arXiv:2302.00814v2 [cs.LG] 3 Feb 2023

长时域奖励的随机上下文赌博机

Yuzhen Qin*, Yingcong Li,* Fabio Pasqualetti,* Maryam Fazel,† Samet Oymak*, ‡

摘要

人们对复杂决策和语言建模问题日益增长的兴趣，凸显了在极长时域内实现样本高效学习的重要性。本文朝着这一方向迈出一步，研究上下文线性赌博机，其中当前奖励最多依赖于 s 个先前动作和上下文（不一定连续），且时域最长为h。为避免对 h的多项式依赖，我们提出了利用稀疏性联合发现依赖模式和臂参数的新算法。我们考虑数据匮乏（$T < h$）和数据丰富（$T \geq h$）两种情形，并分别推导出遗憾上界$\tilde{O}$（$d\sqrt{sT} + \min\{q, T\}$）和 $\tilde{O}$（$\sqrt{sdT}$），其中稀疏度为 s，特征维度为 d，总时域为 T，而 q 则可根据奖励依赖模式自适应调整。除上界之外，我们还表明，在单条轨迹上进行学习会带来固有挑战：尽管依赖模式和臂参数构成一个秩一矩阵，循环矩阵在秩一流形上并不满足等距性，而且样本复杂度确实会受益于稀疏的奖励依赖结构。我们的结果需要新的分析方法，以处理数据中的长程时间依赖关系，并避免对奖励时域 h 的多项式依赖。具体而言，我们利用了与由相依次高斯向量构成的循环矩阵之受限等距性质之间的联系，并建立了具有独立研究价值的新保证。

1 引言

多臂赌博机（MAB）是研究序贯决策中探索与利用权衡的典型模型（例如，参见 Bubeck 等人 [2012]）。智能体需要通过与未知环境交互，反复做出决策，以最大化累积奖励。作为多臂赌博机的推广，上下文赌博机允许智能体根据上下文信息采取动作 Chu 等人 [2007]。鉴于其在临床试验、推荐和广告等领域的广泛应用，针对上下文赌博机已开展了大量研究（例如，参见 Woodroofe [1979], Chu 等人 [2011], Li 等人 [2017, 2010], Qin 等人 [2022b, a]）。

大多数关于上下文赌博机的现有研究都假设，每个奖励仅取决于单个动作及其对应的上下文。这个动作可以是刚刚采取的动作（即时奖励），也可以是若干步之前采取的动作（延迟奖励）。然而，在现实的决策场景中，奖励生成过程可能具有更复杂的非马尔可夫特性。多个先前动作可能共同影响当前奖励。例如，学习者是否会选修在线教育平台推荐的课程，不仅取决于该课程本身，也取决于他们此前修读过的课程组合。要向背景和过往经历各异的用户推荐复杂课程体系中的课程，就需要考虑过去上下文对当前推荐的综合影响。同样，注意力

*加利福尼亚大学河滨分校， †华盛顿大学， ‡密歇根大学。电子邮箱： {yuzhenq, yli692}@ucr.edu, fabiopas@engr.ucr.edu, mfazel@uw.edu, oymak@umich.edu。本文将发表于 AAAI2023。

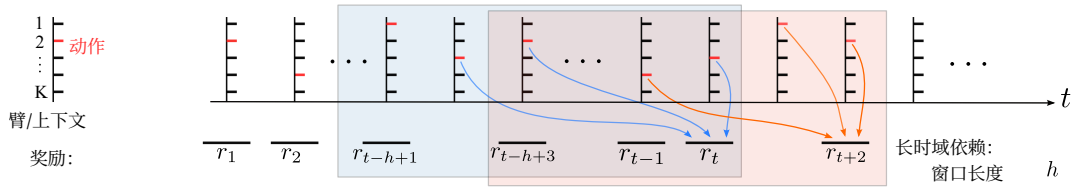

图 1：长时域奖励上下文老虎机。每一轮的奖励 t 取决于与最近 h 个动作相关联的上下文（h 可能非常大）。

机制 [Vaswani 等人， 2017] 在强化学习和自然语言处理应用中日益取得成功 [Chen 等人，2021年，Brown 等人， 2020] 并通过评估当前上下文与过去上下文之间的相似性（例如，句子中的词语或电子游戏中的帧所对应的上下文）来进行预测，进而构建一种基于历史加权的自适应上下文。与此相关的是，利用长上下文历史的益处已在强化学习和控制理论中得到广泛认可（例如，帧/状态堆叠实践 Hessel 等人 [2018]）。这些观察引出了以下核心问题：

问： 我们能否对从长时域奖励中进行学习给出可证明且高效的方法？奖励依赖结构对样本效率有何影响?

在这项工作中，我们针对一种新型的随机线性上下文老虎机问题变体，全面回答了这些问题。在该问题中，当前奖励取决于先前上下文的一个子集，且依赖范围最长可达 h （示意见图 1 ）。具体而言，奖励由一个 滤波上下文 决定，该上下文是先前 h 个选定上下文的线性组合。此外，受实际决策场景的启发，我们还考虑了稀疏交互：只有 s （$s \ll h$）个 h先前上下文实际参与决定当前奖励。这里， $s = 1$ 对应于延迟奖励这一特殊情形。至关重要的是，我们提出了利用奖励函数稀疏依赖结构的策略，并为长时域奖励建立了遗憾保证。

1.1 相关工作

1.1.1 复合匿名奖励。

Pike-Burke 等人 [2018] 研究了复合匿名奖励老虎机，其中：1）智能体在每一轮收到的奖励，是某个未知子集的过去动作所产生的延迟奖励之和；2）过去动作对奖励的各自贡献无法辨别。Cesa-Bianchi 等人 [2018] 将这一设定推广到如下情形：某个动作产生的奖励并非仅在未来某一时刻揭示给智能体，而是分散在多个轮次中。该方向的近期工作还包括Garg 和 Akash [2019], Zhang 等人 [2022],以及 Wang 等人 [2021]。本文考虑一种带有.....的上下文设定，这与上述工作不同，并带来了新的挑战，因为每个臂不再具有固定的奖励分布。

1.1.2 延迟奖励。

延迟奖励老虎机也与我们的工作相关。关于具有随机延迟奖励的随机线性老虎机，已有研究 Vernade 等人 [2020]。 Li 等人 [2019]研究了延迟未知的情形。关于具有随机延迟的广义随机线性老虎机，已有研究Zhou 等人 [2019]和 Howson 等人 [2022]。 Cella 和 Cesa-Bianchi [2020] 以及 Lancewicki 等人 [2021]研究了具有奖励依赖延迟的老虎机问题。在 Lancewicki 等人 [2021] 和 Thune 等人 [2019], 中，延迟可以不受限制。随机老虎机中的臂依赖延迟已有研究

见 Gael 等人 [2020]。对抗多臂老虎机中也考虑了延迟 [Bistritz 等人，2019年，Gyorgy 和 Joulani，2021年，Zimmert 和 Seldin， 2020]。近期，也有研究探讨了具有延迟的非随机协作线性多臂老虎机 [Ito 等人，2020年，Cesa-Bianchi 等人， 2019]。事实上，我们的设定涵盖了未知固定延迟，以及聚合且匿名的延迟奖励。

1.1.3 稀疏参数。

在稀疏多臂老虎机中，特征向量的维度可以很大，为 d，但只有其中一个很小的子集 $s \ll d$ 会影响奖励。关于稀疏线性多臂老虎机的早期研究见 Carpentier 和 Munos[2012] 以及 Abbasi-Yadkori 等人 [2012]。近期研究根据总时域 T 是否小于或大于 d，分别考察了数据匮乏和数据丰富情形。在数据丰富情形下，Lattimore 和 Szepesv´ari [2020] 证明了遗憾下界 $\Omega(\sqrt{sdT})$。在数据匮乏情形下，Hao 等人 [2020] 给出了遗憾下界 $\Omega(s^{\frac{1}{3}}T^{\frac{2}{3}})$。近期的一项工作采用了信息导向采样技术 Hao 等人 [2021]。稀疏上下文线性多臂老虎机也日益受到关注。Kim 和 Paik [2019] 提出了一种结合 Lasso 与双重稳健技术的算法，并给出了上界 $O(s\log(dT)\sqrt{T})$。Bastani 和 Bayati [2020], 以及 Wang 等人 [2018], 研究了扩展设定，其中每个臂都有各自的参数，并分别证明了上界 $O(s^2\log^2(T))$ 和 $O(s^2\log(T))$。Oh 等人 [2021] 提出了一种免探索算法，并得到了上界 $O(s^2\log(d)+s\sqrt{T}\log(dT))$。Ariu 等人 [2022], 提出了一种阈值化 Lasso 算法，得到上界 $O(s^2\log(d)+\sqrt{sT})$。Ren 和 Zhou [2020], 采用动态批量学习方法，得到了上界$O(s\cdot \text{polylog}(d)+\log(T)\sqrt{sT\log(d)})$。相比之下，在我们的情形中，稀疏性源于奖励依赖结构。正如我们将在第 2.2 节讨论的，学习依赖模式颇具挑战，因为测量值具有固有的循环结构。

1.1.4 在线凸优化（OCO）。

另一条与我们相关的研究方向是带记忆的在线凸优化，其中损失取决于从凸集中选取的先前决策 [Anava 等人，2015，Shi 等人，2020年，Kumar 等人， 2022]。

1.2 贡献

本文的贡献总结如下：

1. 我们提出了一种新的上下文老虎机模型，该模型受到现实场景的启发：奖励对先前动作和上下文具有长程且稀疏的依赖。识别奖励参数和稀疏延迟模式的问题具有特殊的低秩稀疏结构。

2. 我们利用稀疏先验，针对数据匮乏和数据丰富情形提出了两种样本高效算法。对于前者，我们证明了遗憾上界 $O\big(d\sqrt{sT}\log(dT)+\min\{q,T\}\big)$，该上界能够适应由 q 描述的奖励依赖模式；对于后者，我们得到遗憾上界 $O(\sqrt{sdT}\log(dT))$。请注意，这两个上界都不以多项式形式依赖于时域 h，因而能够在长时域内实现高效学习；并且二者在 T 意义下均为最优（忽略对数因子）。

3. 我们针对具有块托普利茨/循环矩阵形式的数据中的时间依赖关系，作出了技术贡献。首先，Krahmer 等人 [2014] 关于循环矩阵的受限等距性质 (RIP) 的开创性研究假设上下文向量的元素独立同分布。我们将他们的结果推广到较弱的集中性条件，从而允许存在依赖关系。其次，我们确立了一些结果，揭示了循环测量中特有的低秩估计难题。与理论相符，数值实验表明，我们基于稀疏性的方法确实优于低秩方法。

2 问题设定

记号。 给定 $\boldsymbol{x} = [\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_h^\top]^\top \in \mathbb{R}^{dh}$，其中每个块为 $\boldsymbol{x}_i \in \mathbb{R}^d$，记作 $\|\boldsymbol{x}\|_{2,1}^{(d)} := \sum_{i=1}^h \|\boldsymbol{x}_i\|_2$；对于 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$，$\|\boldsymbol{A}\|_{\mathrm{op}} := \sup_{\|\boldsymbol{x}\|_2\le 1} \|\boldsymbol{A}\boldsymbol{x}\|_2$ 表示其算子范数。令 $[n] = \{1, 2, \ldots, n\}$，其中 n 为任意整数。对于任意 $\mathcal{S} \subset [n]$，$\boldsymbol{x}_{\mathcal{S}}$ 表示 $\boldsymbol{x}$ 中由$\mathcal{S}$索引的子向量。令 $\langle \cdot, \cdot \rangle$ 为内积；对于 $\boldsymbol{A}$ 和 $\boldsymbol{B} \in \mathbb{R}^{m\times n}$，$\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \mathrm{Tr}(\boldsymbol{A}^\top \boldsymbol{B})$。令 $\otimes$ 为克罗内克积。给定 $\boldsymbol{A} \in \mathbb{R}^{m\times p}$，若存在 $\delta \in \in(0,1)$，使得对所有 $\boldsymbol{x} \in \mathbb{R}^p$ 均满足 $(1-\delta)\|\boldsymbol{x}\|_2^2 \le \|\boldsymbol{A}\boldsymbol{x}\|_2^2 \le (1+\delta)\|\boldsymbol{x}\|_2^2$，则称其满足 RIP；满足该不等式的最小 δ 称为限制等距常数（更多详情见附录）。

2.1 随机线性上下文老虎机

本文研究一种随机线性上下文老虎机问题，其奖励取决于过去的动作和上下文（示意见图 1 ）。设 K 为臂数，则动作集为 $[K]$。在每一轮 t，智能体观测到 K 个上下文向量，$\{\boldsymbol{x}_{t,a} \in \mathbb{R}^d : a \in [K]\}$，每个向量对应一个臂，并从未知分布 ν中独立同分布地抽取。随后，智能体选择一个动作 $a_t \in [K]$ ，并获得按以下方式生成的奖励：

$$r_t = \langle \boldsymbol{y}_{t,a_t}, \boldsymbol{\theta} \rangle + \varepsilon_t, \tag{1}$$

其中

$$\boldsymbol{y}_{t,a_t} = \sum_{i=0}^{h-1} w_i \boldsymbol{x}_{t-i,a_{t-i}}, \text{ and } t \in [T].$$

这里，$\boldsymbol{\theta} \in \mathbb{R}^d$ 是系数向量，$\varepsilon_t \in \mathbb{R}$ 是零均值 1-次高斯加性噪声。特别地，向量 $\boldsymbol{y}_{t,\ a_t}$是滤波上下文，由权重向量 $\boldsymbol{w} := [w_0, w_1, \ldots, w_{h-1}]^\top \in \mathbb{R}^h$ 确定，该权重向量描述奖励如何依赖过去和当前选定的上下文（其中 $w_i \ge 0$）。依赖范围 h 可能非常大，这表明奖励可能具有长程上下文依赖。假设 $\boldsymbol{x}_{j,\ a_j} = 0$ 对于 $j \le 0$ ，因为此情形下的 a_j 对应于不存在的动作。

在本文中，我们考虑稀疏上下文依赖，即权重向量 $\boldsymbol{w}$ 是 s-稀疏的（即 $\|\boldsymbol{w}\|_0 \le s$），且满足 $s \ll h$。这与许多现实情形尤为相关，因为通常只有少数过去的“事件”会产生影响。正如前文所述，这种设定涵盖了：a) 延迟未知的老虎机（$\boldsymbol{w}$ 只有一个非零项，且其值为 1）；以及 b) 具有聚合且匿名的奖励的老虎机（ $\boldsymbol{w}$ 的所有非零项的值均为 1）。

系数向量 $\boldsymbol{\theta} \in \mathbb{R}^d$ 和权重向量 $\boldsymbol{w} \in \mathbb{R}^h$ 均未知。不失一般性，我们假设 $\|\boldsymbol{\theta}\|_2 \le 1$ 和 $\|\boldsymbol{w}\|_1 \le 1$。我们还作出了温和的有界性假设，即 $\boldsymbol{x}_{t,\ a}$ 满足 $\|\boldsymbol{x}_{t,\ a}\|_\infty \le 1$ ，对所有 $t \in [T]$ 和 $a \in [K]$ 均成立。

智能体的目标是在 T 轮次内最大化累积奖励，或者等价地，最小化如下定义的伪遗憾：

$$R_T = \left[\sum_{t=1}^{T} \sum_{i=0}^{h-1} w_i \left(\langle \boldsymbol{x}_{t,a_t^*}, \boldsymbol{\theta} \rangle - \langle \boldsymbol{x}_{t,a_t}, \boldsymbol{\theta} \rangle \right) \right], \tag{2}$$

其中，$a_t^* = \arg\max_{a\in[K]} \langle \boldsymbol{x}_{t,a}, \boldsymbol{\theta} \rangle$ 定义了第 t轮次的最优动作。

注 1. 此处遗憾的定义与简单地将 $\langle \boldsymbol{y}_{t,a_t^*} - \boldsymbol{y}_{t,\ a_t}, \boldsymbol{\theta} \rangle$相加略有不同。事实上，这两种遗憾定义本质上是相同的。原因在于，采取一个动作，例如 a_t，会给智能体带来总奖励 $\sum_{i=0}^{h-1} w_i \langle \boldsymbol{x}_{t,a_t}, \boldsymbol{\theta} \rangle$ ，该奖励会分散在接下来的 h 轮次中。

因此，智能体可以在不知道 $\boldsymbol{w}$ 的情况下 $\boldsymbol{\theta}$ 若将其视为先验已知：每轮采用旨在最大化即时奖励的贪心策略，便能从长远来看最大化累积奖励。关于这一点的进一步讨论见附录 F.1。 △

注 2。 尽管仅了解 $\boldsymbol{\theta}$ 似乎足以满足我们的决策需求，但学习 $\boldsymbol{\theta}$ 实际上颇具挑战。这是因为每个奖励都可能以复合形式出现，可能包含最近 h 个动作所产生的贡献。学习 $\boldsymbol{\theta}$ 需要厘清奖励依赖结构。△

2.2 挑战讨论

接下来，我们讨论这个问题中固有的一些技术挑战。为简洁起见，我们将$\boldsymbol{\xi}_t = \boldsymbol{x}_{t,\ a_t}$ 记作每个 t 的所选上下文。

2.2.1 低秩矩阵恢复中的循环设计矩阵。

设 $\boldsymbol{Z}_t = [\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}, \ldots, \boldsymbol{\xi}_{t-h+1}] \in \mathbb{R}^{d\times h}$，则（1）可改写为

$$r_t = \langle \boldsymbol{Z}_t, \boldsymbol{\theta}\boldsymbol{w}^\top \rangle + \varepsilon_t. \tag{3}$$

乍看之下，该问题似乎可归结为重构秩一稀疏矩阵$\boldsymbol{\theta}\boldsymbol{w}^\top \in \mathbb{R}^{d\times h}$，并可应用经典的低秩（及稀疏）矩阵恢复技术（例如，Richard 等人 [2012], Oymak 等人 [2015] Davenport 和 Romberg [2016], 以及 Wainwright [2019,第 10]章）。然而，我们发现事实并非如此，这是由于设计矩阵 Z_t具有托普利茨/循环结构。下文的引理表明，即使循环矩阵的首行由独立同分布元素构成，它们也无法以指数级高的概率满足秩一矩阵的限制等距性质（更多详情见附录C）。

引理 1. 设$\mathbf{C} \in \mathbb{R}^{n\times p}$ 为一个子采样循环矩阵，其第一行包含独立同分布的高斯元素（经过适当归一化），且 $n \le p$。对于任意 $\delta \le 1$，存在一个常数 $c < 1$ ，使得以至少 $1 - c^p$的概率，$\mathbf{C}$ 不满足 $\mathbb{R}^{p_1\times p_2}$ 上的秩一矩阵的限制等距性质，且 $p = p_1 p_2$。

我们还在图 2 中提供了数值实验，说明在处理低秩矩阵恢复时，循环测量确实会带来问题。

这些发现表明，通过低秩矩阵估计来解决我们的问题可能行不通。因此，我们转而采用另一种技术——稀疏性估计——利用奖励依赖模式中的稀疏结构。首先，令 $\boldsymbol{a}_t = \{a_t, a_{t-1}, \ldots, a_1\}$ 表示过去动作序列， $\boldsymbol{z}_{\boldsymbol{a}_t} = [\boldsymbol{\xi}_t^\top, \boldsymbol{\xi}_{t-1}^\top, \ldots, \boldsymbol{\xi}_{t-h+1}^\top]^\top$，且 $\boldsymbol{\phi} = \boldsymbol{w} \otimes \boldsymbol{\theta} \in \mathbb{R}^{dh}$。然后，式（1）可改写为

$$r_t = \langle \boldsymbol{z}_{\boldsymbol{a}_t}, \boldsymbol{\phi} \rangle + \varepsilon_t. \tag{4}$$

由于 $\boldsymbol{w}$ 是 s-稀疏的，重构 $\boldsymbol{\theta}$ 和 $\boldsymbol{w}$ 等价于估计块稀疏向量 $\boldsymbol{\phi}$.

记 $\boldsymbol{r}_t = [r_1, \ldots, r_t]^\top$ 和 $\boldsymbol{\varepsilon}_t = [\varepsilon_1, \ldots, \varepsilon_t]^\top$，由（4）可得

$$\boldsymbol{r}_t = \begin{bmatrix} \boldsymbol{\xi}_1^\top & 0 & \cdots & 0 \\ \boldsymbol{\xi}_2^\top & \boldsymbol{\xi}_1^\top & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_h^\top & \boldsymbol{\xi}_{h-1}^\top & \cdots & \boldsymbol{\xi}_1^\top \\ \boldsymbol{\xi}_{h+1}^\top & \boldsymbol{\xi}_h^\top & \cdots & \boldsymbol{\xi}_2^\top \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_t^\top & \boldsymbol{\xi}_{t-1}^\top & \cdots & \boldsymbol{\xi}_{t-h+1}^\top \end{bmatrix} \boldsymbol{\phi} + \boldsymbol{\varepsilon}_t := \Xi_t \boldsymbol{\phi} + \boldsymbol{\varepsilon}_t. \tag{5}$$

可以看出，上述设计矩阵 Ξ_t 也具有*Toeplitz*/循环 结构。利用这种特殊形式的设计矩阵学习块稀疏 $\boldsymbol{\phi}$ 还面临其他挑战，我们将在下文讨论。

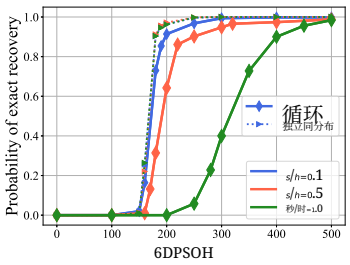

图2：从无噪声测量 $\boldsymbol{y} = \langle \boldsymbol{Z}, \Phi \rangle$ 中通过低秩恢复恢复矩阵 $\Phi = \boldsymbol{\theta}\boldsymbol{w}^\top \in \mathbb{R}^{d\times h}$ 的精确恢复概率（$d = 10$ 和 $h = 100$）。我们比较两种类型的设计矩阵：1）具有独立同分布元素的 $\boldsymbol{Z}$；2）由独立同分布向量生成的循环 $\boldsymbol{Z}$。我们考虑了不同的情形，其中 $\boldsymbol{w}$ 具有不同的稀疏性（由 s量化）。实验表明：1）如果 $\boldsymbol{Z}$ 具有独立同分布元素，所有情形所需的数据量相同；2）而对于循环 $\boldsymbol{Z}$，所需样本数则有显著差异。这些观察结果表明，对于循环测量，恢复低秩矩阵所需的数据量可能并非仅取决于秩，这与独立同分布测量有很大不同。

2.2.2 具有相关元素的循环矩阵。

估计 $\boldsymbol{\phi}$ 是一个稀疏回归问题。RIP及相关的限制特征值条件（REC）广泛用于此类问题 [Candes和Tao，2007年，Bickel等人，2009]。早期研究表明，如果至少有 s-稀疏向量，并且样本数不少于 $\tilde{\Omega}(s\log^2(s))$，那么首行为独立同分布次高斯的子采样循环矩阵满足 RIP [Krahmer 等人，2014]。在我们的情形中，循环矩阵由随机向量生成，其 相关 元素（即每个 $\boldsymbol{\xi}_t$ 中的元素可能相关）。新的挑战是：这类循环测量需要多少样本才能满足 RIP/REC?

