arXiv:2301.00841v1 [stat.ML] 2 Jan 2023

排序差分隐私

Shirong Xu*、Will Wei Sun† 和 Guang Cheng‡,

摘要

排序数据在各种现实场景中被广泛收集，导致个人信息泄露，例如用户对视频或新闻的偏好。为保护排序，现有研究主要针对一组排序中的单个排名，或在 ϵ差分隐私下对排序进行成对比较，以实现隐私保护。本文提出一种名为 ϵ排序差分隐私的新概念，用于保护名次。我们建立了 Mallows 模型（Mallows，1957年）与所提出的 ϵ 排序差分隐私之间的联系。基于此，我们开发了一种多阶段排名算法，在满足所提出的 ϵ排序差分隐私的同时生成合成排序。我们还针对下游任务（包括推断攻击和个性化排名任务）中合成排序的效用，建立了理论结果。对于推断攻击，我们量化了 ϵ 对基于合成排序估计真实排名的影响。对于个性化排名任务，我们考虑用户之间不同的隐私偏好，并量化其隐私偏好对估计最优排序函数的一致性的影响。我们进行了大量数值实验，以验证理论结果并展示所提出的合成排序算法的有效性。

关键词： 差分隐私、学习理论、Mallows 模型、排序数据、合成数据

*加利福尼亚大学洛杉矶分校统计学系。电子邮件：shirong@stat.ucla.edu
†普渡大学克兰纳特管理学院。电子邮件：sun244@purdue.edu。
‡加利福尼亚大学洛杉矶分校统计学系。电子邮件：guangcheng@ucla.edu。

1 引言

排序数据常见于各种商业场景，例如推荐系统（Karatzoglou 等人，2013；Oliveira 等人，2020）和搜索引擎（Dwork 等人，2001；Liu 等人，2007）。通常，人们会收集并利用排序数据，学习用户对各项目的偏好，以便在未来提供恰当的推荐。推荐系统就是一个典型例子：它汇集用户行为的历史信息，为新用户推荐项目。然而，用户的排序通常高度敏感，因为它们会透露用户的购物偏好或政治偏好（Yang 等人，2019；Lee，2015）。因此，在将排序数据共享给外部方之前，亟需开发有效的机制来实现隐私保护。

为确保数据隐私保护，通常会采用差分隐私（DP；Dwork 2006）作为标准度量指标。差分隐私以数学语言形式化了隐私保证，无需对数据作出假设，并具备诸多良好特性，例如对后处理的免疫性，以及对隐私保护机制输出的脱敏数据进行隐私组合。通常，差分隐私可分为两大类：中心化差分隐私（CDP；Dwork 等人 2006）和本地差分隐私（LDP；Wang 等人 2017）。本地差分隐私在用户端扰动数据，并将隐私化数据提交给服务器；而中心化差分隐私则依靠可信数据收集者保护所有数据。本地差分隐私在隐私保护领域日益普及，目前已应用于多种实际场景，包括 Google Chrome 浏览器（Erlingsson 等人，2014）和 macOS（Tang 等人，2017）。

在相关文献中，研究者投入了大量精力，致力于开发有效的机制，以在差分隐私框架下保护排序数据，并建立基于隐私化排序数据进行统计推断的理论结果。根据采用本地差分隐私（LDP）还是中心化差分隐私（CDP），排序数据的隐私保护可分为两大类。在差分隐私的中心化模型中，可信管理者会收集所有用户的非隐私保护排序，并进行差分隐私统计推断（Hay 等人，

2017；Lee，2015；Shang 等人，2014；Lee，2015；S´anchez 等人，2016；Busa-Fekete 等人，2021）。具体而言，Lee（2015）开发了高效算法，用于在中心化差分隐私和策略操纵的条件下，推断项目的真实排名。Shang 等人（2014）提出使用高斯噪声污染收集到的排序直方图，以进行排名聚合。Hay 等人（2017）提出了若干差分隐私排名聚合算法，利用拉普拉斯噪声保护发布输出中的单个排名。Li 等人（2022）将拉普拉斯机制、随机响应机制和指数机制融入孔多塞投票，开发出一类新颖的随机化投票规则，以保护排序隐私。相比之下，在差分隐私的局部模型中，排序数据会先通过某些局部机制进行隐私化，然后再提交给数据管理者（Yang 等人，2019；Song 等人，2022；Yan 等人，2020；Alabi 等人，2022）。Yan 等人（2020）采用拉普拉斯噪声或随机响应机制，对名次的成对比较偏好进行随机置换。Song 等人（2022）提出在本地收集差分隐私排序以进行排名聚合，并通过向项目之间的成对比较添加高斯噪声来随机置换排序。

本文提出一种称为 ϵ-排序差分隐私的新概念，用于保护排序中任意单个项目的位置；其中，ϵ 越小，隐私保护越严格。具体而言，我们首先定义相邻排名：移除一个项目后，若两个排序中的项目具有相同的偏序，则这两个排序被视为相邻。合成排序算法以相邻排名作为输入时，应产生相似的输出分布。其主要动机在于，现有研究主要关注保护项目名次之间的成对比较，或一组排序中的单个排名，并在经典 ϵ-差分隐私框架下形式化隐私保证，而直接保护排序中项目位置的研究却很少。将经典 ϵ-差分隐私应用于排序的一个主要缺点是，对于排序这类多维对象，隐私保护通常依赖组合定理进行隐私度量（Shang 等人，2014；Jeong 等人，2022），这会

导致隐私界较松，且下游隐私-效用分析复杂。相比之下，所提出的 ϵ-排序差分隐私采用单个隐私参数 ϵ 来衡量排序的隐私保护程度，而不依赖组合定理（Kairouz 等人，2015），从而避免隐私预算计算不准确。在所构建的ϵ-排序差分隐私框架下，我们能够分析每个排序的隐私保证 ϵ 如何影响基于隐私保护排名的下游推断任务性能。例如，所提出的 ϵ-排序差分隐私支持个性化排名任务中的个性化隐私偏好，用户可以控制其排序的隐私偏好。此外，拉普拉斯噪声添加等经典数据扰动方法无法匹配名次的序数特性，会隐式地导致下游任务的隐私-效用权衡并非最优。我们进行了大量实验，以验证理论结果并展示合成排序算法的有效性。结果表明，该算法生成的隐私保护排名在下游个性化学习任务中具有更高效用，这隐式地源于更优的隐私-效用权衡。

本文的贡献总结如下。首先，我们开发了一种基于多阶段排名算法（Fligner 和 Verducci，1988年；Critchlow 等人，1991年）的合成排序算法，用于生成隐私保护排名。研究证明，在相同的隐私保证下，与线性拉普拉斯噪声添加相比，该算法能更好地保留排序中有关成对比较的信息。该算法本质上通过 Mallows 模型（Mallows，1957年）生成合成排序，从而建立了 Mallows 模型与 ϵ-排序差分隐私之间出人意料的联系。其次，我们通过两项下游任务分析合成排序的效用，包括针对排序的推断攻击和个性化排名任务。在推断攻击中，我们从理论上量化了 ϵ如何影响基于生成的合成排序对中心排名的估计，并给出了 ϵ 的最优速率；该速率根据合成排序的数量自适应变化，可防止对中心排名的错误推断。对于个性化排名任务，我们允许用户

拥有个性化隐私偏好，并从理论上量化用户的隐私偏好与估计最优排序函数时遗憾之间的关系。具体而言，在保证最优排序函数估计一致性的前提下，我们推导出隐私参数 ϵ 随用户数量自适应变化时的最优阶。理论结果表明，当所有用户选择阶为 $O\big(\sqrt{n^{-1}\log^{1+\zeta}(n)}\big)$ 的自适应隐私保证时，对于任意 $\zeta > 0$ ，都能保证最优排序函数估计的一致性。有趣的是，我们的理论结果与Duchi and Ruan (2018) 的结果相似，都定量说明了 ϵ 对局部隐私模型下估计收敛速率的影响。

本文其余部分安排如下。在第1.1节介绍一些必要的记号后，第2 节介绍差分隐私、排序数据和 Mallows 模型的背景。第3 节形式化定义 ϵ-排序差分隐私，并提出一种满足 ϵ-排序差分隐私的合成排序算法。第4 节建立关于用于生成合成排序的中心排名的推断攻击的理论结果。第5 节考虑将合成排序用于下游个性化排名任务的情形，并建立关于估计最优排序函数时一致性的理论结果，同时量化相应的隐私-效用权衡。第6节开展广泛实验，以验证我们的理论结果。第7 节给出简要总结，所有技术证明均见附录。

1.1 记号

对于正整数 n，记 $[n] = \{1, ..., n\}$ 为 n-集。对于集合 S，记 $|S|$ 为其基数。对于两个正数列 $\{f_n\}_{n=1}^{\infty}$ 和 $\{g_n\}_{n=1}^{\infty}$，若 $\limsup_{n\to\infty} |f_n|/g_n < +\infty$，则记为 $f_n = O(g_n)$。若 $f_n = O(g_n)$ 且 $g_n = O(f_n)$，则记为 $f_n \asymp g_n$。对于随机变量 X_n 和序列 $\{a_n\}_{n=1}^{\infty}$，若 $X_n = o_p(a_n)$ 依概率收敛于零，则记为 X_n/a_n ；若 X_n/a_n 随机有界，则记为 $X_n = O_p(a_n)$。对于整数 K，

