arXiv:2603.09168v1 [cs.LG] 10 Mar 2026

分布式专家问题的更优界

David P. Woodruff* Samson Zhou†

2026年3月11日

摘要

本文研究分布式专家问题，其中 n 位专家分布在 s 台服务器上，共经历 T 个时间步。在每个时间 t，每位专家的损失是一个向量的 ℓ_p 范数，该向量由时间 t时这位专家在 s 台服务器上的损失构成。目标是在所有 T 个时间步内摊销计算，最小化遗憾 R，即分布式协议的损失与最佳专家的损失之差，同时尽可能减少通信量。我们给出一种协议，可实现约为 $R \gtrsim \sqrt{T}$ 的遗憾，使用 $\mathcal{O}\left(\frac{n}{R^2}+\frac{s}{R^2}\right)\cdot\max(s^{1-2/p},1)\cdot$ 多项式对数(nsT)个通信比特数，优于此前的工作。

1 引言

许多现代应用都需要基于大量分布式专家的预测进行序贯决策。例如，在超参数优化或模型选择中，每位专家对应不同的架构或预训练模型，这些架构或模型分别在不同数据集上进行评估，算法必须动态选择部署对象。在推荐系统中，专家编码了不同的排序函数，且只能获得所选动作的反馈；而在强化学习中，专家则对应于在不同环境中训练的策略。在这些应用中，决策者必须汇总来自多个来源的信息，并适应分布式系统中不断变化的损失。专家建议在线预测问题对这一设置进行了形式化： n 位专家在 T 轮中每一步都会产生损失，而算法根据观测到的历史选择一位专家，以最小化累积遗憾；该遗憾是相对于事后最佳专家定义的。经典算法包括指数权重算法（EWA）和乘法权重更新（MWU），它们在全信息设置下实现了最优遗憾 $O\left(\sqrt{\log n}\right)$ [AHK12],；而 EXP3 在多臂老虎机设置下实现了近乎最优的遗憾 $O\left(\sqrt{n\log\frac{n}{T}}\right)$ [ACFS02]。随着专家数量 n 和预测轮次 T 不断增加，集中式运行这些算法的计算成本也会变得高昂，因此有必要研究可扩展模型。先前的研究探索了流式方法 [SWXZ22、PZ23、WZZ23、PR23、ACNS23、BGW+26]，这些方法按顺序处理专家损失，同时仅保留历史数据的精简摘要。

在这项工作中，我们探讨一个互补的挑战：带专家的分布式在线学习。在这种情形下，专家成本分布在 s 台服务器上，而中央协调器旨在实现一种

*卡内基梅隆大学。 dpwoodru@gmail.com
†得克萨斯农工大学。 samsonzhou@gmail.com

低遗憾算法，同时尽量减少与服务器之间的通信。这种分布式形式自然出现在大规模在线优化、跨多个数据集的超参数调优，或每台服务器都持有一部分私有数据的场景中。例如，每位专家可以对应一个基于各服务器上的聚合数据训练的模型，其总损失按每台服务器的损失求和计算，如 HPO-B 基准 [PJWG21]所示。当单台服务器的成本受到限制时，其他聚合函数（例如各服务器上的最大损失）也可能适用，这凸显了该框架能够灵活适应不同实际目标。

我们关注的是理解带专家的分布式在线学习中通信与遗憾之间的权衡。与内存受限的流式处理方法不同，协调器可以掌握所有专家的完整信息，因此主要挑战在于尽可能减少通信。我们在消息传递模型中研究这一问题：协调器在顺序轮次中与每台服务器私下通信。以往工作主要聚焦于 ℓ_1 损失 [JPT+25]，但许多应用——例如风险敏感优化、稳健模型选择和分布式控制——需要一般的 ℓ_p 损失，且 $p > 1$。不同的 p 取值会在稳健性与对大偏差的惩罚之间进行权衡：ℓ_∞ 取值使最大损失最小化，中间取值（如 ℓ_2 ）会适度惩罚大偏差，而接近 ℓ_1 的取值则强调离群值稳健性。例如，ℓ_2 损失因其与熵的联系，在流式设置中得到了广泛研究 [HNO08，WZ21]。因此，我们开发了针对一般 ℓ_p 损失实现近乎最优遗憾，同时确保通信量有严格保证的算法。

1.1 协调器模型下带专家的分布式在线学习

我们研究协调器（消息传递）模型中的专家分布式在线学习问题。共有 s 台服务器，记作 $[s] = \{1, 2, \ldots\ldots, s\}$。在每一轮 $t \in [T]$，每台服务器 $j \in [s]$ 都会收到一个局部损失向量（$\ell_1(j, t)$，……，$\ell_n(j, t)$）$\in \mathbb{R}^n$。目标是在尽量减少服务器间通信的同时，使算法的累积遗憾最小化。

专家在线学习。 在专家在线学习问题中，有 n 位专家在长度为 T 的时间范围内做出预测。对于每个时刻 $t \in [T]$，算法会选择一位专家 $i_t \in [n]$，然后观察损失向量（$L_1(t)$，……，$L_n(t)$），并承担损失 $L_{i_t}(t)$。算法的总损失为 $\sum_{t\in[T]} L_{i_t}(t)$，而最佳专家 i^* 承担的损失为 $\sum_{t\in[T]} L_i(t)$。算法的遗憾定义为

$$R = \frac{1}{T}\left(\sum_{t\in[T]} L_{i_t}(t) - \sum_{t\in[T]} L_{i^*}(t)\right),$$

算法旨在将其最小化。在分布式场景中，损失 $L_i(t)$并非显式给出。相反，每个的损失通过 L_p 损失 $L_i(t) = \left(\sum_{j\in[s]} \ell_i(j,t)^p\right)^{1/p}$在各服务器上计算，其中 $p \geq 1$ 是固定且预先已知的。这种分布式形式刻画了专家评估被划分到多个服务器或数据集上的情形。

通信复杂度。 由于损失 $L_i(t)$并未明确给出，服务器必须共同执行一个分布式协议 Π。每台服务器都拥有私有随机性，且所有服务器共享公共随机性。在协调器模型中，服务器只与协调器通信；而在

消息传递模型中，服务器间直接通信至多带来常数因子改进，因此我们将这两种模型视为可互换。我们假设 Π 是顺序执行且基于轮次的：每一轮中，协调器与部分服务器交互并等待响应。协议是自定界的，因此服务器知道一轮何时结束。令 $\Pi(r)$表示第 r轮的通信记录，其通信成本为 $|\Pi(r)|$ 。总通信量为 $\sum_r |\Pi(r)|$。设计能够兼顾总通信量和低遗憾的协议，是带专家的分布式在线学习面临的核心挑战。

1.2 我们的贡献

本文研究带专家的分布式在线学习问题。作为预热，我们证明存在一种分布式协议，其遗憾为 $\tilde{\mathcal{O}}\left(s^{1/p}\sqrt{\frac{\log n}{T}}\right)$，且近乎最优，通信成本为 $\mathcal{O}(nT+sT)$：

定理 1.1. 设 $b>a>0$ 为固定常数，并假设对于所有 $t\in[T]$、$i\in[n]$和 $j\in[s]$，$\ell_i(j,t)\in[a,b]$。存在一种算法，其期望遗憾至多为 $O\left(s^{1/p}\sqrt{\frac{\log n}{T}}\right)$，且以高概率使用的总通信量至多为 $O\left(sT+nT\right)\cdot$ $\mathrm{polylog}(nsT)$ 比特。

相比之下，[JPT+25] 的工作仅针对一般 ℓ_p 损失，在广播/黑板模型中实现了分布式协议。在该模型中，服务器可以公开广播其他服务器可见的信息，而协议的通信成本则是广播消息的总长度。这种设定比消息传递模型“更容易”，因为在我们的设定中，向所有服务器发送相同消息会产生 s 通信乘法开销。因此，这些技术从根本上有所不同。例如，ℓ_p 由 [JPT+25] 提出的损失算法在广播模型中遍历所有服务器，以找出每个专家对应的 $j\in[s]$ 损失最大的服务器 $i\in[n]$。虽然在广播模型中，每个服务器 j 只需发送每个专家对应的、超过此前广播损失的值，从而使总通信量约为 $\mathcal{O}(s+n)$，但这种方法在我们的设定中需要 $\mathcal{O}(ns)$ 通信。

另一方面，[JPT+25] 在消息传递模型中，为 SUM 问题实现了一种分布式协议，即 ℓ_1 损失，具有 $\tilde{\mathcal{O}}(nT)+\mathcal{O}(Ts)$ 总通信量，遗憾为$\mathcal{O}\left(s^{1/p}\sqrt{\frac{\log n}{T}}\right)$。然而，由于 ℓ_1 损失在各服务器之间是可加的，因此，诸如按损失大小的概率抽样并通信若干损失等标准技术，可用于 ℓ_1 损失，但不适用于次可加或超可加损失，例如 ℓ_p 损失。相比之下，定理 1.1 能够在所有 ℓ_p 损失下达到最优遗憾。

我们指出，[JPT+25] 假设损失经过归一化，使得单个专家在单个时刻的总损失至多为1。相比之下，若根据我们的假设，每台服务器允许有 $[a,b]$ 损失，那么所有专家的损失均为$\Omega(s^{1/p})$，这解释了他们给出的界与我们引用的界为何不同。我们可以对定理 1.1进行参数化，从而得到更一般的通信-遗憾权衡，如下所示：

定理 1.2。令 $b>a>0$ 为固定常数，并假设对于所有 $t\in[T]$、$i\in[n]$和 $j\in[s]$，$\ell_i(j,t)\in[a,b]$。令 $R\geq\frac{1}{\sqrt{T}}$。则存在一种算法，使期望遗憾至多为 $O\left(Rs^{1/p}\sqrt{\log n}\right)$，并且以高概率使用的总通信量至多为 $\binom{n+s}{R^2}\cdot$ 多对数(nsT)比特。

我们强调，寻求遗憾 $R \geq \frac{1}{\sqrt{T}}$ 是标准做法，因为从信息论角度来看，不可能实现遗憾 $R < \frac{1}{\sqrt{T}}$ [Cov66]。