2.3 技术结果

我们首先给出一个关于 RIP 的技术结果，为分析我们的老虎机问题铺平道路；该结果本身也具有独立的研究价值（证明见附录第一部分）。

定理 1. 设 $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n \in \mathbb{R}^d$ 为独立的次高斯各向同性随机向量。假设每个 $\boldsymbol{\xi}_i$ 都满足 Hanson-Wright不等式（HWI）

$$\Pr\left[|\boldsymbol{\xi}_i^\top \boldsymbol{A}\boldsymbol{\xi}_i - \mathbb{E}(\boldsymbol{\xi}_i^\top \boldsymbol{A}\boldsymbol{\xi}_i)| \geq \eta\right] \leq 2\exp\left(-\frac{1}{c}\min\left\{\frac{\eta^2}{k^4\|\boldsymbol{A}\|_F^2}, \frac{\eta}{k^2\|\boldsymbol{A}\|_{\text{op}}}\right\}\right), \qquad \forall \eta > 0, \tag{6}$$

对于任意半正定矩阵 $\boldsymbol{A} \in \mathbb{R}^{d\times d}$，其中 k 是一个常数，且 c 是一个绝对常数。设 $\boldsymbol{\Xi} \in \mathbb{R}^{m\times nd}$ 是通过对子采样分块循环矩阵中的 任意 m 行构成的矩阵：

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{\xi}_n^\top & \boldsymbol{\xi}_{n-1}^\top & \cdots & \boldsymbol{\xi}_1^\top \\ \boldsymbol{\xi}_1^\top & \boldsymbol{\xi}_n^\top & \cdots & \boldsymbol{\xi}_2^\top \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_{n-1}^\top & \boldsymbol{\xi}_{n-2}^\top & \cdots & \boldsymbol{\xi}_n^\top \end{bmatrix}.$$

那么，对于所有 s-稀疏向量，$\boldsymbol{\Xi}$ 的受限等距常数（记为 δ_s）满足 $\delta_s \leq \delta$if $m \geq c_1\delta^{-2}s\log^2(s)\log^2(nd)$，其中 c_1 为某个常数。

备注 3。 尽管在老虎机问题中我们只需重构一个块稀疏向量，但该定理适用于一般的稀疏向量。假设（6）对许多随机向量都成立，例如具有独立元素的次高斯向量，以及满足凸集中性质的随机向量Adamczak [2015]。

算法 1 Doubling Lasso

1：输入：参数 L，**倍增序列** $\{T_i\}$，**其中** $T_i = 4(2^i - 1)L$ **[参见图 3 (a)]**, $\hat{\boldsymbol{\theta}}_0 = 0$。**2：对于** $t = 1$ **：** T **执行 3：观察上下文向量** $\{\boldsymbol{x}_{t,a}$ **：** $a \in [K]\}$。$\hat{}$ **4：采取贪心动作** $a_t \in \sup_{a\in[K]} \langle \boldsymbol{x}_{t,a}, \hat{\boldsymbol{\theta}}_{i-1} \rangle$，**并获得奖励** r_t。**5：如果** $t = T_i$ **则** $\hat{}$ **6：#** *end of the ith epoch*， *estimate a new* $\hat{\boldsymbol{\theta}}$ **7：根据 Lasso (11) 计算** $\phi_{2^{i-1}L}$。$\hat{}$ $\hat{}$ **8：令** $\boldsymbol{\theta}_i$ **为与最大奇异值对应的** $\Phi_{\mathcal{S}_i}$ **的奇异向量。9：结束 if 10：结束 for**

3 算法与主要结果

接下来，我们将介绍针对（1）中所述老虎机问题的一些算法，分别考虑数据贫乏和数据丰富情形，并给出相应的遗憾界。首先，我们作出以下假设。

假设 1. 我们假设分布 ν 满足以下条件：对所有 t而言：(a) 上下文向量$\boldsymbol{x}_{t,a}, a \in [K]$均独立同分布；(b) $\mathbb{E}_{\boldsymbol{x}} = \mathbb{E}[\frac{1}{K}\sum_{a=1}^{K} \boldsymbol{x}_{t,a}\boldsymbol{x}_{t,a}^\top]$ 满足 $\lambda_{\min}(E_{\boldsymbol{x}}) \geq \sigma^2$ 对于某个 σ成立；且 (c)$\boldsymbol{x}_{t,a}$ 满足由（6）给出的 HWI。

备注 4. 若令 ξ_t 为从集合 $\{x_{a,t}, a \in [K]\}$中确定性选取的一个随机向量，则假设 1 保证 ξ_t 也满足 Hanson-Wright 不等式（证明见附录D ）。我们将结合定理 1 利用这一性质分析我们的老虎机算法。 △

3.1 数据匮乏情形

首先，我们考虑权重向量 $\boldsymbol{w}$ 的维度大于轮次数量（即 $T < h$）的情况。在这种数据匮乏情形下，从（5）可以看出，不可能重构 $\boldsymbol{\phi}$。幸运的是，无须完全学习 $\boldsymbol{\phi}$ 来指导决策；只需对 $\boldsymbol{\theta}$ 作出良好估计即可（原因见（2）中的遗憾定义）。因此，我们提出以下方法，部分地并逐步地学习 $\boldsymbol{\phi}$ ，从而使$\boldsymbol{\theta}$ 能够在早期阶段得到估计并加以利用。

方法 1。 回忆一下， $\boldsymbol{\phi} = \boldsymbol{w} \otimes \boldsymbol{\theta}$ 满足 $\boldsymbol{w} \in \mathbb{R}^h$。对于任意整数 $k \leq h$，令 $\boldsymbol{w}_{[k]} = [w_1, \ldots, w_k]^\top$和 $\boldsymbol{\phi}_{\mathcal{K}} = \boldsymbol{w}_{[k]} \otimes \boldsymbol{\theta} \in \mathbb{R}^{kd}$。那么，如果已学习到 $\hat{}$ $\boldsymbol{\phi}_{\mathcal{K}}$ 作为 $\phi_{\mathcal{K}}$的估计， $\boldsymbol{\theta}$ 可以通过以下步骤进行估计：

(a1) 将向量 $\hat{}$ $\boldsymbol{\phi}_{\mathcal{K}}$ 转换为矩阵 $\hat{}\boldsymbol{\Phi}_{\mathcal{K}} \in \mathbb{R}^{d\times k}$ （$\hat{}$ $\Phi_{\mathcal{K}}$ 的第 i 列是 $\hat{}\boldsymbol{\phi}_{\mathcal{K}}$ 中大小为 d 的第 i 个块。

(a2) 令 $\hat{}\boldsymbol{\theta}$ 为与最大奇异值 [1]对应的 $\hat{}$ $\Phi_{\mathcal{K}}$ 的左奇异向量。 △

采用这种方法，我们使用倍增技巧来设计算法（见算法1 和图3）。我们选取一个常数 L 使其满足 $s \leq L < h$，并定义集合序列 $\mathcal{S}_1$， $\mathcal{S}_2$，……， $\mathcal{S}_m$其中元素数量逐渐增加，且 $\mathcal{S}_i = [2^{i-1}Ld]$。接着，我们旨在估计

$$\boldsymbol{\phi}_{\mathcal{S}_1}, \boldsymbol{\phi}_{\mathcal{S}_2}, \boldsymbol{\phi}_{\mathcal{S}_3}, \ldots, \boldsymbol{\phi}_{\mathcal{S}_m},$$

[1]我们指出，这一步估计的并非恰好是 $\boldsymbol{\theta}$，而是它的方向 $\boldsymbol{\theta}/\|\boldsymbol{\theta}\|$。正如后文将会看到的，对方向的良好估计可以确保 $\boldsymbol{\theta}$ 和 $\boldsymbol{\theta}$之间的角度很小，因此足以实现良好的决策。

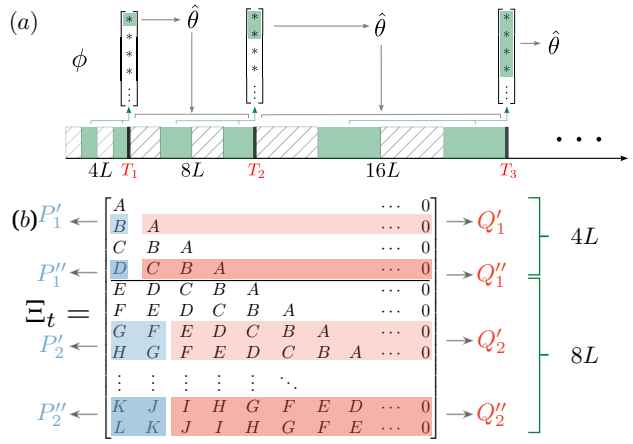

图3：Doubling Lasso 示意图。(a) 在每个轮次 i中，我们执行式（7）中的贪心动作，持续 $2^{i+1}L$ 轮（T_i表示各轮次的结束时刻）。然后，我们使用收集到的数据的第二和第四个四分之一部分（绿色区域）来学习 $2^{i-1}L$ 的第一个块 $\boldsymbol{\phi}$ ，随后估计 $\boldsymbol{\theta}$ （见ˆ方法1）。学习得到的 $\boldsymbol{\theta}$ 用于下一轮次的决策，该轮次的长度是当前轮次的两倍。(b) 式（9）中矩阵的定义示意图。这里，式（5）中定义的矩阵 $\boldsymbol{\Xi}_t$ 以$L \times Ld$维分块矩阵的形式表示。

在连续轮次中，其中 $\boldsymbol{\phi}_{\mathcal{S}_i} \in \mathbb{R}^{2^{i-1}Ld}$ 包含 $\boldsymbol{\phi}$ 的前 $2^{i-1}L$ 个块， m 是满足 $2^{m-1}L \leq h$的最大整数。核心思想是在数据较少时学习 $\boldsymbol{\phi}$ 的一小部分；随着收集到更多数据，我们逐步学习更大的一部分。

在每个轮次 i中，重复执行以下贪心动作 $2^{i+1}L$ 次（倍增技巧，见图3（a））：

$$a_t \in \arg\max_{a\in[K]} \langle \boldsymbol{x}_{t,a}, \hat{\boldsymbol{\theta}}_{i-1} \rangle, \qquad i \geq 1, \tag{7}$$

ˆˆ其中 $\boldsymbol{\theta}_{i-1}$ 是 $\boldsymbol{\theta}$ 在第（$i-1$）轮次的估计值（$\boldsymbol{\theta}_0 = 0$）。如果存在多个贪心动作，智能体会从中均匀随机地选择一个。然后，我们收集 $2^{i+1}L$ 个由（5）生成的数据点ˆ。然而，我们只用其中一半来学习 $\boldsymbol{\theta}_i$。具体而言，将数据划分为四个$2^{i-1}L$维块后，我们使用第二块和第四块（见图3（a））。根据（5），这两个块中的奖励分别由以下方式生成

$$\boldsymbol{r}'[i] = \boldsymbol{\Xi}'[i]\boldsymbol{\phi} + \boldsymbol{\varepsilon}'[i], \quad \boldsymbol{r}''[i] = \boldsymbol{\Xi}''[i]\boldsymbol{\phi} + \boldsymbol{\varepsilon}''[i], \tag{8}$$

其中 $\boldsymbol{r}'[i],\ \boldsymbol{r}''[i] \in \mathbb{R}^{2^{i-1}L}$， $\boldsymbol{\Xi}'[i],\ \boldsymbol{\Xi}''[i] \in \mathbb{R}^{2^{i-1}L\times dh}$ 和 $\boldsymbol{\varepsilon}'[i],\ \boldsymbol{\varepsilon}''[i] \in \mathbb{R}^{2^{i-1}L}$ 分别是相应的奖励向量、上下文矩阵和噪声向量。

将 $\boldsymbol{\phi} = [\boldsymbol{\phi}_{\mathcal{S}_i}^\top, \boldsymbol{\phi}_{\bar{\mathcal{S}}_i}^\top]^\top$ 重写为，其中 $\boldsymbol{\phi}_{\mathcal{S}_i} \in \mathbb{R}^{2^{i-1}Ld}$ 是我们要学习的内容。接下来，可以将（8）改写为

$$\begin{bmatrix} \boldsymbol{r}'[i] \\ \boldsymbol{r}''[i] \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_i' & \boldsymbol{Q}_i' \\ \boldsymbol{P}_i'' & \boldsymbol{Q}_i'' \end{bmatrix} \begin{bmatrix} \boldsymbol{\phi}_{\mathcal{S}_i} \\ \boldsymbol{\phi}_{\bar{\mathcal{S}}_i} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}'[i] \\ \boldsymbol{\varepsilon}''[i] \end{bmatrix}, \tag{9}$$

其中 $\boldsymbol{r}'[i] \in \mathbb{R}^{2^{i-1}L}$，且 $\boldsymbol{\Xi}'[i] = [\boldsymbol{P}_i', \boldsymbol{Q}_i']$ 和 $\boldsymbol{\Xi}''[i] = [\boldsymbol{P}_i'', \boldsymbol{Q}_i'']$（参见图 3（b）中的示意图）。¯ ¯为了学习 $\boldsymbol{\phi}_{\mathcal{S}_i}$，令 $\bar{\boldsymbol{r}}[i] = \boldsymbol{r}''[i] - \boldsymbol{r}'[i]$， $\boldsymbol{P}_i = \boldsymbol{P}_i'' - \boldsymbol{P}_i'$， $\boldsymbol{Q}_i = \boldsymbol{Q}_i'' - \boldsymbol{Q}_i'$，且 $\bar{\boldsymbol{\varepsilon}}[i] = \boldsymbol{\varepsilon}''[i] - \boldsymbol{\varepsilon}'[i]$，于是得到

$$\bar{\boldsymbol{r}}[i] = \bar{\boldsymbol{P}}_i\boldsymbol{\phi}_{\mathcal{S}_i} + \bar{\boldsymbol{Q}}_i\boldsymbol{\phi}_{\bar{\mathcal{S}}_i} + \bar{\boldsymbol{\varepsilon}}[i] := \bar{\boldsymbol{P}}_i\boldsymbol{\phi}_{\mathcal{S}_i} + \boldsymbol{\epsilon}[i], \tag{10}$$

¯其中，将 $\boldsymbol{Q}_i\boldsymbol{\phi}_{\bar{\mathcal{S}}_i} + \bar{\boldsymbol{\varepsilon}}[i]$ 视为新的噪声 $\boldsymbol{\epsilon}[i]$。

然后，通过求解块稀疏恢复 Lasso 来估计 $\boldsymbol{\phi}_{\mathcal{S}_i}$（由于 $\boldsymbol{\phi}$ 是 s，因此它至多是块稀疏的）：

$$\hat{\boldsymbol{\phi}}_{\mathcal{S}_i} = \underset{\tilde{\boldsymbol{\phi}}\in\mathbb{R}^{2^{i-1}Ld}}{\arg\min} \left(\frac{1}{2^i L} \left\| \bar{\boldsymbol{P}}_i \tilde{\boldsymbol{\phi}} - \bar{\boldsymbol{r}}[i] \right\|_2^2 + \lambda_i \|\tilde{\boldsymbol{\phi}}\|_{2,1}^{(d)} \right), \tag{11}$$

其中正则化参数选取为

$$\lambda_i = cd\sqrt{\frac{2\log(2^i dL/\gamma)}{2^{i-1}L}} \tag{12}$$

对于某个 $c > 0$。随后，我们采用方法 1 来估计 $\boldsymbol{\theta}$。算法见算法1。以下定理给出了该算法的遗憾上界。

定理 2。 考虑式（1）中描述的具有长时域奖励的随机上下文线性老虎机模型。在算法 1中，取 $L = csd\log^2(sd)\log^2(hd)$，其中 $c > 0$ 为常数。当 $T < h$时，遗憾满足

$$R_T = O\left(d\sqrt{sT\log(dT)} + \min\{q(\boldsymbol{w}), T\} \right), \tag{13}$$

其中， $q(\boldsymbol{w})$ 是权重向量 $\boldsymbol{w}$ 的函数，用于描述权重在 $\boldsymbol{w}$ 中的分布情况。具体而言， $q(\boldsymbol{w}) := h^{\alpha(\mu)}$，其中 $\mu \in (0,1]$，且 $\alpha(\mu) = \inf_{\alpha\in[0,1]} \|\boldsymbol{w}_{q(\boldsymbol{w})}\|_2 \geq \mu$，满足 $\boldsymbol{w}_{q(\boldsymbol{w})} := \{w_1, w_2, \ldots, w_{\lceil q(\boldsymbol{w})\rceil}\}$ 和 $1/\mu = \Theta(1)$。

备注 5。请注意，我们的算法中以下两个事实对分析至关重要：1）我们使用 $\boldsymbol{P}_i''$ 与 $\boldsymbol{P}_i'$ 之间的差值（即 (10) 中的 $\boldsymbol{P}_i$）作为测量矩阵来学习 $\boldsymbol{\phi}_{\mathcal{S}_i}$，确保 $\boldsymbol{P}_i$ 的均值为零；2）倍增技巧以及对所用数据的选择确保 $\boldsymbol{P}_i''$ 和 $\boldsymbol{P}_i'$ 不重叠且相互独立，并且不同轮次中的 P_i 也互不重叠且相互独立（见图 3 (b)）。我们的分析利用定理 1 证明 (11) 中的每个 $\boldsymbol{P}_i$ 都满足块稀疏向量的限制特征值条件（见附录中的定理 E.1）。然后，我们推导出定理 F.1，它推广了 Wainwright [2019] 中的定理 7.13，从而完成证明。△

注 6。 α 的值描述了权重在 s-稀疏向量$\boldsymbol{w}$中的一种“质量状”分布。较小的 α 意味着非零项出现在 $\boldsymbol{w}$的较早位置，因此相比 α较大的情况，更容易在早期从 $\boldsymbol{\theta}$ 中学习到有用信息。例如，如果 $\alpha \leq \frac{1}{2}\log_h T$（即一半的“质量”位于 $\boldsymbol{w}$最前面的 $\sqrt{T}$ 个位置），那么 $R_h = \tilde{O}(d\sqrt{sT})$。如果 $\alpha = 1$，即$\boldsymbol{w} = [0, 0, \ldots, 1]^\top$，那么 $R_h = O(T)$；这很直观，因为直到时域的最后时刻才能收集到有助于决策的信息。上界 (13) 表明，我们的算法能够适应不同的实例。我们推测，对 q 的依赖是最优的；例如，对于延迟老虎机， q 就变成了延迟，这是不可避免的。

实验。 在图 4中，我们通过考虑两种不同的 $\boldsymbol{w}$，即一种将“质量”分布在较早的位置，另一种将其分布在较晚的位置，来进行一些实验。正如我们的理论所预测的那样，在前一种情况下，我们的算法确实获得了更低的遗憾（见图4 (b)）。

注 7。 除去这一项 $q(\boldsymbol{w}) = h^\alpha$之外，由于它衡量的是问题实例的难度，因此这项可能无法避免；式（13）中的上界不依赖于h的多项式。这意味着，探索奖励依赖模式中的稀疏性确实有益，尤其是在 $sd \ll h$时。 Hao 等人 [2020] 研究了数据匮乏情形下的稀疏线性老虎机问题，并得到了一个最优界，在我们的设定下可具体表示为 $\tilde{\Theta}((sd)^{\frac{2}{3}}T^{\frac{2}{3}})$。由于我们考虑的是不同的设定，而非稀疏臂参数，因此得到了一个不同的界。 △

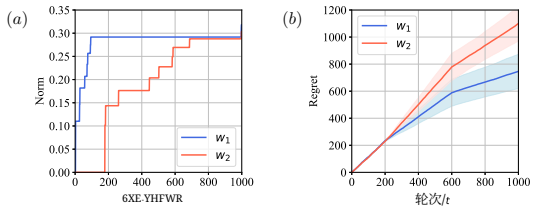

图 4：算法 1 在不同实例下的遗憾比较 $\boldsymbol{w}_*$。（a）由前 ℓ_2 个条目形成的子向量的 k 范数 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$分别为k（x为 横轴）。注意，$\boldsymbol{w}_1$ 的“质量”比 $\boldsymbol{w}_2$更集中于较早的位置。（b）实验表明，对于 $\boldsymbol{w}_1$，算法 1产生的遗憾小于对于 $\boldsymbol{w}_2$的遗憾，正如定理 2所预测。（阴影区域表示 10 次实验中的标准误。参数：$h = 1000$，$T = 999$，以及 $s = 10$。）

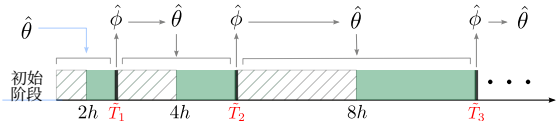

图 5：AD-Lasso 示意图。在 h 轮次的初始阶段，我们在算法 1中使用 Doubling Lasso。对于 $t > h$，我们也使用倍增技巧，但与算法 1略有不同：1）这里 $L = h$，2）我们在每个轮次中估计完整的 $\boldsymbol{\phi}$。

3.2 数据丰富情形

现在，我们考虑轮次数多于权重向量 $\boldsymbol{w}$的维度的情况，即 $T \geq h$。在这种数据丰富情形下，我们介绍一种算法，详见算法 2（另见图 5 示意）。

该算法包含两个阶段，因此具有自适应性：1）在最初的 h 轮次中，我们采用 Doubling Lasso（见算法 1）；2）从第 $h+1$ 轮次开始，我们提出另一种算法。在第二阶段，我们同样采用了类似于算法 1 的倍增技巧。二者唯一的区别在于：1）轮次 i 的长度为 $2^i h$，而不是 $2^{i+1}L$；2）在每个轮次中，我们估计 整个 $\boldsymbol{\phi}$，而不是其中的一部分；3）使用所收集数据的后半部分。

与算法 1 中一样，我们在每个轮次收集 $2^j h$ 个数据点。由（5）可知，后半部分的 $2^{j-1}h$ 奖励（见图 5）由下式生成：

$$\tilde{\boldsymbol{r}}[j] = \tilde{\boldsymbol{\Xi}}[j]\boldsymbol{\phi} + \tilde{\boldsymbol{\varepsilon}}[j],$$

其中，$\tilde{\boldsymbol{r}}[j] \in \mathbb{R}^{2^j h}$、$\tilde{\boldsymbol{\Xi}}[j] \in \mathbb{R}^{2^j h \times dh}$和 $\tilde{\boldsymbol{\varepsilon}}[j] \in \mathbb{R}^{2^j h}$ 分别是轮次 j后半部分中对应的奖励向量、上下文矩阵和噪声向量。

为了学习 $\boldsymbol{\phi}$，我们计算以下 Lasso 程序：

$$\hat{\boldsymbol{\phi}}[j] = \operatorname*{arg\,min}_{\boldsymbol{\phi} \in \mathbb{R}^{hd}} \left(\frac{1}{2^j h} \left\| \tilde{\boldsymbol{\Xi}}[j]\boldsymbol{\phi} - \boldsymbol{r}[j] \right\|_2^2 + \lambda_j \|\boldsymbol{\phi}\|_{2,1}^{(d)} \right), \tag{14}$$

其中正则化参数为

$$\lambda_j = 2\sqrt{\frac{2d\log(2^j h/\gamma)}{2^{j-1}h}}. \tag{15}$$

算法 2 自适应倍增 Lasso（AD-Lasso）

1：输入： L **初始阶段，倍增序列** $\{\tilde{T}_j\}$**，其中** $\tilde{T}_j = (2^{j+1}-1)h$ **[参见图 5]2：对于** $t = 1:$ h **执行 3：使用参数** L**实施算法 1。4：结束循环ˆ ˆ 5：将** θ_0 **重置为最新的** θ**。6：对于** $t = h+1:T$**执行 7：观测上下文向量** $\{\boldsymbol{x}_{t,a} : a \in [K]\}$**。ˆ 8：采取贪心动作** $a_t \in \sup_{a\in[K]} \langle \boldsymbol{x}_{t,a}, \boldsymbol{\theta}_{j-1} \rangle$**，并获得奖励** r_t**。9：如果** $t = T_j$**，则ˆ 10：#** *end of the jth epoch*, *estimate a new* $\boldsymbol{\theta}$ **ˆ 11：根据 Lasso (14) 计算** $\phi[j]$ **。ˆ ˆ 12：令** θ_j **为与最大奇异值对应的** Φ**的奇异向量。13：结束条件 14：结束循环**

定理 3. 考虑式（1）中描述的具有长时域奖励的随机上下文线性老虎机模型。在算法 2 中，令 L 与定理 2 中相同。当 $T \geq h$ 时，遗憾满足以下上界：

$$R_T = O\Big(d\sqrt{sh\log(dh)} + \min\{q(\boldsymbol{w}), h\} + \sqrt{sdT\log(dT)}\Big), \tag{16}$$

其中 $q(\boldsymbol{w})$ 在定理 2 中定义。

注 8. 式（16）中的前两项来自数据较少时的初始阶段（$t \leq h$）。注意，即使它们依赖于 h，也与 T 无关；只有当 T 与 h 同阶，即 $T = \Theta(h)$ 时，它们才会影响上界。在这种情况下，上界变为

$$R_T = O\Big(d\sqrt{sT\log(dT)} + \min\{q(\boldsymbol{w}), T\}\Big).$$

相反，如果 T较大，具体而言， $T \geq \max\{dh, h^{2\alpha(\mu)}/(sd\log(T))\}$，则式（16）中的前两项被最后一项所支配，上界可简化为

$$R_T = O\Big(\sqrt{sdT\log(dT)}\Big).$$

那么，我们的上界是最优的，适用于 d 和 T （忽略对数因子），这一结论可由下界 $\sqrt{dT}$ 得到支持，该下界由Chu 等人 [2011] 针对线性上下文老虎机给出。

关于下界的讨论：Ren 和 Zhou [2020] 针对$\sqrt{sT}$稀疏的 s稀疏上下文线性多臂老虎机得到了一个 $\Omega(\sqrt{(s+d)T}$的下界。考虑到我们问题的低秩和稀疏特性，通过调整他们的证明，可以得出我们的问题存在 $\Omega(\sqrt{(s+d)T}$的下界。因此，定理3 中的上界与该下界之间的差距至多为 $\log(dT)$ $\min\{\sqrt{s}, \sqrt{d}\}$倍。不过，我们认为实际差距要小得多。我们推测可以构造出更紧的下界，因为我们发现，使用循环测量进行低秩估计的采样复杂度并不单纯取决于秩（讨论见第2.2 节）。△

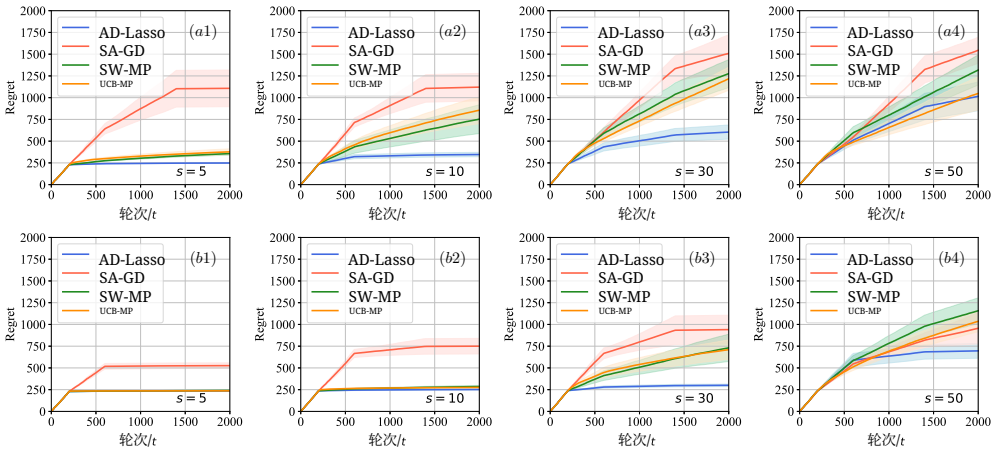

图6：不同算法的性能比较。(a) 平坦 $\boldsymbol{w}$： $\boldsymbol{w}$中的非零项均匀分布。(b) 尖峰型 $\boldsymbol{w}$：大多数权重集中在 $\boldsymbol{w}$中仅占20%的非零位置上。两种情况还考虑了 $\boldsymbol{w}$ 的不同稀疏性。(通用参数： $T = 2000$, $h = 100$， $d = 5$和 $\|\boldsymbol{w}\|_1 = 1$。)

实验。我们进行了一些实验，将我们的算法 AD-Lasso 与以下三种方法进行比较：

1. 稀疏交替梯度下降 (SA-GD)。SA-GD的核心是先进行秩一估计，再进行稀疏矩阵估计。基于(3)，SA-GD通过梯度下降交替重构 $\boldsymbol{\theta}$ 和 $\boldsymbol{w}$ ，并将 $\boldsymbol{w}$ 投影到 s-稀疏空间。2. 单权重匹配追踪(SW-MP)的核心是通过检验奖励向量与上下文矩阵各列之间的相关性，找出中最大的权重 $\boldsymbol{w}$ 。随后，利用该位置信息，通过最小二乘回归直接估计 $\boldsymbol{\theta}$ ，忽略 $\boldsymbol{w}$中的其他权重。

3. 结合匹配追踪的*UCB* （UCB-MP）。该算法与SW-MP相似；不同之处在于，每一轮次中我们使用置信上界更新 $\boldsymbol{\theta}$ 并作出决策。

为确保比较公平，我们对所有算法采用相同的倍增方案，并设置相同的轮次长度。唯一的区别在于我们用于估计 $\boldsymbol{\theta}$ 的方法（有关这些算法的更多细节，请参见附录 G ）。实验中考虑了不同的稀疏性和奖励依赖结构（参见图6的图注）。

当 $\boldsymbol{w}$ 高度稀疏时，我们的算法显著优于SA-GD（参见(a1)、(a2)、(b1)和(b2)）。由于SA-GD主要依赖秩一分解，这表明与低秩性相比，稀疏性对估计质量的影响更为显著，这与我们的理论相符。令人惊讶的是，随着 $\boldsymbol{w}$ 变得不那么稀疏，我们的算法仍然能够优于SA-GD，即使是在 $sd > d + h$的情形下也是如此。这印证了循环测量下低秩矩阵估计的难度，与我们在第2.2节中的讨论一致。不过，要在引理1之外对这些发现进行形式化论证，还需要更深入的理论分析。