$\Upsilon(K)$ 表示 (1，2，……，K) 的排列集合。令 $I(\cdot)$ 为指示函数，若 A 成立，则 $I(A)=1$，否则为 0。对于向量 $\boldsymbol{x}$，记 $\|\boldsymbol{x}\|_2$ 为其 l_2-范数，记 $\|\boldsymbol{x}\|_\infty$ 为其 l_∞-范数。

2 预备知识

本节介绍差分隐私、排序数据和 Mallows 模型（Mallows，1957年；Fligner 和 Verducci，1986年）相关的一些基本概念，为我们提出一种适用于排序数据的新型差分隐私变体奠定基础。

2.1 差分隐私

差分隐私已成为一种严格的框架，用于衡量随机化机制在隐私保护方面的能力；其衡量方式是限定输入中任意单条记录发生变化时输出分布之间的差异。最常见的差分隐私定义是（ϵ，δ）-差分隐私。

定义 1。（（ϵ,δ）-差分隐私）令 $\epsilon>0$ 和 $\delta\in[0,1)$。令 S 和 S' 为两个长度相同的记录集，并定义 d： $\mathcal{S}\times\mathcal{S}\to\mathbb{Z}_{\geq 0}$ 为两个集合之间的汉明距离。我们称随机化机制 $\mathcal{M}$： $\mathcal{S}\to\mathcal{Z}$ 关于 d 满足（ϵ,δ）-差分隐私，若对于任意 $S,S'\in\mathcal{S}$，只要 $d(S,S')=1$，就有

$$\mathbb{P}(\mathcal{M}(D)\in Z)\leq e^{\epsilon}\mathbb{P}(\mathcal{M}(D')\in Z)+\delta,$$

对于任意 $Z\subset\mathcal{Z}$。如果 $\delta=0$，那么 $\mathcal{M}$ 满足纯 ϵ-差分隐私。

差分隐私背后的直觉是，对数据集中的任何单条记录进行推断都不准确，因为输入数据集中任意单条记录的变化对输出分布的影响较小。

2.2 排序数据与 TheMallowsModel

在排序数据中，用户会根据特定标准对一组项目进行排序。偏好排序是典型的排序示例，它根据相对偏好为项目分配序数排名

。令 $\Omega = \{\mathcal{I}_1, \ldots, \mathcal{I}_m\}$ 表示由 m 个项目组成的集合，ϕ_Ω 表示 Ω 中项目的一种排序。排序 ϕ_Ω 通常表示为一个有序列表，用以指示 Ω 中各项目在特定度量指标下的位置，即 $\phi_\Omega(\mathcal{I}_i) = k$ 表示项目 $\mathcal{I}_i$ 的排名为 k。不失一般性，我们令排名为 1 的项目表示最受偏好的项目，因此在我们的设定中，排名越高的项目越不受偏好。例如，偏好排序$\phi_\Omega = (\phi_\Omega(\mathcal{I}_1)\ \phi_\Omega(\mathcal{I}_2)\ \phi_\Omega(\mathcal{I}_3)) = (3, 2, 1)$ 表示项目 3 是最受偏好的项目，而项目 1 是最不受偏好的项目。为简化记号，在不会造成混淆的情况下，后文将 ϕ_Ω 简写为 ϕ 。

令 Φ 表示观测排名 ϕ的随机变量，Φ 是一个取值于 $\Upsilon(|\Omega|)$ 的多元随机变量，其中 $\Upsilon(|\Omega|)$ 表示 $(1, \ldots, |\Omega|)$ 的所有排列构成的集合。自然地，我们假设 $\mathbb{P}(\Phi(\mathcal{I}_i) > \Phi(\mathcal{I}_j)) > \mathbb{P}(\Phi(\mathcal{I}_j) > \Phi(\mathcal{I}_i))$ 当且仅当项目 $\mathcal{I}_j$ 的质量优于项目 $\mathcal{I}_i$ 。这里，$\mathbb{P}(\Phi(\mathcal{I}_i) > \Phi(\mathcal{I}_j))$ 的随机性源于观测排名的随机性。这是一个较弱的假设，适用于多种排序模型，例如随机效用模型（Walker 和 Ben-Akiva，2002；Soufiani 等，2014；Su，2021）和 Mallows 模型（Mallows，1957年；Pearce 和 Erosheva，2022）。

Mallows 模型（Mallows，1957年）是一种常用的排序数据参数模型。具体而言，采用 Kendall-τ 距离的 Mallows 模型（Fligner 和 Verducci，1986年；Mandhani 和 Meila，2009）基于 ϕ_0生成合成排名。具体来说，如果 ϕ 在项目名次的成对比较中与 ϕ_0 更加一致，则排序 ϕ以更高的概率生成。排序的生成遵循如下分布

$$\mathbb{P}_{\theta,\phi_0}(\phi) = \frac{1}{\Psi(\theta)} \exp\Big(\theta T(\phi, \phi_0)\Big), \text{ for any } \phi \in \Upsilon(|\Omega|),$$

其中，ϕ_0 是项目集 Ω 上的中心排名，$\Psi(\theta) = \sum_{\phi \in \Upsilon(|\Omega|)} \exp\big(\theta T(\phi, \phi_0)\big)$，$\theta$ 是离散参数，而 $T(\phi, \phi_0) = \sum_{\mathcal{I}_i,\ \mathcal{I}_j \in \Omega} I\big((\phi(\mathcal{I}_i) - \phi(\mathcal{I}_j))\big(\phi_0(\mathcal{I}_i) - \phi_0(\mathcal{I}_j)\big) > 0\big)$可视为一致对的数量，用于刻画 ϕ 与ϕ_0之间的差异。Mallows 模型广泛用于描述偏好排序的分布（D´esir 等人，2016；Busa-Fekete 等人，2014），而 ϕ_0 通常指项目的真实排名。对于

例如，在推荐系统中，ϕ_0 指一组项目在质量上的真实排名，而 ϕ 是特定用户的观测排名。

3 排序差分隐私

用户的排序数据极其敏感，会泄露其偏好和行为，而这些信息可能对营销用途具有吸引力（Jeckmans 等人，2013）。因此，在将偏好排序分享给外部方之前，保护其隐私至关重要。本节提出 ϵ-排序差分隐私。随后，我们建立 ϵ-排序差分隐私与 Mallows 模型之间的联系，并利用这种联系生成满足所提出的 ϵ-排序差分隐私的隐私保护排序。这种联系本质上源于 Mallows 模型与差分隐私指数机制之间的内在联系（McSherry 和 Talwar，2007；Lantz 等人，2015）。

项目的名次表示其在项目集合中的位置，并隐含地揭示其偏序。直观而言，为了保护排序，项目的名次应根据差分隐私的特性加入噪声。为此，我们提出一种新的排序邻近集合变体。

定义 2. （相邻排名）设 ϕ 和 ϕ' 是同一项目集 Ω上的两个序数排序。如果仅存在一个 $\mathcal{I}_k \in \Omega$，使得对于任意 $\mathcal{I}_i, \mathcal{I}_j \in \Omega \setminus \{\mathcal{I}_k\}$，且 $i \neq j$，则称 ϕ 和 ϕ' 是相邻排名。

$$\big(\phi(\mathcal{I}_i) - \phi(\mathcal{I}_j)\big)\big(\phi'(\mathcal{I}_i) - \phi'(\mathcal{I}_j)\big) > 0.$$

相邻排名的主要思想是，除了某个特定项目外，两个序数排序在项目排名的所有偏序上都一致。接下来，我们提出 ϵ-排序差分隐私。

定义 3. （ϵ-排序差分隐私； ϵ-排序差分隐私）设 ϕ 和 ϕ' 是项目集 Ω 上的两个相邻序数排序，且 $\mathcal{M}(\phi)$ 是一个生成

面向用户的序数排序的算法。如果 $\mathcal{M}$ 满足 ϵ-排序差分隐私，则称

$$\sup_{\widetilde{\phi}\in\Upsilon(|\Omega|)}\left|\log\frac{\mathbb{P}(\mathcal{M}(\phi)=\widetilde{\phi})}{\mathbb{P}(\mathcal{M}(\phi')=\widetilde{\phi})}\right|\leq\epsilon,$$

其中，随机性来自机制 $\mathcal{M}$。

ϵ 排序差分隐私的原理是， $\mathcal{M}(\phi)$ 和 $\mathcal{M}(\phi')$ 的输出分布相似，也就是说，基于输出的合成排名对单个项目的序数排名进行推断在统计上并不准确。换言之， ϵ 排序差分隐私通过确保排名中任意单个排名的变化对排序输出分布几乎没有影响，从而保护排名中的任意单个排名。所提出的 ϵ 排序差分隐私使用单一隐私参数 ϵ 来衡量隐私保护程度。

