遗憾至少为 $\frac{2\log(16|\mathcal{S}|\tau_0^2/\delta)}{\alpha}\,\Delta_a$。

对于 $\tau > \tau_0$，$B_\tau \ge \log(16|\mathcal{S}|\tau^2/\delta)\,\alpha$。对于每个 $a \in \mathcal{S}$，根据定理 6.1，以至少 $1 - \frac{\delta}{2|\mathcal{S}|\tau^2}$ 的概率，

$$|\tilde{\mu}_a - \mu_a| \le \beta_\tau.$$

给定轮次 $\tau > \tau_0$，我们将事件 $\mathcal{E}_\tau$ 记为：对所有 $a \in \mathcal{S}$，均有 $|\tilde{\mu}_a - \mu_a| \le \beta_\tau.$，并记 $\mathcal{E} = \cup_{\tau > \tau_0}\mathcal{E}_\tau$。通过取并集界，我们有

$$\mathbb{P}(\mathcal{E}_\tau) \ge 1 - \frac{\delta}{2\tau^2},$$

and

$$\mathbb{P}(\mathcal{E}) \ge 1 - \frac{\delta}{2}\left(\sum_{\tau > \tau_0} \tau^{-2}\right) \ge 1 - \delta.$$

下文中，我们以良好事件 $\mathcal{E}$ 为条件。首先证明最优臂 a^* 永远不会被淘汰。对于任意轮次 $\tau > \tau_0$，令 $a_\tau = \arg\max_{a \in \mathcal{S}} \widetilde{\mu}_a$。由于

$$(\widetilde{\mu}_{a_\tau} - \widetilde{\mu}_{a^*}) + \Delta_{a_\tau} = |(\widetilde{\mu}_{a_\tau} - \widetilde{\mu}_{a^*}) + \Delta_{a_\tau}| \le |\widetilde{\mu}_{a_\tau} - \mu_{a_\tau}| + |\widetilde{\mu}_{a^*} - \mu_{a^*}| \le 2\beta_\tau,$$

很容易看出，算法不会淘汰 a^*。

接着，我们证明在轮次 $\tau > \tau_0$ 结束时，所有满足 $\Delta_a \ge 4\beta_\tau$ 的臂都会被淘汰。为证明这一点，我们注意到在良好事件 $\mathcal{E}$ 下，

$$\tilde{\mu}_a + \beta_\tau \le \mu_a + 2\beta_\tau < \mu_{a^*} - 4\beta_\tau + 2\beta_\tau \le \tilde{\mu}_{a^*} - \beta_\tau \le \tilde{\mu}_{a_\tau} - \beta_\tau$$

这意味着臂 a 将根据该规则被淘汰。因此，对于每个次优臂 a，令 $\tau(a)$ 为臂 a 尚未被淘汰的最后一轮次。根据上述结果，我们有

$$\Delta_a \le 4\beta_{\tau(a)} = 4\sqrt{\frac{2\log(16|\mathcal{S}|\tau(a)^2/\delta)}{B_{\tau(a)}}} + 8\left(\frac{4\log(16|\mathcal{S}|\tau(a)^2/\delta)}{B_{\tau(a)}\varepsilon}\right)^{1-\frac{1}{k}} + 8(8\alpha)^{1-\frac{1}{k}}.$$

我们将臂 $a \in [K]$ 分为两组：$\mathcal{G}_1 = \{a \in [K] : 16(8\alpha)^{1-\frac{1}{k}} \le \Delta_a\}$ 和 $\mathcal{G}_2 = \{a \in [K] : 16(8\alpha)^{1-\frac{1}{k}} \ge \Delta_a\}$。

第 1 组： 现在，对于所有臂 $a \in \mathcal{G}_1$，我们有

$$\Delta_a \le 8\sqrt{\frac{2\log(16|\mathcal{S}|\tau(a)^2/\delta)}{B_{\tau(a)}}} + 16\left(\frac{4\log(16|\mathcal{S}|\tau(a)^2/\delta)}{B_{\tau(a)}\varepsilon}\right)^{1-\frac{1}{k}}.$$

因此，我们有

$$B_{\tau(a)} \le \max\left\{\frac{128\log(16|\mathcal{S}|\tau(a)^2/\delta)}{\Delta_a^2}, \frac{4\log(16|\mathcal{S}|\tau(a)^2/\delta)}{\varepsilon}\left(\frac{16}{\Delta_a}\right)^{\frac{k}{k-1}}, \frac{\log(16|\mathcal{S}|\tau_0^2/\delta)}{\alpha}\right\}.$$