即使 $\boldsymbol{w}$ 的权重高度集中在少数几个条目上，AD-Lasso 的表现也与 SW-MP 和 UCB-MP 一样出色。当权重更加分散时，AD-Lasso 的效果要好得多，这表明仅仅探索和利用最大权重会变得次优。

4 结论

本文引入了随机上下文老虎机问题的一种新变体，其中奖励取决于最多 s 个先前上下文，时间范围为 h。利用奖励依赖模式中的稀疏性，我们提出了两种同时适用于数据匮乏和数据丰富情形的算法。我们还为这两种算法推导了不依赖于时域（除 log(h) 项外）的遗憾上界，从理论上保证了其样本效率。

我们的工作开辟了许多未来潜在的研究方向。例如，奖励可以以非线性方式依赖于先前的上下文，或者稀疏模式可以以数据依赖方式变化。在这两种情况下，学习奖励依赖模式都会更具挑战性。此外，除了老虎机问题之外，探索具有长期非马尔可夫结构的强化学习和控制场景也很有意义，这类场景需要采用新的策略。

5 致谢

本研究部分由空军资助项目 AFOSR-FA9550-20-1-0140、NSF CCF-2046816、NSF TRIPODS II-DMS 2023166、NSF CCF-2007036、NSF CCF-2212261、NSF DMS-1839371，以及陆军研究办公室资助项目 ARO-78259-NS-MUR 和 W911NF2110312 提供支持。

参考文献

Yasin Abbasi-Yadkori、D´avid P´al 和 Csaba Szepesv´ari。改进线性随机老虎机算法。神经信息处理系统进展，24，2011年。

Yasin Abbasi-Yadkori 等人。在线到置信集合转换及其在稀疏随机老虎机中的应用。载于人工智能与统计，第 1–9 页。PMLR，2012年。

Radoslaw Adamczak。关于具有依赖关系的随机向量的 Hanson-Wright 不等式的注记。*Electronic Communications in Probability*，20:1–13，2015。

Oren Anava 等人。面向具有记忆的对手的在线学习：过去错误的代价。载于神经信息处理系统进展，第 28 卷，2015。Kaito Ariu、Kenshi Abe 和 Alexandre Proutı`ere。阈值化套索老虎机。载于国际机器学习会议，第 878–928 页。PMLR，2022年。Hamsa Bastani 和 Mohsen Bayati。基于高维协变量的在线决策。*Operations Research*，68(1):276–294，2020年。Peter J Bickel、Ya'acov Ritov 和 Alexandre B Tsybakov。Lasso 和 Dantzig 选择器的同步分析。*The Annals of statistics*，37(4):1705–1732，2009年。Ilai Bistritz、Zhengyuan Zhou、Xi Chen、Nicholas Bambos 和 Jose Blanchet。在具有延迟反馈的对抗多臂老虎机中进行在线 Exp3 学习。神经信息处理系统进展，32，2019年。

Tom Brown、Benjamin Mann、Nick Ryder、Melanie Subbiah、Jared D Kaplan、Prafulla Dhariwal、Arvind Neelakantan、Pranav Shyam、Girish Sastry、Amanda Askell 等人。语言模型是少样本学习者。神经信息处理系统进展，33:1877–1901，2020年。

S´ebastien Bubeck、Nicolo Cesa-Bianchi 等人。随机与非随机多臂赌博机问题的遗憾分析。趋势与基础ⓡ 机器学习，5(1):1–122，2012年。Emmanuel Candes 和 Terence Tao。Dantzig 选择器：当 p 远大于 n 时的统计估计。统计年鉴，35(6):2313–2351，2007年。Alexandra Carpentier 和 R´emi Munos。赌博机理论与压缩感知结合，用于高维随机线性赌博机。载于人工智能与统计，第 190–198 页。PMLR，2012年。Leonardo Cella 和 Nicol`o Cesa-Bianchi。具有延迟相关收益的随机老虎机。载于人工智能与统计学国际会议，第 1168–1177 页。PMLR，2020年。Nicolo Cesa-Bianchi、Claudio Gentile 和 Yishay Mansour。具有复合匿名反馈的非随机赌博机。载于学习理论会议，第 750–773 页。PMLR，2018。Nicol‘o Cesa-Bianchi、Claudio Gentile、Yishay Mansour 和 Alberto Minora。非随机赌博机中的延迟与协作。机器学习研究杂志，20:1–38，2019年。Lili Chen、Kevin Lu、Aravind Rajeswaran、Kimin Lee、Aditya Grover、Misha Laskin、Pieter Abbeel、Aravind Srinivas 和 Igor Mordatch。决策变换器：通过序列建模实现强化学习。神经信息处理系统进展，34:15084–15097，2021年。Wei Chu、Lihong Li、Lev Reyzin 和 Robert Schapire。用于上下文多臂赌博机的 epoch-greedy 算法。神经信息处理系统进展，20(1):96–1，2007年。Wei Chu 等人。具有线性收益函数的上下文赌博机。载于第十四届人工智能与统计学国际会议论文集，第 208–214 页。JMLR 研讨会与会议论文集，2011年。Mark A Davenport 和 Justin Romberg。基于不完整观测的低秩矩阵恢复概述。*IEEE* 信号处理精选主题期刊，10(4):608–622，2016年。Simon Foucart 和 Holger Rauhut。压缩感知数学导论，2013。Manegueu Anne Gael、Claire Vernade、Alexandra Carpentier 和 Michal Valko。具有臂依赖延迟的随机老虎机。载于国际机器学习会议，第 3348–3356 页。PMLR，2020年。Siddhant Garg 和 Aditya Kumar Akash。具有延迟复合匿名反馈的随机老虎机。*arXiv* 预印本 *arXiv:1910.01161*，2019年。Andras Gyorgy 和 Pooria Joulani。在对抗性多臂赌博机中适应延迟与数据。载于国际机器学习会议，第 3988–3997 页。PMLR，2021年。Botao Hao、Tor Lattimore 和 Mengdi Wang。高维稀疏线性赌博机。神经信息处理系统进展，33:10753–10763，2020年。Botao Hao、Tor Lattimore 和 Wei Deng。用于稀疏线性赌博机的信息导向采样。神经信息处理系统进展，34，2021年。Matteo Hessel、Joseph Modayil、Hado Van Hasselt、Tom Schaul、Georg Ostrovski、Will Dabney、Dan Horgan、Bilal Piot、Mohammad Azar 和 David Silver。Rainbow：融合深度强化学习中的多项改进。载于第三十二届 *AAAI* 人工智能会议，2018。

Benjamin Howson、Ciara Pike-Burke 和 Sarah Filippi。重新审视广义线性老虎机中的延迟反馈。*arXiv* 预印本 *arXiv:2207.10786*，2022年。Shinji Ito、Daisuke Hatano、Hanna Sumita、Kei Takemura、Takuro Fukunaga、Naonori Kakimura 和 Ken-Ichi Kawarabayashi。非随机线性老虎机中的延迟与协作。神经信息处理系统进展，33：4872–4883，2020年。Chi Jin、Praneeth Netrapalli、Rong Ge、Sham M Kakade 和 Michael I Jordan。关于具有次高斯范数的随机向量集中不等式的简短说明。*arXiv* 预印本 *arXiv:1902.03736*，2019年。Gi-Soo Kim 和 Myunghee Cho Paik。双重稳健 LASSO老虎机。神经信息处理系统进展，32，2019年。Felix Krahmer、Shahar Mendelson 和 Holger Rauhut。混沌过程的上确界与受限等距性质。《纯粹与应用数学通讯》，67(11)：1877–1904，2014年。Raunak Kumar 等人。具有无界记忆的在线凸优化。*arXiv* 预印本 *arXiv:2210.09903*，2022年。Tal Lancewicki、Shahar Segal、Tomer Koren 和 Yishay Mansour。具有不受限延迟分布的随机多臂老虎机。载于国际机器学习会议，第5969–5978页。PMLR，2021年。Tor Lattimore 和 Csaba Szepesv´ari。老虎机算法。剑桥大学出版社，2020年。Bingcong Li、Tianyi Chen 和 Georgios B Giannakis。具有未知延迟的老虎机在线学习。载于第*22*届人工智能与统计学国际会议，第993–1002页。PMLR，2019年。Lihong Li、Wei Chu、John Langford 和 Robert E Schapire。个性化新闻文章推荐的情境老虎机方法。载于第*19*届万维网国际会议论文集，第661–670页，2010年。Lihong Li、Yu Lu 和 Dengyong Zhou。具有可证明最优性的广义线性情境老虎机算法。载于国际机器学习会议，第2071–2080页。PMLR，2017年。Min-hwan Oh、Garud Iyengar 和 Assaf Zeevi。稀疏性无关 LASSO老虎机。载于国际机器学习会议，第8271–8280页。PMLR，2021年。Samet Oymak、Amin Jalali、Maryam Fazel、Yonina C Eldar 和 Babak Hassibi。同时具有结构化特性的模型及其在稀疏和低秩矩阵中的应用。*IEEE* 信息论汇刊，61(5)：2886–2908，2015年。Ciara Pike-Burke、Shipra Agrawal、Csaba Szepesvari 和 Steffen Grunewalder。具有延迟、聚合匿名反馈的老虎机。载于国际机器学习会议，第4105–4113页。PMLR，2018年。Yuzhen Qin、Tommaso Menara、Samet Oymak、ShiNung Ching 和 Fabio Pasqualetti。序列线性老虎机中的非平稳表示学习。*IEEE* 开放控制系统期刊，1：41–56，2022a。doi：10.1109/OJCSYS.2022.3178540。

Yuzhen Qin、Tommaso Menara、Samet Oymak、ShiNung Ching 和 Fabio Pasqualetti。用于上下文相关决策的表示学习。载于*2022*年美国控制会议 *(ACC)*，第2130–2135页，2022b。Holger Rauhut。压缩感知与结构化随机矩阵。稀疏恢复的理论基础与数值方法，9(1)：92，2010年。Zhimei Ren 和 Zhengyuan Zhou。高维稀疏线性情境老虎机中的动态批量学习。*arXiv* 预印本 *arXiv:2008.11918*，2020年。Emile Richard、Pierre-Andr´e Savalle 和 Nicolas Vayatis。同时稀疏且低秩矩阵的估计。*arXiv* 预印本 *arXiv:1206.6474*，2012年。Mark Rudelson 和 Roman Vershynin。Hanson-Wright 不等式与次高斯集中。*Electronic Communications in Probability*，18：1–9，2013年。Guanya Shi 等人。具有记忆的在线优化与竞争控制。载于神经信息处理系统进展，第33卷，第20636–20647页，2020年。Tobias Sommer Thune、Nicol`o Cesa-Bianchi 和 Yevgeny Seldin。具有不受限延迟的非随机多臂老虎机。神经信息处理系统进展，32，2019年。Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Lukasz Kaiser 和 Illia Polosukhin。注意力机制就是一切。神经信息处理系统进展，30，2017年。Claire Vernade、Alexandra Carpentier、Tor Lattimore、Giovanni Zappella、Beyza Ermis 和 Michael Brueckner。具有随机延迟反馈的线性老虎机。载于国际机器学习会议，第9712–9721页。PMLR，2020年。Roman Vershynin。高维概率：数据科学应用导论，第47卷。剑桥大学出版社，2018年。Martin J Wainwright。高维统计：非渐近视角，第48卷。剑桥大学出版社，2019年。Siwei Wang、Haoyun Wang 和 Longbo Huang。具有复合匿名反馈的多臂老虎机自适应算法。载于*AAAI*人工智能会议论文集，第35卷，第10210–10217页，2021年。Xue Wang、Mingcheng Wei 和 Tao Yao。具有高维协变量的极小极大凹惩罚多臂老虎机模型。载于国际机器学习会议，第5200–5208页。PMLR，2018年。Per-Åke Wedin。奇异值分解相关的扰动界。*BIT* 数值数学，12(1)：99–111，1972年。Michael Woodroofe。具有伴随变量的单臂老虎机问题。美国统计学会会刊，74(368)：799–806，1979年。Mengyan Zhang、Russell Tsuchida 和 Cheng Soon Ong。具有聚合反馈的高斯过程老虎机。载于*AAAI*人工智能会议论文集，第36卷，第9074–9081页，2022年。

ZhengyuanZhou、Renyuan Xu 和Jose Blanchet。在具有随机延迟的广义线性上下文老虎机中学习。神经信息处理系统进展，32，2019年。

Julian Zimmert 和 Yevgeny Seldin。《一种适用于任意延迟的对抗多臂老虎机的最优算法》。载于人工智能与统计学国际会议，第 3285–3294 页。PMLR，2020年。

附录

附录分为两部分：

1. 在第一部分中，我们致力于推导一些一般性结果，这些结果不仅对后续分析有用，而且本身也具有研究价值。具体而言，我们证明，从以下分块循环矩阵中子采样 $\tilde{\Omega}(s\log^2(s))$ 行所构成的测量矩阵满足稀疏向量的受限等距性质（RIP）：

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\xi}_n^\top & \boldsymbol{\xi}_{n-1}^\top & \cdots & \boldsymbol{\xi}_1^\top \\ \boldsymbol{\xi}_1^\top & \boldsymbol{\xi}_n^\top & \cdots & \boldsymbol{\xi}_2^\top \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_{n-1}^\top & \boldsymbol{\xi}_{n-2}^\top & \cdots & \boldsymbol{\xi}_n^\top \end{bmatrix},$$

其中，$\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n \in \mathbb{R}^d$ 是相互独立且各向同性的随机向量。作为对Krahmer 等人 [2014], 结果的推广，我们允许 $\boldsymbol{\xi}_i$ 具有相关条目。该结果总结于正文的定理1 中。

2. 在第二部分中，我们为本文的主要问题提供辅助材料：具有长时域奖励的线性上下文老虎机。第一部分中得到的结果将用于证明定理2和3。我们还将展示一些进一步的实验结果。

预备知识

进一步的记号。给定一个向量 $\boldsymbol{x} = [x_1, x_2, \ldots, x_{nd}]^\top$，将 $\|\boldsymbol{x}\|_0 := \sum_{i=1}^{nd} \mathbb{I}(|x_i| \neq 0)$ 记作其ℓ_0范数。我们进一步定义它的一个 $\ell_{2,1}$-范数，它与矩阵版本类似。具体而言，$\|\boldsymbol{x}\|_{2,1}^{(d)} = \sum_{j=1}^{n} \|\boldsymbol{x}_j\|_2$, 其中每个 $\boldsymbol{x}_j$ 都是将 $\boldsymbol{x}$ 划分为 n 个大小相等的块d（即 $\boldsymbol{x} = [\boldsymbol{x}_1^\top, \boldsymbol{x}_2^\top, \ldots, \boldsymbol{x}_n^\top]^\top$ 满足 $\boldsymbol{x}_i \in \mathbb{R}^d$）得到的。类似地，将 $\|\boldsymbol{x}\|_{2,\infty}^{(d)} =$ 记作最大值$_{j\in\{1,\ldots,n\}} \|\boldsymbol{x}_j\|_2$，并将$\|\boldsymbol{x}\|_{2,0}^{(d)} = \sum_{i=1}^{n} \mathbb{I}$记作（$\|\boldsymbol{x}_j\|_2 \neq 0$）。给定一个矩阵 $A \in \mathbb{R}^{m\times n}$，$\|A\|_F$ 表示其弗罗贝尼乌斯范数。我们用 $\lesssim$ 和 $\gtrsim$ 表示忽略常数因子或对数因子后的不等式。

如果 $\|\boldsymbol{a}\|_0 \leq s$，则称向量 $\boldsymbol{a} = [a_1, a_2, \ldots, a_d]^\top \in \mathbb{R}^d$ 为 s-稀疏。如果向量 $\boldsymbol{b} = [\boldsymbol{b}_1^\top, \boldsymbol{b}_2^\top, \ldots, \boldsymbol{b}_h^\top]^\top \in \mathbb{R}^{hd}$的每个块 $\boldsymbol{b}_i \in \mathbb{R}^d$，且 $\boldsymbol{b}_i$中至多有 s 个非零（即如果$\|\boldsymbol{b}\|_{2,0}^{(d)} \leq s$），则称其为 s-块稀疏。

如果随机变量 $x \in \mathbb{R}$ 满足 $\mathbb{E}[x] = 0$，且其矩母函数满足以下条件，则称其为方差代理为 σ^2 的次高斯随机变量（简称 σ^2-次高斯）：

$$\mathbb{E}[\exp(tX)] \leq \exp(\frac{\sigma^2 t^2}{2}), \qquad \forall t \in \mathbb{R}.$$

如果对于任意满足 $\|\mathbf{u}\| \leq 1$ 的 $u \in \mathbb{R}^d$，随机向量 $\boldsymbol{x} \in \mathbb{R}^d$ 都使得随机变量 $\mathbf{u}^\top \boldsymbol{x}$ 是 σ^2-次高斯的，则称随机向量 $\boldsymbol{x} \in \mathbb{R}^d$ 是 σ^2-次高斯的。对于并非零均值的随机向量 $\boldsymbol{x}$，我们滥用记号：若 $\boldsymbol{x} - \mathbb{E}\boldsymbol{x}$ 是 σ^2-次高斯的，则称其为 σ^2-次高斯的。

第一部分：分块循环测量的 RIP 一般结果

A 主要结果

我们首先介绍一个将在后文使用的定义。然后，在本部分中给出主要结果。

定义 A.1. 若矩阵 $\boldsymbol{\Xi} \in \mathbb{R}^{m\times d}$ 满足阶数为 s、水平为 δ 的受限等距性质（简称 (s,δ)-RIP），则称其具有该性质，当且仅当存在 $\delta \in (0,1)$，使得

$$(1-\delta)\|\boldsymbol{a}\|_2^2 \le \|\boldsymbol{\Xi}\boldsymbol{a}\|_2^2 \le (1+\delta)\|\boldsymbol{a}\|_2^2$$

对于所有 s-稀疏向量 $\boldsymbol{a} \in \mathbb{R}^d$均成立。使该不等式成立的最小 δ 记为 δ_s，称为受限等距常数。

在第一部分中，我们关注以下分块循环矩阵：

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{\xi}_n^\top & \boldsymbol{\xi}_{n-1}^\top & \cdots & \boldsymbol{\xi}_1^\top \\ \boldsymbol{\xi}_1^\top & \boldsymbol{\xi}_n^\top & \cdots & \boldsymbol{\xi}_2^\top \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_{n-1}^\top & \boldsymbol{\xi}_{n-2}^\top & \cdots & \boldsymbol{\xi}_n^\top \end{bmatrix} \in \mathbb{R}^{n\times nd}, \tag{A1}$$

其中 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n \in \mathbb{R}^d$是独立随机向量，并且对于每个 i，都有 $\mathbb{E}\boldsymbol{\xi}_i = \boldsymbol{0}$和$\mathbb{E}\boldsymbol{\xi}_i\boldsymbol{\xi}_i^\top = \boldsymbol{I}$。令 $\boldsymbol{\Xi}$ 为对 $\boldsymbol{C}$的任意 m $(m<n)$ 行进行子采样所形成的矩阵，即，

$$\boldsymbol{\Xi} = \frac{1}{\sqrt{m}}\boldsymbol{R}_\Omega \boldsymbol{C}, \tag{A2}$$

其中 $\Omega \subset [n]$ 由 $|\Omega| = m$ 和 $\boldsymbol{R}_\Omega \in \mathbb{R}^{m\times n}$ 选取 $\boldsymbol{C}$ 中由 Ω 索引的 m 行（注意，我们在此处包含项 $\frac{1}{\sqrt{m}}$，以对下采样矩阵~~进~~行归一化）。可以看出，$\boldsymbol{R}_\Omega$ 是从单位矩阵 $\boldsymbol{I}_n$中移除由集合 $[n]\backslash\Omega$索引的行后得到的矩阵。现在，考虑以下测量

$$\boldsymbol{y} = \boldsymbol{\Xi}\boldsymbol{a}. \tag{A3}$$

我们假设未知向量 $\boldsymbol{a} \in \mathbb{R}^{nd}$ 是 s_0-稀疏的，即 $\|\boldsymbol{a}\|_0 \le s_0$。本部分的主要结果证明了下采样块循环矩阵对稀疏向量满足 RIP。

定理 A.1. 考虑设计矩阵 $\boldsymbol{\Xi} \in \mathbb{R}^{m\times nd}$ 由式（A2）给出，其中 Ω 是$[n]$ 的任意子集，其基数为 m。假设每个随机向量 $\boldsymbol{\xi}_i$ 满足 Hanson-Wright 不等式

$$\Pr\left[|\boldsymbol{\xi}_i^\top \boldsymbol{P}\boldsymbol{\xi}_i - \mathbb{E}(\boldsymbol{\xi}_i^\top \boldsymbol{P}\boldsymbol{\xi}_i)| \ge t\right] \le 2\exp\left(-\frac{1}{c}\min\left\{\frac{t^2}{k^4\|\boldsymbol{P}\|_F^2}, \frac{t}{k^2\|\boldsymbol{P}\|_{\mathrm{op}}}\right\}\right) \tag{A4}$$

对于任意半正定矩阵 $\boldsymbol{P} \in \mathbb{R}^{d\times d}$，其中 k 是一个常数，c 是一个绝对常数。那么，对于任意 δ，$\eta \in (0,1)$，存在一个常数 $c_1 > 0$，使得如果

$$m \ge c_1 \frac{1}{\delta^2} s_0 \max\left\{\log^2(s_0)\log^2(nd), \log\left(\frac{1}{\eta}\right)\right\}, \tag{A5}$$

那么，设计矩阵 $\boldsymbol{\Xi}$ 的受限等距常数对于所有 s_0稀疏向量 $\boldsymbol{a} \in \mathbb{R}^{nd}$，记为 δ_{s_0}，以至少 $1-\eta$ 的概率满足 $\delta_{s_0} \le \delta$。

备注 9。该定理与Krahmer 等人 [2014]中的定理 4.1 类似。主要区别如下：1）我们对分块循环矩阵进行子采样，而不是对循环矩阵进行子采样（循环矩阵中的每一行相对于前一行向右循环移位一个元素；而在我们的情形中，分块循环矩阵的每一行向右循环移位一个大小为 d 的块）；2）我们允许设计矩阵 $\boldsymbol{\Xi}$ 的第一行具有相关的元素（即，每个 $\boldsymbol{\xi}_i$ 中的元素可以相关），而Krahmer 等人[2014] 假设这些元素相互独立；3）我们对（A4）中的向量 $\boldsymbol{\xi}_i$ 额外作出一项假设。值得一提的是，这项假设较为宽松，甚至比 Hanson-Wright 不等式的经典假设还要宽松Rudelson 和 Vershynin [2013]。这是因为我们只要求$\boldsymbol{P}$为半正定矩阵，而非任意 $d \times d$矩阵。不过，这项假设能否进一步放宽甚至去除，仍是一个未解决问题。△

这里，我们推导出了一个适用于一般稀疏向量的定理。正如后文所示，在第3节的老虎机问题中，我们的未知向量 $\boldsymbol{\phi} = \boldsymbol{w} \otimes \boldsymbol{\theta}$ 是 s 块稀疏向量，块长度为 d。为了便于分析，我们将根据下面这个直接由定理A.1 推出的定理，推导出适用于块稀疏向量的受限特征值条件。值得一提的是，定理A.1中的结果可应用于更一般的问题。

定理 **A.2.** 考虑设计矩阵 $\boldsymbol{\Xi} \in \mathbb{R}^{m \times hd}$ 由托普利兹矩阵给出

$$\boldsymbol{\Xi} = \frac{1}{\sqrt{m}} \begin{bmatrix} \boldsymbol{\xi}_h^\top & \boldsymbol{\xi}_{h-1}^\top & \cdots & \boldsymbol{\xi}_1^\top \\ \boldsymbol{\xi}_{h+1}^\top & \boldsymbol{\xi}_h^\top & \cdots & \boldsymbol{\xi}_2^\top \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_{m+h-1}^\top & \boldsymbol{\xi}_{m+h-2}^\top & \cdots & \boldsymbol{\xi}_m^\top \end{bmatrix}.$$

假设向量 $\boldsymbol{\xi}$ 满足定理A.1中的假设。那么，对于任意 δ，$\eta \in (0, 1)$，存在一个常数 $c > 0$ 使得，若

$$m \geq c \frac{1}{\delta^2} s_0 \max \left\{ \log^2(s_0) \log^2(hd), \log\left(\frac{1}{\eta}\right) \right\}, \tag{A6}$$

则对于所有 s_0-稀疏向量 $\boldsymbol{a} \in \mathbb{R}^{hd}$，设计矩阵 $\boldsymbol{\Xi}$ 的受限等距常数记为δ_{s_0}，以至少 $1 - \eta$的概率满足$\delta_{s_0} \leq \delta$。

注释 10。 此处，托普利兹矩阵 $\boldsymbol{\Xi}$ 可以看作由对循环矩阵（*A1*）的前 m 行和后 hd 列进行子采样所形成的矩阵。注意，（A6）右侧对 hd 具有多对数依赖，而不是对 nd的依赖，这比（A5）略好。按照与证明定理 A.1中类似的步骤，可以证明这种依赖关系。

B 定理 A.1 的证明

接下来，我们给出定理 A.1 的证明。首先，我们通过将 (A3) 改写为另一种形式来做一些准备工作，以便从不同的角度构造证明。我们将 ($\boldsymbol{a} \in \mathbb{R}^{nd}$ 中的 A3) 向量划分为 n 块，并将其改写为

$$\boldsymbol{a} = [\boldsymbol{a}_1^\top, \boldsymbol{a}_2^\top, \ldots, \boldsymbol{a}_n^\top]^\top, \qquad \text{where } \boldsymbol{a}_i \in \mathbb{R}^d.$$

注意，这里的设计矩阵 $\boldsymbol{\Xi}$ 由 m 的前 $\boldsymbol{C}$ 行子采样得到，取自 (A3) 中。因此，可以将 (A3) 改写为

$$\boldsymbol{y} = \frac{1}{\sqrt{m}}\boldsymbol{R}_\Omega \underbrace{\begin{bmatrix} \boldsymbol{a}_n^\top & \boldsymbol{a}_{n-1}^\top & \cdots & \boldsymbol{a}_2^\top & \boldsymbol{a}_1^\top \\ \boldsymbol{a}_1^\top & \boldsymbol{a}_n^\top & \cdots & \boldsymbol{a}_3^\top & \boldsymbol{a}_2^\top \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \boldsymbol{a}_{n-1}^\top & \boldsymbol{a}_{n-2}^\top & \cdots & \boldsymbol{a}_1^\top & \boldsymbol{a}_n^\top \end{bmatrix}}_{\boldsymbol{V}\in\mathbb{R}^{n\times nd}} \begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \vdots \\ \boldsymbol{\xi}_n \end{bmatrix} := \boldsymbol{A}\boldsymbol{\xi}. \tag{A7}$$