3.1 隐私保护排名

本节将 Mallows 模型（Mallows，1957年；Fligner 和 Verducci，1986年）与所提出的 ϵ 排序差分隐私联系起来，这一联系源于 Mallows 模型属于指数族。基于这一联系，我们提出一种用于生成隐私保护排名的合成排序算法，该算法基于多阶段排名模型（Fligner 和 Verducci，1988年；Critchlow 等人，1991年）。

引理1. 设 ϕ 和 ϕ' 是两个相邻排名。*Mallows* 模型满足 $\left|\log\left(\frac{\mathbb{P}_{\theta,\phi}(\phi)}{\mathbb{P}_{\theta,\phi'}(\widetilde{\phi})}\right)\right|\leq\epsilon$ 对于任意给定的 $\widetilde{\phi}\in\Upsilon(|\Omega|)$ $\theta=2^{-1}\epsilon|\Omega|$。

引理 1 揭示了 Mallows 模型与 ϵ-排序差分隐私之间的联系。具体而言，在适当选择离散参数 $\theta=2^{-1}\epsilon|\Omega|$的情况下，使用 Mallows 模型生成合成排序满足所提出的 ϵ-排序差分隐私。换言之，可以通过生成一个对应的排序来保护观测排名 ϕ ，$\sim\phi$ 该排序由 Mallows 模型生成，且 ϕ 作为输入排名，从而依据所提出的 ϵ-排序差分隐私提供隐私保证。

为了通过 Mallows 模型生成隐私保护排名，我们提出了一种基于多阶段排名模型思想的算法（FlignerandVerducci，1988年；Critchlow 等人，

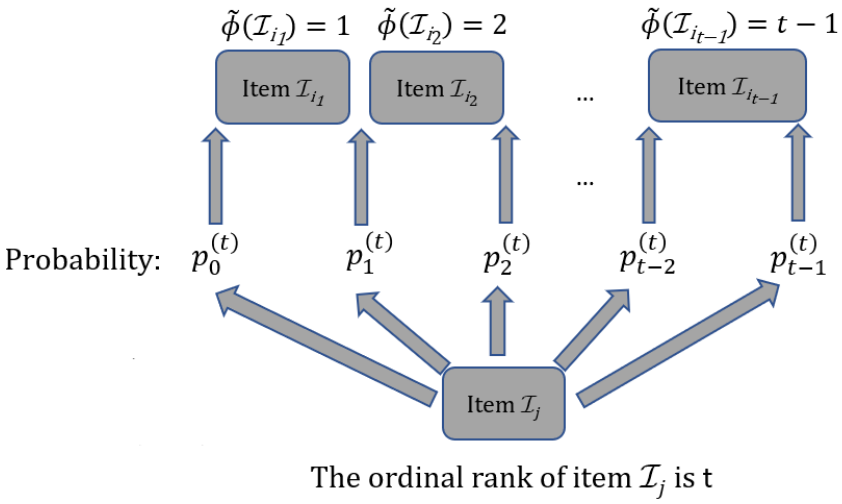


图 1：所提出的合成排序算法的第 t 次迭代。

1991年）。所提出算法的整体思路可视为一种洗牌机制。具体来说，该算法根据项目在 ϕ 中的名次，依次确定其排名。如图 1 所示，项目 $\mathcal{I}_j$ 在合成排序 $\tilde{\phi}$ 中的位置通过抽样过程确定，其概率为 $p_i^{(t)}$，$i = 1$，……，$t-1$。这里需要注意的是，$p_i^{(t)} > p_j^{(t)}$ 对任意 $i > j$ 都成立，因为位置 i 与项目 $\mathcal{I}_j$ 在 ϕ 中的位置更加一致。整体算法总结于算法 1。

~算法 1 将排序 ϕ 作为输入，并输出合成排序 ϕ。在算法 1 中，项目的合成排名以概率方式依次确定，使得与输入排名具有更多相似偏序的位置更有可能被选中。具体而言，在初始化步骤中，首先选取 ϕ 中排名为 1 的项目（记为~$\mathcal{I}_i$），并暂时将其在 ϕ 中的排名设为 1。随后，按照 ϕ中其他项目排名的升序，依次确定合成排序 ϕ 中其他项目~的排名。在第t次迭代中，目标是确定排名为 t 的项目相对于前 $t-1$ 个项目的相对位置，这些项目的相对位置已确定，且其项目索引存储在 $\chi^{(t-1)}$中。由于位置是以概率方式选取的，我们令 $V_\epsilon^{(t)}$表示取值为 $\rho^{(t)}$的、用于表示位置的随机变量，其中每个元素 $k \in \rho^{(t)}$ 都对应一个相对位置。对于每个位置 k，τ（k，$\chi^{(t-1)}$）表示一致的数量

算法 1： 隐私保护排序算法

输入： 项目集 Ω 的排序 ϕ 及隐私保证 ϵ

初始化： 设 ϕ^{-1} 为排序的逆函数，并选择项目 $\mathcal{I}_i$ 满足 $\phi^{-1}(1) = \mathcal{I}_i$，来自 Ω，并设置 $\widetilde{\phi}(\mathcal{I}_i) = 1$ 和 $\chi^{(1)} = \{i\}$；

对于 $t = 2, \ldots, |\Omega|$ **执行**

选择项目 $\mathcal{I}_j$ 从 Ω 中取值，使得 $\mathcal{I}$的排名 j 为 t，并定义

$$\rho^{(t)} = \{0, \ldots, t-1\}.$$

计算取值

$$\tau(k, \chi^{(t-1)}) = \sum_{l \in \chi^{(t-1)}} I\Big(\big(t - \phi(\mathcal{I}_l)\big)\big(k + 0.5 - \widetilde{\phi}(\mathcal{I}_l)\big) > 0\Big), k \in \rho^{(t)}.$$

根据概率，按照以下方式从 $\rho^{(t)}$ 中抽取一个值 k，并将其赋给 $V_\epsilon^{(t)}$

$$p_k^{(t)} = \mathbb{P}(V_\epsilon^{(t)} = k) = \frac{\exp\Big(\epsilon(|\Omega|-1)^{-1}\tau(k, \chi^{(t-1)})\Big)}{\sum_{k \in \rho^{(t)}} \exp\Big(\epsilon(|\Omega|-1)^{-1}\tau(k, \chi^{(t-1)})\Big)}, k \in \rho^{(t)}.$$

定义 $G^{(t)}(V_{(t)\epsilon}) = \{i : \widetilde{\phi}(\mathcal{I}_i) > V_{(t)\epsilon} + 0.5, \ i \in \chi^{(t-1)}\}$ 并将排序重新排列为

$$\begin{aligned} \widetilde{\phi}(\mathcal{I}_i) &\leftarrow \widetilde{\phi}(\mathcal{I}_i) + 1, \text{ 对于 } i \in G^{(t)}(V_\epsilon^{(t)}), \\ \widetilde{\phi}(\mathcal{I}_j) &\leftarrow V_\epsilon^{(t)} + 1. \end{aligned}$$

$\chi^{(t)} \leftarrow \chi^{(t-1)} \cup \{j\}$

结束

输出： 输出合成排序 $\widetilde{\phi}$.

位置为 k 的配对，较大的 $\tau(k, \chi^{(t-1)})$ 表示该位置 $p_k^{(t)}$ 被选中的概率更高。确定 $V_\epsilon^{(t)}$后，前 t 个项目的合成名次将按照步骤 6 所示重新排列。算法 1 的总体计算复杂度为 $O(|\Omega|^2)$

引理 2. 设 $\mathcal{A}_\epsilon$ 为算法 1。给定项目集 $\phi\Omega = \{\mathcal{I}_1, \ldots, \mathcal{I}_m\}$ 的排序 $\mathcal{A}_\epsilon(\phi)$ 合成排序遵循以下分布，

$$\mathbb{P}(\mathcal{A}_\epsilon(\phi) = \widetilde{\phi}) = \frac{\exp\big(2^{-1}\epsilon|\Omega|T(\phi, \widetilde{\phi})\big)}{\sum_{\widetilde{\phi} \in \Upsilon(|\Omega|)} \exp\big(2^{-1}\epsilon|\Omega|T(\phi, \widetilde{\phi})\big)}, \ \textit{for any } \widetilde{\phi} \in \Upsilon(|\Omega|),$$

其中 $T(\phi, \widetilde{\phi}) = \sum_{\mathcal{I}_i, \mathcal{I}_j \in \Omega} I\big(\big(\phi(\mathcal{I}_i) - \phi(\mathcal{I}_j)\big)\big(\widetilde{\phi}(\mathcal{I}_i) - \widetilde{\phi}(\mathcal{I}_j)\big) > 0\big)$。

引理2 表明，合成排序算法本质上等价于从 Mallows 模型中抽样，其基本思想是采用重要性抽样来生成~合成排序，即数值较大的 $T(\phi,\ \phi)$ 对应的排序以更大的概率生成，从而为建立隐私参数 ϵ。