由于对于任意 τ，$|\mathcal{S}| \le K$ 和 $2^\tau \le T$。因此，

$$B_{\tau(a)} \le \max\left\{\frac{128\log(16K\log^2 T/\delta)}{\Delta_a^2}, \frac{4\log(16K\log^2 T/\delta)}{\varepsilon}\left(\frac{16}{\Delta_a}\right)^{\frac{k}{k-1}}, \frac{\log(16K\log^2 T/\delta)}{\alpha}\right\},$$

由于批次大小加倍，对于每个次优臂 a，我们有 $N_a(T) \le 2B_{\tau(a)}$。因此，对于所有臂 $a \in \mathcal{G}_1$,

$$\mathcal{R}_T = \sum_{a\in\mathcal{G}_1} N_a(T)\Delta_a \le 2B_{\tau(a)}\Delta_a.$$

令 η 为 (0,1) 中的一个数。对于所有满足 $\Delta_a \le \eta$ 的臂 $a \in \mathcal{G}_1$，拉取这些臂所造成的遗憾有上界 $T\eta$。对于任意满足 $\Delta_a > \eta$ 的臂 $a \in \mathcal{G}_1$，选择 $\delta = \frac{1}{T}$ 并假设 $T \ge K$，则拉取臂 a 所造成的期望遗憾有上界

$$\begin{aligned}\mathbb{E}\left[\sum_{a\in\mathcal{G}_1,\Delta_a>\eta}\Delta_a N_a(T)\right] &\le \mathbb{P}(\bar{\mathcal{E}})\cdot T + O\left(\sum_{a\in\mathcal{G}_1,\Delta_a>\eta}\left\{\frac{\log T}{\Delta_a} + \frac{\log T}{\varepsilon}\left(\frac{1}{\Delta_a}\right)^{\frac{1}{k-1}} + \frac{\log T}{\alpha}\Delta_a\right\}\right)\\ &\le O\left(\frac{K\log T}{\eta} + \frac{K\log T}{\varepsilon\eta^{\frac{1}{k-1}}} + \frac{K\log T}{\alpha}\right)\end{aligned}$$

其中，最后一个不等式中的最后一项基于以下结果：根据 (1) 中奖励分布的重尾假设，对于任意 $a \in [K]$，$|\mu_a| \le \mathbb{E}_{r_a\sim P_k}|r_a| \le \mathbb{E}_{r_a\sim P_k}|r_a|^k \le 1$，有 $\Delta_a = \mu^* - \mu_a \le 2$。

因此，第 1 组的遗憾至多为

$$T\eta + O\left(\frac{K\log T}{\eta} + \frac{K\log T}{\varepsilon\eta^{\frac{1}{k-1}}} + \frac{K\log T}{\alpha}\right).$$

取 $\eta = \max\left\{\sqrt{\frac{K\log T}{T}}, \left(\frac{K\log T}{T\varepsilon}\right)^{\frac{k-1}{k}}\right\}$，则第 1 组的遗憾至多为

$$O\left(\sqrt{KT\log T} + \left(\frac{K\log T}{\varepsilon}\right)^{\frac{k-1}{k}}T^{\frac{1}{k}} + \frac{K\log T}{\alpha}\right)$$

组 2： 对于所有其他臂 $a \in \mathcal{G}_2$，总遗憾至多为 $O(T\Delta_a) = O(T\alpha^{1-\frac{1}{k}})$。

合并两组，选择 $\delta = \frac{1}{T}$ 并假设 $T \ge K$，我们有期望遗憾满足：

$$\mathcal{R}_T \le O\left(\sqrt{KT\log T} + \left(\frac{K\log T}{\varepsilon}\right)^{\frac{k-1}{k}}T^{\frac{1}{k}} + \frac{K\log T}{\alpha} + T\alpha^{1-\frac{1}{k}}\right)$$

我们还给出了该算法的隐私保证。根据定义3.5 中的拉普拉斯机制以及引理 A.1中的后处理，我们可以得到算法 1 满足 ε-差分隐私。 □

D 第6.2节的证明

定理 6.5的证明。步骤 1： 我们将证明以高概率 $1-\delta/2$, $|J-\mu| \leq 2r$。为此，我们首先研究私有直方图。注意 $\mathbb{E}[\mathbb{1}(X_i \in B_j)] = P_{\alpha,\ k}(B_j)$，那么