与定理 1.2, [JPT+25] 相比，ℓ_1 针对 SUM 问题（即 ˜损失）实现了一个分布式协议，在遗憾为 R 时，其总通信量为 $\binom{n}{R^2} + \mathrm{O}\, Ts\mathrm{O}\,(Ts)$。因此，定理 1.2 不仅改进了 [JPT+25] 针对一般 ℓ_p 损失的结果，还改进了 [JPT+25] 的结果中关于参数化的一般遗憾 R的 O (Ts) 依赖关系。例如，当遗憾为 $R = \mathrm{O}\,(1)$ 时， [JPT+25] 的协议仍具有 O (Ts) 的依赖关系，而我们的协议具有 O (s) 的依赖关系；对于较大的时间范围 T，后者显著更小。我们在图 1 中总结了这一讨论。

参考文献	损失函数	通信成本（比特）
[JPT+25]	ℓ_1 损失	$\binom{n}{R^2} + Ts) \cdot$ 多对数(nsT)
定理 1.2	ℓ_p 损失	$\binom{n+s}{R^2} \cdot$ 多对数(nsT)

图 1：我们的工作首次研究了协调器模型中的 ℓ_p 损失；对于 $p = 1$这一特殊情形，我们针对遗憾 R取得了更优的遗憾-通信权衡。

我们通过以下方式实现对 ℓ_p 损失的处理：将其嵌入 ℓ_p 中， ℓ_∞ 方法是使用指数随机变量，类似于 [JPT+25]的黑板协议。然而，在协调器模型中，服务器无法轻易找到最大值，因此我们需要仔细分析子采样和阈值化，以确保找到最大值，同时避免太多服务器发送较小的值。此外，所得估计量的方差无界，因此我们采用几何平均估计量。这些创新构成了我们的主要算法贡献。尽管类似方法已用于其他场景，例如流模型中的范数估计 [Li08、WZ21], ，但据我们所知，这是首次将这些技术应用于分布式场景下的在线学习。因此，我们能够处理 ℓ_p 损失，而以往的工作无法处理 [JPT+25]。更多细节请参见第 3 节 和第 4 节。最后，我们去除了损失必须处于某个常数范围内的假设 $[a, b]$。

定理 1.3。 假设对于所有 $t \in [T]$，我们有 $\ell_i(j, t) \leq 1$。存在一种算法，其期望遗憾至多为 $O\left(Rs^{1/p}\sqrt{\text{对数}}\, n\right)$，并且以高概率使用的总通信量至多为$\binom{n+s}{R^2} \cdot \max(s^{1-2/p}, 1) \cdot$ 多对数(nsT) 比特。

最后，为补充我们的理论结果，我们进行了一系列实证评估，作为简单的概念验证。这些结果见第 6 节。

算法与技术创新。 主要挑战在于， ℓ_p 损失比 ℓ_1 损失复杂得多。注意到，由于 ℓ_1 损失具有可加性，因此总损失是各项单项损失之和。因此，按照单项损失的大小成比例地进行采样是自然的。对于 ℓ_p 损失，可以进行某种更精细的采样，但目前尚不清楚如何在分布式模型中跨所有专家高效地执行采样。为克服这一障碍，我们使用随机指数缩放，将 ℓ_p 损失嵌入 ℓ_∞ 框架，使得最大的缩放贡献能够近似原始的 ℓ_p 损失。然而，由于

指数随机变量的概率密度函数，所得分布具有无界方差，因此我们对数量相互独立的随机缩放因子取几何平均，以获得具有有界方差的无偏估计量。随后，我们将这些估计损失与乘法权重更新（MWU）算法结合，用于在线学习。值得注意的是，使用几何平均估计量来降低基于指数随机变量的估计量的方差，对于专家分布式在线学习问题而言，既是一项算法创新，也是一项技术创新；我们认为，它可能对其他应用具有独立的研究价值。原则上，几何平均估计量可以由其他方差缩减技术替代；此类算法的设计是未来工作中一个有趣的研究方向。

1.3 预备知识

我们简要介绍一些与算法和分析相关的必要预备知识。首先回顾集中不等式中切尔诺夫界的一种标准表述。

定理 1.4 (切尔诺夫界**)**。 设 $X_1, X_2, \ldots, X_n$ 为相互独立的伯努利随机变量，且 $\mathbf{Pr}[X_i = 1] = p_i$ 且 $\mathbf{Pr}[X_i = 0] = 1 - p_i$。定义 $X = \sum_{i=1}^{n} X_i$，并令 $\mu = \mathbb{E}[X] = \sum_{i=1}^{n} p_i$。那么对于任意 $\delta > 0$，有 $\Pr(|X - \mu| \geq \delta\mu) \leq 2\exp\left(-\frac{\delta^2\mu}{3}\right)$。

接下来定义指数随机变量。

定义 1.5 (指数随机变量**)**。 速率参数为 $\lambda > 0$ 的指数随机变量 是一个连续随机变量 X，其概率密度函数（概率密度函数）定义如下：

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

我们证明指数随机变量具有以下最大值稳定性。

引理 1.6。 *[*指数随机变量的最大值稳定性*]* 令 $X_i = \frac{f_i}{e_i^{1/p}}$，其中 $i \in [n]$，$e_1, \ldots, e_n$ 是速率为1的独立指数随机变量。那么 $\max_{i\in[n]} \frac{f_i}{e_i^{1/p}}$ 服从与 $\frac{\|f\|_p}{e^{1/p}}$ 相同的分布，其中 e 是速率为1的指数随机变量。

证明。 令 $X = \max_{i\in[n]} \frac{f_i}{e_i^{1/p}}$。那么对于任意 $t > 0$，有

$$\begin{aligned} \mathbf{Pr}[X \leq t] &= \mathbf{Pr}[X^p \leq t^p] \\ &= \prod_{i=1}^{n} \mathbf{Pr}\left[\frac{f_i^p}{e_i} \leq t^p\right] \\ &= \prod_{i=1}^{n} \exp\left(-t^p \cdot f_i^p\right) \\ &= \exp\left(-t^p \cdot \|f\|_p^p\right). \end{aligned}$$

相比之下，若 $Y = \frac{\|f\|_p}{e^{1/p}}$，其中 e 是速率为 1 的指数随机变量，则

$$\mathbf{Pr}[Y \leq t] = \mathbf{Pr}[Y^p \leq t^p] = \exp\left(-t^p \cdot \|f\|_p^p\right).$$

因此，$\max_{i\in[n]} \frac{f_i}{e_i^{1/p}}$ 与 $\frac{\|f\|_p}{e^{1/p}}$ 服从相同的分布。 □

最后，我们回顾一下乘法权重更新（MWU）算法。对于 η 大约 $\frac{1}{\sqrt{T}}$，关于 算法 1中的乘法权重更新算法，已知以下保证：

定理 1.7. 对于一个包含 n 个专家的集合，若每一轮 T 中损失的二阶矩至多为 ρ，则乘法权重更新（*MWU*）算法的期望遗憾为 $O\left(\sqrt{\rho \log_T n}\right)$。

算法 1 乘法权重更新算法

输入： 学习率 η，所有时刻 $t \in [T]$ 的损失 $\{\ell_i(t)\}$，专家 $i \in [n]$
输出： 每天要选择的专家序列

```
w_i ← 0 对于所有 i ∈ [n]
对于 每个时刻 t ∈ [T] 执行
    对每个 i ∈ [n] 执行
        w_i ← w_i + ℓ_i(t)
    结束循环
    按与 exp(−ηw_i) 成正比的概率采样 i ∈ [n]
结束循环
```

2 相关工作

本节简要概述专家学习问题的相关早期文献。在 第 2.1 节中，我们首先介绍中心化设定下该问题的若干重要研究成果。在这一设定中，所有 n 专家及其预测都可供算法使用，且对算法的内存界限没有任何限制。由于所有信息都集中在一处，因此也不存在通信限制，这与我们研究的设定不同。随后，我们将在 第 2.2 节概述大数据模型中专家学习问题的相关早期研究。

2.1 专家学习经典方法

专家在线学习问题历史悠久，与强化学习和优化有着广泛联系。早期研究试图理解这一问题时，首先假设存在一位“完美专家”——其预测总是正确。在这一设定下，一个广为人知的“折半”算法最多会犯 $\log_2 n$ 次错误（给定 n 位专家）：该算法保留尚未犯错的专家集合，并根据多数投票进行预测；每当预测错误时，就淘汰所有意见不一致的专家。

随机加权多数。然而，完美专家这一假设往往并不现实。确定性加权多数算法通过为每位专家分配权重、根据加权多数进行预测，并将预测错误的专家的权重按乘法方式降低 $(1-\varepsilon)$（其中参数 $\varepsilon \in (0,1)$固定）来放宽这一假设。从定量角度看，确定性加权多数算法所犯的错误数至多为 $2(1+\varepsilon)m^* + \mathcal{O}\left(\log_\varepsilon n\right)$，其中 m^* 是最佳专家所犯的错误数 [LW94]。事实上，一个简单的例子就能说明这一点：两位专家每次总是选择不同的结果，这足以迫使任何确定性算法

每次都犯错，而每位专家只在一半的情况下出错。因此，与最佳专家相比，任何确定性算法都无法将错误数控制在两倍以下；换言之，任何确定性算法都无法实现次线性遗憾。

为克服确定性算法在最坏情况下的局限性，Littlestone 和 Warmuth 提出了随机加权多数 [LW94]。随机加权多数并不基于严格的加权多数进行预测，而是根据专家上的概率分布进行预测，其中选择某位专家的概率与其权重成正比。当某位专家犯错时，其权重会再次乘法降低 $(1-\varepsilon)$，因此该算法总共会犯 $(1+\varepsilon)m^* + \mathcal{O}\left(\log_{\varepsilon} n\right)$ 次错误，这相当于 $\mathcal{O}\left(\sqrt{T \log n}\right)$ 的遗憾（对于 $\varepsilon = \Theta\left(\sqrt{\frac{\log n}{T}}\right)$），并且已知其在渐近意义下是最优的 [Cov66]。