定理 1。 算法1 具有以下性质：

(1) 排序最优性的一致性：对于任意满足 $i \neq j$ 的 $\mathcal{I}_i, \mathcal{I}_j \in \Omega$，有

$$\Big(\mathbb{P}\big(\Phi(\mathcal{I}_i) > \Phi(\mathcal{I}_j)\big) - \mathbb{P}\big(\Phi(\mathcal{I}_i) < \Phi(\mathcal{I}_j)\big)\Big)\Big(\mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i) > \widetilde{\Phi}(\mathcal{I}_j)\big) - \mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i) < \widetilde{\Phi}(\mathcal{I}_j)\big)\Big) > 0,$$

where $\widetilde{\Phi} = \mathcal{A}_\epsilon(\Phi)$ *denotes the random synthetic ranking output by Algorithm 1 with* Φ *being input ranking.*

(2) 排序分布的影响：对于任意满足 $\mathcal{I}_i \neq \mathcal{I}_j$ 的 $\mathcal{I}_i, \mathcal{I}_j \in \Omega$，有 t

$$\frac{\exp\left((|\Omega|-1)^{-1}\epsilon\right)-1}{\exp\left((|\Omega|-1)^{-1}\epsilon\right)+1} \leq \frac{|2\widetilde{\eta}_{ij}-1|}{|2\eta_{ij}-1|} \leq \frac{\exp\left(\frac{2|\Omega|-3}{|\Omega|-1}\epsilon\right)-1}{\exp\left(\frac{2|\Omega|-3}{|\Omega|-1}\epsilon\right)+1},$$

其中 $\eta_{ij} = \mathbb{P}\big(\Phi(\mathcal{I}_i) > \Phi(\mathcal{I}_j)\big)$ 且 $\widetilde{\eta}_{ij} = \mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i) > \widetilde{\Phi}(\mathcal{I}_j)\big)$。

(3) 在迭代 t 处，$V_\epsilon^{(t)}$ 的期望和方差为

$$\mathbb{E}\big(V_\epsilon^{(t)}\big) = \frac{(t-1)q_\epsilon^t}{q_\epsilon^t-1} - \frac{q_\epsilon^t - q_\epsilon}{(q_\epsilon-1)(q_\epsilon^t-1)} \xrightarrow{\epsilon\to+\infty} t-1$$

$$Var\big(V_\epsilon^{(t)}\big) = \frac{(t-1)^2 q_\epsilon^t}{q_\epsilon^t-1} - \frac{2\mathbb{E}\big(V_\epsilon^{(t)}\big)}{q_\epsilon-1} + \frac{q_\epsilon^t-q_\epsilon}{(q_\epsilon^t-1)(q_\epsilon-1)} - \Big(\mathbb{E}\big(V_\epsilon^{(t)}\big)\Big)^2 \xrightarrow{\epsilon\to+\infty} 0,$$

where $q_\epsilon = \exp\left(\epsilon(|\Omega|-1)^{-1}\right)$.

在定理1中，我们介绍了所提出的合成排序算法的若干性质。在某种程度上，定理1中的这些性质定量地确立了 Mallows 模型对排名分布的影响。具体而言，性质 (1) 表明，经过算法1处理后，任意两个项目的名次之间的偏序关系在总体层面保持不变。这意味着，如果在 $\mathcal{I}_i$ 中，$\mathcal{I}_j$ 的名次高于 Φ，那么在合成排序中，

基于项目质量值的真实排名也具有不变性。性质 (2) 刻画了隐私参数 ϵ 对排名分布的影响。当 ϵ 趋近于零时， $\widetilde{\eta}_{ij}$趋近于 1/2，这意味着 $\mathcal{I}_i$ 与 $\mathcal{I}_j$ 之间的偏序关系无法区分。性质 (3) 刻画了 ϵ 与 $V_\epsilon^{(t)}$之间的关系，其中 $V_\epsilon^{(t)}$表示在第 t次迭代中，合成排序保留的正确偏序关系数量。当 ϵ 趋近于无穷大时， $V_{(t)\epsilon}$按预期收敛至常数 $t-1$。这是自然的，因为根据差分隐私的定义， $\epsilon=\infty$ 指的是非隐私情形；在$\widetilde{}$这种情形下，输出排名 ϕ 与输入排名 ϕ 完全相同。

3.2 与拉普拉斯噪声的比较

本节将展示合成排序算法相较于拉普拉斯机制（Dwork，2006年）在所提出的 ϵ-排序差分隐私下的有效性。在差分隐私领域，拉普拉斯机制因其计算简单，已成为保护数值数据隐私的一种常用技术。为了安全地发布项目集 Ω ϕ的排序 $=\{\mathcal{I}_1,\ldots,\mathcal{I}_m\}$，拉普拉斯机制会为项目排名添加逐元素噪声，其定义如下：

$$\mathcal{M}_\lambda^{lap}(\phi)=\big(\phi(\mathcal{I}_i)+\xi_i\big)_{i\in[m]},$$

其中， ξ_i 是从均值为 0、尺度参数为 λ 的拉普拉斯分布中独立同分布抽取的随机样本。

引理 3. 设 ϕ和 ϕ'是项目集 $\Omega=\{\mathcal{I}_1,\ldots,\mathcal{I}_m\}$上的两个相邻排序。给定 $\lambda=2(m-1)\epsilon^{-1}$，有：

$$\Big|\log\frac{\mathbb{P}(\mathcal{M}_\lambda^{lap}(\phi)=\boldsymbol{r})}{\mathbb{P}(\mathcal{M}_\lambda^{lap}(\phi')=\boldsymbol{r})}\Big|\le\epsilon,$$

对任意 $\boldsymbol{r}\in\mathbb{R}^m$ 均成立。

引理 3 表明，对项目排名添加拉普拉斯噪声，并以适当的尺度 ϵ，也能实现所提出的 $\lambda=2(m-1)\epsilon^{-1}$。

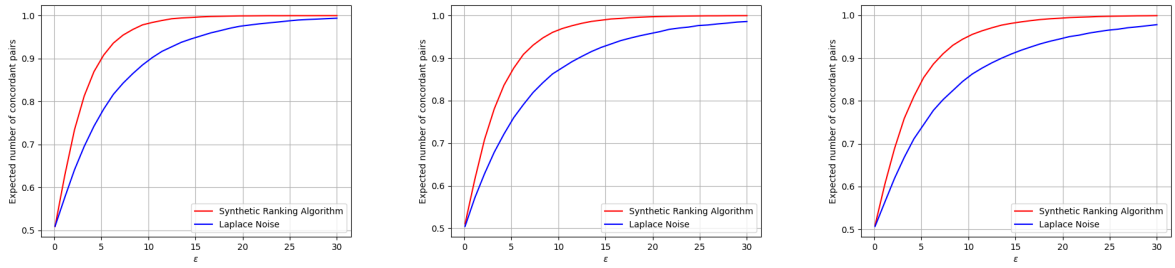
引理 4. 令 $\mathcal{A}_\epsilon$ 和 $\mathcal{M}_2^{lap}{}_{(m-1)\epsilon^{-1}}$ 分别表示满足ϵ-排序差分隐私的算法 *1*和拉普拉斯机制。对于任意长度为 m 的排序 ϕ ，对任意 $\epsilon > 0$

$$\mathbb{E}\Big[T\big(\phi, \mathcal{A}_\epsilon(\phi)\big)\Big] > \mathbb{E}\Big[T\big(\phi, \mathcal{M}^{lap}_{2(m-1)\epsilon^{-1}}(\phi)\big)\Big],$$

其中 $T(\phi, \widetilde{\phi}) = \sum_{\mathcal{I}_i,\ \mathcal{I}_j \in \Omega} I\big((\phi(\mathcal{I}_i) - \phi(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i) - \widetilde{\phi}(\mathcal{I}_j)) > 0\big)$。

~ ~对于两个排序 ϕ 和 ϕ， $T(\phi, \phi)$ 计算

~ ~ϕ 和 ϕ之间一致对的数量。显然， $T(\phi, \phi)$ 在 ϕ 和 ϕ 在项目间具有相同偏序时达到最大值。在引理 4中，我们使用一致对的期望数量来衡量所提出的合成排序算法和满足相同 ϵ-排序差分隐私的拉普拉斯机制的效用。结果表明，尽管拉普拉斯机制能够实现相同的隐私保证，但在保留原始排序中项目相对位置方面，其效果不如我们的算法。这是因为差分隐私框架基于最坏情况施加隐私保证。



(a) m=4 (b) m=5 (c) m=6

图 2：这些图展示了在 20,000 次重复实验中， $T\big(\phi, \mathcal{A}_\epsilon(\phi)\big)$ (红色)和 $T\big(\phi, \mathcal{M}_2^{lap}{}_{(m-1)\epsilon^{-1}}(\phi)\big)$(蓝色)的平均值，其中 ϕ 是项目集 $\Omega = \{\mathcal{I}_1, \ldots, \mathcal{I}_m\}$ 上的一种排序， $\epsilon \in [0.1, 30]$，且 $\phi(\mathcal{I}_i) = i, i = 1, \ldots, m$。

为了进一步验证引理4中的理论结果，我们进行了一个简单的实验，以展示所提出的合成排序算法相较于拉普拉斯机制的改进。如图2所示，所提出的算法在所有情况下生成的合成排序都保留了更多项目之间的偏序。特别是，随着