$$\begin{aligned}
\mathbb{P}\left(|\tilde{p}_j - P_{\alpha,k}(B_j)| > t\right) &= \mathbb{P}\left(\left|\frac{\sum_{i=1}^n \mathbb{1}(X_i \in B_j)}{n} + \mathrm{Lap}\left(\frac{2}{n\varepsilon}\right) - P_{\alpha,k}(B_j)\right| > t\right) \\
&\leq \mathbb{P}\left(\left|\frac{\sum_{i=1}^n \mathbb{1}(X_i \in B_j)}{n} - P_{\alpha,k}(B_j)\right| > t/2\right) + \mathbb{P}\left(\left|\mathrm{Lap}\left(\frac{2}{n\varepsilon}\right)\right| > t/2\right) \\
&\leq 2\exp\left(-\frac{nt^2}{2}\right) + \exp\left(-\frac{n\varepsilon t}{4}\right),
\end{aligned}$$

其中，最后一个不等式由引理 A.7 和引理 A.6得出。对 j使用并集界，我们进一步得到

$$\begin{aligned}
\mathbb{P}\left(\max_{j\in\mathcal{J}} |\tilde{p}_j - P_{\alpha,k}(B_j)| > t\right) &\leq \frac{2D}{r}\left(2\exp\left(-\frac{nt^2}{2}\right) + \exp\left(-\frac{n\varepsilon t}{4}\right)\right) \\
&\leq 2D\left(2\exp\left(-\frac{nt^2}{2}\right) + \exp\left(-\frac{n\varepsilon t}{4}\right)\right)
\end{aligned}$$

因此，以概率 $1-\delta/2$，我们有：

$$\max_{j\in\mathcal{J}} |\tilde{p}_j - P_{\alpha,k}(B_j)| \leq \max\left\{\sqrt{\frac{2\ln\frac{16D}{\delta}}{n}}, \frac{4\ln\frac{16D}{\delta}}{n\varepsilon}\right\} := C_1.$$

接下来，我们以事件成立为条件。根据引理 A.5 中的切比雪夫不等式以及 $\mathcal{P}_k$ 的假设，即 k阶中心矩小于 1，我们有

$$\begin{aligned}
\mathbb{P}_{X\sim\mathcal{P}_{\alpha,k}}\left(|X-\mu| \geq r\right) &\leq \alpha\mathbb{P}_{X\sim\mathcal{G}}(|X-\mu| \geq r) + (1-\alpha)\mathbb{P}_{X\sim\mathcal{P}_k}(|X-\mu| \geq r) \\
&\leq \alpha + (1-\alpha)(1/r)^k := C_2
\end{aligned} \tag{5}$$

令 j^* 为包含真实均值 μ 的区间的索引，并考虑三个连续区间$A_{j^*} = B_{j^*-1} \cup B_{j^*} \cup B_{j^*+1}$

$$\begin{aligned}
P_{\alpha,k}(A_{j^*}) &= P_{\alpha,k}(B_{j^*-1}) + P_{\alpha,k}(B_{j^*}) + P_{\alpha,k}(B_{j^*+1}) \\
&\geq P_{\alpha,k}\left((\mu-r, \mu+r)\right) \\
&\geq 1 - C_2.
\end{aligned}$$

其中，第一个不等式由不等式（5）得出。现在，对于任意 $j \notin \{j^*-1, j^*, j^*+1\}$，当 $D \geq 2r$

$$\tilde{p}_j \leq P_{\alpha,k}(B_j) + C_1 \leq 1 - P_{\alpha,k}(A_{j^*}) + C_1 \leq C_2 + C_1.$$

另一方面，由于 $P_{\alpha,\ k}(A_{j^*}) \geq 1 - C_2$，因此必定存在某个 $j \in \{j^*-1, j^*, j^*+1\}$ 使得 $P_{\alpha,\ k}(B_j) \geq \frac{1-C_2}{3}$。因此，对于这个 j，我们有

$$\tilde{p}_j \geq P_{\alpha,k}(B_j) - C_1 \geq \frac{1-C_2}{3} - C_1.$$

因此，如果 n (取决于 α， ε， r) 使得 $\frac{1-C_2}{3} - C_1 > C_2 + C_1$，真实均值 μ 位于算法 3 第 3 行所选的区间或其相邻区间中，这意味着其概率至少为 $1-\delta/2$，$|J-\mu| \le 2r$。