回顾一下，$\Omega \subset [n]$ with $|\Omega| = m$ and $\boldsymbol{R}_\Omega \in \mathbb{R}^{m\times n}$ 选取了 $\boldsymbol{V}$ 中由 Ω 索引的 m 行。回顾一下，我们的目标是证明随机矩阵 $\boldsymbol{\Xi}$ 满足针对 s_0-稀疏向量 $\boldsymbol{a}$的限制等距条件。等价地，只需对以下量给出上界

$$C_{\mathcal{A}}(\boldsymbol{\xi}) := \sup_{\boldsymbol{A}\in\mathcal{A}} \left| \|\boldsymbol{A}\boldsymbol{\xi}\|_2^2 - \mathbb{E}\|\boldsymbol{A}\boldsymbol{\xi}\|_2^2 \right|, \tag{A8}$$

其中， $\mathcal{A}$ 是由 $\boldsymbol{a} \in \mathcal{D}_{s_0}$ 生成的矩阵集合 $\boldsymbol{A}$ ，满足

$$\mathcal{D}_{s_0} := \left\{ \boldsymbol{a} \in \mathbb{R}^{nd} : \|\boldsymbol{a}\|_2 \le 1, \|\boldsymbol{a}\|_0 \le s_0 \right\}. \tag{A9}$$

对于矩阵 $\boldsymbol{A}$，我们将由第 $((i-1)d+1)$ 列至第 id 列组成的子矩阵称为 i-th 块列 $\boldsymbol{A}$ ，并记作 $\boldsymbol{A}^i$。至于式 (A7) 中的 $\boldsymbol{A}$ ，图 7 展示了块列的定义。

$$\mathbf{A} = \begin{bmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & * \end{bmatrix}, \text{ where } \mathbf{A}^i \in \mathbb{R}^{m\times d}.$$
$$\mathbf{A}^1 \quad \mathbf{A}^2 \quad \mathbf{A}^3$$

图7：块列定义示意图。

接下来定义以下量

$$N_{\mathcal{A}}(\boldsymbol{\xi}) := \sup_{\boldsymbol{A}\in\mathcal{A}} \|\boldsymbol{A}\boldsymbol{\xi}\|_2, \quad O_{\mathcal{A}}(\boldsymbol{\xi}) := \sup_{\boldsymbol{A}\in\mathcal{A}} \left| \sum_{i,j=1, i\neq j}^{n} \boldsymbol{\xi}_i^\top (\boldsymbol{A}^i)^\top \boldsymbol{A}^j \boldsymbol{\xi}_j \right|,$$
$$D_{\mathcal{A}}(\boldsymbol{\xi}) := \sup_{\boldsymbol{A}\in\mathcal{A}} \left| \sum_{i=1}^{n} \left(\boldsymbol{\xi}_i^\top (\boldsymbol{A}^i)^\top \boldsymbol{A}^i \boldsymbol{\xi}_i - \langle \boldsymbol{A}^i, \boldsymbol{A}^i \rangle \right) \right|,$$

其中 $\langle \boldsymbol{A}^i, \boldsymbol{A}^i \rangle = \mathrm{Tr}\big((\boldsymbol{A}^i)^\top \boldsymbol{A}^i\big)$。注意，

$$C_{\mathcal{A}}(\boldsymbol{\xi}) = \sup_{\boldsymbol{A}\in\mathcal{A}} \left| \sum_{i,j=1, i\neq j}^{n} \boldsymbol{\xi}_i^\top (\boldsymbol{A}^i)^\top \boldsymbol{A}^j \boldsymbol{\xi}_j + \sum_{i=1}^{n} \boldsymbol{\xi}_i^\top (\boldsymbol{A}^i)^\top \boldsymbol{A}^i \boldsymbol{\xi}_i - \langle \boldsymbol{A}^i, \boldsymbol{A}^i \rangle \right| \le O_{\mathcal{A}}(\boldsymbol{\xi}) + D_{\mathcal{A}}(\boldsymbol{\xi}). \tag{A10}$$

注意到 $O_{\mathcal{A}}(\boldsymbol{\xi})$ 和 $D_{\mathcal{A}}(\boldsymbol{\xi})$ 分别量化了矩阵 $\boldsymbol{A}^\top\boldsymbol{A}$ 对 $C_{\mathcal{A}}$中非块对角项和块对角项的贡献。接下来需要推导非块对角项和分块对角项的界。这些项的界分别见引理B.1 和

定理B.2 分别位于小节B.1和B.2中。借助这些结果，我们即可构造证明。

定理 A.1 的证明：我们利用引理B.1 和定理 B.2 分别对以下量进行界定：$O_{\mathcal{A}}(\xi)$ 和 $D_{\mathcal{A}}(\xi)$。

步骤 1：由于 $\boldsymbol{a}$ 是 s_0-稀疏的，按照Krahmeretal. [2014],中的类似步骤，可以推导出引理 B.1中定义的量满足

$$d_F(\mathcal{A}) \le 1, \quad d_{2\to 2}(\mathcal{A}) \le \sqrt{\frac{s_0}{m}}, \quad \text{and } \gamma_2(\mathcal{A}, \|\cdot\|_{\text{op}}) \lesssim \sqrt{\frac{s_0}{m}} \log(s_0)\log(nd).$$

由于 $m \ge m_1 = c_1 \frac{1}{\delta^2}\left(s_0 \log^2(s_0)\log^2(nd) + \log(\frac{1}{\eta_1})\right)$，由式（A13）可推导出，对于任意 δ和 $\eta_1 \in (0,1)$，存在 $c_1 > 0$使得

$$\Pr\left[O_{\mathcal{A}}(\boldsymbol{\xi}) \ge \frac{\delta}{2}\right] \le \eta_1. \tag{A11}$$

步骤 2：为了利用定理B.2界定分块对角项，还需推导集合 $\mathcal{A}$关于度量 $\|\cdot\|_F$的覆盖数。对于任意 $\boldsymbol{A}_1$，$\boldsymbol{A}_2 \in \mathcal{A}$由$\boldsymbol{a}_1$，$\boldsymbol{a}_2 \in \mathcal{D}_{s_0}$分别构成的量，根据式（A7）可得

$$\|\boldsymbol{A}_1 - \boldsymbol{A}_2\|_F = \frac{1}{\sqrt{m}}\sqrt{m\|\boldsymbol{a}_1 - \boldsymbol{a}_2\|_2^2} \le \|\boldsymbol{a}_1 - \boldsymbol{a}_2\|_2.$$

因此，有 $N(\mathcal{A}, \|\cdot\|_F, u) \le N(\mathcal{D}_{s_0}, \|\cdot\|_2, u)$。根据 Foucart 和 Rauhut 中的命题 C.3[2013],，有 $N(\mathcal{D}_{s_0}, \|\cdot\|_2, u) \le (1+\frac{2}{u})^{s_0}$。将此 $N(\mathcal{A}, \|\cdot\|_F, u)$ 代入以下不等式（由定理 B.2 得出）

$$\Pr\left[\sup_{A\in\mathcal{A}} D_{\mathcal{A}}(\boldsymbol{\xi}) \ge t\right] \le 2N\left(\mathcal{A}, \|\cdot\|_F, \frac{t}{6nd}\right)\exp\left(-\frac{1}{ck^2}\min(mt^2, mt)\right),$$

可以证明，对于任意 δ，$\eta_2 \in (0,1)$，存在 $c_2 > 0$使得，当 $m \ge m_2 = c_2 \frac{1}{\delta^2}\left(s_0 \log(nd) + \log(\frac{1}{\eta_2})\right)$时，有

$$\Pr\left[D_{\mathcal{A}}(\boldsymbol{\xi}) \ge \frac{\delta}{2}\right] \le \eta_2. \tag{A12}$$

结合（A11）和（A12），并选择 $m \ge \max\{m_1, m_2\}$，可得 $\Pr[C_{\mathcal{A}}(\boldsymbol{\xi}) \ge \delta] \le \eta$，其中 $\eta = \eta_1 + \eta_2$，证明至此完成。□

B.1 非块对角项的界

为了证明非块对角项的界，分析过程与Krahmer 等人 [2014]的分析基本相同。主要区别在于，我们在定理B.1 中推导了一个新的解耦不等式（关于我们贡献的讨论见该定理）。为使内容自足，我们在此给出完整证明。

引理 B.1。非块对角项 $O_{\mathcal{A}}(\boldsymbol{\xi})$满足

$$\Pr\left[O_{\mathcal{A}}(\boldsymbol{\xi}) \ge c_1 E + t\right] \le 2\exp\left(-c_2 \min\left\{\frac{t^2}{V^2}, \frac{t}{U}\right\}\right), \tag{A13}$$

其中

$$E = \gamma_2(\mathcal{A}, \|\cdot\|_{\mathrm{op}})\Big(\gamma_2(\mathcal{A}, \|\cdot\|_{\mathrm{op}}) + d_F(\mathcal{A}) + \sqrt{p}d_{2\to 2}(\mathcal{A})\Big), \tag{A14}$$

$$V = d_{2\to 2}(\mathcal{A})\Big(d_F(\mathcal{A}) + \gamma_2(\mathcal{A}, \|\cdot\|_{\mathrm{op}})\Big), \tag{A15}$$

$$U = d^2_{2\to 2}(\mathcal{A})。 \tag{A16}$$

这里，量 $d_F(\mathcal{A})$、 $d_{2\to 2}(\mathcal{A})$ 和 $\gamma_2(\mathcal{A}, \|\cdot\|_{\mathrm{op}})$ 满足

$$\begin{aligned} d_F(\mathcal{A}) &= \sup_{A\in\mathcal{A}} \|\boldsymbol{A}\|_F, \\ d_{2\to 2}(\mathcal{A}) &= \sup_{\boldsymbol{A}\in\mathcal{A}} \|\boldsymbol{A}\|_{\mathrm{op}}, \\ \gamma_2(\mathcal{A}, \|\cdot\|_{\mathrm{op}}) &\le c\int_0^{d_{2\to 2(\mathcal{A})}} \sqrt{\log N(\mathcal{A}, \|\cdot\|_{\mathrm{op}}, u)du}, \end{aligned}$$

其中， $N(\mathcal{A}, \|\cdot\|_{\mathrm{op}}, u)$ 是相对于度量 $(\|\cdot\|_{\mathrm{op}}, u)$ 的 $\mathcal{A}$ 覆盖数。

为了构造引理 B.1 的证明，我们需要引用 Krahmer 等人 [2014]。

引理 B.2. 设 $\mathcal{A}$ 是一个矩阵集合， $\boldsymbol{\xi}$ 是一个次高斯随机向量，且 $\boldsymbol{\xi}'$ 是 $\boldsymbol{\xi}$ 的一个独立副本。那么对于任意 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$，有

$$\left\|\sup_{\boldsymbol{A}\in\mathcal{A}} \langle \boldsymbol{A}\boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}'\rangle\right\|_{L_1} \lesssim \gamma_2(\mathcal{A}, \|\cdot\|_{\mathrm{op}}) \cdot \|N_{\mathcal{A}}(\boldsymbol{\xi})\|_{L_1} + \sup_{\boldsymbol{A}\in\mathcal{A}} \|\langle \boldsymbol{A}\boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}'\rangle\|_{L_1},$$

其中 $\|X\|_{L_1} := \mathbb{E}|X|$。

引理 B.3。 若 $\boldsymbol{\xi}'$ 是 $\boldsymbol{\xi}$的独立副本，则

$$\sup_{\boldsymbol{A}\in\mathcal{A}} \|\langle \boldsymbol{A}\boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}'\rangle\|_{L_1} \lesssim \sqrt{p}d_F(\mathcal{A})d_{2\to 2}(\mathcal{A}) + pd^2_{2\to 2}(\mathcal{A}).$$

引理 B.4。 设 z 是满足以下条件的随机变量

$$(\mathbb{E}|z|^p)^{1/p} \le \alpha + \beta\sqrt{p} + \gamma p, \text{for all } p \ge p_0$$

对于某些 α， β， γ， $p_0 > 0$。那么，对于 $u \ge p_0$,

$$\Pr\left(|z| \ge e(\alpha + \beta\sqrt{u} + \gamma y)\right) \le e^{-u}.$$

引理 B.1 的证明： 将定理 B.1 中的解耦不等式应用于 $O_{\mathcal{A}}(\boldsymbol{\xi})$，我们有

$$\begin{aligned} \|O_{\mathcal{A}}(\boldsymbol{\xi})\|_{L_1} &\le 4\mathbb{E}\sup_{\boldsymbol{A}\in\mathcal{A}}\left|\sum_{i,j=1}^n \boldsymbol{\xi}_i^\top (\boldsymbol{A}^i)^\top \boldsymbol{A}^j \boldsymbol{\xi}'_j\right| = 4\left\|\sup_{\boldsymbol{A}\in\mathcal{A}} \langle \boldsymbol{A}\boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}'\rangle\right\|_{L_1} \\ &\lesssim \gamma_2(\mathcal{A}, \|\cdot\|_{2\to 2}) \cdot \|N_{\mathcal{A}}(\boldsymbol{\xi})\|_{L_1} + \sup_{\boldsymbol{A}\in\mathcal{A}} \|\langle \boldsymbol{A}\boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}'\rangle\|_{L_1}. \end{aligned} \tag{A17}$$

按照与Krahmer 等人 [2014], 类似的步骤，可以证明

$$\|N_{\mathcal{A}}(\boldsymbol{\xi})\|_{L_1} \lesssim \gamma_2(\mathcal{A}, \|\cdot\|_{\mathrm{op}}) + d_F(\mathcal{A}) + d_{2\to 2}(\mathcal{A}),$$

并由引理B.3 可得

$$\sup_{\boldsymbol{A}\in\mathcal{A}} \|\langle \boldsymbol{A}\boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}'\rangle\|_{L_1} \lesssim d_F(\mathcal{A})d_{2\to 2}(\mathcal{A}) + d^2_{2\to 2}(\mathcal{A}).$$

将它们代入（A17），得到 $\|O_{\mathcal{A}}(\boldsymbol{\xi})\|_{L_1} \lesssim E+U+V$。应用引理 B.4即可完成证明。□

定理 B.1 （解耦不等式）。设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$是满足条件 $\mathbb{E}\boldsymbol{x}_i = \boldsymbol{0}$的独立随机向量。设 $\boldsymbol{x} = (\boldsymbol{x}_1^\top, \boldsymbol{x}_2^\top, \ldots, \boldsymbol{x}_n^\top)^\top$。定义 $\mathcal{B}_o := \{\boldsymbol{B} \in \mathbb{R}^{dn\times dn}\}$，其中

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \ldots & \boldsymbol{B}_{1n} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \ldots & \boldsymbol{B}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{B}_{n1} & \boldsymbol{B}_{n2} & \ldots & \boldsymbol{B}_{nn} \end{bmatrix}, \quad \text{and} \quad \boldsymbol{B}_{ij} \in \mathbb{R}^{d\times d} \text{ for all } i, j.$$

设 $\boldsymbol{x}'$ 是 $\boldsymbol{x}$ 的一个独立副本。那么，有

$$\mathbb{E}_{\boldsymbol{x}} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i,j=1, i\neq j}^{n} \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j \right| \le \mathbb{E}_{\boldsymbol{x},\boldsymbol{x}'} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| 4\sum_{i,j=1}^{n} \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j' \right|.$$

注 11。在 Rauhut [2010], 中给出了一个类似的解耦不等式，其中 $x_1, x_2, \ldots, x_n$是一维独立随机变量。这里，我们证明该结论可以推广到多元独立随机向量，而不要求每个向量中的各个分量也相互独立。该定理的证明核心沿用Rauhut [2010]。

证明。考虑一列相互独立的随机变量 δ_1， δ_2，…， δ_n，每个 δ_i 以概率 $_{12}$取值 0 或 1。对于 $j \neq k$, $\mathbb{E}\delta_j(1-\delta_k) = \frac{1}{4}$。令 $\delta = (\delta_1, \cdots, \delta_n)$。

由于对于任意 $\boldsymbol{B} \in \mathcal{B}$，对所有 i，我们有

$$\begin{aligned} \mathbb{E} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i,j=1, i\neq j}^{n} \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j \right| &\le 4\cdot \mathbb{E}_{\boldsymbol{x}} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i,j=1, i\neq j}^{n} \mathbb{E}[\delta_i(1-\delta_j)] \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j \right| \\ &\le 4\cdot \mathbb{E}_{\boldsymbol{x},\delta} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i,j=1, i\neq j}^{n} \delta_i(1-\delta_j) \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j \right|. \end{aligned}$$

接下来，我们定义集合 $\mathcal{I}(\delta) := \{k \in [n]\colon \delta_k = 1\}$。于是，根据 Fubini 定理可得

$$\mathbb{E} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i,j=1, i\neq j}^{n} \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j \right| \le 4\cdot \mathbb{E}_\delta \mathbb{E}_{\boldsymbol{x}} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i\in\mathcal{I}(\delta)} \sum_{j\notin\mathcal{I}(\delta)} \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j \right|.$$

对于固定的 δ，可以看出 $\{\boldsymbol{x}_i\}_{i\in\mathcal{I}(\delta)}$ 和 $\{\boldsymbol{x}_j\}_{j\notin\mathcal{I}(\delta)}$ 相互独立。因此，我们得到

$$\mathbb{E} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i,j=1, i\neq j}^{n} \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j \right| \le 4\cdot \mathbb{E}_\delta \mathbb{E}_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{x}'} \sup_{\boldsymbol{B}\in\mathcal{B}_o} \left| \sum_{i\in\mathcal{I}(\delta)} \sum_{j\notin\mathcal{I}(\delta)} \boldsymbol{x}_i^\top \boldsymbol{B}_{ij}\boldsymbol{x}_j' \right|.$$

那么，存在 δ_0 使得

$$\mathbb{E}\sup_{\boldsymbol{B}\in\mathcal{B}_o}\left|\sum_{i,j=1,i\neq j}^{n}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\boldsymbol{x}_j\right|\leq 4\cdot\mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{x}'}\sup_{\boldsymbol{B}\in\mathcal{B}_o}\left|\sum_{i\in\mathcal{I}(\delta_0)}\sum_{j\notin\mathcal{I}(\delta_0)}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\boldsymbol{x}_j'\right|.$$

由于 $\mathbb{E}[\boldsymbol{x}_i]=\mathbb{E}[\boldsymbol{x}_j']=\mathbf{0}$，我们观察到

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{x}'}\sup_{\boldsymbol{B}\in\mathcal{B}_o}\left|\sum_{i\in\mathcal{I}(\delta_0)}\sum_{j\notin\mathcal{I}(\delta_0)}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\boldsymbol{x}_j'\right|\\&=\mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{x}'}\sup_{\boldsymbol{B}\in\mathcal{B}_o}\left|\sum_{i\in\mathcal{I}(\delta_0)}\left(\sum_{j\notin\mathcal{I}(\delta_0)}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\boldsymbol{x}_j'+\sum_{j\in\mathcal{I}(\delta_0)}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\mathbb{E}\boldsymbol{x}_j'\right)+\sum_{i\notin\mathcal{I}(\delta_0)}\sum_{j=1}^{n}\mathbb{E}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\boldsymbol{x}_j'\right|.\end{aligned}$$

应用 Fubini 定理，可得

$$\mathbb{E}_{\boldsymbol{x}}\sup_{\boldsymbol{B}\in\mathcal{B}_o}\left|\sum_{i,j=1,i\neq j}^{n}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\boldsymbol{x}_j\right|\leq 4\cdot\mathbb{E}_{\boldsymbol{x},\boldsymbol{x}'}\sup_{\boldsymbol{B}\in\mathcal{B}_o}\left|\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{x}_i^\top\boldsymbol{B}_{ij}\boldsymbol{x}_j'\right|,$$

由此完成证明。 □

B.2 分块对角项的界

现在，我们来界定分块对角项 $D_{\mathcal{A}}(\boldsymbol{\xi})$。定义 $\mathcal{B}:=\{\boldsymbol{B}=\boldsymbol{A}^\top\boldsymbol{A}:\ \boldsymbol{A}\in\mathcal{A}\in\mathbb{R}^{m\times nd}\}$。回忆一下，$\mathcal{A}$ 是在（A8）之后定义的。对于任意 $\boldsymbol{A}\in\mathcal{A}$，我们有

$$\boldsymbol{B}=\boldsymbol{A}^\top\boldsymbol{A}=\begin{bmatrix}\boldsymbol{B}_{11}&\boldsymbol{B}_{12}&\cdots&\boldsymbol{B}_{1n}\\\boldsymbol{B}_{21}&\boldsymbol{B}_{22}&\cdots&\boldsymbol{B}_{2n}\\\vdots&\vdots&\ddots&\vdots\\\boldsymbol{B}_{n1}&\boldsymbol{B}_{n2}&\cdots&\boldsymbol{B}_{nn}\end{bmatrix}.$$

可以看出 $\boldsymbol{B}_{ij}=\boldsymbol{B}_{ji}$。由于 $\boldsymbol{B}$ 由 $\boldsymbol{A}$确定，我们将与 $\boldsymbol{B}=\boldsymbol{A}^\top\boldsymbol{A}$ 相关联的分块对角矩阵 $\mathbb{D}(\boldsymbol{A})=\text{Blkdiag}(\boldsymbol{B}_{11},\ \boldsymbol{B}_{22},\ \ldots\ldots,\ \boldsymbol{B}_{nn})$ 记为。于是，$D_{\mathcal{A}}(\boldsymbol{\xi})$ 可以改写为

$$D_{\mathcal{A}}(\boldsymbol{\xi})=\sup_{\boldsymbol{A}\in\mathcal{A}}\left|\sum_{i=1}^{n}\left(\boldsymbol{\xi}_i^\top(\boldsymbol{A}^i)^\top\boldsymbol{A}^i\boldsymbol{\xi}_i-\langle\boldsymbol{A}^i,\boldsymbol{A}^i\rangle\right)\right|=\sup_{\boldsymbol{A}\in\mathcal{A}}\left|\boldsymbol{\xi}^\top\mathbb{D}(\boldsymbol{A})\boldsymbol{\xi}-\text{Tr}(\mathbb{D}(\boldsymbol{A}))\right|.$$

我们有关于 $D_{\mathcal{A}}(\boldsymbol{\xi})$ 的如下结果。

定理 B.2。设 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\ldots,\boldsymbol{\xi}_n$ 是 $\mathbb{R}^d$中的独立随机向量。设 $\boldsymbol{\xi}=[\boldsymbol{\xi}_1^\top,\boldsymbol{\xi}_2^\top,\ldots,\boldsymbol{\xi}_n^\top]^\top$。假设 $\boldsymbol{\xi}$ 满足 Hanson-Wright 不等式（A4）。那么，存在一个常数 k 使得

$$\Pr\left[\sup_{A\in\mathcal{A}}D_{\mathcal{A}}(\boldsymbol{\xi})\geq t\right]\leq 2N\left(\mathcal{A},\|\cdot\|_F,\frac{t}{6nd}\right)\exp\left(-\frac{1}{ck^2}\min(mt^2,mt)\right),\tag{A18}$$

其中 $N(\mathcal{A},\|\cdot\|_F,u)$ 是相对于度量 $\|\cdot\|_F$的 $\mathcal{A}$ 的 u-覆盖的覆盖数。

证明。我们将证明分为两部分：

1. 在第一部分中，我们证明

$$\Pr\left[\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A})\boldsymbol{\xi} - \operatorname{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \geq t\right] \leq 2\exp\left(-\frac{1}{ck_1^2}\min(mt^2, mt)\right). \tag{A19}$$

由 Hanson-Wright 不等式（A4）可知，存在 $k_1 > 0$ 使得

$$\Pr\left[\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A})\boldsymbol{\xi} - \operatorname{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \geq t\right] \leq 2\exp\left(-\frac{1}{ck_1^2}\min\left(\frac{t^2}{\|\mathbb{D}(\boldsymbol{A})\|_F^2\|}, \frac{t}{\|\mathbb{D}(\boldsymbol{A})\|_{\text{op}}}\right)\right), \tag{A20}$$

我们只需证明

$$\sup_{\boldsymbol{A}\in\mathcal{A}} \|\mathbb{D}(\boldsymbol{A})\|_F^2 \leq \frac{1}{m}, \qquad \sup_{\boldsymbol{A}\in\mathcal{A}} \|\mathbb{D}(\boldsymbol{A})\|_{\text{op}} \leq \frac{1}{m}.$$

2. 在第二部分中，我们按照与Foucart 和 Rauhut [2013,定理 98]的证明类似的覆盖论证完成证明。

第一部分-步骤 *1*： 在此步骤中，我们对 $\sup_{\boldsymbol{A}\in\mathcal{A}} \|\mathbb{D}(\boldsymbol{A})\|_F^2$进行界定。对于每个 $\boldsymbol{A} \in \mathcal{A}$，均有$\|\mathbb{D}(\boldsymbol{A})\|_F^2 = \sum_{i=1}^n \|B_{ii}\|_F^2$。由于 $\boldsymbol{B} = \boldsymbol{A}^\top\boldsymbol{A}$ 且 $\boldsymbol{A} = \frac{1}{\sqrt{m}}\boldsymbol{R}_\Omega\boldsymbol{V}$满足 $\boldsymbol{V}$中给出的条件（A7），因此有

$$\boldsymbol{B}_{ii} = \frac{1}{m}(\boldsymbol{V}^i)^\top \boldsymbol{R}_\Omega^\top \boldsymbol{R}_\Omega \boldsymbol{V}^i, \tag{A21}$$

其中 $\boldsymbol{V}^i$ 是 $\boldsymbol{V}$的第 i 个块列。回顾 $\boldsymbol{R}_\Omega$ 的定义，可推得

$$\boldsymbol{B}_{ii} = \frac{1}{m}\sum_{j\in\Omega}(\boldsymbol{V}_j^i)^\top \boldsymbol{V}_j^i, \tag{A22}$$

其中 $\boldsymbol{V}_j^i$是 $\boldsymbol{V}^i$的第 j 行。由于 $\boldsymbol{B}$ 具有分块对角结构，因此有 $\|\mathbb{D}(\boldsymbol{A})\|_F^2 = \sum_{i=1}^n \|\boldsymbol{B}_{ii}\|_F^2$。注意到

$$\sum_{i=1}^n \|\boldsymbol{B}_{ii}\|_F^2 = \frac{1}{m^2}\sum_{i=1}^n \left(\left\|\sum_{j\in\Omega}(\boldsymbol{V}_j^i)^\top \boldsymbol{V}_j^i\right\|_F^2\right).$$

利用 $\boldsymbol{V}$在（A7）中的定义式，可得

$$\begin{aligned}\sum_{i=1}^n \left(\left\|\sum_{j\in\Omega}(\boldsymbol{V}_j^i)^\top \boldsymbol{V}_j^i\right\|_F^2\right) &\leq m\left(\sum_{i=1}^n \|\boldsymbol{a}_i\boldsymbol{a}_i^\top\|_F^2\right) + 2m\sum_{i\neq j}\|\boldsymbol{a}_i\boldsymbol{a}_i^\top\|_F\|\boldsymbol{a}_j\boldsymbol{a}_j^\top\|_F \\ &= m(\|\boldsymbol{a}_1\|^2 + \cdots + \|\boldsymbol{a}_n\|^2)^2,\end{aligned}$$

其中，最后一个不等式用到了 $\|\boldsymbol{a}_i\boldsymbol{a}_i\|_F = \|\boldsymbol{a}_i\|^2$。由于 $\|\boldsymbol{a}_1\|^2 + \cdots + \|\boldsymbol{a}_n\|^2 \leq 1$，因此可得

$$\sum_{i=1}^n \|\boldsymbol{B}_{ii}\|_F^2 \leq \frac{1}{m^2}\cdot m = \frac{1}{m},$$

这证明了 $\sup_{\boldsymbol{A}\in\mathcal{A}} \|\mathbb{D}(\boldsymbol{A})\|_F^2 \leq \frac{1}{m}$。–

第一部分-步骤 *2*： 在此步骤中，我们对 $\sup_{A\in\mathcal{A}} \|D(\boldsymbol{A})\|_{\text{op}}$ 进行界定。由于对于任意 $\boldsymbol{A} \in \mathcal{A}$，与 $\mathbb{D}(A)$ 相关联的 $\mathbb{D}(\boldsymbol{A})$ 是对称的，根据算子范数的定义，我们知道 $\|\mathbb{D}(\boldsymbol{A})\|_{\text{op}} =$

$\lambda_{\max}(\mathbb{D}(\boldsymbol{A}))$。由于 $\mathbb{D}(\boldsymbol{A})$ 具有对角形式 $\mathrm{Blkdiag}(\boldsymbol{B}_{11}, \boldsymbol{B}_{22}, \ldots, \boldsymbol{B}_{nn})$，因此可推得 $\lambda_{\max}(\mathbb{D}(\boldsymbol{A})) = \max\{\lambda_{\max}(\boldsymbol{B}_{ii}): \ i = 1, 2, \ldots, p\}$。根据式 (A22) 中给出的 $\boldsymbol{B}_{ii}$ 的表达式，由引理 B.