其他随机化变体。 除了这些基础方法之外，还发展出了许多其他算法，每种算法都有各自的优势。例如，广为人知的乘法权重更新（MWU）框架推广了权重更新规则，并且能够适应各种损失函数 [Bro51]。值得注意的是，MWU 已应用于 AdaBoost 等提升算法 [FS97]以及零和博弈的近似求解 [FS99]，，还能够高效地近似求解一大类线性规划和半定规划 [Cla95]，，从而为旅行商问题、调度问题和多商品流等各种 NP完全问题提供快速近似算法 [PST95，GK07]。

基于“跟随领导者”原则（该原则会直接选择迄今累积损失最低的专家），跟随扰动领导者（FTPL）在做出选择前向各专家的累积损失添加扰动，引入了随机性 [KV05]。这种随机化有助于避免确定性算法经常遭遇的最坏情况，从而获得渐近最优的遗憾界，与随机加权多数和乘法权重更新类似。此外，对于某些结构化问题，FTPL 的计算效率高于乘法权重更新。相比之下，跟随正则化领导者（FTRL）在决策过程中加入正则化项 [McM11]，，以惩罚剧烈变化、增强对噪声的稳定性；同时，通过选择正则化函数，它还能够灵活地纳入先验知识。

2.2 大数据场景下的专家学习

我们指出，第 2.1 节讨论的所有算法都需要在算法运行期间存储所有专家在各个时刻的表现。然而，在许多现实应用中，专家数量 n 和时间范围 T 可能都相当庞大，因此存储和处理所有专家的表现会带来很高的计算开销，需要占用大量内存或进行大量通信。因此，近期有一系列研究致力于开发内存使用量相对于专家数量呈次线性的专家在线学习算法 n。

数据流上的专家学习。 [SWXZ22] 证明了，任何使用 S 总空间的算法都必须达到遗憾 $\tilde{\mathcal{O}}\left(\sqrt{\frac{nT}{S}}\right)$，这意味着若要达到最优的 $\mathcal{O}\left(\sqrt{T \log n}\right)$ 遗憾，所用空间必须大致呈线性增长，即 $S = \mathcal{O}(n)$。 [SWXZ22] 还首次提出了一种用于随机顺序模型下专家在线学习的算法，实现了 $\tilde{\mathcal{O}}\left(\frac{nT}{R}\right)$ 空间

对于遗憾 R。随后， [PZ23、PR23] 改进了这些界，最终得到了 [PR23]的算法，该算法在任意顺序模型中使用多对数空间即可实现 $\tilde{\mathcal{O}}\left(\sqrt{nT}\right)$ 遗憾。

多臂老虎机。 多臂老虎机是一类相关的强化学习问题，其中各臂具有固定的奖励分布，但每次只能获得单个臂的反馈；其空间复杂度也已得到研究。 [LSPY18] 证明了实现近乎最优的遗憾只需常数空间。然而，由于专家预测具有对抗性，而老虎机中的奖励是静态的，且专家算法每天都会观察所有“臂”，这种方法无法直接应用于我们的设定。因此，专家算法与老虎机算法不能直接比较。 [AW20] 证明了 $\Omega(k)$ 空间是识别前 k 个臂所必需的，条件是在流式老虎机模型中。这些结果表明，专家问题本质上比老虎机问题更难，因为低遗憾的老虎机解可以具有常数空间复杂度。

流式学习。 关于独立同分布流式模型下统计学习与估计的空间和样本复杂度权衡，也有大量研究，包括矩阵行推断方面的研究 [Raz17、GRT18、GRT19] 以及奇偶学习方面的研究 [KRT17、Raz19]。另一条活跃的研究方向则关注特定的流式学习问题，例如相关性查找 [DS18], 碰撞概率估计、图连通性和秩估计 [CMVW16]。然而，我们对专家问题的研究不对数据分布作任何假设，并且关注预测而非推断。因此，这些研究与我们的目标有显著差异：我们旨在证明专家算法的一般空间复杂度界，同时保持相对于最佳专家的竞争性遗憾。

3 预热：SimpleAlgorithm

我们提出一种简单的分布式算法，它能达到接近最优的遗憾，最多使用 $O(sT) + nT \cdot$ 多对数(nsT) 的总通信比特数。该算法的直觉如下。假设对所有专家 $i \in [n]$ 以及时刻 $t \in [T]$，我们都有 i 在时刻 t的损失， $L_i(t) = \left(\sum_{j\in[s]} \ell_i(j,t)^p\right)^{1/p}$。然后，通过设定 $w_i(t) = \sum_{t'<t} L_i(t)$，我们可以运行标准的乘法权重更新（乘法权重更新）框架，权重为 $w_i(t)$，即以与 $\exp(-\eta w_i(t))$ 成正比的概率选择专家 i ，学习率 $\eta > 0$固定。遗憾的是，我们并没有这些损失 $\{L_i(t)\}$。相反，对于时刻$t \in [T]$、每个专家 $i \in [n]$以及每台服务器 $j \in [s]$，我们生成一个指数随机变量$e_i(j,t)$。根据指数随机变量的最大稳定性，参见引理 1.6，我们有

$$\max_{j\in[s]} \frac{\ell_i(j,t)}{(e_i(j,t))^{1/p}} \sim \frac{\left(\sum_{j\in[s]} \ell_i(j,t)^p\right)^{1/p}}{e^{1/p}},$$

其中， e 是另一个指数随机变量。注意，右侧恰好是 i 在时间 t的 ℓ_p损失。此外，我们可以计算 $\frac{1}{e^{1/p}}$的期望值，因此，如果服务器只需发送缩放后的损失的最大值$\ell_i(j,t)/(e_i(j,t))^{1/p}$，那么我们就可以计算出 ℓ_p 损失 $\left(\sum_{j\in[s]} \ell_i(j,t)^p\right)^{1/p}$的无偏估计。

遗憾的是，这种方法存在两个挑战。首先，服务器不知道如何发送最大值。其次，由此产生的随机变量的方差是无界的。为处理第一个问题，我们首先证明，以高概率，最大值高于某个

阈值，该阈值大致为 $s^{1/p}$ 此外，高于该阈值的缩放后的损失数量很少。因此，我们可以对通信总量给出上界。为解决第二个问题，我们使用标准估计量 [Li08] ，该估计量取这些估计量中常数个的几何平均，从而将方差降低到较小的程度。我们再次指出，与 [JPT+25] 等先前工作相比，这一流程构成了我们的主要算法贡献；该 $\widehat{\ }$ 算法完整地出现在 算法 2 中，其中 $s_i(t)$ 是损失的几何平均估计量$L_i(t)$，对应于专家 i在时间 t，而 $w_i(t)$ 是作为 i 输入 MWU 的专家权重。

算法 2 具有近乎最优遗憾的分布式协议

输入： 所有时间 $t \in [T]$、专家 $i \in [n]$ 和服务器 $j \in [s]$, L_p 的损失 $\{\ell_i(j,t)\}$ 损失参数 p

输出： 每天要进行选择的专家序列

$B \leftarrow \left\lfloor \frac{3}{p} \right\rfloor$

对于每个时间 $t \in [T]$ **执行**

  对于每个 $i \in [n]$ **执行**

    对于每个服务器 $j \in [s]$ **执行**

      令 $\ell_i(j,\ t)$ 表示服务器 j 上时间 t专家 i 的损失

      令 $e_i^{(b)}(j,\ t)$ 对所有 $b \in [B]$均为独立指数随机变量

      $q_i^{(b)}(j,t) \leftarrow \frac{\ell_i(j,\ t)}{C_{3.2}(e_i^{(b)}(j,\ t))^{1/p}}$ 对于 $b \in [B]$ ▷引理 3.2

      如果 $q_i^{(b)}(j,\ t) \geq \frac{s^{1/p}}{100\log(nsT)}$ **then**

        将 $q_i^{(b)}(j,\ t)$ 发送给协调器

      结束 if

    结束循环

  结束循环

  对于每个 $i \in [n]$，**执行**

    $\widehat{s_i}(t) \leftarrow \prod_{b\in[B]} \left(\max_{j\in[s]} q_i^{(b)}(j,t)\right)^{1/B}$

    $w_i(t) \leftarrow w_i(t-1) + \widehat{s_i(t)}$

  结束循环

  对 $\{w_i(t)\}$ 执行乘法权重更新

结束循环

我们首先回顾一个关于随机变量的多项式的概率密度函数的事实。

事实 3.1. 设随机变量 X 的概率密度函数为 $f(x)$，其支撑集为非负实数。则随机变量 $X^{-1/p}$ 的概率密度函数为 $px^{-p-1}\cdot f(x^{-p})$。

我们现在计算 计算期望，并为我们的几何平均估计量的方差提供上界 r.

引理 3.2. *[几何平均估计量的期望和方差]* 设 $p > 0$ 为固定值，且$R \geq 3p$ 为整数。设 $e_1, \ldots, e_B$ 为速率为 1 的独立指数随机变量。令$Z_b = \frac{1}{e_b^{1/p}}$ 对所有 $b \in [B]$，并令 $Z = \left(\prod_{b\in[B]} Z_b\right)^{1/B}$ 为它们的几何平均。则存在一个普适常数 $C_{3.2} \in (0, 2^B]$ 使得 $\mathbb{E}[Z] = C_{3.2}$ 且 $\mathbb{E}[Z^2] \leq 3^B$。

证明。 速率为 1 的指数随机变量的概率密度函数为 $f(x) = e^{-x}$。因此，根据 事实 3.1，每一项 Z_b 的概率密度函数为 $px^{-p-1} \cdot e^{-x^{-p}}$，类似地，几何平均中每一项 $Z_b^{1/B}$ 的概率密度函数为 $p(x) = Bpx^{-Bp-1} \cdot e^{-x^{-Bp}}$。因此，我们有 $\mathbf{Pr}\left[Z_b^{1/B} > t\right] \leq \frac{1}{t^{Bp}}$，所以当 $B = \left\lceil \frac{3}{p} \right\rceil$。

$$\begin{aligned}
\mathbb{E}\left[Z_b^{1/B}\right] &= \int_0^\infty \mathbf{Pr}\left[Z_b^{1/B} > t\right] dt \\
&= \int_0^1 \mathbf{Pr}\left[Z_b^{1/B} > t\right] dt + \int_1^\infty \mathbf{Pr}\left[Z_b^{1/B} > t\right] dt \\
&\leq 1 + \int_1^\infty \frac{1}{t^{Bp}} dt \\
&\leq 1 + \int_1^\infty \frac{1}{t^3} dt \leq 2.
\end{aligned}$$