步骤 2： 利用上述结果，我们旨在证明截断能够处理集中性中的重尾、隐私和鲁棒性。

$$
\begin{aligned}
|\tilde{\mu}-\mu| &= |J + \frac{1}{n}\sum_{i=n+1}^{2n}(X_i - J)\mathbb{1}(|X_i - J| \le M) + \mathrm{Lap}\left(\frac{2M}{n\varepsilon}\right) - \mu| \\
&= \left| \frac{1}{n}\sum_{i=n+1}^{2n}(X_i - J)\mathbb{1}(|X_i - J| \le M) + \mathrm{Lap}\left(\frac{2M}{n\varepsilon}\right) \right. \\
&\left. + \frac{1}{n}\sum_{i=n+1}^{2n}(J-\mu)\left\{\mathbb{1}(|X_i - J| \le M) + \mathbb{1}(|X_i - J| > M)\right\} \right| \\
&= \left| \frac{1}{n}\sum_{i=n+1}^{2n}(X_i - J + J - \mu)\mathbb{1}(|X_i - J| \le M) + \mathrm{Lap}\left(\frac{2M}{n\varepsilon}\right) + \frac{1}{n}\sum_{i=n+1}^{2n}(J-\mu)\mathbb{1}(|X_i - J| > M) \right| \\
&\le \left| \frac{1}{n}\sum_{i=n+1}^{2n}(X_i - \mu)\mathbb{1}(|X_i - J| \le M) \right| + \left| \mathrm{Lap}\left(\frac{2M}{n\varepsilon}\right) \right| + \left| \frac{1}{n}\sum_{i=n+1}^{2n}(J-\mu)\mathbb{1}(|X_i - J| > M) \right|
\end{aligned}
$$

我们首先关注最后一个不等式右侧的第一项。令 N_G 为根据 $G \in \mathcal{G}$ 分布的 n 个样本的索引集合，令 N_{P_k} 为根据 $P_k \in \mathcal{P}_k^c$ 分布的 n 个样本的索引集合。于是，我们有

$$
\begin{aligned}
&\left| \frac{1}{n}\sum_{i=n+1}^{2n}(X_i - \mu)\mathbb{1}(|X_i - J| \le M) \right| \\
&\le \underbrace{\left| \frac{1}{n}\sum_{i\in N_G}(X_i - \mu)\mathbb{1}(|X_i - J| \le M) \right|}_{T_1} + \underbrace{\left| \frac{1}{n}\sum_{i\in N_{P_k}}(X_i - \mu)\mathbb{1}(|X_i - J| \le M) \right|}_{T_2}.
\end{aligned}
$$

为了控制 T_1，我们可以将其写为

$$
\begin{aligned}
T_1 &= \left| \frac{1}{n}\sum_{i\in N_G}(X_i - \mu)\mathbb{1}(|X_i - J| \le M) \right| \\
&\le \frac{1}{n}\sum_{i\in N_G}|(X_i - \mu)|\mathbb{1}(|X_i - J| \le M) \\
&\le \frac{1}{n}\sum_{i\in N_G}|(X_i - \mu)|\mathbb{1}(|X_i - \mu| \le M + 2r) \\
&\le \frac{|N_G|}{n}(M + 2r).
\end{aligned}
$$

那么$\frac{|N_G|}{n}$可以视为伯努利分布的均值估计 $Ber(\alpha)$。然后根据引理 A.9中的伯恩斯坦不等式，我们以概率 $1-\delta/8$，

$$\left|\frac{|N_G|}{n}-\alpha\right| \le \sqrt{\frac{2\alpha(1-\alpha)\log(16/\delta)}{n}}+\frac{2\log(16/\delta)}{3n}.$$

因此，

$$T_1 \le \left(\alpha+\sqrt{\frac{2\alpha\log(16/\delta)}{n}}+\frac{2\log(16/\delta)}{3n}\right)(M+2r), \quad \text{with probability} 1-\delta/8$$

· 因此，如果 n 满足 $\frac{\sqrt{2\alpha}\log(16/\delta}{n}+\frac{2)\log(16/\delta)}{3n}=O(\alpha)$，那么我们有 $T_1=O(\alpha(M+2r))$。现在，我们对 T_2 给出界，

$$\begin{aligned}
T_2 &= \left|\frac{1}{n}\sum_{\substack{i\in N_G\cup N_{P_k}\\ X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)-\frac{1}{n}\sum_{\substack{i\in N_G\\ X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)\right| \\
&\le \left|\frac{1}{n}\sum_{\substack{i\in N_G\cup N_{P_k}\\ X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)\right|+\left|\frac{1}{n}\sum_{\substack{i\in N_G\\ X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)\right| \\
&\le \left|\frac{1}{n}\sum_{\substack{i\in [n]\\ X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)\right|+T_1.
\end{aligned}$$

现在我们关注 $\left|\frac{1}{n}\sum_{\substack{i\in[n]\\X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)\right|$的上界。以 $1-\delta/8$ 的概率,