排名 m 的规模增加，这种改进会更加显著，表明所提出的合成排序算法在应用于大规模排序时极具竞争力。

4 排序推断攻击

隐私保护面临的一项关键挑战是防范推断攻击（Williams 和 McSherry，2010年；Sun 等人，2018；Wu 等人，2022），其目标是根据已发布数据恢复真实数据集中包含的某些敏感信息。对于排序数据，获取项目的排序通常是攻击者感兴趣的内容。例如，揭示用户偏好的偏好排序可用于广告（Mayer 和 Mitchell，2012年；Chen 和 Stallaert，2014）。

本节旨在研究所提出的合成排序算法在抵御基于合成排序的推断攻击方面的有效性。具体而言，如图 3所示，我们假设攻击者能够频繁调用合成排序算法，要求针对同一个中心排名生成多个合成排序 ϕ_0。此外，假设攻击者完全了解 Mallows 模型和隐私参数 ϵ，并试图估计观测排名。

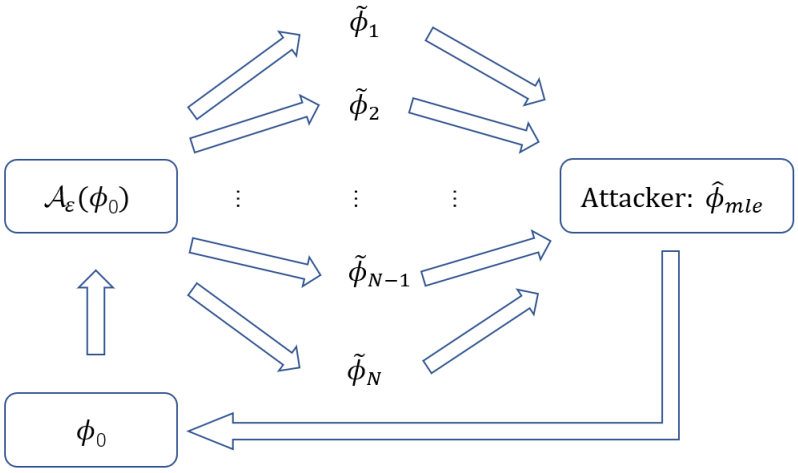


图3：排序推断攻击框架。

令 $\mathcal{S} = \{\phi_i\}_{i=1}^N$ 表示由算法1 在输入排名 ϕ_0 和项目集 Ω 上生成的一组合成排名。我们首先考虑一个简单的情形，即隐私参数

ϵ 已知。因此，接下来需要为攻击者估计观测排名 ϕ_0 。 $\mathcal{S}$ 的对数似然函数可写为

$$\mathcal{L}(\phi) = N \log \Big\{ \sum_{\widetilde{\phi} \in \Upsilon(|\Omega|)} \exp\left(2^{-1}\epsilon|\Omega| T(\phi, \widetilde{\phi})\right) \Big\} + \frac{\epsilon|\Omega|}{2} \sum_{i=1}^{N} T(\phi, \widetilde{\phi}_i).$$

令 $\widehat{\phi}_{mle}$ 表示 ϕ_0 的最大似然估计（MLE），其定义为

$$\widehat{\phi}_{mle} = \underset{\phi \in \Upsilon(|\Omega|)}{\operatorname{argmin}} \mathcal{L}(\phi). \tag{1}$$

对于任意固定的 ϕ_{mle} ， ϕ_0 ， ϵ 由 Choirat 和 Seri (2012年) 中的一般性结果可保证 M-估计量在离散参数模型中具有一致性。

实际上，估计量 $\widehat{\phi}_{mle}$ 难以求得，因为 $\widehat{\phi}_{mle}$ 的可行域是一个离散空间。因此，求解 $\widehat{\phi}_{mle}$ 的问题已知是 NP 难的（Meil˘a 和 Bao，2010年；Young，1986年）。这从计算角度明确展现了所提出的排名算法的优点：它避免了对观测排名 ϕ_0 进行估计。具体而言，即使攻击者已知合成排序算法和隐私参数 ϵ，通过最大似然估计获得正确排名仍然需要大量计算。

定理 2。 设 ϕ_0 表示观测排名， $\mathcal{S} = \{\phi_i\}_{i=1}^{N}$ 表示由 $\mathcal{A}_\epsilon(\phi_0)$生成的一组合成排序。对于任意 $\epsilon > 0$ 和 Ω，存在一个正的常数 C_0 ，使得

$$\mathbb{P}_{\mathcal{S}}\Big(\widehat{\phi}_{mle} \neq \phi_0\Big) \geq C_0 \frac{\exp(\epsilon(|\Omega|-1)^{-1}) + 1}{\exp(\epsilon(|\Omega|-1)^{-1}) - 1} \sqrt{\frac{2}{\pi N}} \exp\Big\{ - N \Big(\exp(\epsilon(|\Omega|-1)^{-1}) - 1 \Big)^2 \Big\}.$$

此外，如果 $\epsilon = O\big((|\Omega|-1)\sqrt{N^{-1}}\big)$，则 $\mathbb{P}_{\mathcal{S}}\big(\widehat{\phi}_{mle} \neq \phi_0\big)$ 远离 *0*。 ——

定理2 给出了 $\widehat{\phi}_{mle}$的估计误差下界，量化了隐私参数对 ϵ 收敛的影响 $\widehat{\phi}_{mle}$。 定理2 有双重含义。首先，当排序的长度 $|\Omega|$ 增加时，下界也随之增大，这表明排序长度会加剧 ϕ_0的估计难度。此外，将 ϵ 设为

使其以 $\epsilon = O\big((|\Omega| - 1)\sqrt{N^{-1}}\big)$ 的阶数适应合成排序的数量，会导致$\widehat{\ }$ ϕ_{mle}的不一致性；无论请求生成多少合成排序，攻击者都会以固定概率不准确地估计 ϕ_0 。

5 个性化排名中的排序差分隐私

本节旨在探讨将 ϵ-排序差分隐私应用于个性化排名，并建立理论结果，阐明基于合成排序的估计排序函数的一致性与 ϵ-排序差分隐私之间的定量关系。

5.1 个性化排名

排序问题中的典型数据集由一组三元组（$\boldsymbol{x}_u$， $\boldsymbol{y}_i$， ϕ_u（$\mathcal{I}_i$）构成，其中$\boldsymbol{x}_u \in \mathbb{R}^p$ 表示用户或查询 p的 u维特征向量， $\boldsymbol{y}_i \in \mathbb{R}^q$ 表示项目q的 i维特征向量，而 ϕ_u（$\mathcal{I}_i$）表示用户 u 对项目 i的偏好或相关性排名。我们假设项目集固定，个性化排名问题（Balakrishnan andChopra，2012年）的目标是根据新用户的偏好或相关性，为其选出最受偏好或最相关的项目。

为了估计排序函数，常见做法是采用成对排名损失（Rendle 等人，2012年；Balakrishnan and Chopra，2012年）。

$$R_n(f) = \frac{1}{nm(m-1)} \sum_{u=1}^{n} \sum_{i \neq j} I\Big(\phi_u(\mathcal{I}_i) > \phi_u(\mathcal{I}_j)\Big) I\Big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) \leq f(\boldsymbol{x}_u, \boldsymbol{y}_j)\Big). \tag{2}$$

从（2）可以看出，当 f 与两个观测排名的相对顺序不一致时，就会产生误差，因此最小化（2）会促使 f 尽可能与排序中的观测排名保持一致。

令 ϕ_u 表示 Φ_u 的一次实现，并令 $\eta_{uij} = \mathbb{P}\big(\Phi_u(\mathcal{I}_i) > \Phi_u(\mathcal{I}_j)\big)$ 表示用户 u 在 $\mathcal{I}_i$ 上的排名高于 $\mathcal{I}_j$的概率。排名风险可写为

$$R(f) = \frac{1}{m(m-1)} \sum_{i<j} \mathbb{E}\Big[\eta_{uij} I\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) \leq f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) + (1 - \eta_{uij}) I\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) \geq f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big)\Big],$$

其中，期望是针对用户随机性取的。

令 $f^* = \operatorname{argmin}_f R(f)$ 为最优排序函数。具体而言， f^*在项目对（i， j）上的行为由 η_{uij} 和 η_{uji} 的相对顺序刻画，如引理5所示。需要注意的是，由于最优性的平移不变性， f^* 并不唯一，因此，具有正确偏好排序的有偏估计量在实践中同样能够实现最优性能。具体而言，对于任意常数 $f^* + C$， C 仍然是一个最优排序函数。

引理5。 *(Dai 等人，2021)* 最优排序函数 f^* 满足：对于任意用户 u和 i， $j \in [m]$,

$$\Big(f^*(\boldsymbol{x}_u, \boldsymbol{y}_i) - f^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\Big)\Big(\eta_{uij} - \eta_{uji}\Big) \geq 0, \tag{3}$$