5 $\lambda_{\max}\ \boldsymbol{B}_{ii}\ \leq \frac{1}{m}$ $i = 1\ 2\ \ldots\ n$ $_{\boldsymbol{A}\in\mathcal{A}}\|\mathbb{D}\ \boldsymbol{A}\ \|_{\mathrm{op}} =$ 可得，对所有 , , , , 均成立，这意味着 sup ( ) (( )) 。 —

第二部分： 令 $\mathcal{A}_u$ 为集合 $\mathcal{A}$ 的最小 u-覆盖，其中 $\mathcal{A}$ 关于度量 $\|\cdot\|_F$的基数为 $N(\mathcal{A}, \ \|\cdot\|_F, u)$。对于所有 $\boldsymbol{A} \in \mathcal{A}$，均有 $\min_{A_u\in\mathcal{A}_u}\|\boldsymbol{A} - \boldsymbol{A}_u\|_F \leq u$ $\boldsymbol{A} \in \mathcal{A}$。由不等式（A19）可得

$$\begin{aligned}\Pr\left[\exists \boldsymbol{A}_u \in \mathcal{A}_u : \left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A}_u)\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}_u))\right| > \eta\right] &\leq \sum_{\boldsymbol{A}_u\in\mathcal{A}_u} \Pr\left[\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A}_u)\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}_u))\right| > \eta\right] \\ &\leq 2N(\mathcal{A}, \|\cdot\|_F, u) \exp\left(-\frac{1}{ck_1^2}\min(m\eta^2, m\eta)\right),\end{aligned}$$

其中第二个不等式由引理B.6得出。于是可得

$$\Pr\left[\sup_{\boldsymbol{A}_u\in\mathcal{A}_u}\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A}_u)\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}_u))\right| \leq \eta\right] \geq 1 - 2N(\mathcal{A}, \|\cdot\|_F, u) \exp\left(-\frac{1}{ck_1^2}\min(m\eta^2, m\eta)\right).$$

现在注意到，对于任意 $\boldsymbol{A} \in \mathcal{A}$，均有

$$\begin{aligned}&\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A})\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \\ &= \left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A}_u)\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}_u)) + \boldsymbol{\xi}^\top\big(\mathbb{D}(\boldsymbol{A}) - \mathbb{D}(\boldsymbol{A}_u)\big)\boldsymbol{\xi} + \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}_u)) - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \\ &\leq \left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A}_u)\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}_u))\right| + \left|\boldsymbol{\xi}^\top\big(\mathbb{D}(\boldsymbol{A}) - \mathbb{D}(\boldsymbol{A}_u)\big)\boldsymbol{\xi}\right| + \left|\mathrm{Tr}\big(\mathbb{D}(\boldsymbol{A}_u) - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}))\big)\right|.\end{aligned}$$

假设 $\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A}_u)\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \leq \eta$，则可得

$$\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A})\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \leq \eta + \|\mathbb{D}(\boldsymbol{A}) - \mathbb{D}(\boldsymbol{A}_u)\|_2\|\boldsymbol{\xi}\|_2^2 + nd\|\mathbb{D}(\boldsymbol{A}) - \mathbb{D}(\boldsymbol{A}_u)\|_2,$$

其中 $\boldsymbol{A}_u$ 的选取满足 $\|\boldsymbol{A} - \boldsymbol{A}_u\|_F \leq u$。现在注意到

$$\|\mathbb{D}(\boldsymbol{A}) - \mathbb{D}(\boldsymbol{A}_u)\|_2 \leq \|\mathbb{D}(\boldsymbol{A}) - \mathbb{D}(\boldsymbol{A}_u)\|_F \leq \|\boldsymbol{B} - \boldsymbol{B}_u\|_F,$$

and

$$\boldsymbol{B} - \boldsymbol{B}_u = (\boldsymbol{A} - \boldsymbol{A}_u)^\top(\boldsymbol{A} - \boldsymbol{A}_u) + \boldsymbol{A}_u^\top(\boldsymbol{A} - \boldsymbol{A}_u) + (\boldsymbol{A} - \boldsymbol{A}_u)^\top\boldsymbol{A}_u.$$

利用 $\|\boldsymbol{A}_u\|_F \leq 1$ 这一事实，可得

$$\|\mathbb{D}(\boldsymbol{A}) - \mathbb{D}(\boldsymbol{A}_u)\|_2 \leq \|\boldsymbol{B} - \boldsymbol{B}_u\|_F \leq u^2 + 2u$$

由于 $\boldsymbol{\xi}$ 满足 Hanson-Wright 不等式（A4），$\|\boldsymbol{\xi}\|_2^2 \leq nd$ 以至少 $1 - 2\exp(-nd)$ 的概率成立。因此，$\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A})\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \leq \eta + 2nd(u^2 + 2u)$。取 $\eta = \frac{t}{2}$和 $u = \frac{t}{6nd}$，利用 $u^2 + 2u < 3u$ 对 $u < 1$成立这一事实，我们可以推得

$$\Pr\left[\sup_{\boldsymbol{A}\in\mathcal{A}}\left|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{A})\boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{A}))\right| \geq t\right] \leq 2N\left(\mathcal{A}, \|\cdot\|_F, \frac{t}{6nd}\right)\exp\left(-\frac{1}{4ck_1^2}\min(mt^2, mt)\right),$$

至此，证明完成。 □

B.3 有用引理与定理

引理 B.5。设 $\boldsymbol{a} = [\boldsymbol{a}_1^\top, \boldsymbol{a}_2^\top, \ldots, \boldsymbol{a}_n^\top]^\top$ 是满足 $\|\boldsymbol{a}\| \leq 1$ 的向量，其中 $\boldsymbol{a}_i \in \mathbb{R}^d$。考虑矩阵 $\boldsymbol{P} = \sum_{i=1}^n k_i \boldsymbol{a}_i \boldsymbol{a}_i^\top$，其中 $k_i \in \{0,1\}$ $i = 1, 2, \ldots\ldots, n$。那么，$\boldsymbol{P}$ 的最大特征值满足 $\lambda_{\max}(\boldsymbol{P}) \leq 1$，对于 k_i 的任意可能组合均成立。

证明。可以看出，矩阵 $\boldsymbol{P}$满足 $\boldsymbol{P}^\top = \boldsymbol{P}$。因此，$\lambda_{\max}(\boldsymbol{P}) = \|\boldsymbol{P}\|$。由此可得

$$\lambda_{\max}(\boldsymbol{P}) = \|\boldsymbol{P}\| \leq \sum_{i=1}^n k_i \|\boldsymbol{a}_i \boldsymbol{a}_i^\top\| \leq \sum_{i=1}^n k_i \|\boldsymbol{a}_i \boldsymbol{a}_i^\top\|_F \leq \sum_{i=1}^n \|\boldsymbol{a}_i \boldsymbol{a}_i^\top\|_F.$$

注意到 $\|\boldsymbol{a}_i \boldsymbol{a}_i^\top\|_F = \|\boldsymbol{a}_i\|^2$。因此，我们有

$$\lambda_{\max}(\boldsymbol{P}) \leq \sum_{i=1}^n \|\boldsymbol{a}_i \boldsymbol{a}_i^\top\|_F = \sum_{i=1}^n \|\boldsymbol{a}_i\|^2 = \|\boldsymbol{a}\|^2 \leq 1,$$

至此，证明完成。 □

引理 B.6。假设 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_n \in \mathbb{R}^d$ 是相互独立的各向同性次高斯随机向量，且每个 $\boldsymbol{\xi}_i$ 都满足

$$\Pr\left[|\boldsymbol{\xi}_i^\top \boldsymbol{P} \boldsymbol{\xi}_i - \mathbb{E}(\boldsymbol{\xi}_i^\top \boldsymbol{P} \boldsymbol{\xi}_i)| \geq \eta\right] \leq 2\exp\left(\frac{1}{c}\min\left\{\frac{\eta^2}{k^4\|\boldsymbol{P}\|_F^2}, \frac{\eta}{k^2\|\boldsymbol{P}\|_{\mathrm{op}}}\right\}\right)$$

对于任意半正定矩阵 $\boldsymbol{P} \in \mathbb{R}^{d\times d}$。考虑半正定矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_n \in \mathbb{R}^{d\times d}$。令 $\mathbb{D}(P) := \mathrm{Blkdiag}(\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_n)$，且 $\boldsymbol{\xi} = [\boldsymbol{\xi}_1^\top, \boldsymbol{\xi}_2^\top, \ldots, \boldsymbol{\xi}_n^\top]^\top$。那么，存在 $k_1 > 0$ 使得

$$\Pr\left[|\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{P}) \boldsymbol{\xi} - \mathbb{E}(\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{P}) \boldsymbol{\xi})| \geq \eta\right] \leq 2\exp\left(\frac{1}{c}\min\left\{\frac{\eta^2}{k^4\|\mathbb{D}(\boldsymbol{P})\|_F^2}, \frac{\eta}{k^2\|\mathbb{D}(\boldsymbol{P})\|_{\mathrm{op}}}\right\}\right). \tag{A23}$$

证明。由于 $\mathbb{D}(\boldsymbol{P})$ 具有特殊的分块对角结构，因此可以看出 $\boldsymbol{\xi}^\top \mathbb{D}(\boldsymbol{P}) \boldsymbol{\xi} - \mathrm{Tr}(\mathbb{D}(\boldsymbol{P}))$ $= \sum_{i=1}^n \left(\boldsymbol{\xi}_i^\top \boldsymbol{P}_i \boldsymbol{\xi}_i - \mathrm{Tr}(\boldsymbol{P}_i)\right)$。每个 $\boldsymbol{\xi}_i^\top \boldsymbol{P}_i \boldsymbol{\xi}_i - \mathrm{Tr}(\boldsymbol{P}_i)$ 都是均值为零的次指数随机变量。应用伯恩斯坦不等式 [Vershynin, 2018, 第 2], 章，可以推导出存在 $k_1 > 0$ 使得

$$\Pr\left[\left|\sum_{i=1}^n (\boldsymbol{\xi}_i^\top \boldsymbol{P}_i \boldsymbol{\xi}_i - \mathrm{Tr}(\boldsymbol{P}_i))\right| \geq t\right] \leq 2\exp\left(-\frac{1}{ck_1^2}\min\left(\frac{t^2}{\sum_{i=1}^n \|\boldsymbol{P}_i\|_F^2}, \frac{t}{\sup_i \|\boldsymbol{P}_i\|_{\mathrm{op}}}\right)\right).$$

利用以下事实 $\sum_{i=1}^n \|\boldsymbol{P}_i\|_F^2 = \|\mathbb{D}(\boldsymbol{P})\|_F^2$ 以及 $\sup_i \|\boldsymbol{P}_i\|_{\mathrm{op}} = \|\mathbb{D}(\boldsymbol{P})\|_{\mathrm{op}}$ 即可完成证明。 □

定理 B.3。设 $\boldsymbol{\xi} \in \mathbb{R}^d$ 是满足 Hanson-Wright 不等式的各向同性 1-次高斯随机向量

$$\Pr\left[|\boldsymbol{\xi}^\top \boldsymbol{A} \boldsymbol{\xi} - \mathbb{E}(\boldsymbol{\xi}^\top \boldsymbol{A} \boldsymbol{\xi})| \geq \eta\right] \leq 2\exp\big(-\frac{1}{c}\min\big\{\frac{\eta^2}{k^4\|\boldsymbol{A}\|_F^2}, \frac{\eta}{k^2\|\boldsymbol{A}\|_{\mathrm{op}}}\big\}\big), \qquad \forall \eta > 0, \tag{A24}$$

对于任意半正定矩阵 A。令 $\boldsymbol{\xi}'$ 为 $\boldsymbol{\xi}$ 的一个独立副本。那么，$\boldsymbol{\xi} - \boldsymbol{\xi}'$ 也满足 Hanson-Wright 不等式。

定理的证明 B.3： 为了构造证明，我们将使用以下引理，该引理来自Krahmer 等人 [2014]。

引理 B.7. 设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_d \in \mathbb{C}^m$ 和 $\mathbb{T} \in \mathbb{C}^m$。若 $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_d]^\top$ 是各向同性的 1-次高斯随机向量，且 $\boldsymbol{y} = \sum_{i=1}^d \xi_i \boldsymbol{x}_i$，则对每个 $p \geq 1$，

$$\left(\mathbb{E} \sup_{\boldsymbol{h} \in \mathbb{T}} |\langle \boldsymbol{h}, \boldsymbol{y} \rangle|^p\right)^{1/p} \leq c \left(\mathbb{E} \sup_{\boldsymbol{h} \in \mathbb{T}} |\langle \boldsymbol{h}, \boldsymbol{g} \rangle| + \sup_{\boldsymbol{h} \in \mathbb{T}} (\mathbb{E} |\langle \boldsymbol{h}, \boldsymbol{y} \rangle|^p)^{1/p}\right), \tag{A25}$$

其中 c 是一个常数，$\boldsymbol{g} = \sum_{j=1}^d g_j \boldsymbol{x}_j$ 是 $g_1, g_2, \ldots, g_d$ 个相互独立的标准正态随机变量。△

注意，$(\boldsymbol{\xi} - \boldsymbol{\xi}')^\top \boldsymbol{A} (\boldsymbol{\xi} - \boldsymbol{\xi}') = \boldsymbol{\xi}^\top \boldsymbol{A} \boldsymbol{\xi} + (\boldsymbol{\xi}')^\top \boldsymbol{A} \boldsymbol{\xi}' - (\boldsymbol{\xi}')^\top \boldsymbol{A} \boldsymbol{\xi} - \boldsymbol{\xi}^\top A \boldsymbol{\xi}'$。由于前两项满足汉森-赖特不等式，只需证明

$$\Pr\left[|\boldsymbol{\xi}^\top \boldsymbol{A} \boldsymbol{\xi}'| \geq \eta\right] \leq 2 \exp\left(-\frac{1}{c_1} \min\left\{\frac{\eta^2}{k_1^4 \|\boldsymbol{A}\|_F^2}, \frac{\eta}{k_1^2 \|\boldsymbol{A}\|_{\mathrm{op}}}\right\}\right), \qquad \forall \eta > 0. \tag{A26}$$

定义集合 $\mathbb{S} = \{\boldsymbol{A}\boldsymbol{x} : \ \|\boldsymbol{x}\| \leq 1, \ \boldsymbol{x} \in \mathbb{R}^d\}$。由于 $\boldsymbol{\xi}$ 是 1-次高斯的，随机变量 $\langle \boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}' \rangle$ 在给定 $\boldsymbol{\xi}'$ 的条件下也是次高斯的。因此，

$$\begin{aligned}
\left(\mathbb{E} |\langle \boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}' \rangle|^p\right)^{1/p} &= \left(\left(\mathbb{E}_{\boldsymbol{\xi}'} (\mathbb{E}_{\boldsymbol{\xi}} |\langle \boldsymbol{\xi}, \boldsymbol{A}\boldsymbol{\xi}' \rangle|^p)^{1/p}\right)^p\right)^{1/p} \overset{(a)}{\lesssim} \left(\mathbb{E}_{\boldsymbol{\xi}'} (\sqrt{p})^p \|\boldsymbol{A}\boldsymbol{\xi}'\|_2^p\right)^{1/p} \\
&= \sqrt{p} \left(\mathbb{E}_{\boldsymbol{\xi}'} \sup_{\boldsymbol{h} \in \mathbb{S}} |\langle \boldsymbol{h}, \boldsymbol{\xi}' \rangle|^p\right)^{1/p} \\
&\overset{(b)}{\lesssim} c\sqrt{p} \left(\mathbb{E} \sup_{\boldsymbol{h} \in \mathbb{S}} |\langle \boldsymbol{h}, \boldsymbol{g} \rangle| + \sup_{\boldsymbol{h} \in \mathbb{S}} (\mathbb{E} |\langle \boldsymbol{h}, \boldsymbol{\xi}' \rangle|^p)^{1/p}\right),
\end{aligned}$$

其中，不等式 (a) 用到了不等式 $\sup_{\theta \in \mathcal{S}^{d-1}} (\mathbb{E} |\langle X, \theta \rangle|^p)^{1/p} \leq L\sqrt{p}$，该不等式适用于一个 L-次高斯随机向量 X，其中 $\mathcal{S}^{n-1}$ 是 $\mathbb{R}^d$ 中的单位球面（参见 Krahmer 等人 [2014]）；不等式 (b) 用到了引理 B.7。接下来，我们对上述不等式右侧的两项进行界定。对于第一项，可以推导出

$$\mathbb{E} \sup_{\boldsymbol{h} \in \mathbb{S}} |\langle \boldsymbol{h}, \boldsymbol{g} \rangle| = \mathbb{E} \|\boldsymbol{A}\boldsymbol{g}\|_2 \leq (\mathbb{E} \|A\boldsymbol{g}\|_2^2)^{1/2} = \|\boldsymbol{A}\|_F.$$

对于第二项，有

$$\sup_{\boldsymbol{h} \in \mathbb{S}} (\mathbb{E} |\langle \boldsymbol{h}, \boldsymbol{\xi}' \rangle|^p)^{1/p} = \sup_{\boldsymbol{z}, \|\boldsymbol{z}\| \leq 1} (\mathbb{E} |\langle A\boldsymbol{z}, \boldsymbol{\xi}' \rangle|^p)^{1/p} \lesssim \sqrt{p} \sup_{\boldsymbol{z}, \|\boldsymbol{z}\| \leq 1} \|A\boldsymbol{z}\|_2 \leq \sqrt{p} \|A\|_{\mathrm{op}}.$$

现在，可以使用引理B.4来证明不等式 (A26)。 □

C 循环矩阵不满足矩阵限制等距性质

令 $\boldsymbol{\xi} \in \mathbb{R}^p$ 为服从复正态分布 $\mathcal{N}_{\mathcal{C}}(0, 1)$ 的向量。令 $\boldsymbol{C}_{\boldsymbol{\xi}}$ 为由 $\boldsymbol{\xi}$生成的循环矩阵，即:

$$\boldsymbol{C}_{\boldsymbol{\xi}} = \begin{bmatrix} \xi_p & \xi_{p-1} & \cdots & \xi_1 \\ \xi_1 & \xi_p^\top & \cdots & \xi_2 \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{p-1} & \xi_{p-2} & \cdots & \xi_p \end{bmatrix} \in \mathbb{R}^{p \times p}.$$

注意，$\boldsymbol{C}_{\boldsymbol{\xi}} = \frac{1}{\sqrt{p}}\boldsymbol{F}_p^*\mathrm{diag}(\boldsymbol{d})\boldsymbol{F}_p$，其中 $\boldsymbol{F}_p$ 为离散傅里叶变换，$\boldsymbol{F}_p[i,j] = \omega^{ij}$ 适用于 $0 \le i, j < p$ 和 $\omega = e^{-\iota 2\pi/p}$。此处，对角 $\boldsymbol{d}$ 与 $\boldsymbol{C}_{\boldsymbol{\xi}}$ 的奇异值相关，并且服从 $\boldsymbol{d} = \frac{1}{\sqrt{p}}\boldsymbol{F}_p\boldsymbol{\xi}$ 独立同分布$\sim$ $\mathcal{NC}(0, 1)$。

假设 $p = p_1p_2$，其中 p_1 为正整数，p_2。令 f_k 表示 $\boldsymbol{F}_p$ 的第 k 行。注意，对于所有 $1 \le k \le p$，f_k 对应于秩一矩阵 $\mathrm{mtx}(f_k) = \boldsymbol{a}\boldsymbol{b}^\top \in \mathbb{R}^{p_1\times p_2}$，因为它可以写成克罗内克积

$$f_k = \boldsymbol{a}\otimes\boldsymbol{b} \in \mathbb{R}^{p_1p_2}, \quad \text{where} \quad \boldsymbol{a} = \begin{bmatrix} \omega^0 \\ \omega^{p_2k} \\ \vdots \\ \omega^{(p_1-1)p_2k} \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} \omega^0 \\ \omega^{k} \\ \vdots \\ \omega^{(p_2-1)k} \end{bmatrix}.$$

引理 C.1. 考虑向量集合 $\mathcal{M} = (f_k^*/\sqrt{p})_{k=1}^p$，它们等价于 $\mathbb{R}^{p_1\times p_2}$ 中弗罗贝尼乌斯范数为单位的秩一矩阵。令 $\bar{\boldsymbol{C}} = \frac{1}{\sqrt{n}}\mathrm{Mask}_n(\boldsymbol{C}_{\boldsymbol{\xi}})$ 为具有 n 行的子采样归一化循环矩阵。令 $M_p = \max_{1\le k\le p} ||d_k|^2 - 1|$。我们有

$$M_p = \sup_{\boldsymbol{v}\in\mathcal{M}} \left|\|\bar{\boldsymbol{C}}\boldsymbol{v}\|_2^2 - 1\right|. \tag{A27}$$

具体而言，对于所有 $\delta \le 1$，存在一个常数 $0 < c < 1$，使得以至少 $1 - c^p$ 的概率，$M_p > \delta$ 和 $\boldsymbol{C}$ 不满足秩一矩阵上的 δ-限制等距性质。

引理 1 的证明： 设 $\boldsymbol{e}_k$ 为标准基的第 k 个元素。设 $\boldsymbol{A} = \frac{1}{\sqrt{n}}\mathrm{Mask}_n(\boldsymbol{F}_p^*)$ 为子采样归一化离散傅里叶变换矩阵。注意，矩阵 $\boldsymbol{a}_k$ 的每一列 A 都具有单位欧几里得范数，因为它包含 n 个具有 ω 的幂次的元素。现在，注意到

$$\frac{1}{\sqrt{n}}\mathrm{Mask}_n(\boldsymbol{C}_{\boldsymbol{\xi}})\frac{f_k^*}{\sqrt{p}} = A\mathrm{diag}(\boldsymbol{d})\frac{\boldsymbol{F}_p f_k^*}{p} = A\mathrm{diag}(\boldsymbol{d})\boldsymbol{e}_k = d_k\boldsymbol{a}_k.$$

因此，$\|\boldsymbol{C}f_k^*\|_{\ell_2} = |d_k|$ 当我们对 A27) 取最大值/最小值时，便得到结论 $1 \le k \le p$。

现在，设 $c = \Pr(||d_k|^2 - 1| \le 1)$。显然，由于 $d_k \sim \mathcal{CN}(0, 1)$ 无界，因此 $c < 1$。利用 d_k 的独立性，我们得到 $\Pr(M_p > 1) \ge 1 - c^p$。最后，注意到 $M_p > 1$ 意味着对于任意 $\delta \le 1$，在秩一矩阵上都不满足δ-RIP。□

第二部分：Bandit 问题的补充材料

在第一部分中，我们推导出了关于分块循环测量的 RIP 的一些一般性结果。在本部分中，我们利用所得结果来分析我们的老虎机问题。本部分的组织结构如下。

1. 在第D节中，我们证明，如果每一轮的随机上下文向量都满足 Hanson-Wright 不等式，那么任何确定性选择的上下文向量也满足该不等式。这一性质确保我们算法中所选的上下文向量始终满足 Hanson-Wright 不等式。

2. 在第E节中，我们使用第一部分 B 节中的 RIP 结果， 证明由非各向同性随机向量生成的循环矩阵满足块稀疏向量的受限特征值条件 (RE)。

3. 在第F节中，我们使用第E 节中的结果，分析正文第3 节中算法的遗憾上界。

4. 最后，在第G节中，我们提供正文中实验的一些细节。此外，我们还展示了一些额外的实验结果。

D 确定性选择的上下文向量满足 Hanson-Wright 不等式

在数据贫乏和数据丰富情形下的算法中，我们在每一轮次选择如下贪心动作：

$$a_t \in \arg\max_{a\in[K]} \langle \boldsymbol{x}_{t,a}, \hat{\boldsymbol{\theta}}_{i-1} \rangle, \qquad i \geq 1.$$

下面的定理保证，如果每个 $x_{t,a}, a \in [K]$ 都满足 Hanson-Wright 不等式，那么每个被选中的上下文向量$\boldsymbol{x}_{t,\ a_t}$ 也满足该不等式。这一性质将在我们算法的分析中发挥非常重要的作用。