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{\substack{i\in[n]\\X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)\right|\\
\le&\left|\frac{1}{n}\sum_{\substack{i\in[n]\\X_i\sim P_k}}(X_i-\mu)\mathbb{1}(|X_i-J|\le M)-\mathbb{E}[(X_1-\mu)\mathbb{1}(|X_1-J|\le M)]\right|\\
&+|\mathbb{E}[(X_1-\mu)\mathbb{1}(|X_1-J|\le M)]-\mathbb{E}[(X_1-\mu)]|\\
\le&\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4(M+2r)\log(16/\delta)}{3n}+|\mathbb{E}[(X_i-\mu)\mathbb{1}(|X_i-J|\ge M)]|\\
\le&\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4(M+2r)\log(16/\delta)}{3n}+\left(\mathbb{E}[|X_i-\mu|^k]\right)^{\frac{1}{k}}\left(\mathbb{P}(|X_i-\mu|\ge M-2r)\right)^{\frac{k-1}{k}}]\\
\le&\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4(M+2r)\log(16/\delta)}{3n}+\frac{1}{(M-2r)^{k-1}}\\
\le&\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4(M+2r)\log(16/\delta)}{3n}+\left(\frac{2}{M}\right)^{k-1}
\end{aligned}
$$

其中最后一个不等式由 $M\ge 4r$ 得出，第三个不等式由引理 ¨ 中的 Hölder 不等式得出，A.8 而第二个不等式由引理 A.9 中的 Bernstein不等式得出。也就是说，令

$$Y_i=(X_i-\mu)\mathbb{1}(|X_i-J|\le M),$$

then

$$
\begin{aligned}
|Y_i-\mathbb{E}[Y_i]|&\le|Y_i|+|\mathbb{E}[Y_i]|\\
&\le|X_i-\mu|\mathbb{1}(|X_i-\mu|\le M+2r)+\mathbb{E}[|X_i-\mu|\mathbb{1}(|X_i-\mu|\le M+2r)]\\
&\le 2(M+2r)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{Var}(Y_i-\mathbb{E}[Y_i])&=\mathbb{E}(Y_i-\mathbb{E}[Y_i])^2\le\mathbb{E}[Y_i^2]\\
&\le\mathbb{E}_{X_i\sim P_k}[(X_i-\mu)^2\mathbb{1}(|X_i-\mu|\le M+2r)]\\
&\le\mathbb{E}_{X_i\sim P_k}[(X_i-\mu)^2]\le 1.
\end{aligned}
$$

因此，以概率 $1-3\delta/8$,

$$T_2\le\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4(M+2r)\log(16/\delta)}{3n}+\left(\frac{2}{M}\right)^{k-1}+T_1.$$

现在，我们关注 $T_3:=\left|\frac{1}{n}\sum_{i=n+1}^{2n}(J-\mu)\mathbb{1}(|X_i-J|>M)\right|$的上界。

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{i=n+1}^{2n}(J-\mu)\mathbb{1}(|X_i-J|>M)\right| \\
\leq&\frac{1}{n}\sum_{i=n+1}^{2n}|J-\mu|\,\mathbb{1}(|X_i-J|>M) \\
\leq&2r\frac{\sum_{i=n+1}^{2n}\mathbb{1}(|X_i-J|>M)}{n} \\
\leq&2r\frac{\sum_{i=n+1}^{2n}\mathbb{1}(|X_i-\mu|>M-2r)}{n}
\end{aligned}
$$

其中

$$
\begin{aligned}
\mathbb{E}_{X_i\sim P_{k,\alpha}}[\mathbb{1}(|X_i-\mu|>M-2r)] &= \mathbb{P}_{X_i\sim P_{k,\alpha}}(|X_i-\mu|>M-2r) \\
&\leq \alpha+(1-\alpha)\mathbb{P}_{X_i\sim P_k}(|X_i-\mu|>M-2r) \\
&\leq \alpha+\frac{1}{(M-2r)^k}\leq\alpha+\left(\frac{2}{M}\right)^k
\end{aligned}
$$

根据霍夫丁不等式，以概率 $1-\delta/8$,

$$
\frac{\sum_{i=n+1}^{2n}\mathbb{1}(|X_i-\mu|>M-2r)}{n}\leq\mathbb{P}_{X_i\sim P_{k,\alpha}}(|X_i-\mu|>M-2r)+\sqrt{\frac{\log(16/\delta)}{2n}}.
$$

因此，我们有

$$
T_3=\left|\frac{1}{n}\sum_{i=n+1}^{2n}(J-\mu)\mathbb{1}(|X_i-J|>M)\right|\leq 2r\left(\alpha+\left(\frac{2}{M}\right)^k+\sqrt{\frac{\log(16/\delta)}{2n}}\right).
$$