其中，等式成立当且仅当 $\eta_{uij} = \eta_{uji}$ ，且最小风险为

$$R^* = \mathbb{E}\Big[\frac{1}{m(m-1)} \sum_{i<j} \min\left(\eta_{uij}, \eta_{uji}\right)\Big].$$

由于指示函数不可微，最小化（2）在计算上难以处理。一种常见方法是采用替代损失函数来替代指示函数。具体而言，配备替代损失的经验风险可写为

$$R_{n,\upsilon}(f) = \frac{1}{nm(m-1)} \sum_{u=1}^{n} \sum_{i \neq j} I\Big(\phi_u(\mathcal{I}_i) > \phi_u(\mathcal{I}_j)\Big)\upsilon\Big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\Big), \tag{4}$$

其中， υ表示替代损失函数，例如逻辑损失 $\upsilon(x) = \log(1+\exp(-x))$（Rendle 等人，2012年；Zhu和Hastie，2005），指数损失 $\upsilon(x) = \exp(-x)$（Schapire，2003），以及 ψ-损失 $\upsilon(x) = \min(1, (1-x)_+)$（Shen等人，2003；Dai 等人，2021）。需要注意的是， υ不能是合页损失，因为一般而言，合页损失下的最优排序函数无法达到（Gao 和 Zhou，2015；Dai 等人，2021）。

5.2 差分隐私个性化排序

为了保护用户 $\{\phi_u\}_{u=1}^n$的观测排序，可以使用所提出的合成排序算法 $\widetilde{}$ 生成合成排序 $\{\phi_u\}_{u=1}^n$ ，供后续个性化

~学习。令 $\phi_u = \mathcal{A}_{\epsilon_u}(\phi_u)$ 表示算法 1 生成的合成排序，且满足用户 ϵ_u 的排序差分隐私 u，其中 ϵ_u 表示用户 u 的个性化隐私偏好。然后，差分隐私成对排序任务可表示为

$$\widetilde{R}_{n,\upsilon}(f) = \frac{1}{nm(m-1)} \sum_{u=1}^{n} \sum_{i \neq j} I\Big(\widetilde{\phi}_u(\mathcal{I}_i) > \widetilde{\phi}_u(\mathcal{I}_j)\Big) \upsilon\Big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\Big). \tag{5}$$

~令 $\mathcal{F}$ 表示所关注的排序函数类。估计得到的排序函数 f 为

$$\widetilde{f} = \underset{f \in \mathcal{F}}{\operatorname{argmin}}\, \widetilde{R}_{n,\upsilon}(f) + \lambda_n J(f), \tag{6}$$

其中 λ_n 是一个随 n 趋于零的调节参数，而 $J(\cdot)$ 是一个正则化项 .

记 $\widetilde{\Phi}_u = \mathcal{A}_{\epsilon_u}(\Phi_u)$ 为用户 u的合成排序随机变量。令$\widetilde{\eta}_{uij} = \mathbb{P}\big(\widetilde{\Phi}_u(\mathcal{I}_i) > \widetilde{\Phi}_u(\mathcal{I}_j)\big)$ 表示用户 u 在合成排序 ϕ_u中为 $\mathcal{I}_i$~赋予高于 $\mathcal{I}_j$ 的排名的概率。相应地，合成排序的排名风险可写为

$$\widetilde{R}(f) = \frac{1}{m(m-1)} \sum_{i<j} \mathbb{E}\Big[\widetilde{\eta}_{uij} I\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) \leq f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) + (1 - \widetilde{\eta}_{uij}) I\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) \geq f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big)\Big].$$

~ ~ ~令 $f^* = \operatorname{argmin}_f R(f)$ 表示使 $R(f)$ 最小化的最优排序函数，并令 $R^* = \sim \sim R(f^*)$ 表示最小合成排序排名风险。

引理6. 对于任意 $\epsilon_u > 0$，有

$$\big(f^*(\boldsymbol{x}_u, \boldsymbol{y}_i) - f^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\big)\big(\widetilde{f}^*(\boldsymbol{x}_u, \boldsymbol{y}_i) - \widetilde{f}^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) > 0,$$

对于任意用户 u 和 $i, j \in [m]$。特别地，当 $\epsilon_u = \epsilon$ 时，对于任意 u，有 t

$$\big(R(f) - R^*\big) \frac{\exp((|\Omega|-1)^{-1}\epsilon) - 1}{\exp((|\Omega|-1)^{-1}\epsilon) + 1} \leq \widetilde{R}(f) - \widetilde{R}^* \leq \big(R(f) - R^*\big) \frac{\exp\left(\frac{2|\Omega|-3}{|\Omega|-1}\epsilon\right) - 1}{\exp\left(\frac{2|\Omega|-3}{|\Omega|-1}\epsilon\right) + 1}. \tag{7}$$

~引理6表明 f^*与 f^*在任何隐私保证下对项目进行排序ϵ是一致的，这表明在个性化学习中使用合成排序具有最优性，不存在偏差。此外，（7）确立了超额风险之间的定量关系

在原始排序分布与合成排序分布下，由此我们可以推导出~ ~ ~ ~
的收敛速率 $R(f) - R^*$根据 $R(f) - R^*$。根据（7）中的上界，我们可以~ ~看出 $R(f) - R^*$ 对任何 f，只要 ϵ 趋于0，就会趋于0。这正如预期，因为 $\epsilon = 0$ 指的是隐私性最强的情况，此时项目的合成排名是随机确定的，所有排序函数都会同时退化。

5.3 差分隐私个性化排序的一致性

~本节将建立关于 f的渐近行为的理论结果，~量化 ϵ-排序差分隐私如何影响 $R(f) - R^*$的收敛速率。我们首先定义关于合成排序的排序 υ-风险：

$$\widetilde{R}_\upsilon(f) = \frac{1}{m(m-1)} \sum_{i \neq j} \mathbb{E}\Big[I\big(\widetilde{\Phi}_u(\mathcal{I}_i) - \widetilde{\Phi}_u(\mathcal{I}_i) > 0\big) \upsilon\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) \Big],$$

其中，期望是相对于用户和合成~ ~排名算法所产生的随机性而言的。我们用 $f_\upsilon^* = \operatorname{argmin}_f R_\upsilon(f)$表示在~ ~$\upsilon$-损失下的最优排序函数。需要注意的是，$f^*$并不唯一，而且 f_υ^*也可以是在适当选择~的情况下，使 $R(f)$最小化的最优排序 υ 函数，当$\cdot(\cdot)$被适当选择时，例如逻辑损失（Gao 和 Zhou，2015）、指数损失（Gao 和 Zhou，2015）以及 ψ-损失（Dai 等人，2021）。

~ ~ ~ ~ ~本文仅考虑满足 $R(f_\upsilon^*) = R(f^*)$ 的损失函数，并且下文将 f^* ~ ~ ~ ~ ~ ~ ~称为 f_υ^*。我们将 $e(f, f^*) = R(f) - R(f^*)$ 和 $e_\upsilon(f, f^*) =$~ ~ ~ ~$R_\upsilon(f) - R_\upsilon(f^*)$ 分别记为其在 0-1 损失和 υ-损失下的超额风险。为量化 ~ ~ϵ-排序差分隐私对超额风险的影响，我们首先推导 $e(f, f^*)$ 相对于 ~ ~ ~$e_\upsilon(f, f^*)$ 的收敛速率；结合式（7），这刻画了 $R(f) - R^*$ 的渐近行为。已知，根据替代损失 υ，从 $e_\upsilon(f, f^*)$ 推导出的 $e(f, f^*)$ 的收敛行为取决于该替代损失（Zhang，2004；Gao 和 Zhou，2015）。在介绍主要理论之前，我们先列出一些假设。

~**假设1**。令 $\mathcal{F}_\delta = \{f \in \mathcal{F} : \ e_\upsilon(f, f^*) \leq \delta\}$ 表示 $\mathcal{F}$的一个子集，使得任何 $f \in \mathcal{F}_\delta$ 的超额υ风险都小于 δ。假设存在一个常数 $\alpha > 0$ 和一个~足够小的常数 $\delta > 0$，使得 $\sup_{f \in \mathcal{F}_\delta} e(f, f^*) \leq \delta^\alpha$。

~ ~假设1建立了 $e_v(f, f^*)$ 与 $e(f, f^*)$ 之间的转换关系，这~会影响 $e(f, f^*)$ 的渐近行为。具体而言， α 取决于损失函数的选择，尤其是对于 ψ-损失（Dai 等人，2021）的 $\alpha = 1$ ，以及对于指数损失和逻辑损失（Zhang，2004；Bartlett 等人，2006；Gao 和 Zhou，2015）的 $\alpha = 1/2$ 。此外，~ ~假设1表明，可以根据~ ~ ~ ~ ~ ~$e_v(f, f^*)$ 的收敛速率推导出 $e(f, f^*)$ 的收敛速率。令 $f_{\mathcal{F}}^* = \text{argmin}_{f \in \mathcal{F}} R_v(f)$ 为 $\mathcal{F}$中逼近 f^*的最佳排序函数。~ ~超额风险 $e_v(f, f^*)$ 可分解为

$$e_v(\widetilde{f}, \widetilde{f}^*) = e_v(\widetilde{f}, \widetilde{f}_{\mathcal{F}}^*) + e(\widetilde{f}_{\mathcal{F}}^*, \widetilde{f}^*) = \widetilde{R}_v(\widetilde{f}) - \widetilde{R}_v(\widetilde{f}_{\mathcal{F}}^*) + \widetilde{R}_v(\widetilde{f}_{\mathcal{F}}^*) - \widetilde{R}_v(\widetilde{f}^*),$$