定理 D.1. 假设 $\boldsymbol{x}_1,\ \boldsymbol{x}_2, \ldots, \boldsymbol{x}_K \in \mathbb{R}^d$ 是独立同分布随机向量，并且 $\boldsymbol{A} \in \mathbb{R}^{d\times d}$。令 $E_{\boldsymbol{x}} = \mathbb{E}[\boldsymbol{x}_i^\top \boldsymbol{A}\boldsymbol{x}_i]$。假设 x_i 满足 Hanson-Wright 不等式（以下将以等价形式给出）

$$\Pr\Big(|\boldsymbol{x}_i^\top \boldsymbol{A}\boldsymbol{x}_i - E_{\boldsymbol{x}}| \geq t\Big) \leq 2\exp\Big(-C\min\big\{\frac{t^2}{\|\boldsymbol{A}\|_F^2}, \frac{t}{\|\boldsymbol{A}\|_{\mathrm{op}}}\big\}\Big) \tag{A28}$$

$$\Longleftrightarrow \Pr\Big(|\boldsymbol{x}_i^\top \boldsymbol{A}\boldsymbol{x}_i - E_{\boldsymbol{x}}| \geq \sqrt{\tau}\|\boldsymbol{A}\|_F/\sqrt{C} + \tau\|\boldsymbol{A}\|/C|\Big) \geq 2e^{-\tau}. \tag{A29}$$

令 $\boldsymbol{\xi}$ 为从集合 $\{\boldsymbol{x}_a, a \in [K]\}$ 中确定性选取的随机向量。相同的 Hanson-Wright 界成立，至多相差 $\mathcal{O}(\log(2K))$，即

$$\Pr\Big(|\boldsymbol{\xi}^\top \boldsymbol{A}\boldsymbol{\xi} - \mathbb{E}[\boldsymbol{\xi}^\top \boldsymbol{A}\boldsymbol{\xi}]| \geq c\log(2K)t\Big) \leq 2\exp\Big(-C\min\big(\frac{t^2}{\|\boldsymbol{A}\|_F^2}, \frac{t}{\|\boldsymbol{A}\|_{\mathrm{op}}}\big)\Big).$$

证明。 通过并合界，我们有

$$\Pr\Big(|\boldsymbol{\xi}^\top \boldsymbol{A}\boldsymbol{\xi} - E_{\boldsymbol{x}}| \geq t\Big) \leq Q(t) := \min\left(1, 2K\exp\Big(-C\min\big\{\frac{t^2}{\|\boldsymbol{A}\|_F^2}, \frac{t}{\|\boldsymbol{A}\|_{\mathrm{op}}}\big\}\Big)\right).$$

记作 $R = |\boldsymbol{x}_i^\top A \boldsymbol{x}_i - E_{\boldsymbol{x}}|$。注意到

$$\begin{aligned}
\mathbb{E}[R] &= -\int_0^\infty t dQ(t) \\
&= -[tQ(t)]_0^\infty + \int_0^\infty Q(t) dt \\
&= \int_0^\infty \min\left(1, 2K \exp\left(-C \min\left\{\frac{t^2}{\|\boldsymbol{A}\|_F^2}, \frac{t}{\|\boldsymbol{A}\|}\right\}\right)\right) dt \\
&\leq \int_0^\infty \min\left(1, 2Ke^{-C\frac{t^2}{\|\boldsymbol{A}\|_F^2}}\right) dt + \int_0^\infty \min\left(1, 2Ke^{-C\frac{t}{\|\boldsymbol{A}\|_{\mathrm{op}}}}\right) dt.
\end{aligned}$$

令 $T_1 = \sqrt{\log(2K)/C}\|\boldsymbol{A}\|_F$。我们可以推导出

$$\begin{aligned}
\int_0^\infty \min\left(1, 2Ke^{-C\frac{t^2}{\|\boldsymbol{A}\|_F^2}}\right) dt &= T_1 + \int_{T_1}^\infty 2Ke^{-C\frac{t^2}{\|\boldsymbol{A}\|_F^2}} dt \\
&= T_1 + \int_0^\infty 2Ke^{-C\frac{(t+T_1)^2}{\|\boldsymbol{A}\|_F^2}} dt \\
&\leq T_1 + \int_0^\infty e^{-C\frac{t^2}{\|\boldsymbol{A}\|_F^2}} dt \\
&= T_1 + \frac{\sqrt{\pi}}{2}\|\boldsymbol{A}\|_F/\sqrt{C} \leq T_1 + \|\boldsymbol{A}\|_F/\sqrt{C}.
\end{aligned}$$

同样，令 $T_2 = \log(2K)\|\boldsymbol{A}\|/C$，则有

$$\begin{aligned}
\int_0^\infty \min\left(1, 2Ke^{-C\frac{t}{\|\boldsymbol{A}\|}}\right) dt &= T_2 + \int_{T_1}^\infty 2Ke^{-C\frac{t}{\|\boldsymbol{A}\|}} dt \\
&= T_2 + \int_0^\infty 2Ke^{-C\frac{t+T_1}{\|\boldsymbol{A}\|}} dt \\
&= T_2 + \int_0^\infty e^{-C\frac{t}{\|\boldsymbol{A}\|}} dt \\
&= T_2 + \|\boldsymbol{A}\|/C.
\end{aligned}$$

结合上述界，并利用 $\log(2K) \geq 1$ ($K > 1$，因为它是臂数)，我们得到

$$\begin{aligned}
\mathbb{E}[R] = \mathbb{E}[|\boldsymbol{x}_i^\top A \boldsymbol{x}_i - E_{\boldsymbol{x}}|] &\leq \left(\sqrt{\log(2K)} + 1\right)\|\boldsymbol{A}\|_F/\sqrt{C} + (\log(2K) + 1)\|\boldsymbol{A}\|/C \\
&\leq 4\log(2K)\left(\|\boldsymbol{A}\|_F/\sqrt{C} + \|\boldsymbol{A}\|/C\right).
\end{aligned}$$

现在，应用等价界（A29），我们有

$$\begin{aligned}
&\Pr\left(|\boldsymbol{\xi}^\top A \boldsymbol{\xi} - E_{\boldsymbol{x}}| \geq \sqrt{\tau}\|A\|_F/\sqrt{C} + \tau\|A\|/C\right) \leq 2Ke^{-\tau} \\
\iff &\Pr\left(|\boldsymbol{\xi}^\top A \boldsymbol{\xi} - E_{\boldsymbol{x}}| \geq \sqrt{\tau}\sqrt{\log(K)}\|A\|_F/\sqrt{C} + \tau\log(K)\|A\|/C\right) \\
\implies &\Pr\left(|\boldsymbol{\xi}^\top A \boldsymbol{\xi} - E_{\boldsymbol{x}}| \geq \sqrt{\tau}\sqrt{\log(2K)}\|A\|_F/\sqrt{C} + \tau\log(2K)\|A\|/C\right). \qquad \text{(A30)}
\end{aligned}$$

令 $E_{\boldsymbol{\xi}} = \mathbb{E}[\boldsymbol{\xi}^\top \boldsymbol{A} \boldsymbol{\xi}]$，我们可以推导出

$$|E_{\boldsymbol{\xi}} - E_{\boldsymbol{x}}| \leq |\boldsymbol{x}_i^\top A \boldsymbol{x}_i - E_{\boldsymbol{x}}| \leq 4\log(2K)\left(\|\boldsymbol{A}\|_F/\sqrt{C} + \|\boldsymbol{A}\|/C\right).$$

将其与（A30）结合起来，我们得到，以概率至少为 $1-2e^{\tau}$，

$$|\boldsymbol{\xi}^{\top}\boldsymbol{A}\boldsymbol{\xi}-E_{\boldsymbol{\xi}}| \leq \underbrace{\sqrt{\tau}\log(2K)\|A\|_F/\sqrt{C}+\tau\log(2K)\|A\|/C}_{\text{Probability}}+\underbrace{4\log(2K)\big(\|A\|_F/\sqrt{C}+\|A\|/C\big)}_{\text{Expectation}}$$
$$\leq 2\log(2K)\Big(\sqrt{\tau+16}\|A\|_F/\sqrt{C}+(\tau+16)\|A\|/C\Big).$$

这意味着，以至少 $1-2e^{\tau}$ 的概率，对于某个绝对常数 $c>0$，我们有

$$|\boldsymbol{\xi}^{\top}\boldsymbol{A}\boldsymbol{\xi}-E_{\boldsymbol{\xi}}| \leq c\log(2K)\Big(\sqrt{\tau}\|A\|_F/\sqrt{C}+\tau\|A\|/C\Big).$$

现在，利用 (A28) 与 (A29) 的等价性，我们证明

$$\Pr\left(|\boldsymbol{\xi}^{\top}\boldsymbol{A}\boldsymbol{\xi}-E_{\boldsymbol{\xi}}| \geq c\log(2K)t\right) \leq 2\exp\Big(-C\min\Big\{\frac{t^2}{\|\boldsymbol{A}\|_F^2},\frac{t}{\|\boldsymbol{A}\|_{\text{op}}}\Big\}\Big).$$

□

E 从限制等距性质到受限特征值条件（RE）

在定理 A.2中，我们已证明，在某个假设下，以下形式的块托普利兹矩阵

$$\boldsymbol{\Xi}=\frac{1}{\sqrt{m}}\begin{bmatrix}\boldsymbol{\xi}_h^{\top} & \boldsymbol{\xi}_{h-1}^{\top} & \cdots & \boldsymbol{\xi}_1^{\top}\\ \boldsymbol{\xi}_{h+1}^{\top} & \boldsymbol{\xi}_h^{\top} & \cdots & \boldsymbol{\xi}_2^{\top}\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{\xi}_{m+h-1}^{\top} & \boldsymbol{\xi}_{m+h-2}^{\top} & \cdots & \boldsymbol{\xi}_m^{\top}\end{bmatrix}. \tag{A31}$$

当 m 时满足限制等距性质 $\geq O(s_0\log^2(s_0)\log^2(hd))$，其中每个 ξ_i 都假设为各向同性次高斯
。然而，对于我们的上下文老虎机问题，每一步的上下文向量 $\boldsymbol{x}_{t,\,a},\ a\in[K]$并非各向同性。根据假设1 以及第 3节中的算法，可以由定理 D.1 推导出，在我们的老虎机问题中，由所选上下文向量生成的测量矩阵具有以下形式

$$\boldsymbol{\Psi}=\boldsymbol{\Xi}\begin{bmatrix}\sqrt{\Sigma} & & & \\ & \sqrt{\Sigma} & & \\ & & \ddots & \\ & & & \sqrt{\Sigma}\end{bmatrix}+\mathbf{1}_m\boldsymbol{\mu}^{\top}, \tag{A32}$$

其中 $\lambda_{\max}(\Sigma)\geq\lambda_{\min}(\Sigma)>\sigma>0$。该测量矩阵 $\boldsymbol{\Psi}$ 考虑了所选上下文向量的协方差矩阵 $\boldsymbol{x}_{t,\,a_t}$具有 Σ （它并非简单的单位矩阵），且它们具有非零均值 $\boldsymbol{\mu}$。如式（5）所示，我们的老虎机问题的核心变成了通过以下含噪测量来学习 s-块稀疏向量 $\boldsymbol{\phi}$

$$\boldsymbol{y}=\boldsymbol{\Psi}\boldsymbol{\phi}+\boldsymbol{\varepsilon}. \tag{A33}$$

本节将证明，如果 $\boldsymbol{\Xi}$满足 RIP，那么 $\boldsymbol{\Psi}$ 对于 s-块稀疏向量满足受限特征值条件。首先，我们来定义块稀疏向量的受限特征值条件。

定义 E.1 (块稀疏向量的受限特征值条件[2])。 定义集合 $\mathbb{C}_{\alpha}^{(d)}(S) := \{\boldsymbol{a} \in \mathbb{R}^{hd} : \|\boldsymbol{a}_{S^c}\|_{2,1}^{(d)} \leq \alpha\|\boldsymbol{a}_S\|_{2,1}^{(d)}\}$，其中 $\mathcal{S}$ 是 s-块稀疏向量 $\boldsymbol{a} \in \mathbb{R}^{hd}$ 的块支撑集， $\mathcal{S}^c$ 是 S的补集。矩阵 $\boldsymbol{M} \in \mathbb{R}^{m\times d}$ 在 $\mathcal{S}$上满足参数为 κ 和 α 的受限特征值 (RE)条件，当且仅当

$$\frac{1}{m}\|\boldsymbol{M}\Delta\|_2^2 \geq \kappa\|\Delta\|_2^2, \qquad \text{for all } \Delta \in \mathbb{C}_{\alpha}^{(d)}(\mathcal{S}).$$

注释 12。 在 Wainwright [2019], 第 7 章中给出了稀疏向量的受限特征值条件定义。这里，我们将该定义推广到块稀疏向量。

以下定理是本节的主要结果。

定理 E.1。 假设矩阵 $\boldsymbol{\Xi}$ 在（A31）中定义，满足块 RIP[3]，其受限等距常数为 δ_{2s}，概率为 $1-p$（对于某个 $p>0$），即

$$(1-\delta_{2s})\|\boldsymbol{a}\|_2^2 \leq \|\boldsymbol{\Xi}\boldsymbol{a}\|_2^2 \leq (1+\delta_{2s})\|\boldsymbol{a}\|_2^2,$$

对于任意 $2s$-块稀疏向量 $\boldsymbol{a} \in \mathbb{R}^{hd}$。设 $m > \frac{s(\sqrt{d}+\sqrt{h}+t\sqrt{2\log})(1+9\lambda_{\max}^2/\lambda_{\min}^2)}{1-\delta_{2s}}$，其中 $\lambda_{\max}^2$和 $\lambda_{\min}^2$分别是 Σ的最大和最小特征值。则以概率 $1-p-2e^{-ct^2}$，在（A32）中定义的矩阵满足块稀疏支撑集 $\mathcal{S}$上的受限特征值条件，即存在 $\kappa>0$ 使得

$$\|\boldsymbol{\Psi}\Delta\|_2^2 \geq \kappa\|\Delta\|_2^2, \qquad \text{for all } \Delta \in \mathbb{C}_{3}^{(d)}(\mathcal{S}). \tag{A34}$$

证明。 令 $\boldsymbol{\Psi}' = \boldsymbol{\Psi} - \mathbf{1}_m\boldsymbol{\mu}^\top$ 为零均值分量。接下来，我们分两步进行证明：

1. 首先， 我们证明存在 $\kappa'>0$，使得对于所有 $\Delta \in \mathbb{C}_3^{(d)}$ 都有 $\|\boldsymbol{\Psi}'\Delta\|_2^2 \geq \kappa'\|\Delta\|_2^2$ （$\mathcal{S}$ ）；
2. 然后，考虑非零均值部分，完成证明 $\mathbf{1}_m\boldsymbol{\mu}^\top$ 并证明 (A34)。

步骤 1： 我们先引入一些记号。记 $\mathcal{H} := \{1,2,\ldots,h\}$。令 $\mathcal{S}_0 = \mathcal{S}$为由 $\boldsymbol{a}$ 中 s 个范数最大的块所对应的块索引集合， $\mathcal{S}_1$ 为由 $\Delta_{\mathcal{H}\setminus\mathcal{S}_0}$ 中 s 个范数最大的块所对应的块索引集合。其他集合 $\mathcal{S}_2$、.....、$\mathcal{S}_k$ 以相同方式定义（$\mathcal{S}_k$ 可能包含少于 s 个块）。将 $\boldsymbol{\Xi} = [\boldsymbol{\Xi}_1, \boldsymbol{\Xi}_2, \ldots, \boldsymbol{\Xi}_h]$ 改写为 $\boldsymbol{\Xi}_i \in \mathbb{R}^{m\times d}$。我们以类似方式定义 $\boldsymbol{\Xi}_{\mathcal{S}_i}$、 $i=0$、1、.....、 k。

注意到

$$\|\boldsymbol{\Psi}'\Delta\|_2^2 = \boldsymbol{\Xi}\begin{bmatrix}\sqrt{\Sigma} & & & \\ & \sqrt{\Sigma} & & \\ & & \ddots & \\ & & & \sqrt{\Sigma}\end{bmatrix}\Delta := \boldsymbol{\Xi}\tilde{\Delta}.$$

由于 $\lambda_{\min}^2$ 一个 d $\lambda_{\max}^2$ 分别是 Σ的最小和最大特征值，则有

$$\lambda_{\min}\|\Delta_{\mathcal{S}^c}\|_{2,1}^{(d)} \leq \|\tilde{\Delta}_{\mathcal{S}^c}\|_{2,1}^{(d)} \leq \lambda_{\max}\|\Delta_{\mathcal{S}^c}\|_{2,1}^{(d)},$$

[2]回忆一下，给定 $\boldsymbol{x} = [\boldsymbol{x}_1^\top, \boldsymbol{x}_2^\top, \ldots, \boldsymbol{x}_h^\top]^\top$，其中每个 $\boldsymbol{x}_i \in \mathbb{R}^d$， $\ell_{2,1}^{(d)}$,范数定义为 $\|\boldsymbol{x}\|_{2,1}^{(d)} = \sum_{i=1}^h \|\boldsymbol{x}_i\|_2$。 $\boldsymbol{x}$ 的块支撑集 $\mathcal{S}$ 定义为 $\mathcal{S} := \{i \in \{1,2,\ldots,h\} : \|\boldsymbol{x}_i\|_2 \neq 0\}$。此外，给定集合 $Q \subseteq \{1,2,\ldots,h\}$, $\boldsymbol{a}_Q = [\tilde{\boldsymbol{a}}_1^\top, \tilde{\boldsymbol{a}}_2^\top, \ldots, \tilde{\boldsymbol{a}}_h^\top]$ $\tilde{\boldsymbol{a}}_i = \boldsymbol{a}_i$ $i \in Q$ $\tilde{\boldsymbol{a}}_i = 0$ i ，对所有情况均成立，且对任意其他情况也成立。[3]满足一般 sd-稀疏向量的 RIP 的设计矩阵，必然也满足块大小为 d的 s-块稀疏向量的 RIP。

and

$$\lambda_{\min}\|\Delta_{\mathcal{S}}\|_{2,1}^{(d)} \le \|\tilde{\Delta}_{\mathcal{S}}\|_{2,1}^{(d)} \le \lambda_{\max}\|\tilde{\Delta}_{\mathcal{S}}\|_{2,1}^{(d)},$$

进而可得

$$\tilde{\Delta} \in \{\boldsymbol{a} \in \mathbb{R}^{hd} : \|\boldsymbol{a}_{\mathcal{S}^c}\|_{2,1}^{(d)} \le \alpha\|\boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)}\} := \mathbb{C}_{\alpha}^{(d)}(\mathcal{S}). \tag{A35}$$

其中 $\alpha = 3\lambda_{\max}/\lambda_{\min}$。现在，只需证明对于所有 $\tilde{\Delta} \in \mathbb{C}_{\alpha}^{(d)}(\mathcal{S})$，$\|\boldsymbol{\Xi}\tilde{\Delta}\|_2^2 \ge \kappa'\|\tilde{\Delta}\|_2^2$。

接下来，可以推导出

$$\begin{aligned}
\|\boldsymbol{\Xi}\tilde{\Delta}\|_2^2 =& \langle \boldsymbol{\Xi}\tilde{\Delta}, \boldsymbol{\Xi}\tilde{\Delta}\rangle = \langle \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle + 2\langle \boldsymbol{\Xi}\tilde{\Delta}_{(\mathcal{S}_0\cup\mathcal{S}_1)^c}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle + \langle \boldsymbol{\Xi}\tilde{\Delta}_{(\mathcal{S}_0\cup\mathcal{S}_1)^c}, \boldsymbol{\Xi}\tilde{\Delta}_{(\mathcal{S}_0\cup\mathcal{S}_1)^c}\rangle \\
=& \langle \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle + 2\langle \sum_{i=2}^{k}\boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle + \langle \sum_{i=2}^{k}\boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}, \sum_{i=2}^{k}\boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}\rangle \\
=& \langle \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle + \sum_{j=2}^{k}\langle \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}\rangle + 2\langle \sum_{i=2}^{k}\boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle + \sum_{i\ne j}\tilde{\Delta}_{\mathcal{S}_i}^{\top}\boldsymbol{\Xi}^{\top}\boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_j}
\end{aligned} \tag{A36}$$

由于 $\boldsymbol{\Xi}$ 满足 $(2s, \delta_{2s})$-限制等距性质，因此有

$$\langle \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle \ge (1-\delta_{2s})\|\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\|_2^2, \tag{A37}$$

$$\langle \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}\rangle \ge (1-\delta_{2s})\|\tilde{\Delta}_{\mathcal{S}_i}\|_2^2. \tag{A38}$$

此外，$\boldsymbol{\Xi}$ 满足 $(2s, \delta_{2s})$-限制等距性质，这意味着对于任意不相交集合 $\mathcal{Q}_1$ 和 $\mathcal{Q}_2$，都有 $\|\boldsymbol{\Xi}_{\mathcal{Q}_1}^{\top}\boldsymbol{\Xi}_{\mathcal{Q}_2}\| \le \delta_{2s}$。因此可得

$$|\langle \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_i}, \boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle| = |\langle \boldsymbol{\Xi}_{\mathcal{S}_i}\tilde{\Delta}_{\mathcal{S}_i}, \boldsymbol{\Xi}_{\mathcal{S}_0\cup\mathcal{S}_1}\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\rangle| \le \delta_{2s}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\|\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\|_2, \tag{A39}$$

and

$$|\tilde{\Delta}_{\mathcal{S}_i}^{\top}\boldsymbol{\Xi}^{\top}\boldsymbol{\Xi}\tilde{\Delta}_{\mathcal{S}_j}| = |\tilde{\Delta}_{\mathcal{S}_i}^{\top}\boldsymbol{\Xi}_{\mathcal{S}_i}^{\top}\boldsymbol{\Xi}_{\mathcal{S}_j}\tilde{\Delta}_{\mathcal{S}_j}| \le \delta_{2s}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\|\tilde{\Delta}_{\mathcal{S}_j}\|_2. \tag{A40}$$

将 (A37)-(A40 代入 (A36) 可得

$$\begin{aligned}
\|\boldsymbol{\Xi}\tilde{\Delta}\|_2^2 \ge& (1-\delta_{2s})\|\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\|_2^2 + (1-\delta_{2s})\|\tilde{\Delta}_{\mathcal{S}_i}\|_2^2 \\
&- 2\delta_{2s}\big(\sum_{i=2}^{k}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\big)\|\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\|_2 - \delta_{2s}\big(\sum_{i=2}^{k}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\big)\big(\sum_{i=2}^{k}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\big) \\
\ge& (1-\delta_{2s})\|\tilde{\Delta}\|_2^2 \\
&- 2\delta_{2s}\big(\sum_{i=2}^{k}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\big)\|\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\|_2 - \delta_{2s}\big(\sum_{i=2}^{k}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\big)\big(\sum_{i=2}^{k}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2\big).
\end{aligned} \tag{A41}$$

注意，对于 $i \ge 2$，我们有

$$\|\tilde{\Delta}_{\mathcal{S}_i}\|_2 \le \sqrt{s}\|\tilde{\Delta}_{\mathcal{S}_i}\|_{2,\infty}^{(d)} \le \frac{1}{\sqrt{s}}\|\tilde{\Delta}_{\mathcal{S}_{i-1}}\|_{2,1}^{(d)},$$

随后，可以推导出

$$\begin{aligned}
\sum_{i=2}^{k}\|\tilde{\Delta}_{\mathcal{S}_i}\|_2 \le& \frac{1}{\sqrt{s}}\sum_{j=1}^{k-1}\|\tilde{\Delta}_{\mathcal{S}_j}\|_{2,1}^{(d)} \le \frac{1}{\sqrt{s}}\|\tilde{\Delta}_{\mathcal{S}_0^C}\|_{2,1}^{(d)} \\
\overset{(a)}{\le}& \frac{1}{\sqrt{s}}\alpha\|\tilde{\Delta}_{\mathcal{S}_0}\|_{2,1}^{(d)} \\
\le& \alpha\|\tilde{\Delta}_{\mathcal{S}_0}\|_2,
\end{aligned} \tag{A42}$$

where 不等式 (a) 可由（A35）得到。将（A42）代入（A41）可得

$$\begin{aligned}\|\boldsymbol{\Xi}\tilde{\Delta}\|_2^2 \geq & (1-\delta_{2s})\|\tilde{\Delta}\|_2^2 - 2\delta_{2s}\alpha\|\tilde{\Delta}_{\mathcal{S}_0}\|_2\|\tilde{\Delta}_{\mathcal{S}_0\cup\mathcal{S}_1}\|_2 - \delta_{2s}\alpha^2\|\tilde{\Delta}_{\mathcal{S}_0}\|_2^2 \\ \overset{(a)}{\geq} & \big(1-\delta_{2s}(1+\alpha)^2\big)\|\tilde{\Delta}\|_2^2,\end{aligned}$$

其中，不等式 (a) 用到了以下事实：$\|\Delta_{\mathcal{S}_0}\|_2 \leq \|\Delta\|_2$ 和 $\|\Delta_{\mathcal{S}_0\cup\mathcal{S}_1}\|_2 \leq \|\Delta\|_2$。令$\kappa' := 1-\delta_{2s}(1+\alpha)^2$，步骤 1 即完成。

步骤 2： 首先，我们有以下直接的引理。

引理 E.1. 以至少 $1-2e^{-ct^2}$ 的概率，对于所有 $\Delta \in \mathbb{C}_3^{(d)}(\mathcal{S})$，存在 $C>0$ 使得

$$|\mathbf{1}_m^\top \Psi' \boldsymbol{a}| \leq C\sqrt{s}(\sqrt{d}+\sqrt{\log h}+t).$$

证明。令 $\boldsymbol{z}^\top = \mathbf{1}_m^\top \Psi'$。注意到 i 个块 $\boldsymbol{z}_i \in \mathbb{R}^d$ 恰好是以下各项之和：m 个定义在 $\mathbb{R}^d$上的次高斯向量。通过标准集中性论证（参见 Vershynin [2018]），可知以概率 $1-2e^{-ct^2}$，

$$\|\boldsymbol{z}_i\|_2 \leq c(\sqrt{d}+t).$$

由于有 h 个这样的块，应用并合界可得，以相同的概率，

$$\|\boldsymbol{z}\|_{2,\infty}^{(d)} = \sup_{i\in[h]}\|\boldsymbol{z}_i\|_2 \leq c(\sqrt{d}+\sqrt{\log h}+t).$$

现在注意到，对于任意满足 ℓ_2 范数为 1 的 $\Delta \in \mathbb{C}_3^{(d)}(\mathcal{S})$，我们有

$$|\boldsymbol{z}^\top \Delta| \leq \|\Delta\|_{2,1}^{(d)}\|\boldsymbol{z}\|_{2,\infty}^{(d)}.$$

由于 $\Delta \in \mathbb{C}_3^{(d)}(\mathcal{S})$，可以推出 $\|\Delta\|_{2,1}^{(d)} \leq (1+3)\|\Delta_{\mathcal{S}}\|_{2,1}^{(d)} \leq 4\sqrt{s}$。于是可得

$$|\boldsymbol{z}^\top \Delta| \leq 4c\sqrt{s}(\sqrt{d}+\sqrt{\log h}+t),$$

由此证明了所需的结论。 □

回顾一下，$m \gtrsim m_0 = 4s(\sqrt{d}+\sqrt{\text{对数}}\ h+t)^2$，其中 $\mu = 1-\delta_{2s}(1+9\lambda_{\max}^2/\lambda_{\min}^2)$。接下来，我们声称，对于 $\boldsymbol{a}$的所有取值，$\mathbf{1}_m$ 与 $\Psi'^{\mu}\boldsymbol{a}$ 之间的相关系数 ρ 均不超过 0.5。为此，

$$\begin{aligned}\rho(\mathbf{1}_m, \Psi'\boldsymbol{a}) &= \frac{\mathbf{1}_m^\top \Psi'\boldsymbol{a}}{\sqrt{m}\|\Psi'\boldsymbol{a}\|_2} \\ &\overset{(a)}{\leq} \frac{C\sqrt{s}(\sqrt{d}+\sqrt{\log h}+t)}{\sqrt{m}\sqrt{\mu}} \\ &\leq 0.5,\end{aligned}$$

其中，不等式 (a) 用到了步骤 1 和引理 E.1的结果。

由此，我们可以如下完成证明。首先注意，若 $\boldsymbol{a},\boldsymbol{b}$ 是两个相关性至多为 0.5 的向量，则对于某个常数 $c>0$，有 $\|\boldsymbol{a}+\boldsymbol{b}\|_2 \geq c(\|\boldsymbol{a}\|_2+\|\boldsymbol{b}\|_2)$。基于这一点，我们写出

$$\begin{aligned}\|\Psi\boldsymbol{a}\|_2^2 &= \|\Psi'\boldsymbol{a}+\mathbf{1}_m\boldsymbol{\mu}^\top\boldsymbol{a}\|_2^2 \\ &\geq c(\|\Psi'\boldsymbol{a}\|_2^2 + \|\mathbf{1}_m\boldsymbol{\mu}^\top\boldsymbol{a}\|_2^2) \\ &\geq c\|\Psi'\boldsymbol{a}\|_2^2\end{aligned}$$

代入我们在步骤 1 中得到的不等式，即完成证明。 □

F 第 3 节的补充材料3

首先，我们介绍一些对分析有用的定义和结果。考虑如下块稀疏线性回归模型：

$$\boldsymbol{y} = \boldsymbol{\Psi}\boldsymbol{a} + \boldsymbol{\varepsilon},$$

其中，$\boldsymbol{a} \in \mathbb{R}^{hd}$和 $\|\boldsymbol{a}\|_0^{(d)} = s \ll d$、$\boldsymbol{\Psi} \in \mathbb{R}^{m\times hd}$，以及 $\boldsymbol{\varepsilon} = [\varepsilon_1, \ldots, \varepsilon_m]$ 中的噪声 ε_i 相互独立，且服从 1-次高斯分布。令 $\mathcal{S}$ 为 $\boldsymbol{a}$的块支撑集。定义 Lasso 程序：

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}\in\mathbb{R}^d}{\arg\min} \left(\frac{1}{2m} \|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{a}\|_2^2 + \lambda_m \|\boldsymbol{a}\|_{2,1}^{(d)} \right) \tag{A43}$$

其中正则化参数为 λ_m。

以下定 理给出了解 $\hat{\boldsymbol{a}}$ 与真实 $\boldsymbol{a}$之间的 ℓ_2-误差界 .