综上所述，我们有

$$
\begin{aligned}
|\tilde{\mu}-\mu| =&O\left(\left(\alpha+\sqrt{\frac{2\alpha\log(16/\delta)}{n}}+\frac{2\log(16/\delta)}{3n}\right)(M+2r)\right) \\
&+O\left(\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{4(M+2r)\log(16/\delta)}{3n}+\left(\frac{2}{M}\right)^{k-1}\right) \\
&+O\left(2r\left(\alpha+\left(\frac{2}{M}\right)^k+\sqrt{\frac{\log(16/\delta)}{2n}}\right)\right) \\
&+O\left(\frac{M\log(1/\delta)}{n\varepsilon}\right)
\end{aligned}
$$

情形 I： $\alpha=0$，**未污染集中界**。我们想证明，我们的集中性界优于 [KSU20]（定理 3.5）中的均值中位数法。也就是说，我们相对于其中的第三项（即$\log(D)+\log(1/\delta)$）是加性的，而他们是乘性的。

在这种情况下，我们有 $C_2=(1/r)^k$，根据关于 n 的第一个条件，它需要满足 $6C_1+4C_2<1$。这意味着 $C_2<1/4$。因此，令 $r=10^{1/k}$ 就足够了。于是，我们有 $C_1<0.1$，这要求 n 满足 $n\geq 200\log(16D/\delta)$ 和 $n\geq 20\log(16D/\delta)/\varepsilon$。我们可以放心地设为 $n\geq 200\log(16D/\delta)/\varepsilon$。

在 $\alpha=0$的情况下，T_1 不是问题，它只会额外引入一个 $O(M\log(1/\delta)/n)$。T_3 也不是问题，因为它被 $O((2/M)^{k-1}+\sqrt{\log(1/\delta)/\sqrt{n}})$ 所支配。

总结所有取值：当 $\alpha=0, r=10^{1/k}$ 和 $n\geq 200\log(16D/\delta)/\varepsilon$ 时，我们有

$$\begin{aligned}|\tilde{\mu}-\mu| =& O\left(\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{M\log(16/\delta)}{3n}+\left(\frac{2}{M}\right)^{k-1}\right)\\ &+O\left(\frac{M\log(1/\delta)}{n\varepsilon}\right)\\ =& O\left(\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{M\log(16/\delta)}{\varepsilon n}+\left(\frac{2}{M}\right)^{k-1}\right)\end{aligned}$$

现在，我们需要选择 M，使上式最小，同时满足 $M\geq 4r$。按照标准选取方式，令 $M=4\left(\frac{n\varepsilon}{\log(1/\delta)}\right)^{1/k}$，当 $n\geq \frac{10}{\log(1/\delta)\varepsilon}$ 时，这满足 $M\geq 4r$。

情况 II： $\alpha>0$。**污染集中界**。我们希望最小化项 $\mathcal{T}(\alpha,\varepsilon)$，同时尽可能扩大 α的范围。

在这种情况下，$C_2=\alpha+(1-\alpha)(1/r)^k$，并且我们仍需满足 $6C_1+4C_2<1$，这首先意味着 α 需要为 $\alpha<1/4$。令 $r=\iota^{1/k}$，则有 $C_2=\alpha+\frac{1}{\iota}(1-\alpha)$，它需要小于 $1/4$。令 $\iota=\frac{1-\alpha}{0.249-\alpha}$（因此 $\alpha<0.249$），则存在一个绝对常数 c_1，使得当 $n\geq c_1\log(16D/\delta)/\varepsilon$ 时，我们可以保证 $6C_1+4C_2<1$。

现在，我们转向 T_1。如果 $n\geq \frac{\log(16/\delta)}{\alpha}$ 且 $M\geq 4r$，则有 $T_1=O(\alpha M)$。

对于 T_3，有

$$T_3=2r\left(\alpha+\left(\frac{2}{M}\right)^k+\sqrt{\frac{\log(16/\delta)}{2n}}\right)$$

一种简单的方法是令 $n\geq\log(16/\delta)/\alpha^2$。于是有 $T_3=O(\alpha M+(1/M)^{k-1})$。

总结如下。对于任意 $\alpha\in(0,0.249)$，令 $r=\left(\frac{1-\alpha}{0.249-\alpha}\right)^{1/k}$。那么，对于所有 $n\geq\max\{c_1\log(16D/\delta)/\varepsilon,\log(16/\delta$