~ ~ ~ ~其中， $e_v(f, f_{\mathcal{F}}^*)$ 和 $e_v(f_{\mathcal{F}}^*, f^*)$ 通常分别称为估计误差和近似误差。

~本文主要关注 f 一致性中的估计误差。因此，在假设2中，我们假设排序函数类设定正确，且近似误差可以忽略。

~ ~**假设2**。排序函数类 $\mathcal{F}$选择适当，使得 $e_v(f_{\mathcal{F}}^*, f^*) = 0$。**假设3**。我们假设每个用户 u 都有个人隐私偏好 ϵ_u ，并且 $\Psi_u = \exp\left((|\Omega|-1)^{-1}\epsilon_u\right)+1 \ \exp\left((|\Omega|-1)^{-1}\epsilon_u\right)-1$ 是一个次高斯随机变量。

假设 3 考虑了用户隐私偏好的差异，这在现实应用中通常成立（Watson 等人，2015）。此外，假设 3 对用户隐私偏好的分布施加了约束，即 Ψ_u 服从次高斯分布，这意味着追求更强排名隐私保护的用户较少。

假设 4。 对于任意用户 u，我们假设存在一些常数 $C_2 > 0$ 和$0 < \gamma \leq +\infty$ ，使得 $\mathbb{P}\big(|2\eta_{uij} - 1| \leq \beta\big) \leq C_2\beta^\gamma$ 对于 $0 \leq \beta \leq 1$成立。

假设 4称为低噪声假设（Bartlett等人，2006年；Shen等人，2003），用于刻画 $2\eta_{uij} - 1$ 在决策边界附近的行为，这会影响~的渐近行为。 f特别是， $\gamma = +\infty$ 意味着偏好的偏序~ ~是确定性的，从而使 e（f， f^*）的收敛速率达到最优。

引理 7。 在假设 4 下，对于任意用户 u, $\mathbb{P}\big(|2\widetilde{\eta}_{uij}-1|\leq\beta\big)\leq C_2\Psi_u^{\gamma}\beta^{\gamma}$，有 $\Psi_u=\exp\big((|\Omega|-1)^{-1}\epsilon_u\big)+1\exp\big((|\Omega|-1)^{-1}\epsilon_u\big)-1$。

假设 5。 令 Θ 表示 $\mathcal{F}$中函数的所有参数。我们假设 $\|\Theta\|_\infty\leq C_{\mathcal{F}}$，并且存在一些正的常数 C_1，使得对于任意 f_1, $f_2\in\mathcal{F}$

$$\big(f_1(\boldsymbol{x}_u,\boldsymbol{y}_i)-f_2(\boldsymbol{x}_u,\boldsymbol{y}_i)\big)^2\leq C_1\big(\|\boldsymbol{x}_u\|_2^2+\|\boldsymbol{y}_i\|_2^2\big)\|\Theta_{f_1}-\Theta_{f_2}\|_\infty^2,$$

其中 Θ_f 表示 f的参数。

假设 5 是一个温和的假设，它确保 $\mathcal{F}$ 中的函数关于输入特征和参数具有光滑性，并建立了 $\mathcal{F}$ 的度量熵与相应参数空间之间的联系。

定理 3。 在假设1-5下，对于任意在 f 中定义的最小化解 (6)，存在一些常数 C_3, $C_4>0$，使得

$$\mathbb{P}\big(e(\widetilde{f},f^*)\geq\delta_n^{\alpha}\big)\leq 8C_3\exp\Big(-C_4n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^{\gamma})\delta_n^{\frac{\gamma+2}{\gamma+1}}\Big),$$

其中 $\mathbb{E}(\Psi_u^{\gamma})=O(\delta_n)$， $\mathbb{E}^{\frac{1}{\gamma+2}}(\Psi_u^{\gamma})\big(|\Theta|n^{-1}\log(n/|\Theta|)\big)^{\frac{\gamma+1}{\gamma+2}}=O(\delta_n)$，且 $|\Theta|$ 是 $f\in\mathcal{F}$的参数数量，并且 $\lambda_nJ_0\asymp\delta_n$满足 $J_0=$最大值$\{J(f_{\mathcal{F}}^*)\ 1\}$。

定理 3 描述了 $e(f,f^*)$ 的渐近行为，以及它与隐私保证之间的关系。具体而言，收敛速率 δ_n^{α} 取决于 γ、 α、平均隐私偏好 $\mathbb{E}(\Psi_u^{\gamma})$，以及 $\mathcal{F}$ 的复杂度。特别是，当 $\alpha=1$ 且 $\gamma=\epsilon=\infty$ 时，可得到最佳收敛速率 $O_p\Big(|\Theta|n^{-1}\log(n/|\Theta|)\Big)$，这与 Dai 等人（2021）的现有理论结果一致。此外，值得注意的是，当 $\mathbb{E}^{-1}(\Psi_u^{\gamma})=o(1)$ 时，与非隐私情形相比，速率 δ_n 会变慢。例如，当所有用户的 $\epsilon_u=\epsilon$ 时， $\mathbb{E}^{-1}(\Psi_u^{\gamma})$ 的阶为 ϵ^{γ}，其中 $\epsilon=o(1)$，这表明，当所有用户都追求更高的排序隐私保护（对应于更小的 ϵ）时，收敛到最优排序函数的速度会变慢。

推论 1. 在定理 3的假设下，我们有

$$R(\widetilde{f}) - R(f^*) = O_p\Big(\mathbb{E}(\Psi_u)\mathbb{E}^{\frac{\alpha}{\gamma+2}}(\Psi_u^{\gamma})\big(|\Theta|n^{-1}\log(n/|\Theta|)\big)^{\frac{\alpha(\gamma+1)}{\gamma+2}}\Big).$$

特别是，当 $\epsilon_u = \epsilon = o(1)$ 对任意用户 u 都成立，且 $\upsilon(\cdot)$ 的选取满足 $\alpha = 1$ 时，我们有

$$R(\widetilde{f}) - R(f^*) = O_p\Big(\big(|\Theta|(|\Omega|-1)^2 n^{-1}\epsilon^{-2}\log(n/|\Theta|)\big)^{\frac{\gamma+1}{\gamma+2}}\Big).$$

此外，如果对某个$\zeta > 0$， $|\Omega|\sqrt{|\Theta|n^{-1}\log^{1+\zeta}(n/|\Theta|)} = O(\epsilon)$，则有 $R(\widetilde{f}) - R(f^*) = o_p(1)$。

推论 1 给出了差分隐私排序~函数 f的超额风险收敛速率，并量化了合成排序算法的影响。首先，如果所有用户都选择相同的隐私常数，与非隐私情形相比，收敛速率会慢一个乘法常数。推论 1 还揭示了在成对学习一致性的条件下，所有用户所能达到的最佳隐私保证速率。具体而言，当所有用户的隐私保证 ϵ 满足 $|\Omega|\sqrt{|\Theta|n^{-1}\log^{1+\zeta}(n/|\Theta|)} = O(\epsilon)$ 对某个 $\zeta > 0$ 且对任意 $\zeta > 0$，在适当选择参数空间 f^* 和损失函数 $\mathcal{F}$ 后，成对学习仍能很好地估计最优排序函数 $\upsilon(\cdot)$。

---

6 实验

在本节中，我们将在模拟数据集上进行一系列数值实验，以验证我们的理论结果并展示所提出算法的有效性。

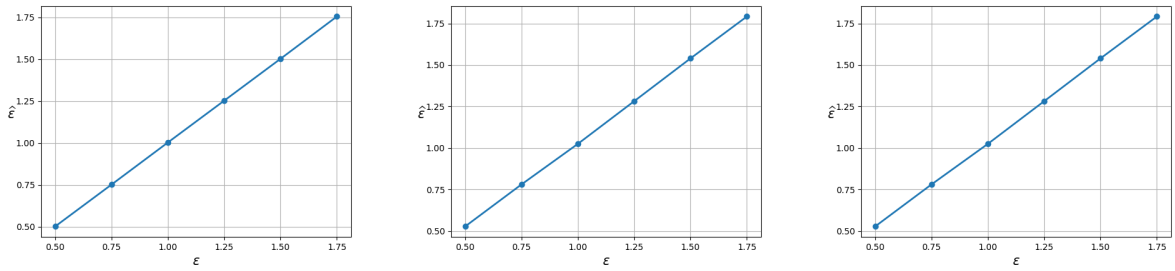
6.1 隐私保证

在第一次模拟中，我们旨在对引理 1中所述的所提出合成排序算法的隐私保证进行经验验证。为此，我们考虑一个大小为 σ 的排序 m，满足 $\phi(\mathcal{I}_i) = i$，其中 $i = 1, \ldots, m$。令 $\mathcal{S}$表示所有可能的 ϕ 相邻排名的集合，并令 $\mathcal{C}$ 表示 ϕ 的所有可能排列的集合。对于每个 $\phi' \in \mathcal{S}$，我们在 ϕ 上执行所提出的算法，其隐私保证为 ϵ，并且 ϕ'