由于 $\mathbb{E}[Z] = \prod_{b=1}^{B} \mathbb{E}\left[Z_b^{1/B}\right]$，因此我们有 $\mathbb{E}[z] \leq 2^B$。此外，几何平均中每一项 $Z_b^{1/B}$ 的概率密度函数满足 $p(x) = Bpx^{-Bp-1} \cdot e^{-x^{-Bp}} > Bp \cdot e^{-2^{Bp}}$ 对所有 $x \in \left[\frac{1}{2}, 1\right]$ 均成立。

$$\begin{aligned}
\mathbb{E}\left[Z_b^{1/B}\right] &= \int_0^\infty p(t) \cdot t \, dt \\
&> \int_{1/2}^1 p(t) \cdot t \, dt \\
&\geq \frac{1}{2} \cdot \min_{t \in \left[\frac{1}{2}, 1\right]} p(t) \cdot t \\
&\geq \frac{1}{2} \cdot \min_{t \in \left[\frac{1}{2}, 1\right]} Bp \cdot e^{-2^{Bp}} \cdot \frac{1}{2} = \Omega(1),
\end{aligned}$$

从而 $\mathbb{E}\left[Z_b^{1/B}\right] = \Omega(1)$ 与 0 保持有界距离。因此，$\mathbb{E}[Z]$ 等于某个常数 $C_{3.2} \in (0, 2^B]$。通过类似的推理，我们有

$$\begin{aligned}
\mathbb{E}[Z_b^{2/R}] &= \int_0^\infty \mathbf{Pr}\left[Z_b^{2/R} > t\right] dt \\
&\leq \int_0^1 \mathbf{Pr}\left[Z_b^{2/R} > t\right] dt + \int_1^\infty \mathbf{Pr}\left[Z_b^{2/R} > t\right] dt \\
&\leq 1 + \int_1^\infty \frac{1}{t^{Bp/2}} dt \\
&\leq 1 + \int_1^\infty \frac{1}{t^{3/2}} dt \leq 3.
\end{aligned}$$

因此，我们有 $\mathbb{E}[Z^2] = \prod_{b=1}^{B} \mathbb{E}\left[Z_b^{2/R}\right]$，从而 $\mathbb{E}[Z^2] \in [0, 3^B]$。

□

我们现在证明一个结构性质，说明如何利用对损失 $L_i(t) = \left(\sum_{j \in [s]} \ell_i(j, t)^p\right)^{1/p}$ 近似无偏且具有较小二阶矩的估计值 $\widehat{s_i t}()$，实现具有有界遗憾的乘法权重更新算法。

引理 3.3。 假设 $\mathbb{E}\left[\widehat{s_i\ t()}\right] = L_i(t)$ ，且该随机变量对所有 $i \in [n]$ 和 $t \in [T]$ 的二阶矩至多为 ρ。假设乘法权重更新（*MWU*）算法在所有 $i \in [n]$ 和 $t \in [T]$ 中使用损失序列 $s_i(t)$ 而不是 $L_i(t)$ 执行。则在损失序列 $\{L_i(t)\}_{i\in[n],\ t\in[T]}$ 上，由此产生的选择序列具有期望遗憾 $O\left(\sqrt{\rho \log_T n}\right)$。

证明。对于每个 $t \in [T]$，令 $w_t = (w_{t,1}, \ldots, w_{t,n})$ 为 MWU 维护的专家权重向量。则概率分布 p_t 满足 $p_t = \frac{w_t}{\sum_{j=1}^{n} w_{t,j}}$。观测到损失估计 $\widehat{s_i(t)} \geq 0$ 后，权重更新为 $w_{t+1,i} = w_{t,i}\exp(-\eta \widehat{s_i(t)})$。根据假设，$\widehat{s_i(t)}$ 是 $L_i(t)$ 的无偏估计，且独立于 MWU，因此对于 p_t 的每种取值以及每个 i, t，有

$$\mathbb{E}\left[\widehat{s_i(t)} \mid p_t\right] = L_i(t), \qquad \mathbb{E}\left[\widehat{s_i(t)}^2 \mid p_t\right] \leq \rho.$$

对于 $y \geq 0$, $e^{-y} \leq 1 - y + y^2$，因此

$$\mathbb{E}\left[e^{-\eta \widehat{s_i(t)}} \mid p_t\right] \leq 1 - \eta L_i(t) + \eta^2 \rho.$$

For $W_t = \sum_i w_{t,i}$ 以及 $L_t = (L_1(t) \ldots, L_n(t))$，并对 i 按权重 $w_{t,i}$ 求和，

$$\mathbb{E}\left[W_{t+1} \mid \mathcal{F}_t\right] = \sum_i w_{t,i} \mathbb{E}\left[e^{-\eta \widehat{s_i(t)}} \mid p_t\right] \leq W_t(1 - \eta\, p_t \cdot L_t + \eta^2 \rho).$$

迭代并取期望，

$$\mathbb{E}\left[W_{T+1}\right] \leq n \prod_{t=1}^{T} (1 - \eta\, \mathbb{E}\left[p_t \cdot L_t\right] + \eta^2 \rho).$$

对于任意专家 i^*，$w_{T+1,i^*} = \exp(-\eta \sum_t s_{i^*}(t)) \leq W_{T+1}$，因此

$$\mathbb{E}\left[e^{-\eta \sum_t \widehat{s_{i^*}(t)}}\right] \leq n \prod_{t=1}^{T} (1 - \eta\, \mathbb{E}\left[p_t \cdot L_t\right] + \eta^2 \rho).$$

取对数并利用 $\ln(1 + u) \leq u$，

$$\ln \mathbb{E}\left[e^{-\eta \sum_t \widehat{s_{i^*}(t)}}\right] \leq \ln n - \eta \sum_t \mathbb{E}\left[p_t \cdot L_t\right] + \eta^2 \rho T.$$

根据詹森不等式，

$$-\eta \sum_t L_{i^*}(t) = \ln e^{-\eta \sum_t L_{i^*}(t)} \leq \ln \mathbb{E}\left[e^{-\eta \sum_t \widehat{s_{i^*}(t)}}\right].$$

结合这两个不等式可得

$$-\eta \sum_t L_{i^*}(t) \leq \ln n - \eta \sum_t \mathbb{E}\left[p_t \cdot L_t\right] + \eta^2 \rho T,$$

整理后得到

$$\sum_t \mathbb{E}\left[p_t \cdot L_t\right] - \sum_t L_{i^*}(t) \le \frac{\ln n}{\eta} + \eta\rho T.$$

因此，乘法权重更新的损失至多为 $\frac{\ln n}{\eta} + \eta\rho T$。取 $\eta = \sqrt{\frac{\ln n}{\rho T}}$ 并按时间范围 T 归一化，可得遗憾上界至多为 $\mathcal{O}\left(\sqrt{\rho \log n / T}\right)$。 □

遗憾的是，我们无法立即应用 引理 3.3，因为我们没有 $\mathbb{E}\left[\widehat{s_i}(t)\right] = L_i(t)$。也就是说，当 $\max_{j\in[s]} q_i^{(b)}(j,t) < 100 s^{1/p} \log(nsT)$ 时，该值不会发送给协调器。幸运的是，由于指数随机变量的分布，我们证明这种情况发生的概率仅为 $1/\mathrm{poly}(nT)$。因此，总体遗憾仅受到低阶项的影响。

引理 3.4。令 $q_i^{(b)}(j,t)$ 对于 $j \in [s]$ 如 算法 2 中所定义，并假设 $L_i(t) \ge s^{1/p}$。那么

$$\mathbf{Pr}\left[\max_{j\in[s]} q_i^{(b)}(j,t) < \frac{100\log(nsT)}{s^{1/p}}\right] \le \frac{1}{\mathrm{poly}(nST)}.$$

证明。根据 引理 1.6 和 引理 3.2，最大值 $\max_{j\in[s]} q_i^{(b)}(j,t)$ 服从分布 $C_{3.2}\, e^{1/p} L_i(t)$，其中 e 是速率为1的指数随机变量。因此

$$\begin{aligned}\mathbf{Pr}\left[\max_{j\in[s]} q_i^{(b)}(j,t) < \frac{100\log(nsT)}{s^{1/p}}\right] &= \mathbf{Pr}\left[C_{3.2}\, e^{1/p} L_i(t) < \frac{100\log(nsT)}{s^{1/p}}\right] \\ &= \mathbf{Pr}\left[e > \left(\frac{C_{3.2}\, s^{1/p}}{100\, L_i(t)\, \log(nsT)}\right)^p\right].\end{aligned}$$

由于 $L_i(t) \ge s^{1/p}$，我们有

$$\mathbf{Pr}\left[e > \left(\frac{C_{3.2}\, s^{1/p}}{100\, L_i(t)\, \log(nsT)}\right)^p\right] \le \exp\left(-\frac{(100\log(nsT))^p}{C_{3.2}^p}\right) \le \frac{1}{\mathrm{poly}(nsT)}.$$

□

不失一般性，我们将固定常数 $b > a > 0$ 设为 $[1,5]$；我们的证明很容易推广到区间 $[a,b]$。利用标准的集中不等式，例如 定理 1.4，我们现在给出简单算法的期望遗憾以及预备分布式协议的总通信量的上界。

引理 3.5。假设 $\ell_i(j,t) \in [1,5]$ 对所有 $t \in [T]$、$i \in [n]$ 和 $j \in [s]$ 均成立。那么 算法 2 的期望遗憾至多为 $O\left(s^{1/p}\sqrt{\text{对数}\ n/T}\right)$。

证明。由于每台服务器上的损失都在范围 $[1,5]$，内，因此每天每位专家的损失 $L_i(t)$ 至多为 $O\left(s^{1/p}\right)$。根据引理 1.6，我们有 $\widehat{s_i}(t)$ 与 $L_i(t)\, e^{1/p}$ 的分布相同，其中 e 是一个指数随机变量。根据引理 3.2，$\widehat{s_i}(t)$ 是 $L_i(t) + {}^{1}$ 的 $\mathrm{poly}(nT)$ 无偏估计，其二阶矩为 $O\left(s^{1/p}\right)$。因此，根据引理 3.3，算法 2 的期望遗憾至多为

$$\mathcal{O}\left(s^{1/p}\sqrt{\frac{\log n}{T}}\right) + \frac{1}{\mathrm{poly}(nT)} \cdot T = \mathcal{O}\left(s^{1/p}\sqrt{\frac{\log n}{T}}\right).$$