有

$$|\tilde{\mu}-\mu|=O\left(\sqrt{\frac{2\log(16/\delta)}{n}}+\frac{M\log(16/\delta)}{\varepsilon n}+\left(\frac{2}{M}\right)^{k-1}+\alpha M\right)$$

现在，我们需要选择 M，使上述表达式最小，同时满足 $M\geq 4r$。按照标准选择，我们令$M=\min\{4\left(\frac{n\varepsilon}{\log(1/\delta)}\right)^{1/k},4\alpha^{-1/k}\}$，当 n 和 α 满足时，该选择满足 $M\geq 4r$

$$n\geq\frac{\iota\log(1/\delta)}{\varepsilon}\quad\text{and}\quad\frac{1}{\alpha}\geq\iota,$$

这里回忆一下 $\iota=\frac{1-\alpha}{0.249-\alpha}$。因此，我们只有在 $\alpha\in(0,0.133)$处才有有效集中界。

□

定理 6.10的证明。令 τ_0 为满足 $B_\tau<\frac{200}{\log(16D|\mathcal{S}|\tau^2/\delta)\varepsilon}$ 的最大轮次。

对于所有轮次 $\tau\leq\tau_0$，批次大小均小于 2^{τ_0}。由于批次大小会翻倍，直到轮次 τ_0，每条臂的拉取次数 $a\in[K]$ 小于 $2\cdot 2^{\tau_0}\leq 2\frac{200}{\log(16D|\mathcal{S}|\tau_0^2/\delta)\varepsilon}$。因此，每个 $a\in[K]$ 都必须承受$\frac{400\log(16D|\mathcal{S}|\tau_0^2/\delta)}{\varepsilon}\Delta_a$ 的遗憾。

对于 $\tau > \tau_0, B_\tau \geq \frac{200}{\log(16D|\mathcal{S}|\tau^2/\delta)\,\varepsilon}$。对于每个 $a \in \mathcal{S}$，根据推论 6.6，我们以至少 $1 - \frac{\delta}{2|\mathcal{S}|\tau^2}$ 的概率，——

$$|\tilde{\mu}_a - \mu_a| \leq \beta_\tau.$$

给定轮次 $\tau > \tau_0$，我们用 $\mathcal{E}_\tau$ 表示对所有 $a \in \mathcal{S}$ 均满足 $|\tilde{\mu}_a - \mu_a| \leq \beta_\tau$. 的事件，并记为 $\mathcal{E} = \cup_{\tau > \tau_0} \mathcal{E}_\tau$。利用并集界可得

$$\mathbb{P}(\mathcal{E}_\tau) \geq 1 - \frac{\delta}{2\tau^2},$$

and

$$\mathbb{P}(\mathcal{E}) \geq 1 - \frac{\delta}{2}\left(\sum_{\tau > \tau_0} \tau^{-2}\right) \geq 1 - \delta.$$

下文中，我们以良好事件 $\mathcal{E}$ 为条件。首先证明最优臂 a^* 永远不会被淘汰。对于任意轮次 $\tau > \tau_0$，令 $a_\tau = = \arg\max_{a \in \mathcal{S}} \widetilde{\mu}_a$。由于

$$(\widetilde{\mu}_{a_\tau} - \widetilde{\mu}_{a^*}) + \Delta_{a_\tau} = |(\widetilde{\mu}_{a_\tau} - \widetilde{\mu}_{a^*}) + \Delta_{a_\tau}| \leq |\widetilde{\mu}_{a_\tau} - \mu_{a_\tau}| + |\widetilde{\mu}_{a^*} - \mu_{a^*}| \leq 2\beta_\tau,$$

很容易看出，该算法不会淘汰 a^*。

接下来，我们证明在轮次 $\tau > \tau_0$ 结束时，所有满足 $\Delta_a \geq 4\beta_\tau$ 的臂都会被淘汰。为此，在良好事件 $\mathcal{E}$ 下，我们有

$$\tilde{\mu}_a + \beta_\tau \leq \mu_a + 2\beta_\tau < \mu_{a^*} - 4\beta_\tau + 2\beta_\tau \leq \tilde{\mu}_{a^*} - \beta_\tau \leq \tilde{\mu}_{a_\tau} - \beta_\tau$$

这意味着臂 a 将按该规则被淘汰。因此，对于每个次优臂 a，令 $\tau\,(a)$ 为臂 a 尚未被淘汰的最后一个轮次。根据上述结果，我们有

$$\Delta_a \leq 4\beta_{\tau(a)} = O\left(\sqrt{\frac{\log(|\mathcal{S}|\tau(a)^2/\delta)}{B_{\tau(a)}}} + \left(\frac{\log(|\mathcal{S}|\tau(a)^2/\delta)}{B_{\tau(a)}\varepsilon}\right)^{1-\frac{1}{k}}\right).$$