定理 F.1. 假设矩阵 Ξ 在支撑集 $\mathcal{S}$ 上满足受限特征值条件，参数为 $(\kappa,3)$。那么，任何满足正则化参数下界为 $\lambda_m \geq 2\|\frac{\boldsymbol{\Psi}^\top \boldsymbol{\varepsilon}}{m}\|_2^{(d)}{}_{,\infty}$ 的 Lasso 程序 (A43) 的解都满足

$$\|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_2 \leq \frac{3}{\kappa}\sqrt{s}\lambda_m, \qquad \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_1 \leq 4\sqrt{s}\|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_2.$$

注 13. 定理 7.13 中也有类似的结果，见 Wainwright [2019], ，其中给出了稀疏向量恢复的一个界限。在我们的情形中，目标是重构一个块稀疏向量。我们的定理与 Wainwright [2019] 中的定理仅有以下区别：1）在我们的情形中，$\boldsymbol{\Psi}$ 需要满足针对块稀疏向量的受限特征值条件；2）正则化参数需要由一个 $\ell_{2,\infty}$ 范数界定，而不是仅由 ℓ_∞ 范数界定，$\frac{\boldsymbol{\Psi}^\top \boldsymbol{\varepsilon}}{n}$。下面给出的证明是对 Wainwright [2019]中定理 7.13 的推广。

证明。记 $\Delta = \hat{\boldsymbol{a}} - \boldsymbol{a}$。第一步是证明，在条件 $\Delta \in \mathbb{C}_3^{(d)}$ ($\mathcal{S}$下，误差向量满足 $\lambda_m \geq 2\|\frac{\boldsymbol{\Psi}^\top \boldsymbol{\varepsilon}}{m}\|_2^{(d)}{}_{,\infty}$。为此，定义 $L(\boldsymbol{a}, \lambda_m) = \frac{1}{2m}\|\boldsymbol{y} - \boldsymbol{\Psi}\boldsymbol{a}\|_2^2 + \lambda_m \|\boldsymbol{a}\|_{2,1}^{(d)}$。由于 $\hat{\boldsymbol{a}}$ 满足 (A43)，因此有

$$L(\hat{\boldsymbol{a}}, \lambda_m) \leq L(\boldsymbol{a}, \lambda_m) = \frac{1}{2m}\|\boldsymbol{\varepsilon}\|_2^2 + \lambda_m \|\boldsymbol{a}\|_{2,1}^{(d)}.$$

由于 $L(\hat{\boldsymbol{a}}, \lambda_m) = \frac{1}{2m}\|\boldsymbol{y} - \boldsymbol{\Psi}\hat{\boldsymbol{a}}\|_2^2 + \lambda_m \|\hat{\boldsymbol{a}}\|_{2,1}^{(d)}$，，因此可以推导出

$$\frac{1}{2m}\|\boldsymbol{\Psi}\Delta\|_2^2 \leq \frac{\boldsymbol{\varepsilon}^\top \boldsymbol{\Psi}\Delta}{m} + \lambda_m \left(\|\boldsymbol{a}\|_{2,1}^{(d)} - \|\hat{\boldsymbol{a}}\|_{2,1}^{(d)} \right). \tag{A44}$$

由于 $\boldsymbol{a}$ 是 s-块稀疏的，因此有 $\|\boldsymbol{a}\|_{2,1}^{(d)} = \|\boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)}$，，以及 $\|\hat{\boldsymbol{a}}\|_{2,1}^{(d)} = \|\hat{\boldsymbol{a}} - \boldsymbol{a} + \boldsymbol{a}\|_{2,1}^{(d)} = \|\Delta_{\mathcal{S}} + \Delta_{\mathcal{S}^c} + \boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)} = \|\Delta_{\mathcal{S}} + \boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)} + \|\Delta_{\mathcal{S}^c}\|_{2,1}^{(d)}$。随后可得

$$\|\boldsymbol{a}\|_{2,1}^{(d)} - \|\hat{\boldsymbol{a}}\|_{2,1}^{(d)} = \|\boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)} - \|\Delta_{\mathcal{S}} + \boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)} - \|\Delta_{\mathcal{S}^c}\|_{2,1}^{(d)}.$$

然后，由（A44）可得

$$\begin{aligned} \frac{1}{m}\|\boldsymbol{\Psi}\Delta\|_2^2 &\leq 2\frac{\varepsilon^\top \boldsymbol{\Psi}\Delta}{m} + 2\lambda_m \left(\|\boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)} - \|\Delta_{\mathcal{S}} + \boldsymbol{a}_{\mathcal{S}}\|_{2,1}^{(d)} - \|\Delta_{\mathcal{S}^c}\|_{2,1}^{(d)} \right) \\ &\leq 2\|\boldsymbol{\Psi}^\top \varepsilon/m\|_{2,\infty}^{(d)} \|\Delta\|_{2,1}^{(d)} + 2\lambda_m \left(\|\Delta_{\mathcal{S}}\|_{2,1}^{(d)} - \|\Delta_{\mathcal{S}^c}\|_{2,1}^{(d)} \right). \end{aligned}$$

在条件 $\lambda_m \geq 2\|\boldsymbol{\Psi}_m^\top \boldsymbol{\epsilon}\|_{2,\infty}^{(d)}$ 下，我们得到

$$\frac{1}{m}\|\boldsymbol{\Psi}\Delta\|_2^2 \leq \lambda_m\Big(3\|\Delta_{\mathcal{S}}\|_{2,1}^{(d)} - \|\Delta_{\mathcal{S}^c}\|_{2,1}^{(d)}\Big),$$

这意味着 $\Delta \in \mathbb{C}_3^{(d)}(\mathcal{S})$。应用受限特征值条件 $\frac{1}{m}\|\boldsymbol{\Psi}\Delta\|_2^2 \geq \kappa\|\Delta\|_2^2$，我们得到

$$\kappa\|\Delta\|_2^2 \leq \frac{1}{m}\|\boldsymbol{\Psi}\Delta\|_2^2 \leq 3\lambda_m\|\Delta_{\mathcal{S}}\|_{2,1}^{(d)} \leq 3\lambda_m\sqrt{s}\|\Delta_{\mathcal{S}}\|_2 \leq 3\lambda_m\sqrt{s}\|\Delta\|_2.$$

根据 $\mathbb{C}_3^{(d)}(\mathcal{S})$的定义，不等式 $\|\hat{\boldsymbol{a}}-\boldsymbol{a}\|_1 \leq 4\sqrt{s}\|\hat{\boldsymbol{a}}-\boldsymbol{a}\|_2$ 可直接得出。 □

定理 F.2 (Wedin [1972])。考虑两个矩阵 M 和 $M \in \mathbb{R}^{m\times n}$。考虑它们的如下奇异值分解（SVD）

$$\boldsymbol{M} = \begin{bmatrix}\boldsymbol{U}_1 & \boldsymbol{U}_2\end{bmatrix}\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_1^\top\\ \boldsymbol{V}_2^\top\end{bmatrix}, \qquad \hat{\boldsymbol{M}} = \begin{bmatrix}\hat{\boldsymbol{U}}_1 & \hat{\boldsymbol{U}}_2\end{bmatrix}\begin{bmatrix}\hat{\Sigma}_1 & 0\\ 0 & \hat{\Sigma}_2\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{V}}_1^\top\\ \hat{\boldsymbol{V}}_2^\top\end{bmatrix},$$

其中 Σ_1 和 $\Sigma_1 \in \mathbb{R}^{r\times r}$。令 σ_1、…、σ_k 和 $\hat{\sigma}_1$、…、$\hat{\sigma}_k$，其中 $k = \min\{m,n\}$，分别为 M 和 M的奇异值。若 $\min_{1\leq i\leq r}\sigma_i \geq \mu$ 和 $\min_{1\leq i\leq r,\ r+1\leq j\leq k}|\sigma_i - \hat{\sigma}_j| \geq \mu$，则

$$\|\sin\boldsymbol{\theta}(\boldsymbol{U}_1,\hat{\boldsymbol{U}}_1)\|_F^2 + \|\sin\boldsymbol{\theta}(\boldsymbol{V}_1,\hat{\boldsymbol{V}}_1)\|_F^2 \leq \frac{\|\boldsymbol{U}_1^\top\Delta\|_F^2 + \|\Delta\boldsymbol{V}_1\|_F^2}{\mu^2},$$

其中 $\Delta = \boldsymbol{M} - \hat{\boldsymbol{M}}$。

F.1 伪遗憾的定义

为说明伪遗憾为何可以定义为

$$R_T = \left[\sum_{t=1}^{T}\sum_{i=0}^{h-1} w_i\langle \boldsymbol{x}_{t,a_t^*} - \boldsymbol{x}_{t,a_t}, \theta\rangle\right], \tag{A45}$$

我们为该问题设定提供一种替代性解释。

令 $\bar{r}(x_{t,a_t}) := \sum_{i=0}^{h-1} w_i \boldsymbol{x}_{t,a_t}^\top\theta$，可以看出，$\bar{r}(\boldsymbol{x}_{t,a_t})$ 是动作 a_t 产生的总奖励。然而，智能体在第 t 轮次采取该动作后，并不会立即收到全部奖励。相反，奖励会分散在至多 h 轮次内。在第 $k \in \{t, t+1, \ldots, t+h-1\}$ 轮次，智能体从动作 a_t 中收到的奖励部分为 $w_{k-t}\boldsymbol{x}_{t,a_t}^\top\theta$。因此，可以看出，在 T 轮次期间的总奖励 $\sum_{t=1}^{T} r_t$ 可计算为

$$\sum_{t=1}^{T} r_t = \sum_{t=1}^{T}\bar{r}(\boldsymbol{x}_{t,a_t}) - r, \tag{A46}$$

其中，r是动作 a_{T-h+1}、a_{T-h+2}、……、a_T在时域 T之后应为智能体带来的奖励。由于 r有界且独立于 T，因此，要最大化 $\sum_{t=1}^{T} r_t$，等价于在每一轮次选择使 $a_t \in \mathcal{A}$ 最大化的动作 $\bar{r}(\boldsymbol{x}_{t,\ a_t})$。由于 $\sum_{i=0}^{h-1} w_i$是一个常数，因此每一轮次的最优动作就是 $a_t = \text{argmax}_{a\in\mathcal{A}}\boldsymbol{x}_{t,\ a}^\top\theta$。因此，遗憾可定义为（A45）。

F.2 定理 2

定理 F.3。在算法1中，令 $L = csd\log^2(sd)\log^2(hd)$。回顾一下，$\boldsymbol{\phi}_{\mathcal{S}_i}$ 是由 $2^{i-1}L$ 的前 ϕ 个条目组成的向量，而 $\hat{\boldsymbol{\theta}}_i$ 是在每个倍增周期 $\boldsymbol{\theta}_i$ 对 i的估计。则以至少 $1-\gamma$的概率，有：

$$\left|\sin\boldsymbol{\Theta}\left(\hat{\boldsymbol{\theta}}_i, \frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|}\right)\right| \leq \begin{cases} \frac{6\kappa d\sqrt{2s}}{\|\boldsymbol{w}_{\mathcal{S}_i}\|_2}\left(\sqrt{\frac{2\log(2^iL/\gamma)}{2^{i-1}L}}\right), & \text{if } \|\boldsymbol{w}_{\mathcal{S}_i}\|_2^2 \neq 0, \\ \frac{\pi}{2} & \text{if } \|\boldsymbol{w}_{\mathcal{S}_i}\|_2^2 = 0. \end{cases} \tag{A47}$$

注释 14. 请注意，本定理中的常数 c 对 K 的依赖至多为多对数级；具体而言，$c \lesssim \log^2(K)$，其中 K 是臂数。这是因为，根据定理 D.1，我们的算法所选取的上下文向量满足 Hanson-Wright 不等式，但 (C （A28）中的常数至多会缩小 $c'\log^2(K)$ 倍。

证明。 我们分两步构造证明：

1. 首先，我们使用定理 F.1 证明

$$\|\hat{\boldsymbol{\phi}}_{\mathcal{S}_i} - \boldsymbol{\phi}_{\mathcal{S}_i}\|_2 \lesssim 6\kappa d\sqrt{s}\left(\sqrt{\frac{2\log(2^iL/\gamma)}{2^{i-1}L}}\right). \tag{A48}$$

2. 然后，我们将定理 F.2 应用于（A48），并证明（A47）。

步骤 1：在假设 1 下，由于 $L \gtrsim sd\log^2(sd)\log^2(hd)$，给定于 (9) 中的 P_i' 和 P_i'' 的每个元素都满足 Hanson-Wright 不等式。此外，我们对 P_i' 和 P_i'' 的选择（相隔 $2^{i-1}L$ 行）确保 P_i'' 是 P_i' 的独立副本。接着，可以利用定理 B.3 和 E.1 证明，矩阵 $\boldsymbol{P}_i = P_i'' - P_i$ 满足参数为 (κ, 3) 的块稀疏向量的受限特征值条件。然后，只需证明 $\lambda_i \geq 2\|\frac{1}{2^{i-1}L}\bar{\boldsymbol{P}}_i^\top\boldsymbol{\epsilon}[i]\|_{2,\infty}$，其中 $\boldsymbol{\epsilon}[i] = \bar{\boldsymbol{Q}}_i\boldsymbol{\phi}_{\bar{\mathcal{S}}_i} + \boldsymbol{\varepsilon}[i]$ 定义于 (10)。为此，我们将证明

$$\|\frac{1}{2^{i-1}L}\bar{\boldsymbol{P}}_i^\top\boldsymbol{\epsilon}[i]\|_{2,\infty} \lesssim d\sqrt{\frac{2\log(2^idL/\gamma)}{2^{i-1}L}}. \tag{A49}$$

注意到

$$\left\|\frac{1}{2^{i-1}L}\bar{\boldsymbol{P}}_i^\top\boldsymbol{\epsilon}[i]\right\|_{2,\infty}^{(d)} \leq \frac{1}{2^{i-1}L}\left(\|\bar{\boldsymbol{P}}_i^\top\bar{\boldsymbol{Q}}_i\boldsymbol{\phi}_{\bar{\mathcal{S}}_i}\|_{2,\infty}^{(d)} + \|\bar{\boldsymbol{P}}_i^\top\boldsymbol{\varepsilon}[i]\|_{2,\infty}^{(d)}\right). \tag{A50}$$

为简化记号，令 $m_i := 2^{i-1}L$。回忆一下，$\bar{\boldsymbol{P}}_i^\top \in \mathbb{R}^{m_id\times m_i}$，将 $\bar{\boldsymbol{P}}_i^\top$ 的第 k 个块行[4] 记作 $\bar{\boldsymbol{P}}_i^{[k]} = [\boldsymbol{x}_1^{[k]}, \boldsymbol{x}_2^{[k]}, \ldots, \boldsymbol{x}_{m_i}^{[k]}]$，其中每个 $\boldsymbol{x}_j^{[k]} \in \mathbb{R}^d$（由于 $\bar{\boldsymbol{P}}_i$ 中的每个元素都是两个选定上下文向量之差，因此可以推得 $\|\boldsymbol{x}_j^{[k]}\|_\infty \leq 2$）。有

$$\begin{aligned}
\|\bar{\boldsymbol{P}}_i^\top\boldsymbol{\varepsilon}[i]\|_{2,\infty}^{(d)} &= \max_k\left\{\|\bar{\boldsymbol{P}}_i^{[k]}\boldsymbol{\varepsilon}[i]\|_2\right\} = \max_k\left\{\Big\|\sum_{j=1}^{m_i}\boldsymbol{x}_j^{[k]}\boldsymbol{\varepsilon}_j[i]\Big\|_2\right\} \\
&\leq \max_k\left\{\sqrt{d}\Big\|\sum_{j=1}^{m_i}\boldsymbol{x}_j^{[k]}\boldsymbol{\varepsilon}_j[i]\Big\|_\infty\right\} \\
&= \max_k\left\{\sqrt{d}\max_{\ell\in[d]}\Big\{\Big|\sum_{j=1}^{m_i}[\boldsymbol{x}_j^{[k]}]_\ell\cdot\boldsymbol{\varepsilon}_j[i]\Big|\Big\}\right\}.
\end{aligned} \tag{A51}$$

[4]给定一个矩阵 $A \in \mathbb{R}^{md\times n}$，可以将其行划分为 m 个大小相等的块 d，第 j 个块行 $j = 1, 2, \ldots\ldots, m$ 是由第 j 个块构成的矩阵。

由于 $\varepsilon_1[i]$, $\varepsilon_2[i]$, ..., $\varepsilon_{m_i}[i]$ 是相互独立的次高斯随机变量，利用如下事实：对于任意 j, K 和 ℓ，有 $\left|[\boldsymbol{x}_j^{[k]}]_\ell\right| < 2$，可以推得

$$\Pr\Big(\big|\sum_{j=1}^{m_i}[\boldsymbol{x}_j^{[k]}]_\ell \cdot \varepsilon_j[i]\big| \geq \eta\Big) \leq 2\exp\big(-\frac{\eta^2}{8m_i}\big).$$

令$\eta = 2\sqrt{2m_i\log(2dm_i/\gamma)}$且 $0 < \gamma < 1$，则可以推得，以至少$1-\gamma$的概率，有

$$\max_k \max_{\ell\in[d]} \big\{\big|\sum_{j=1}^{m_i}[\boldsymbol{x}_j^{[k]}]_\ell \cdot \varepsilon_j[i]\big|\big\} \leq 2\sqrt{2m_i\log(2dm_i/\gamma)}.$$

Su将此不等式代入（A51）可得，以至少 $1-\gamma$ ，

$$\|\bar{\boldsymbol{P}}_i^\top \varepsilon[i]\|_{2,\infty}^{(d)} \leq 2\sqrt{2^i dL\log(2^i dL/\gamma)}. \tag{A52}$$

¯进一步，将$\boldsymbol{Q}_i\phi_{\bar{\mathcal{S}}_i}$ 的第 j行记为ν_j，则可得

$$\|\bar{\boldsymbol{P}}_i^\top \bar{\boldsymbol{Q}}_i\phi_{\bar{\mathcal{S}}_i}\|_{2,\infty}^{(d)} = \max_k\Big\{\Big\|\sum_{j=1}^{m_i}\boldsymbol{x}_j^{[k]}\nu_j\Big\|_2\Big\}.$$

由于 $\|\boldsymbol{x}_j^{[k]}\|_\infty \leq 2$ 对所有 j 以及 k均成立，随机向量 $\boldsymbol{x}_1^{[k]}, \ldots, \boldsymbol{x}_{m_i}^{[k]}$ 是范数次高斯的，满足 Pr($\|x_j^{[k]}\|_2 \geq t] \leq 2e^{-\frac{t^2}{4d}}$。根据 Jin等人 的推论7 [2019]，存在一个绝对常数 $c > 0$ 使得

$$\Big\|\sum_{j=1}^{m_i}\boldsymbol{x}_j^{[k]}\nu_j\Big\|_2 \leq c\sqrt{d\sum_{j=1}^{m_i}\nu_j^2\log(4d/\gamma)} \tag{A53}$$

以至少 $1-\gamma$ 的概率成立。由于 $\boldsymbol{\phi} = \boldsymbol{w}\otimes\boldsymbol{\theta}$ 和 $\boldsymbol{w}$ 是 s-稀疏的，因此 $\boldsymbol{\phi}$ 至多是 s-块稀疏的。于是，$\boldsymbol{\phi}_{\bar{\mathcal{S}}_i}$ 至多是 s-块稀疏的。记 $\boldsymbol{\phi}_{\bar{\mathcal{S}}_i} := [\psi_1^\top, \psi_2^\top, \ldots, \psi_r^\top]^\top$，其中 $\psi_k \in \mathbb{R}^d$，且 r 是适当的维度。于是，$\bar{\boldsymbol{Q}}_i\boldsymbol{\phi}_{\bar{\mathcal{S}}_i}$ 的每一行都具有如下形式：$\sum_{\ell=1}^r \boldsymbol{\xi}_\ell^\top\psi_\ell$，其中 $\boldsymbol{\xi}_\ell$ 是在 $\boldsymbol{P}_i''$ 和 $\boldsymbol{P}_i'$ 中选定上下文向量之间相关的¯ ¯差值（见（9））。由于$\|\boldsymbol{x}_{t,\ a}\|_\infty \leq 1$，可以推得$|\boldsymbol{\xi}_\ell^\top\psi_\ell| \leq \|\boldsymbol{\xi}_\ell\|_\infty\|\psi_\ell\|_1 \leq \|\psi_\ell\|_1$。于是，我们有

$$\Big|\sum_{\ell=1}^r \boldsymbol{\xi}_\ell^\top\psi_\ell\Big| \leq \sum_{\ell=1}^r\|\psi_\ell\|_1 \leq \sqrt{d},$$

其中，第二个不等式用到了这一事实：$\sum_{\ell=1}^r\|\psi_\ell\|_1 \leq \sum_{i=1}^h w_i\sqrt{d}\|\theta\|_2 \leq \sqrt{d}$。¯ 随后，我们得到

$$|\nu_j| \leq \sqrt{d}, \qquad \forall j = 1, \ldots, m_i.$$

将此不等式代入（A53），并采用 Jin等人 [2019], 中的类似论证，可得

$$\|\bar{\boldsymbol{P}}_i^\top \bar{\boldsymbol{Q}}_i\phi_{\bar{\mathcal{S}}_i}\|_{2,\infty}^{(d)} \leq cd\sqrt{m_i\log(2dm_i/\gamma)} \leq cd\sqrt{2^i L\log(2^i dL/\gamma)}. \tag{A54}$$

以至少 $1-\gamma$ 的概率成立。

将 (A52) 和 (A54) 代入 (A50，可得 (A49)。然后，应用定理 F.1，可以证明 Lasso 程序 (11) 的解满足 (A48)。

步骤 2： 根据 $\Phi_{\mathcal{S}_i}$ （分别为 $\Phi_{\mathcal{S}_i}$）的定义，其 ith 列由 $((i-1)d+1)$-th 到 id-th 个条目构成，取自 ϕ （分别为 $\phi_{\mathcal{S}_i}$）。不等式 (A48) 表明

$$\|\hat{\Phi}_{\mathcal{S}_i}-\Phi_{\mathcal{S}_i}\|_F \le 6\kappa d\sqrt{s}\left(\sqrt{\frac{2\log(2^i dL/\gamma)}{2^{i-1}L}}\right). \tag{A55}$$

考虑 $\Phi_{\mathcal{S}_i}$ 和 $\Phi_{\mathcal{S}_i}$ 的奇异值分解

$$\Phi_{\mathcal{S}_i}=\begin{bmatrix}\boldsymbol{U}_1 & \boldsymbol{U}_2\end{bmatrix}\begin{bmatrix}\sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_1^\top\\ \boldsymbol{V}_2^\top\end{bmatrix}, \qquad \hat{\Phi}_{\mathcal{S}_i}=\begin{bmatrix}\hat{\boldsymbol{U}}_1 & \hat{\boldsymbol{U}}_2\end{bmatrix}\begin{bmatrix}\hat{\sigma}_1 & 0\\ 0 & \hat{\Sigma}_2\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{V}}_1^\top\\ \hat{\boldsymbol{V}}_2^\top\end{bmatrix}.$$