□

引理 3.6。 假设 $\ell_i(j,t) \in [1,5]$ 对所有 $t \in [T]$， $i \in [n]$ 且 $j \in [s]$。那么以高概率，算法 2 的总通信量至多为 $\mathcal{O}(sT) + nT \cdot \mathrm{polylog}(nsT)$ 比特。

证明。考虑算法 2。注意，在时间 $t \in [T]$，服务器 $j \in [s]$ 只有在 $q_i^{(b)}(j,t) \geq \frac{s^{1/p}}{100}\log(nsT)$ 时才会与专家 $i \in [n]$ 通信。根据假设，对于所有 $t \in [T]$、$i \in [n]$ 和 $j \in [s]$，都有 $\ell_i(j,t) \leq 5$。由于对于 $b \in [B]$，$q_i^{(b)}(j,t) = \frac{\ell_i(j,t)}{(e_i^{(b)}(j,t))^{1/p}}$，因此对于固定的 $b \in [B]$，服务器 $j \in [s]$只有在 $(e_i^{(b)}(j,t))^{1/p} \leq \frac{500}{s^{1/p}}\log(nsT)$ 时才会进行通信，等价地说，只有在 $e_i^{(b)}(j,t) \leq \frac{500^p}{s}\log^p(nT)$ 时才会进行通信。由于 $(e_i^{(b)}(j,t))$ 是速率为 1 的独立指数随机变量，因此我们有

$$\mathbf{Pr}\left[(e_i^{(b)}(j,t))^{1/p} \leq \frac{500\log(nsT)}{s^{1/p}}\right] \lesssim \frac{\log^p(nT)}{s}.$$

现在，对于固定的 $i \in [n]$，令 $Y_1, \ldots, Y_S$ 表示指示随机变量，用以指示 $q_i^{(b)}(j,t)$ 是否触发从服务器发往协调器的消息。换言之，若服务器 j 因 i 发送消息，则取 $Y_j = 1$，否则取 $Y_j = 0$。于是我们有

$$\mathbb{E}[Y_1 + \ldots + Y_S] \lesssim \log^p(nsT).$$

根据标准切尔诺夫界，参见 定理 1.4，我们有 $\mathbf{Pr}\left[Y_1 + \ldots + Y_s \gtrsim \log^{p+1}(nsT)\right] \leq \frac{1}{(nsT)^{10}}$。通过对所有 $b \in [B]$、$i \in [n]$、$t \in [T]$ 使用并集界，可知发送到协调器的样本总数至多为 $nT \cdot$ 多对数(nsT)。另一方面，协调器需要对所有 $t \in [T]$ 与每台服务器同步，从而产生 $\mathcal{O}(sT)$ 通信。因此，以高概率，总通信量至多为 $\mathcal{O}(sT) + nT \cdot$ 多对数(nsT)。

□

结合引理 3.5和引理 3.6，我们得到定理 1.1。

4 通信-遗憾权衡

在本节中，我们对预热算法进行参数化，以实现通信-遗憾权衡。具体而言，我们证明，如果目标是实现遗憾 R，那么存在一种分布式协议，其总通信量至多为 $\binom{n+s}{R^2} \cdot \mathrm{polylog}(nsT)$ 比特。

该算法与算法 2颇为相似。主要区别在于，现在每个时刻都会对每台服务器进行抽样，以决定是否运行先前的协议。也就是说，以概率 ϱ，每台服务器独立地执行与之前相同的协议。否则，以概率 $1-\varrho$，服务器不发言。值得注意的是，协调器也可以利用公共随机性获知这些事件的结果，因此无需逐一与服务器通信来确定该服务器是否被抽中。该算法的完整描述见算法 3。

我们首先证明指数随机变量具有以下性质。

事实 4.1。 设 e 为速率为1的指数随机变量。那么，对于所有 $x > 2$，都有$\mathbf{Pr}\left[\frac{1}{e} \in (x, 2x]\right] \in \left[\frac{1}{4x}, \frac{1}{2x}\right]$。

证明。 根据 事实 3.1，$\frac{1}{e}$ 的概率密度函数为 $p(x) = \frac{1}{x^2} \cdot e^{-1/x}$。我们有

$$\mathbf{Pr}\left[\frac{1}{e} \in (x, 2x]\right] = \int_x^{2x} p(x)\,dx = \int_x^{2x} \frac{1}{x^2} \cdot e^{-1/x}\,dx.$$

算法 3 带通信的分布式协议 -遗憾权衡 -off

输入：目标遗憾 R**，对于所有时间** $t \in [T]$**、专家** $i \in [n]$ **和服务器** $j \in [s]$**，损失** $\{\ell_i(j,t)\}$**，** L_p**损失参数** p**输出：每天要进行的专家序列 1：** $B \leftarrow \left\lceil \frac{3}{p} \right\rceil$**,** $\varrho \leftarrow \frac{1}{R^2 T}$**2：对于每个时间** $t \in [T]$ **，执行 3：以概率** $1-\varrho$**，服务器不发送任何内容，且权重** $\{w_i(t)\}$ **不更新 4：否则，以概率** ϱ **执行以下操作：5：对于每个** $i \in [n]$ **6：对于每个服务器** $j \in [s]$ **7：令** $\ell_i(j,t)$ **表示专家** i **在服务器** j **上于时间** t **的损失 8：令** $e_i^{(b)}(j,t)$ **对所有** $b \in [B]$ **均为独立指数随机变量 9：** $q_i^{(b)}(j,t) \leftarrow \frac{\ell_i(j,t)}{C_{3.2}(e_i^{(b)}(j,t))^{1/p}}$ **对于** $b \in [B]$ ▷ **引理 3.2 10：如果** $q_i^{(b)}(j,t) \geq \frac{s^{1/p}}{100} \log(nsT)$**，则 11：向协调器发送** $q_i^{(b)}(j,t)$ **12：结束条件 13：结束循环 14：结束循环 15：对于每个** $i \in [n]$ **16：** $\widehat{s_i\ t}() \leftarrow \frac{1}{\varrho} \cdot \prod_{b \in [B]} \left(\mathbf{max}_{j \in [s]}\, q_i^{(b)}(j,t) \right)^{1/B}$ **17：** $w_i(t) \leftarrow w_i(t-1) + s_i(t)$ **18：结束循环 19：在** $\{w_i(t)\}$ **上运行乘法权重更新 20：结束循环**

注意，对于 $x > 2$，我们有 $e^{1/x} \in \left(1, \frac{1}{2}\right]$。因此，

$$\mathbf{Pr}\left[\frac{1}{e} \in (x, 2x]\right] \leq \int_x^{2x} \frac{1}{x^2}\, dx = \frac{1}{x} - \frac{1}{2x} = \frac{1}{2x},$$

and

$$\mathbf{Pr}\left[\frac{1}{e} \in (x, 2x]\right] \geq \int_x^{2x} \frac{1}{2}\frac{1}{x^2}\, dx = \frac{1}{2x} - \frac{1}{4x} = \frac{1}{4x}.$$

□

现在，我们利用指数随机变量的性质，分析该通信-遗憾权衡协议的期望并给出方差的上界。

引理 4.2. 我们有 $\mathbb{E}\left[\widehat{s_i\ t}()\right] = L_i(t) + \frac{1}{\mathrm{poly}(nT)}$ 并且 $\mathbb{E}\left[(\widehat{s_i\ t}())^2\right] \leq \mathcal{O}\left(TR^2\right) \cdot 3^B \cdot (L_i(t))^2$。

证明。 设每次抽样的概率为 ϱ $t \in [T]$。注意到，以概率 $1-\varrho$，我们有 $s_i(t) = 0$。否则，我们像之前一样继续。

令 e、e_1、……、e_B 为独立随机变量。根据 引理 1.6，我们有

$$\begin{aligned}
\mathbb{E}\left[\widehat{s_i(t)}\right] &= \varrho \cdot \frac{1}{\varrho} \cdot \mathbb{E}\left[\prod_{b=1}^{B}\left(\max_{j\in[s]} q_i^{(b)}(j,t)\right)^{1/B}\right] \\
&= \varrho \cdot \frac{1}{\varrho} \cdot \mathbb{E}\left[\prod_{b=1}^{B}\left(\frac{L_i(t)}{C_{3.2}e_b^{1/p}}\right)^{1/B}\right] \\
&= \varrho \cdot \frac{1}{\varrho} \cdot \prod_{b=1}^{B}\mathbb{E}\left[\left(\frac{L_i(t)}{C_{3.2}e^{1/p}}\right)^{1/B}\right] \\
&= \frac{L_i(t)}{C_{3.2}} \cdot \left(\mathbb{E}\left[\frac{1}{e^{1/Bp}}\right]\right)^{B} \\
&= L_i(t),
\end{aligned}$$

其中最后一个等式由 $C_{3.2} = \left(\mathbb{E}\left[\frac{1}{e^{1/Bp}}\right]\right)^{B}$ 在 引理 3.2中的定义得出。同样地，我们有

$$\begin{aligned}
\mathbb{E}\left[(\widehat{s_i(t)})^2\right] &= p \cdot \frac{1}{\varrho^2} \cdot \mathbb{E}\left[\prod_{b=1}^{B}\left(\frac{(L_i(t))^2}{(C_{3.2})^2 e_b^{2/p}}\right)^{1/B}\right] \\
&= \frac{1}{\varrho} \cdot \prod_{b=1}^{B}\mathbb{E}\left[\left(\frac{(L_i(t))^2}{(C_{3.2})^2 e_b^{2/p}}\right)^{1/B}\right] \\
&= \frac{1}{\varrho} \cdot \prod_{b=1}^{B}\cdot\frac{(L_i(t))^2}{(C_{3.2})^2} \cdot \mathbb{E}\left[\frac{1}{e_b^{2/Bp}}\right]^{B}.
\end{aligned}$$