因此，我们有

$$B_{\tau(a)} \leq O\left(\frac{\log(|\mathcal{S}|\tau(a)^2/\delta)}{\Delta_a^2} + \frac{\log(|\mathcal{S}|\tau(a)^2/\delta)}{\varepsilon}\left(\frac{1}{\Delta_a}\right)^{\frac{k}{k-1}} + \frac{\log(D|\mathcal{S}|\tau_0^2/\delta)}{\varepsilon}\right).$$

由于对于任意 τ，$|\mathcal{S}| \leq K$ 和 $2^\tau \leq T$。因此，

$$B_{\tau(a)} \leq O\left(\frac{\log(K\log^2 T/\delta)}{\Delta_a^2}, \frac{\log(K\log^2 T/\delta)}{\varepsilon}\left(\frac{1}{\Delta_a}\right)^{\frac{k}{k-1}}, \frac{\log(DK\log^2 T/\delta)}{\varepsilon}\right),$$

由于批次大小翻倍，对于每个次优臂 a，我们有 $N_a(T) \leq 2B_{\tau(a)}$。因此，对于所有臂 $a \in [K]$,

$$\mathcal{R}_T = \sum_{a \in [K]} N_a(T)\Delta_a \leq 2B_{\tau(a)}\Delta_a.$$

设 η 为 (0,1) 中的一个数。对于满足 $\Delta_a \le \eta$ 的所有臂 $a \in [K]$，拉动这些臂所产生的遗憾上界为 $T\eta$。对于满足 $\Delta_a > \eta$ 的任意臂 $a \in [K]$，取 $\delta = \frac{1}{T}$ 并假设 $T \ge K$，则拉动臂 a 所产生的期望遗憾上界为

$$\mathbb{E}\left[\sum_{a\in[K],\Delta_a>\eta} \Delta_a N_a(T)\right] \le \mathbb{P}(\bar{\mathcal{E}}) \cdot T + O\left(\sum_{a\in[K],\Delta_a>\eta} \left\{\frac{\log T}{\Delta_a} + \frac{\log T}{\varepsilon}\left(\frac{1}{\Delta_a}\right)^{\frac{1}{k-1}} + \frac{\log DT}{\varepsilon}\Delta_a\right\}\right)$$
$$\le O\left(\frac{K\log T}{\eta} + \frac{K\log T}{\varepsilon\eta^{\frac{1}{k-1}}} + \frac{KD\log(DT)}{\varepsilon}\right)$$

其中，最后一个不等式中的最后一项依据以下结果得出：根据定义 3.2 中奖励分布的重尾假设，对于任意 $a \in [K]$，$\mu_a \in [-D, D]$，因此 $\Delta_a = \mu^* - \mu_a \le 2D$。

因此，遗憾至多为

$$T\eta + O\left(\frac{K\log T}{\eta} + \frac{K\log T}{\varepsilon\eta^{\frac{1}{k-1}}} + \frac{KD\log(DT)}{\varepsilon}\right).$$

取 $\eta = \max\left\{\sqrt{\frac{K\log T}{T}}, \left(\frac{K\log T}{T\varepsilon}\right)^{\frac{k-1}{k}}\right\}$，遗憾至多为

$$O\left(\sqrt{KT\log T} + \left(\frac{K\log T}{\varepsilon}\right)^{\frac{k-1}{k}} T^{\frac{1}{k}} + \frac{DK\log(DT)}{\varepsilon}\right).$$

对于隐私保证，根据定义 3.5 中的拉普拉斯机制、 [KV17 的直方图学习器的隐私保证（引理 2），3], 引理 A.3中的并行组合定理以及引理 A.1中的后处理，我们可以得到该结果。

□

定理 6.12 的证明。该定理的证明与定理 6.2 的证明类似，现在开始臂消除所需的批次大小条件变为 $\max\{\frac{\iota}{\log(16/\delta)\varepsilon}, \frac{c_1\log(16D/\delta)\varepsilon}{\log(16/\delta)\alpha^2}\}$，并且每个 $a \in [K]$ 的 Δ_a 上界为 $2D$。由此可以得到遗憾上界的结果。

根据定义中的拉普拉斯机制，3.5，以及 [KV17 中引理 2 对直方图学习器的隐私保证3], 、引理 A.3中的并行组合定理和引理 A.1中的后处理，我们可以得到该结果。

□