Since $\Phi_{\mathcal{S}_i}=\boldsymbol{\theta}\boldsymbol{w}_{\mathcal{S}_i}^\top$，可以看出 $U_1=\frac{\theta}{\|\theta\|}$。应用定理 F.2 可得

$$\begin{aligned}\left\|\sin\boldsymbol{\Theta}\Big(\hat{\theta},\frac{\theta}{\|\theta\|}\Big)\right\|_F^2 &\le \|\sin\boldsymbol{\theta}(\hat{\boldsymbol{U}}_1,\boldsymbol{U}_1)\|_F^2+\|\sin\boldsymbol{\theta}(\hat{\boldsymbol{V}}_1,\boldsymbol{V}_1)\|_F^2\\ &\le \frac{\|\boldsymbol{U}_1^\top\Delta\|_F^2+\|\Delta\boldsymbol{V}_1\|_F^2}{\mu^2}\\ &\le \frac{2\|\Delta\|_F^2}{\mu^2},\end{aligned}$$

其中 $\mu=\min\{\sigma_1,\min_{2\le j\le\min\{d,2^{i-1}L\}}|\sigma_1-\hat{\sigma}_j|\}$ 且 $\Delta=\Phi_{\mathcal{S}_i}-\Phi_{\mathcal{S}_i}$。根据 Weyl 定理（见 Wedin [1972] 中的定理 1），对于任意 $j\ge 2$，$|\sigma_j-\sigma_j|\le\|\Phi_{\mathcal{S}_i}-\Phi_{\mathcal{S}_i}\|_2$。$j\le\{d,2^{i-1}L\}$ 当 min 时，有 $|\sigma_j-\sigma_j|=|\sigma_j|\le\|\Phi_{\mathcal{S}_i}-\Phi_{\mathcal{S}_i}\|_2$。因此，可以推导出 $\mu\ge\sigma_1-\|\Phi_{\mathcal{S}_i}-\Phi_{\mathcal{S}_i}\|_2$。然后，由 (A55) 可得

$$\left\|\sin\boldsymbol{\Theta}\Big(\hat{\theta},\frac{\theta}{\|\theta\|}\Big)\right\|_F^2\le\frac{72\kappa^2sd^2}{\sigma_1^2}\left(\frac{2\log(2^idL/\gamma)}{2^{i-1}L}\right).$$

由于 $\Phi_{\mathcal{S}_i}=\boldsymbol{\theta}\boldsymbol{w}_{\mathcal{S}_i}^\top$，因此有 $\sigma_1=\|\boldsymbol{\theta}\|_2\|\boldsymbol{w}_{\mathcal{S}_i}\|_2$。若 $\|\boldsymbol{w}_{\mathcal{S}_i}\|_2\neq 0$，则有

$$\left|\sin\boldsymbol{\Theta}\Big(\hat{\boldsymbol{\theta}},\frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|}\Big)\right|\le\frac{6d\sqrt{2}\kappa\sqrt{s}}{\|\boldsymbol{\theta}\|_2\|\boldsymbol{w}_{\mathcal{S}_i}\|_2}\left(\sqrt{\frac{2\log(2^idL/\gamma)}{2^{i-1}L}}\right), \tag{A56}$$

其中用到了 $\sin\boldsymbol{\Theta}(\hat{\theta},\frac{\theta}{\|\theta\|})$ 是标量这一事实。若 $\|\boldsymbol{w}_{\mathcal{S}_i}\|_2=0$，$\sin\boldsymbol{\Theta}(\hat{\theta},\frac{\theta}{\|\theta\|})$ 可达到 $\frac{\pi}{2}$，由此完成证明。□

现在，我们可以开始证明定理 2。

定理 2 的证明： 证明分为两步：

1. 我们证明，在每一轮 t 中，选择次优动作 $a_t=\arg\max_{a\in[K]}\langle x_{t,a},\boldsymbol{\theta}_t\rangle$ 会导致总遗憾仅取决于 $\boldsymbol{\theta}_t$ 与真实 $\boldsymbol{\theta}$ 之间的角距离，即 $\sin\boldsymbol{\Theta}(\hat{\theta},\frac{\theta}{\|\theta\|})$。

2. 我们应用定理 F.3，并将倍增方案中每个轮次累积的遗憾相加，进而说明向量 $\boldsymbol{w}$ 中权重的分布如何影响最终遗憾。

步骤 1： 由于 $w_i > 0$ 对所有 $i = 1, 2, \ldots\ldots, h$ 均成立，因此有 $\|\boldsymbol{w}\|_1 = \sum_{i=1}^{h} w_i$。回想一下，$a_t^* = \arg\max_{a\in[K]} \langle \boldsymbol{x}_{t,a}, \boldsymbol{\theta}\rangle$，且动作 $a_t \in [K]$ 在轮次 t 产生的总奖励对累积奖励的贡献为 $\bar{r}(\boldsymbol{x}_{t,a_t}) = \|\boldsymbol{w}\|_1 \boldsymbol{x}_{t,a_t}^\top \theta$。因此，在轮次 t 采取动作 $a_t \in [K]$ 所产生的遗憾可计算为

$$\bar{r}(\boldsymbol{x}_{t,a_t^*}) - \bar{r}(\boldsymbol{x}_{t,a_t}) = \|\boldsymbol{w}\|_1 \boldsymbol{x}_{t,a_t^*}^\top \theta - \|\boldsymbol{w}\|_1 \boldsymbol{x}_{t,a_t}^\top \theta.$$

令 $\bar{\theta} = \frac{\theta}{\|\theta\|_2}$，于是可得

$$\bar{r}(\boldsymbol{x}_{t,a_t^*}) - \bar{r}(\boldsymbol{x}_{t,a_t}) = \|\boldsymbol{w}\|_1 \|\theta\|_2 (\boldsymbol{x}_{t,a_t^*}^\top \bar{\theta} - \boldsymbol{x}_{t,a_t}^\top \bar{\theta}).$$

注意到

$$\boldsymbol{x}_{t,a_t^*}^\top \bar{\boldsymbol{\theta}} - \boldsymbol{x}_{t,a_t}^\top \bar{\boldsymbol{\theta}} = \boldsymbol{x}_{t,a_t^*}^\top (\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}) - \boldsymbol{x}_{t,a_t}^\top (\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}) + (\boldsymbol{x}_{t,a_t^*} - \boldsymbol{x}_{t,a_t})^\top \hat{\boldsymbol{\theta}}.$$

由于 $\boldsymbol{x}_{t,a_t} = \arg\max_{a\in[K]} \langle \boldsymbol{x}_{t,a}, \theta\rangle$ 是估计得到的最优动作，因此有 $\langle \boldsymbol{x}_{t,a_t}, \boldsymbol{\theta}\rangle \geq \langle \boldsymbol{x}_{t,a_t^*}, \boldsymbol{\theta}\rangle$，这意味着

$$\boldsymbol{x}_{t,a_t^*}^\top \boldsymbol{\theta} - \boldsymbol{x}_{t,a_t}^\top \boldsymbol{\theta} \leq x_{t,a_t^*}^\top (\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}) - \boldsymbol{x}_{t,a_t}^\top (\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}) = (\boldsymbol{x}_{t,a_t^*} - \boldsymbol{x}_{t,a_t})^\top (\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}).$$

根据假设，$\|\boldsymbol{x}_{t,a}\|_2 \leq 1$ 对于所有 $a \in [K]$、$\|\boldsymbol{\theta}\|_2 \leq 1$ 和 $\|\boldsymbol{w}\|_1 \leq 1$。因此可得

$$\bar{r}(\boldsymbol{x}_{t,a_t^*}) - \bar{r}(\boldsymbol{x}_{t,a_t}) \leq 2\|\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}\|_2.$$

接下来，我们为 $\|\boldsymbol{\theta} - \boldsymbol{\theta}\|_2$ 提供了界。令 $\alpha =: \sin\Theta(\boldsymbol{\theta}, \boldsymbol{\theta})$，则可以推导出 that

$$\|\bar{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}\|_2 \leq 2\left|\sin\left(\frac{\arcsin\alpha}{2}\right)\right| \leq c|\alpha|,$$

可以推导出，当 $|\alpha|$ 较小时，$2\left|\sin\left(\frac{\arcsin\alpha}{2}\right)\right| \leq c|\alpha|$，其中 c 为某个常数；当 $|\alpha| = \frac{\pi}{2}$ 时，$2\left|\sin\left(\frac{\arcsin\alpha}{2}\right)\right| = \sqrt{2}$。因此，我们有

$$\bar{r}(\boldsymbol{x}_{t,a_t^*}) - \bar{r}(\boldsymbol{x}_{t,a_t}) \leq \begin{cases} 2c|\sin\Theta(\hat{\boldsymbol{\theta}}, \bar{\boldsymbol{\theta}})|, & \text{if } \alpha \text{ is small,} \\ 2\sqrt{2} & \text{if } \alpha = \frac{\pi}{2}. \end{cases} \tag{A57}$$

步骤 2： 回顾算法 1 中的倍增序列为

$$T_1, T_2, \ldots, T_m, \qquad \text{where } T_i = 4(2^i - 1)L,$$

且每个 T_i 都会估计 $\boldsymbol{\theta}_i$。回顾一下，遗憾满足

$$R_T = \left[\sum_{t=T_1+1}^{T} \sum_{i=0}^{h-1} w_i \langle \boldsymbol{x}_{t,a_t^*} - \boldsymbol{x}_{t,a_t}, \theta\rangle\right] \leq \sum_{t=1}^{T} \left(\bar{r}(\boldsymbol{x}_{t,a_t^*}) - \bar{r}(\boldsymbol{x}_{t,a_t})\right).$$

令 $T_0 = 0$ 并记 $\tilde{R}_k = \sum_{T_{k-1}+1}^{T_k} (\bar{r}(\boldsymbol{x}_{t,a_t^*}) - \bar{r}(\boldsymbol{x}_{t,a_t}))$ 对于 $k = 1, 2, \ldots, m$。可以看出 $R_1 \leq 4L$。对于 $k = 2, 3, \ldots, m$，由定理 F.3 和（A57）可知，若 $\|\boldsymbol{w}_{\mathcal{S}_{k-1}}\|_2 \neq 0$，

$$\begin{aligned} \tilde{R}_k &\leq \frac{6\kappa d\sqrt{2s}}{\|\boldsymbol{w}_{\mathcal{S}_{k-1}}\|_2} \left(\sqrt{\frac{2\log(2^{k-1}L/\gamma)}{2^{k-2}L}}\right)(T_k - T_{k-1}) \\ &= \frac{24\kappa d\sqrt{2s}}{\|\boldsymbol{w}_{\mathcal{S}_{k-1}}\|_2} \left(\sqrt{2\log(2^{k-1}dL/\gamma)}\right)\sqrt{2^k L}, \end{aligned}$$

并且 $\tilde{R}_k \le 2^{k+1}L$ 若 $\|\boldsymbol{w}_{\mathcal{S}_{k-1}}\|_2 = 0$。假设 k_1 是第一个满足 $\|\boldsymbol{w}_{\mathcal{S}_k}\|_2 \ge \mu$ 且 $0 \le \mu \le 1$的 k。那么，遗憾满足

$$\begin{aligned} R_T = \sum_{k=1}^{m} \tilde{R}_k &\le \sum_{k=1}^{k_1} 2^{k+1}L + \frac{24\kappa d\sqrt{2s}}{\mu} \sum_{k=k_1+1}^{m} \left(\sqrt{2\log(2^{k-1}dL/\gamma)}\right)\sqrt{2^k L} \\ &\le 4(2^{k_1}-1)L \cdot 2\sqrt{2} + \frac{24\kappa d\sqrt{2s}}{\mu}\left(\sqrt{2\log(T/\gamma)}\right)\frac{(\sqrt{2})^{m+1}-(\sqrt{2})^{k_1+1}}{\sqrt{2}-1}\sqrt{L}. \end{aligned}$$

由于 $T_i = 4(2^m - 1)L = T$，可以计算得出 $m = \log_2(\frac{T}{4L} + 1)$，这意味着

$$R_T \le 8\sqrt{2}(2^{k_1}-1)L + \frac{24\kappa d\sqrt{s}}{(\sqrt{2}-1)\mu}\left(\sqrt{2\log(dT/\gamma)}\right)\sqrt{T}.$$

对于任意 $\mu \in (0,1)$ 以及满足 $4(2^{k_1}-1)L \le T$ 的任意 k_1，总存在 $\alpha(\mu) \in (0,1)$，使得

$$h^{\alpha(\mu)} = 4(2^{k_1}-1)L.$$

因此，有

$$R_T \le 2\sqrt{2}\min\{h^{\alpha(\mu)}, T\} + \frac{24\kappa d}{(\sqrt{2}-1)\mu}\left(\sqrt{2sT\log(dT/\gamma)}\right), \tag{A58}$$

由此完成证明。 □

F.3 定理 3 的证明

定理 F.4. 在算法 2 中，令 $\phi[j]$ 为 ϕ 在 j 阶段的第 j 轮次中，当 $t > h$ 时的估计（即 $\phi[j]$ 是 Lasso 程序 14 的解）。令 $\boldsymbol{\theta}_j$ 为 $\boldsymbol{\theta}$ 在同一轮次的估计。则以至少 $1-\gamma$ 的概率，有

$$\left|\sin\boldsymbol{\Theta}\left(\hat{\boldsymbol{\theta}}_j, \frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|}\right)\right| \le 6\kappa\sqrt{2sd}\sqrt{\frac{2\log(2^j dh/\gamma)}{2^{j-1}h}}. \tag{A59}$$

该定理的证明与定理 F.3 的证明非常相似，因此这里省略 .

定理 3 的证明： 回顾一下，在算法2 中，倍增序列为

$$\tilde{T}_1, \tilde{T}_2, \ldots, \tilde{T}_p, \qquad \text{where } \tilde{T}_j = 2(2^j - 1)h,$$

和 $\boldsymbol{\theta}_j$ 在每个 T_j处进行估计。按照定理2证明中步骤 1 的类似推理，对于每个轮次 t 中的任意 j，均有

$$\bar{r}(\boldsymbol{x}_{t,a_t^*}) - \bar{r}(\boldsymbol{x}_{t,a_t}) \le 2c|\sin\boldsymbol{\Theta}(\hat{\boldsymbol{\theta}}_{j-1}, \bar{\boldsymbol{\theta}})|.$$

遗憾可分为两部分：1） h 轮次的初始阶段，以及 2）后续阶段，其中 $t > h$。初始阶段的遗憾可由定理2得出，即

$$R_h \le 2\sqrt{2}h^{\alpha(\mu)} + \frac{12\kappa}{(\sqrt{2}-1)\mu}\left(\sqrt{2sdh\log(dh/\gamma)}\right). \tag{A60}$$

ˆ初始阶段之后，根据定理F.3，估计的 $\boldsymbol{\theta}$ 满足

$$\left|\sin\boldsymbol{\Theta}\left(\hat{\boldsymbol{\theta}}, \frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|}\right)\right| \leq 6\kappa\sqrt{2sd}\left(\sqrt{\frac{2\log(2^m dL/\gamma)}{2^{m-1}L}}\right).$$

由于 m 是满足 $T_m = 2(2^m - 1)L \leq h$ 的最大整数，因此有

$$\left|\sin\boldsymbol{\Theta}\left(\hat{\boldsymbol{\theta}}, \frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|}\right)\right| \leq 12\kappa\sqrt{sd}\left(\sqrt{\frac{2\log(dh/\gamma)}{h}}\right).$$

因此，第二阶段中前 $2h$ 轮次累积的遗憾，记为 $R_{h+1:3h}$，满足

$$R_{h+1:3h} \leq 2h \cdot 12\kappa\sqrt{sd}\left(\sqrt{\frac{2\log(dh/\gamma)}{h}}\right) = 24\kappa\sqrt{2sdh\log(dh/\gamma)}. \tag{A61}$$

应用定理 3，时域 T 上的总遗憾可界定为 Then，应用定理3，可得

$$\begin{aligned}
R_T &\leq R_h + R_{h+1:3h} + \sum_{j=2}^{p} 2(\tilde{T}_j - \tilde{T}_{j-1})2c|\sin\boldsymbol{\Theta}(\hat{\boldsymbol{\theta}}_{j-1}, \bar{\boldsymbol{\theta}})| \\
&\leq R_h + R_{h+1:3h} + \sum_{j=2}^{p} 2^j h \cdot 6\kappa\sqrt{2sd}\sqrt{\frac{2\log(2^j dh/\gamma)}{2^{j-2}h}}, \\
&\leq R_h + R_{h+1:3h} + 48\kappa\sqrt{sd}\sum_{j=2}^{p}\sqrt{(2^{j-1})h\log(2^j dh/\gamma)}, \\
&\leq R_h + R_{h+1:3h} + 48\kappa\sqrt{sd\log(dT/\gamma)}\sum_{j=2}^{p}\sqrt{2^{j-1}h}.
\end{aligned}$$

由于 p 是满足 $\tilde{T}_j = (2^{p-1} + p)h \leq T$ 的最大整数，因此可得 $\sum_{j=2}^{p}\sqrt{2^{j-1}h} \leq \sqrt{T}$。结合式(A60) 和式 (A61)，有

$$\begin{aligned}
R_T \leq & 2\sqrt{2}h^{\alpha(\mu)} + \frac{12\kappa}{(\sqrt{2}-1)\mu}\sqrt{2sdh\log(h/\gamma)} \\
& + 24\kappa\sqrt{2sdh\log(dh/\gamma)} + 48\kappa\sqrt{sdT\log(dT/\gamma)},
\end{aligned}$$

由此完成证明。 □

G 实验：详情与更多结果

G.1 算法详情

本节首先进一步介绍第3.2节中提出的稀疏交替梯度下降（SA-GD）、单权重匹配追踪（SW-MP）和结合匹配追踪的 UCB（UCB-MP）。正如我们在第3.2节中所述，所有算法都采用与自适应 Lasso 相同的倍增技巧，以便进行公平比较。它们唯一的区别在于学习 $\boldsymbol{\theta}$ 和做出决策的方式不同。

SA-GD. 为详细说明 SA-GD 的工作原理，我们将（4）改写为

$$r_t = \boldsymbol{\theta}^\top Z_t \boldsymbol{w} + \varepsilon_t,$$

其中 $Z_t = [\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}, \ldots, \boldsymbol{\xi}_{t-h+1}]$。回顾一下，第 i 轮次从 $t = T_{i-1} + 1$ 开始，到 T_i 结束。接着，我们将第 i 轮次中的损失函数定义为

$$f_i(\boldsymbol{\theta}, \boldsymbol{w}) = \sum_{t=\tilde{T}_{i-1}+1}^{\tilde{T}_j} (r_t - \boldsymbol{\theta}^\top Z_t \boldsymbol{w})^2.$$

然后，$\boldsymbol{\theta}$ 和 $\boldsymbol{\phi}$ 估计为

$$\hat{\boldsymbol{\theta}}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \beta \nabla f_{i,\boldsymbol{\theta}}(\boldsymbol{\theta}^{(k)}, \boldsymbol{w}^{(k)}),$$
$$\hat{\boldsymbol{w}}^{(k+1)} = \boldsymbol{w}^{(k)} - \beta \nabla f_{i,\boldsymbol{w}}(\boldsymbol{\theta}^{(k)}, \boldsymbol{w}^{(k)}),$$

或者，其中 $f_{i,\boldsymbol{\theta}}$ 和 $f_{i,\boldsymbol{w}}$ 分别是 f_i 关于 $\boldsymbol{\theta}$ 和 $\boldsymbol{w}$ 的梯度。我们选择一个阈值 ϵ。当损失 $f_i(\hat{\boldsymbol{\theta}}^{(k^*)}, \hat{\boldsymbol{w}}^{(k^*)}) \le \epsilon$ 时，替代梯度下降算法停止，并将 $\boldsymbol{w}^{(k^*)}$ 投影到 s-稀疏子空间。

我们想指出，由于我们的问题设置较为复杂，大多数现有算法的表现都会逊于我们的方法，因为它们不会学习长时域奖励模式。例如，对于特殊情形 $w_0 = 0$（其中 $\boldsymbol{w} = [w_0, w_1, \ldots, w_h]^\top$），不难看出，经典的置信上界（UCB）和 OFUL Chu 等人 [2011], Abbasi-Yadkori 等人 [2011] 提出的方法都会失效，因为无法从中学到任何东西。为了进行更公平的比较，我们为经典算法增加了估计 $\boldsymbol{w}$中最大权重位置的能力，具体将在下文详述。

SW-MP。单权重匹配追踪算法采用了与经典匹配追踪相同的思路。我们不再估计向量中的所有非零值，而只寻找其中最大的非零值，因此将该算法称为单权重匹配追踪。

在每个轮次中，奖励由以下方式生成

$$\underbrace{\begin{bmatrix} r_{\tilde{T}_{i-1}+1} \\ r_{\tilde{T}_{i-1}+2} \\ \vdots \\ r_{\tilde{T}_i} \end{bmatrix}}_{\mathbf{r}} = \underbrace{\begin{bmatrix} \boldsymbol{\xi}_{\tilde{T}_{i-1}+1}^\top & \boldsymbol{\xi}_{\tilde{T}_{i-1}}^\top & \cdots & \boldsymbol{\xi}_{\tilde{T}_{i-1}-h+2}^\top \\ \boldsymbol{\xi}_{\tilde{T}_{i-1}+2}^\top & \boldsymbol{\xi}_{\tilde{T}_{i-1}+1}^\top & \cdots & \boldsymbol{\xi}_{\tilde{T}_{i-1}-h+3}^\top \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_{\tilde{T}_i}^\top & \boldsymbol{\xi}_{\tilde{T}_i-1}^\top & \cdots & \boldsymbol{\xi}_{\tilde{T}_i-h+1}^\top \end{bmatrix}}_{\Xi} \mathbf{w} \otimes \boldsymbol{\theta} + \boldsymbol{\varepsilon}.$$

SW-MP 的第一步是估计权重向量 $\boldsymbol{w}$ 中最大 w_i 的位置。为此，它会检验哪些块列（块列的示例见上方阴影区域）与奖励向量 $\boldsymbol{r}$ 的相关性最高。令 Ξ^k 为第 k 个块列。然后，我们在每个轮次结束时求解以下优化问题

$$k^* = \underset{k\in\{1,2,\ldots,h\}}{\arg\max} \frac{1}{\|\Xi^k\|}\|\boldsymbol{r}^\top \Xi^k\|$$

以找出 $\boldsymbol{w}$ 中最大权重的索引。然后，我们通过求解以下最小二乘问题来估计 $\boldsymbol{\theta}$

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\mathbb{R}^d}{\arg\min} \|\boldsymbol{r} - \Xi^{k^*}\boldsymbol{\theta}\|_2^2.$$

ˆ该估计的 $\boldsymbol{\theta}$ 将用于下一轮次的决策。值得注意的是，正如预期的那样，当 $\boldsymbol{w}$中非零元素较多时，该算法可能不准确，这一点也从图 6 中的 (a3)、(a4)、(b3) 和 (b4) 得到了印证。

UCB-MP。 该算法与 SW-MP 有一些相似之处。在每个轮次 i结束时，我们使用与 SW-MP 相同的方法来估计权重向量$\boldsymbol{w}$中最大 w_i 的位置，即

$$k^* = \underset{k\in\{1,2,\ldots,h\}}{\arg\max} \frac{1}{\|\Xi^k\|}\|\boldsymbol{r}^\top \Xi^k\|,$$

利用学习到的位置，我们使用 UCB Chu 等人 [2011] 来学习 $\boldsymbol{\theta}$ 并在下一轮次进行决策。具体而言，在每一轮中，我们仅使用 轮之前所选上下文向量及其对应的奖励来更新置信界。前。

G.2 更多结果

We 我们还进行了更多实验，以补充第 3.2 节中的实验。

首先，我们考虑权重向量 $\boldsymbol{w}$ 均匀随机生成的情形。我们研究了 $\boldsymbol{w}$ 的不同稀疏性。如图 8所示，我们的算法在大多数情况下仍表现最佳。有趣的是，即使 sd远大于 $h+d$，它的表现仍优于 SA-GD。这表明利用低秩结构可能存在问题。这些观察结果再次支持了我们在第 2.2 节和 3.2节关于低秩矩阵恢复中循环矩阵的讨论。

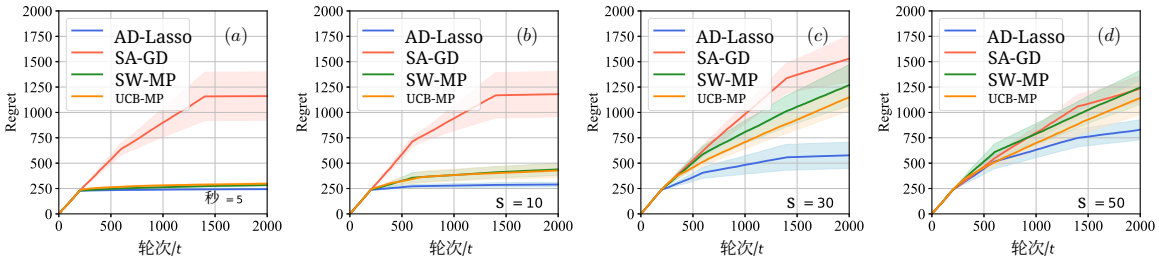

图 8：不同算法的性能比较。这里， s-稀疏向量 $\boldsymbol{w}$ 通过均匀随机生成，然后进行归一化。（通用参数： $T = 2000$， $h = 100$， $d = 5$和$\|\boldsymbol{w}\|_1 = 1$。阴影区域表示标准误。）

此外，我们研究了 $\boldsymbol{w}$中只有一个非零元素的情形。此情形对应于具有未知延迟的上下文线性赌博机。从图 9 可以看出，我们的算法表现几乎与 SW-MP 和 UCB-MP 一样好，这表明我们的算法在定位未知延迟方面能力出色。正如预期的那样，低秩恢复算法表现不佳。

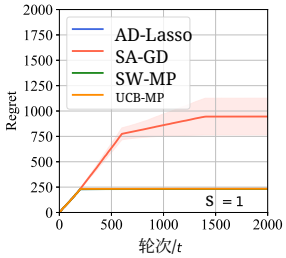

图 9：不同算法的性能比较。这里，$\boldsymbol{w}$ 仅包含一个非零项。($T = 2000$, $h = 100$, $d = 5$和 $\|\boldsymbol{w}\|_1 = 1$。)