根据 引理 3.2，我们有 $C_{3.2} = \Theta(1)$ 以及 $\mathbb{E}\left[\frac{1}{e_b^{2/Bp}}\right]^{B} \le 3^B$。此外，我们令 $\varrho = \frac{1}{TR^2}$，从而 $\frac{1}{\varrho} = TR^2$。因此，

$$\mathbb{E}\left[(\widehat{s_i(t)})^2\right] \le TR^2 \cdot 3^B \cdot \mathcal{O}(1) \cdot (L_i(t))^2,$$

如所需。 □

再次强调，我们无法得到 $\mathbb{E}\left[\widehat{s_i\,t()}\right] = L_i(t)$ 的理想化期望，原因在于$\max_{j\in[s]} q_i^{(b)}(j,t)$若小于 $\frac{1}{100^{\,s^{1/p}\log(nsT)}}$，则可能不会被发送给协调器，这会改变最终的期望计算。然而，由于该事件发生的概率仅为 $\frac{1}{\mathrm{poly}(nT)}$，因此可得出相应结论。随后，我们对算法 3 的遗憾和总通信量进行如下上界估计：

引理 4.3. 设 $\ell_i(j,t) \in [1,5]$ 对所有 $t \in [T]$, $i \in [n]$ 且 $j \in [s]$。算法 *3* 的期望遗憾至多为 $\mathcal{O}\left(Rs^{1/p}\sqrt{\log n}\right)$。

证明。 由于每台服务器上的损失都在范围 $[1,5]$, 内，因此每位专家每天的损失 $L_i(t)$ 至多为 $\mathcal{O}\left(s^{1/p}\right)$。根据引理 5.1，我们有 $\mathbb{E}\left[\widehat{s_i\,t()}\right] = L_i(t) + \frac{1}{\mathrm{poly}(nT)}$ 且

$\mathbb{E}\left[(\widehat{s_i\ t}())^2\right] = \mathcal{O}(TR^2)\cdot 3^B\cdot(L_i(t))^2$。因此，根据 引理 3.3，算法 3 的期望遗憾至多为 $\mathcal{O}\left(Rs^{1/p}\sqrt{3^B \log\ n}\right) + \frac{1}{\text{poly}(nT)}\cdot T = \mathcal{O}\left(Rs^{1/p}\sqrt{3^B \log\ n}\right)$。

□

引理 4.4. 假设对所有 $t\in[T]$, $i\in[n]$ 和 $j\in[s]$, $\ell_i(j,t)\in[1,5]$。那么，以高概率，总通信量至多为 $\binom{n+s}{R^2}\cdot$ 多对数(nsT) 比特。

证明。考虑算法 3。注意，在时刻 $t\in[T]$，服务器 $j\in[s]$ 仅当先以概率 ϱ 被采样，且随后 $q_i^{(b)}(j,t) \geq 100^{s^{1/p}}\log(nsT)$ 时，才会与专家 $i\in[n]$ 通信。根据假设，对所有~~$t\in[T]$~~, $i\in[n]$ 和 $j\in[s]$，有 $\ell_i(j,t)\leq 5$。由于对于 $b\in[B]$，$q_i^{(b)}(j,t) = \ell_i(j,t)(e_i^{(b)}(j,t))^{1/p}$，~~因此~~对于固定的 $b\in[B]$，服务器 $j\in[s]$ 仅当 $(e_i^{(b)}(j,t))^{1/p}\leq 500\log(nsT)\, s^{1/p}$ 时才会进行通信；等价地，当且仅当 $e_i^{(b)}(j,t)\leq 500^p\log^p(nT)\, s$。由于 $(e_i^{(b)}(j,t))$ 是一个速率为 1 的独立指数随机变量，因此我们有

$$\mathbf{Pr}\left[(e_i^{(b)}(j,t))^{1/p}\leq\frac{500\log(nsT)}{s^{1/p}}\right]\lesssim\frac{\log^p(nT)}{s}.$$

现在，对于固定的 $i\in[n]$，令 $Y_1,\ldots,Y_S$ 表示指示随机变量，用于判断 $q_i^{(b)}(j,t)$ 是否会触发从服务器向协调器发送消息。换言之，$Y_j=1$ 若服务器因 i 而发送消息，则 j；否则 $Y_j=0$。那么，对于每个 $j\in[s]$，我们有

$$\mathbb{E}[Y_j]\lesssim\varrho\cdot\frac{\log^p(nT)}{s},$$

因为概率为 ϱ 的事件必须先发生。由期望的线性性可得

$$\mathbb{E}[Y_1+\ldots+Y_s]\lesssim\varrho\cdot\log^p(nsT).$$

类似地，在所有时刻和所有专家上，由这些事件产生的期望通信总量至多为 $\mathcal{O}(\varrho\cdot nT\cdot\log^p(nsT))$。回忆一下，算法 3 设置 $\varrho=\frac{1}{R^2T}$。根据标准的切尔诺夫界（参见 定理 1.4），以高概率，发送给协调器的样本总数至多为 $\frac{n}{R^2}\cdot$ 多对数(nsT)。另一方面，对于所有选定日期，协调器都需要与每台服务器同步，以防服务器没有发送任何通信。通过类似的切尔诺夫界论证，以高概率，这一过程会产生 $\frac{s}{R^2}\cdot$ 多对数(nsT) 通信。因此，如所述，以高概率，总通信量至多为$\binom{n+s}{R^2}\cdot$ 多对数(nsT)。

□

结合引理 4.3 和引理 4.4，可得定理 1.2。

5 完整算法

在本节中，我们基于先前的算法给出主要结果，其实现的遗憾大致为 R，同时使用 $\binom{n+s}{R^2}\cdot\max(s^{1-2/p},1)\cdot\text{polylog}(nsT)$ 通信总比特数。与前几节不同，我们不要求每个损失都位于区间 $[a,b]$ 内，其中 $b>a>0$是固定常数。 算法 4 与先前算法的不同之处在于，它允许服务器

算法 4 专家问题的分布式协议

输入：目标遗憾 R**，对所有时间** $t \in [T]$**、专家** $i \in [n]$ **和服务器** $j \in [s]$**,** L_p **的损失 {**$\ell_i(j,t)$**}，损失参数** p**输出：每天要使用的专家序列 1:** $B \leftarrow \left\lceil \frac{3}{p} \right\rceil$**,** $\varrho \leftarrow \frac{1}{R^2T} \cdot$ **max(**$s^{1-2/p}$**, 1)，**$A \leftarrow \lceil 10$**log(**nsT**)⌉2: 对每个时间** $t \in [T]$ **执行 3: 以概率** $1-\varrho$**，服务器不发送任何内容，且权重 {**$w_i(t)$**} 不更新 4: 否则，以概率** $\frac{\varrho}{2^a}$ **选择一个整数** $a > 0$ **5: 所有服务器向协调器发送值** $q_i^{(b)}(j,t)$**，这些值至少为** $\frac{s^{1/p}}{100\ 2^{a\ 1/p}} \cdot$**( ) log(**$nsT$**)▷引理 3.2 6: 若协调器收到，则令** $\widehat{q_i^{(b)}\ j}$ **(,** t**)** $\frac{q_i^{(b)}(j,t)}{\neq B}$ $C_{3.2}$**，否则设为零 7: 对每个** $i \in [n]$ **执行 8:** $\widehat{s_i\ t}$**( )** $\leftarrow \cdot \prod_{b\in[B]} \left(\max_{j\in[s]} \widehat{q_i^{(b)}\ j}\ (\ ,t) \right)$ **9: 令** $a^* \geq 0$ **为满足** $\widehat{s_i\ t}$**( )** $\geq \left(\frac{s}{2^{a^*}}\right)^{1/p}$ **的最小非负整数 10: 如果** $s_i(t) \neq 0$ **且** $a^* \leq a$**，则 11:** $w_i(t) \leftarrow w_i(t-1) + 2^{a^*} R^2 T \cdot \widehat{s_i\ t}$**( ) 12: 结束 if 13: 结束 for 14: 对 {**$w_i(t)$**} 使用乘法权重更新 15: 结束 for**

通过发送更小的值来更积极地通信，尤其是那些大于较小阈值的值。不过，只有在服务器被恰当地抽样时，这种情况才会发生。

我们首先考察算法 4的期望和方差。我们利用指数随机变量的性质证明，除去一个较小的扰动，$s_i(t)$ 的期望值恰好等于专家 $L_i(t)$ 在时间 i 的损失 t。对于 $s_i(t)$的二阶矩，我们采用类似的论证。

引理 5.1. 假设对于所有 $b \in [B]$，$q_i^{(b)}(j,t) \geq \frac{1}{100\sqrt{T}}$。那么我们有 $\mathbb{E}\left[\widehat{s_i\ t}()\right] = L_i(t) + \frac{1}{\text{poly}(nT)}$且 $\mathbb{E}\left[(\widehat{s_i\ t}())^2\right] \leq s^{2/p} \cdot R^2T$。

证明。令 $a^* \geq 0$ 为满足 $\widehat{s_i\ t}$**( )** $\geq \left(\frac{s}{2^{a^*}}\right)^{1/p}$ 的最小整数。注意，以概率 $\varrho_{a^*} := \frac{\varrho}{2^{a^*}}$，我们有 $a^* \leq a$。此外，我们要求 $a^* \leq a$，以便设置 $w_i(t) = w_i(t-1) + 2^{a^*} \cdot \widehat{s_i\ t}$**( )**。考虑一个假设过程 $\mathcal{P}$：若 $a^* \leq a$，则所有服务器都报告 $q_i^{(b)}(j,t)$。我们最终将证明 $\mathcal{P}$ 以高概率成立，因此实际过程对期望遗憾的影响也很小。

现在，在此过程中，$\widehat{s_i\ t}()$ 以概率 ϱ_{a*}非零。回顾 $B = \left\lceil \frac{3}{p} \right\rceil$。于是，根据引理 1.6，

$$\mathbb{E}\left[\widehat{s_i(t)}\right] = \mathbb{E}\left[\varrho_{a*} \cdot \frac{1}{\varrho_{a^*}} \cdot \prod_{b=1}^{B} \left(\max_{j\in[s]} q_i^{(b)}(j,t) \right)^{1/B}\right]$$

$$
\begin{aligned}
&= \prod_{b=1}^{B} \mathbb{E}\left[\left(\max_{j\in[s]} q_i^{(b)}(j,t)\right)^{1/B}\right] \\
&= \left(\mathbb{E}\left[\left(\max_{j\in[s]} q_i^{(b)}(j,t)\right)^{1/B}\right]\right)^{B} \\
&= \left(\mathbb{E}\left[\left(\frac{L_i(t)}{C_{3.2}e^{1/p}}\right)^{1/B}\right]\right)^{B} \\
&= \frac{L_i(t)}{C_{3.2}} \cdot \left(\mathbb{E}\left[\frac{1}{e^{1/Bp}}\right]\right)^{B} \\
&= L_i(t),
\end{aligned}
$$

其中，最后一个等式由 $C_{3.2} = \left(\mathbb{E}\left[\frac{1}{e^{1/Bp}}\right]\right)^{B}$ 在引理 3.2中的定义得出。同理，

$$
\begin{aligned}
\mathbb{E}\left[(\widehat{s_i(t)})^2\right] &= \mathbb{E}\left[\varrho_{a^*} \cdot \frac{1}{\varrho_{a^*}^2} \cdot \prod_{b=1}^{B}\left(\max_{j\in[s]} q_i^{(b)}(j,t)\right)^{2/B}\right] \\
&= \mathbb{E}\left[\frac{1}{\varrho_{a^*}} \prod_{b=1}^{B}\left(\max_{j\in[s]} q_i^{(b)}(j,t)\right)^{2/B}\right].
\end{aligned}
$$

我们有

$$
\prod_{b\in[B]}\left(\max_{j\in[s]} \widehat{q_i^{(b)}(j,t)}\right)^{1/B} \in \left[\left(\frac{s}{2^{a^*}}\right)^{1/p}, \left(\frac{s}{2^{a^*-1}}\right)^{1/p}\right].
$$

我们还有 $\varrho_{a^*} = \frac{\varrho}{2^{a^*}\circ}$ 因此，

$$
\begin{aligned}
\mathbb{E}\left[(\widehat{s_i(t)})^2\right] &= \mathbb{E}\left[\frac{1}{p_{a^*}} \cdot \prod_{b=1}^{B}\left(\max_{j\in[s]} q_i^{(b)}(j,t)\right)^{2/B}\right] \\
&\le \mathbb{E}\left[\frac{2^{a^*}}{\varrho} \cdot \left(\frac{s}{2^{a^*}}\right)^{2/p}\right] \\
&\le \frac{(2^{a^*})^{1-2/p} \cdot s^{2/p}}{\varrho}.
\end{aligned}
$$

现在，对于 $p \le 2$，我们有 $(2^{a^*})^{1-2/p} \le 1$ 以及 $\varrho = \frac{1}{R^2T} \cdot \max(s^{1-2/p}, 1) = \frac{1}{R^2T}$，因此 $\mathbb{E}\left[(\widehat{s_i t()})^2\right] \le R^2T$。

类似地，对于 $p > 2$，我们有 $\varrho = \frac{1}{R^2T} \cdot \max(s^{1-2/p}, 1) = \frac{s^{1-2/p}}{R^2T}$ 因此

$$
\mathbb{E}\left[(\widehat{s_i(t)})^2\right] \le \frac{(2^{a^*})^{1-2/p} \cdot s^{2/p} \cdot R^2T}{s^{1-2/p}}.
$$

由于 $2^{a^*} \le s$，因此可得 $\mathbb{E}\left[(\widehat{s_i t()})^2\right] \le s^{2/p} \cdot R^2T$，正如所需。 □

接下来，我们对算法 4的遗憾和总通信量进行如下上界估计：

引理 5.2。 假设对于所有 $\ell_i(j,t) \le 1$，都有 $t \in [T]$。算法算法 4 的期望遗憾至多为$\mathcal{O}\left(Rs^{1/p}\sqrt{\log n}\right)$。

证明。 由于每台服务器上的损失至多为 1，因此每一天每位专家的损失 $L_i(t)$至多为$\mathcal{O}\left(s^{1/p}\right)$。根据引理 5.1，我们有 $\mathbb{E}\left[\widehat{s_i\ t}(\ )\right] = L_i(t) + \quad {}^{1}{}_{\mathrm{poly}(nT)}$ 且 $\mathbb{E}\left[(\widehat{s_i\ t}(\ ))^2\right] \le s^{2/p}\cdot R^2T$。因此，根据引理 3.3，算法 4 的期望遗憾至多为$\mathcal{O}\left(Rs^{1/p}\sqrt{\log\ n}\right) + \quad {}^{1}{}_{\mathrm{poly}(nT)} \cdot T = \mathcal{O}\left(Rs^{1/p}\sqrt{\log n}\right)$。 □

接下来需要给出算法 4的总通信量上界。

引理 5.3。 假设我们有 $\ell_i(j,t) \le 1$ 对所有 $t \in [T]$均成立。那么以高概率，算法 4 的总通信量至多为 $\binom{n+s}{R^2}\cdot\max(s^{1-2/p},1)\cdot\mathrm{polylog}(nsT)$ 比特。

证明。考虑算法 4 和一个固定的 $a \in [A]$。在时间 $t \in [T]$，服务器 $j \in [s]$ 只有在以概率 $\frac{\varrho}{2^a}$ 首先被选中后，才会与专家 $i \in [n]$通信；此时 $q_i^{(b)}(j,t) \ge {}_{100}\ {}_{2^a}\frac{s^{1/p}}{1/p}\cdot(\ )\log(nsT)$。根据假设，对于所有 $t \in [T]$、$i \in [n]$和 $j \in [s]$，都有 $\ell_i(j,t) \le 1$。由于当 $b \in [B]$时，$q_i^{(b)}(j,t) = \quad {}^{\ell_i}(j,t)\ (e_i^{(b)}(j,t))^{1/p}$，对于固定的 $b \in [B]$，服务器 $j \in [s]$只有在 $(e_i^{(b)}(j,t))^{1/p} \le {}^{100}\cdot(2^a)^{1/p}\log(nsT)\ s^{1/p}$ 时才会通信，等价地说，当且仅当 $e_i^{(b)}(j,t) \le {}^{100^p}\cdot 2^a\cdot\log^p(nT)\ s$ 时才会通信。由于 $(e_i^{(b)}(j,t))$ 是速率为 1 的独立指数随机变量，因此我们有

$$\mathbf{Pr}\left[(e_i^{(b)}(j,t))^{1/p} \le \frac{100\cdot(2^a)^{1/p}\log(nsT)}{s^{1/p}}\right] \lesssim \frac{2^a\cdot\log^p(nT)}{s}$$

现在，对于固定的 $i \in [n]$，令 $Y_1,\ldots,Y_S$表示指示随机变量，用于指示 $q_i^{(b)}(j,t)$ 是否会触发从服务器发送给协调器的消息。换言之，若服务器 j 因 i发送消息，则 $Y_j = 1$；否则 $Y_j = 0$。那么，对于每个 $j \in [s]$，我们有

$$\mathbb{E}[Y_j] \lesssim \varrho\cdot\frac{\log^p(nT)}{s},$$

因为该事 件必须先以概率 $\frac{\varrho}{2^a}$ 发生。根据期望的线性性，我们有

$$\mathbb{E}[Y_1+\ldots+Y_s] \lesssim \varrho\cdot\log^p(nsT).$$

类似地，对于所有时刻和专家，由这些事件产生的期望总通信量至多为 $O(\varrho\cdot nT\cdot\log^p(nsT))$。回顾一下，算法 4 将 $\varrho = \frac{1}{R^2T}$设为 $\cdot\max(s^{1-2/p},1)$。根据标准的切尔诺夫界，如定理 1.4 所示，发送给协调器的样本总数以高概率至多为 $\frac{n}{R^2}\cdot\mathrm{polylog}(nsT)\cdot\max(s^{1-2/p},1)$。另一方面，为防止服务器不发送任何消息，协调器必须在所有选定日期与每台服务器同步。通过类似的切尔诺夫界论证，该过程以高概率产生 $\frac{s}{R^2}\cdot\mathrm{polylog}(nsT)\cdot\max(s^{1-2/p},1)$ 的通信量。因此，以高概率，总通信量至多为$\binom{n+s}{R^2}\cdot\max(s^{1-2/p},1)\cdot\mathrm{polylog}(nsT)$，符合要求。 □

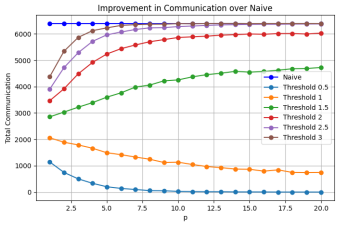
综合引理 5.2和引理 5.3，我们得到主要结果：

定理 1.3。 假设对于所有 $t \in [T]$，我们都有 $\ell_i(j,t) \le 1$。存在一种算法，其期望遗憾至多为 $O\left(Rs^{1/p}\sqrt{\log n}\right)$，并且以高概率使用的总通信量至多为$\binom{n+s}{R^2}\cdot\max(s^{1-2/p},1)\cdot\mathrm{polylog}(nsT)$ 比特。

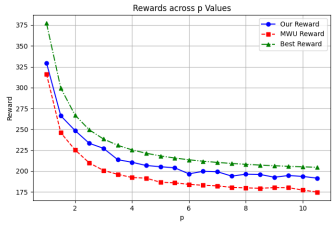
6 实证评估

本节中，我们对所提出的分布式协议进行实证评估，并将其与标准离线乘法权重更新算法以及 [JPT+25]先前提出的协议进行比较。具体而言，我们研究了 HPO-B 数据集 [PJWG21], ，这是一个基准数据集，旨在评估和比较超参数优化（HPO）算法在各种机器学习任务中的表现。我们使用 Python 3.11.5，在一台配备 Intel(R) Core(TM) i7-3770 处理器的 64 位操作系统上开展实证评估；该处理器具有 4 个核心、3.4GHz 基础时钟频率和 16GB 内存。该数据集涵盖了多种数据集和学习任务，例如分类和回归，用于评估不同超参数优化方法的有效性、效率和计算成本。为确保透明度并便于复现结果，我们提供了所有算法、实验设置和数据预处理流程的完整实现，可在 https://github.com/samsonzhou/distributed-experts获取。

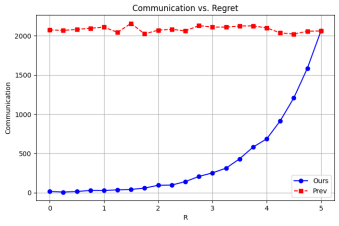
作为超参数优化的黑盒基准，我们可以将 HPO-B 基准中的各种模型视为分布式专家问题中彼此不同的专家，并将不同的数据集分配给不同的服务器。对于所有模型类别，每个搜索步骤都是随机搜索，可视为分布式专家场景中的一天。成本向量表示搜索步骤期间模型在不同数据集上的归一化负准确率。因此，我们将算法 2 中的简单分布式协议 与乘法权重更新以及 [JPT+25]的算法进行了比较。我们的结果见图 2a ，结果表明，对于 $p > 1$，总通信量会随着阈值的增加而增加；但对于 $p < 1$，趋势则相反。图 2b 显示，我们的算法所得奖励出乎意料地高于乘法权重更新算法；我们认为，后者的问题在于乘法权重更新的学习率调优不够精细。最后，图 2c 显示，对于特殊情形$p = 1$，我们的算法比 [JPT+25]的算法通信量更低。此外，图 2c 展现了理论分析所预期的通信-遗憾权衡。



(a) 通信与 p 的关系



(b) 奖励与 p 的关系



(c) 通信与遗憾的关系

结论

在本文中，我们研究了协调器模型下的专家分布式在线学习问题，其中每个专家损失隐式地定义在 s 个服务器上。我们提出了针对一般 ℓ_p损失的首个协议，以通信O $\left(\left(\frac{n}{R^2} + \frac{s}{R^2}\right) \cdot \max(s^{1-2/p}, 1) \cdot$ 多对数$(nsT)\right)$实现遗憾$R \gtrsim \sqrt{T}$。对于非零服务器损失有界这一重要特例，我们的协议以总通信量$\left(\frac{n+s}{R^2}\right) \cdot$ 多对数(nsT)达到遗憾O $\left(Rs^{1/p}\sqrt{\log n}\right)$，改进了仅处理 $p = 1$的先前界 [JPT+25]。一个关键

技术贡献在于我们针对 ℓ_p 损失的算法框架，这类损失比 ℓ_1要微妙得多。通过使用随机指数缩放将 ℓ_p 损失嵌入 ℓ_∞ 结构，我们将该问题归结为跟踪所有专家损失的最大缩放贡献。为控制指数随机变量固有的无界方差，我们在独立缩放上引入几何平均估计量，从而得到方差可证明有界的无偏估计。该估计量在算法和技术上都具有新颖性，并且在带专家的分布式在线学习之外也可能具有更广泛的适用性。尽管几何平均非常有效，也可探索其他方差缩减技术，这代表未来工作中一个有前景且富有成果的方向。最后，我们的结果凸显了具有一般 ℓ_p 损失的分布式在线学习中通信与遗憾之间的基本权衡。将这些技术扩展到其他结构化损失函数，例如次模目标或 ℓ_∞ (max)损失，为进一步推进分布式序贯决策提供了一条令人振奋的途径。

致谢

David P. Woodruff 获得美国海军研究办公室项目（编号 N000142112647）和西蒙斯研究员奖的部分资助。Samson Zhou 获得美国国家科学基金会 CCF-2335411 项目的部分资助。Samson Zhou 谨此感谢橡树岭联合大学（ORAU）Ralph E. Powe 青年教师发展奖提供的资助。

参考文献

[ACFS02] Peter Auer、Nicolò Cesa-Bianchi、Yoav Freund 和 Robert E. Schapire。非随机多臂老虎机问题。《SIAM计算杂志》，32(1):48–77，2002年。1[ACNS23] Anders Aamand、Justin Y. Chen、Huy Lê Nguyen 和 Sandeep Silwal。改进专家学习的空间界。CoRR, abs/2303.01453，2023年。1[AHK12] Sanjeev Arora、Elad Hazan 和 Satyen Kale。乘法权重更新方法：一种元算法及其应用。《Theory Comput.》，8(1):121–164，2012年。1[AW20] Sepehr Assadi 和 Chen Wang。有限内存下的探索：用于抛硬币、含噪比较和多臂老虎机的流式算法。载于第52届ACM SIGACT理论计算研讨会论文集（STOC），第1237–1250页，2020年。8[BGW+26] Vladimir Braverman、Sumegha Garg、Chen Wang、David P. Woodruff 和 Samson Zhou。滑动窗口模型下的有限信息在线学习。载于ACM-SIAM离散算法研讨会论文集（SODA），2026年。1[Bro51] George W Brown。通过虚拟博弈迭代求解博弈。《Act. Anal. Prod Allocation》，13(1):374，1951年。7[Cla95] Kenneth L. Clarkson。低维情形下线性规划与整数规划的拉斯维加斯算法。《J. ACM》，42(2):488–499，1995年。7

[CMVW16] Michael S. Crouch、Andrew McGregor、Gregory Valiant 和 David P. Woodruff。随机流：样本复杂度与空间复杂度之争。载于第24届欧洲算法年会（ESA），第32:1–32:15页，2016年。8[Cov66] Thomas M Cover。二进制序列序贯预测器的行为。斯坦福电子实验室技术报告第7002号。斯坦福大学系统理论，1966年。4, 7[DS18] Yuval Dagan 和 Ohad Shamir。以极少内存和通信检测相关性。载于学习理论会议（COLT），第1145–1198页，2018年。8[FS97] Yoav Freund 和 Robert E Schapire。在线学习的决策理论推广及其在提升中的应用。《计算机与系统科学杂志》，55(1):119–139，1997年。7[FS99] Yoav Freund 和 Robert E Schapire。使用乘法权重的自适应博弈。《博弈与经济行为》，29(1-2):79–103，1999年。7[GK07] Naveen Garg 和 Jochen Könemann。用于多商品流及其他分数背包装箱问题的更快、更简单算法。《SIAM计算杂志》，37(2):630–652，2007年。7[GRT18] Sumegha Garg、Ran Raz 和 Avishay Tal。基于提取器的学习时间—空间下界。载于第50届ACM SIGACT计算理论年会论文集（STOC），第990–1002页，2018年。8[GRT19] Sumegha Garg、Ran Raz 和 Avishay Tal。双遍学习的时间—空间下界。载于第34届计算复杂性会议（CCC），第22:1–22:39页，2019年。8[HNO08] Nicholas J. A. Harvey、Jelani Nelson 和 Krzysztof Onak。通过逼近理论进行草图化和流式熵计算。载于第49届IEEE计算机科学基础年会（FOCS），第489–498页，2008年。2[JPT+25] Zhihao Jia、Qi Pang、Trung Tran、David Woodruff、Zhihao Zhang 和 Wenting Zheng。分布式专家问题的通信下界。arXiv预印本 arXiv:2501.03132，2025年。2, 3, 4, 9, 20[KRT17] Gillat Kol、Ran Raz 和 Avishay Tal。学习稀疏奇偶校验的时间—空间困难性。载于第49届ACM SIGACT理论计算研讨会论文集（STOC），第1067–1080页，2017年。8[KV05] Adam Tauman Kalai 和 Santosh S. Vempala。在线决策问题的高效算法。《J. Comput. Syst. Sci.》，71(3):291–307，2005年。7[Li08] Ping Li。用于降维的估计量与尾界 ℓ_α ($0 < p \leq 2$)通过稳定随机投影实现。载于第19届ACM-SIAM离散算法年会论文集（SODA），第10–19页，2008年。4, 9[LSPY18] David Liau、Zhao Song、Eric Price 和 Ger Yang。常数空间下的随机多臂老虎机。载于国际人工智能与统计会议（AISTATS），第386–394页，2018年。8

[LW94] Nick Littlestone 和 Manfred K. Warmuth。加权多数算法。《信息与计算》，108(2):212–261，1994。6，7[McM11] H. Brendan McMahan。跟随正则化领导者与镜像下降：等价性定理与 L1正则化。载于《第十四届国际人工智能与统计会议论文集》，AISTATS，第 525–533 页，2011。7[PJWG21] Sebastian Pineda-Arango、Hadi S. Jomaa、Martin Wistuba 和 Josif Grabocka。HPO-B：基于 OpenML 的大规模可复现黑盒超参数优化基准。载于 Joaquin Vanschoren 和 Sai-Kit Yeung 编，《神经信息处理系统数据集与基准专题论文集》，2021。2，20[PR23] Binghui Peng 和 Aviad Rubinstein。在线学习的近最优内存-遗憾权衡。载于第 64 届 IEEE计算机科学基础年会，FOCS，第 1171–1194 页，2023。1，8[PST95] Serge A. Plotkin、David B. Shmoys 和 Éva Tardos。分数打包与覆盖问题的快速近似算法。《数学运筹研究》，20(2):257–301，1995。7[PZ23] Binghui Peng 和 Fred Zhang。亚线性空间中的在线预测。载于《2023 年 ACM-SIAM 离散算法研讨会论文集》，SODA，第 1611–1634 页，2023。1，8[Raz17] Ran Raz。一大类学习问题的时间-空间下界。载于第 58 届 IEEE计算机科学基础年会，FOCS，第 732–742 页，2017。8[Raz19] Ran Raz。快速学习需要良好的内存：奇偶学习的时间-空间下界。《ACM 杂志》，66(1):3:1–3:18，2019。8[SWXZ22] Vaidehi Srinivas、David P. Woodruff、Ziyu Xu 和 Samson Zhou。专家问题的内存界限。载于 STOC ’22：第 54 届 ACM SIGACT 理论计算研讨会，第 1158–1171 页，2022。1，7[WZ21] David P. Woodruff 和 Samson Zhou。通过差分估计器获得对抗鲁棒流与滑动窗口的紧确界。载于第 62 届 IEEE计算机科学基础年会，FOCS 2021，第 1183–1196 页，2021。2，4[WZZ23] David P. Woodruff、Fred Zhang 和 Samson Zhou。面向专家学习的鲁棒流式处理：算法与下界。载于《神经信息处理系统进展 36：神经信息处理系统年会》，NeurIPS，2023。1