进行 N 次，然后通过以下方式估计隐私保证

$$\widehat{\epsilon} = \max_{\phi' \in \mathcal{S}} \max_{\widetilde{\phi} \in \mathcal{C}} \left| \log \left(\frac{\sum_{i=1}^{N} I(\mathcal{A}_\epsilon(\phi) = \widetilde{\phi})}{\sum_{i=1}^{N} I(\mathcal{A}_\epsilon(\phi') = \widetilde{\phi})} \right) \right|, \tag{8}$$

其中 $\mathcal{A}_\epsilon(\phi)$ 表示算法在第 i 次重复实验中的输出排名。我们考虑 (m, ϵ) $\in \{3, 4, 5\} \times \{0.5 + 0.25 * i, i = 0, \ldots, 6\}$ 且 $N = 10^7$ 的情形，结果见图4。



(a) m=3 (b) m=4 (c) m=5

图 4：估计的隐私保证 $\widehat{\epsilon}$ 与预先指定的隐私保证

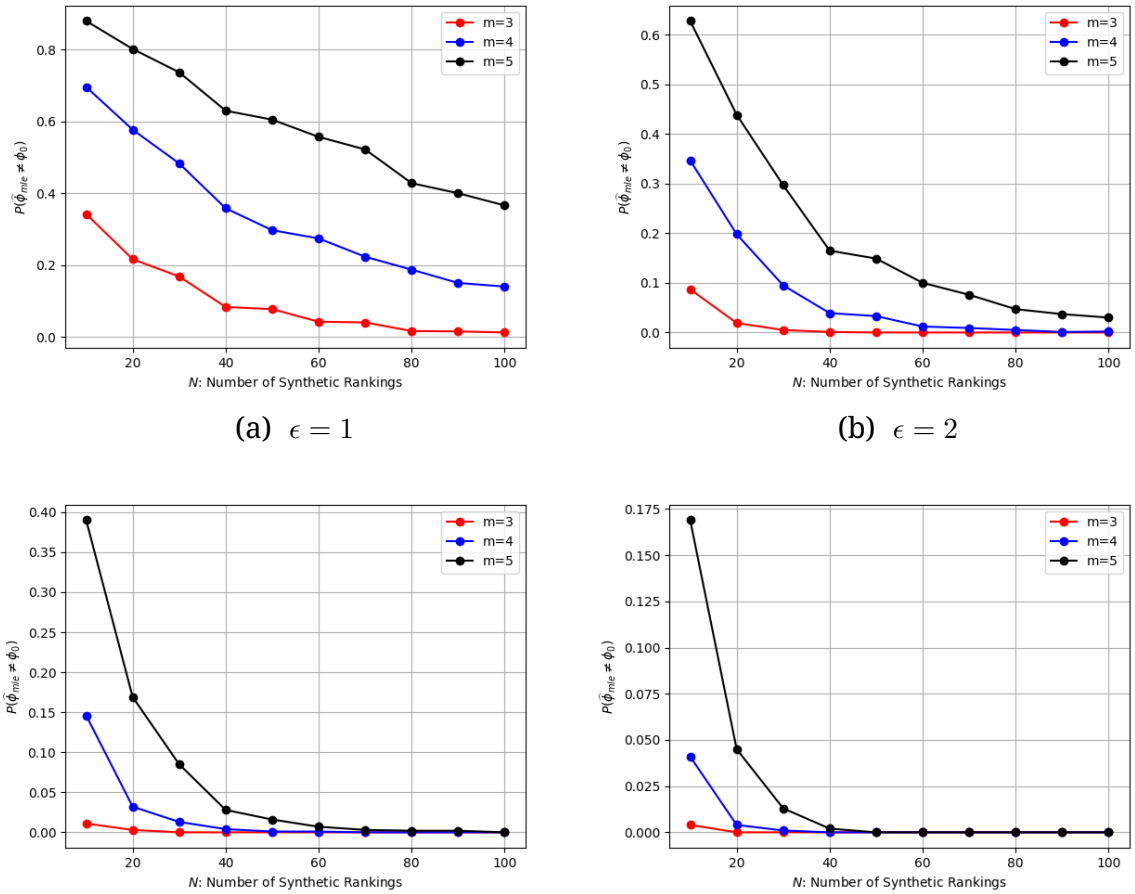
如图 4 所示，估计得到的隐私保证 $\widehat{\epsilon}$ 与预先指定的隐私保证 ϵ完全吻合，这与我们在引理 1中确立的理论结果一致。

6.2 推断攻击

本模拟旨在验证定理2 中关于自适应方案的理论结果ϵ：合成排序的数量会导致排名估计性能下降。为此，模拟设置如下。首先，对于一个大小为 ϕ_0 的$\widetilde{}$排序 m，我们生成一组合成排序 $\mathcal{S}_m = \{\phi_i\}_{i=1}^{N}$。其次，基于 ϕ_0，按照式（ $\mathcal{S}_m$ （1）通过最大似然估计实现对 R 的估计。上述步骤重复 $\widehat{\phi}_{mle}^{(i)}$ 次。令 i表示第 i次重复实验中的最大似然估计量。然后，我们估计最大似然估计不一致的概率，如下所示：

$$\widehat{\mathbb{P}}(\widehat{\phi}_{mle} \neq \phi_0) = \frac{1}{R} \sum_{i=1}^{R} I(\widehat{\phi}_{mle}^{(i)} \neq \phi_0).$$

在第一个示例中，我们旨在验证在隐私保证固定的情况下， ϕ_{mle} 依概率收敛到 ϕ_0 。具体而言，我们考虑大小为 $m = 3$、4、5的排序，其中 $\phi_0(\mathcal{I}_i) = i$ 取值为 i =1、.....、 m ，且 $R = 1$取值为000。合成排序的数量 N 设为 $\{10, 20, 30, \ldots, 100\}$。隐私保证设为 $\epsilon \in \{1, 2, 3, 4\}$。



(a) $\epsilon = 1$ (b) $\epsilon = 2$

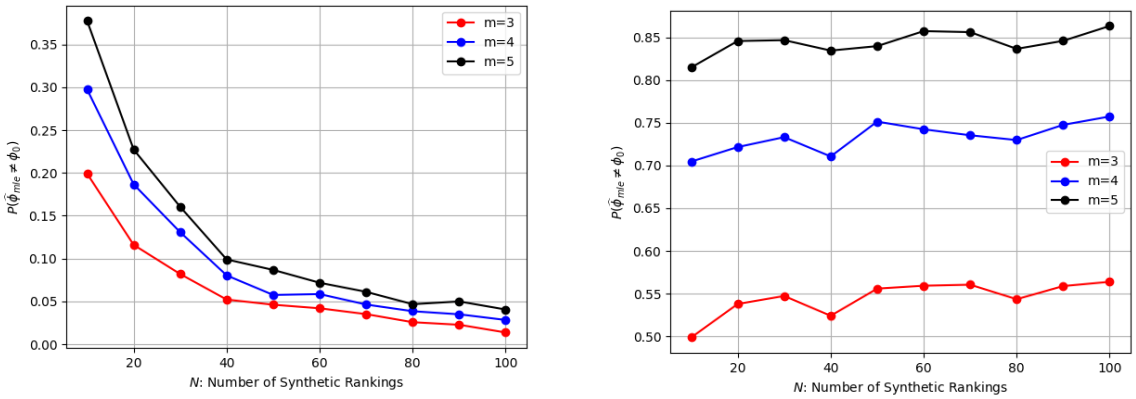
(c) $\epsilon = 3$ (d) $\epsilon = 4$

图 5：在隐私保证为 $\phi_{mle} = \phi_0$ 、排序长度为 $m = 3$的情况下，1,000 次重复实验中事件的经验概率，2，3，4，4，5。

错误估计的经验概率 ϕ_{mle} 在不同隐私和排序设置下的结果如图5所示。显然，ϕ_{mle} 会随着合成排序数量的增加，在概率意义上收敛到真实排名 ϕ_0 ，这表明，只要攻击者获得足够多的合成排序，就能准确估计用于生成合成排序的真实排名。此外，不难发现，较长的排名

长度会降低真实排名估计的准确性。这是因为所提出的 ϵ-排序差分隐私机制会保护每个排名，从而随着排序长度增加，估计过程也变得更加困难。

在第二个示例中，我们旨在验证在自适应隐私保证下，$\hat{\phi}_{mle}$ 到 ϕ_0 的收敛可以被阻止，且最优自适应隐私保证方案为 $\epsilon \asymp (m-1)/\sqrt{N}$。为此，我们沿用第一个示例中的模拟设置，但将 ϵ 设置为随合成排序的数量自适应，即$\epsilon \asymp (m-1)/\sqrt{N}$和 $\epsilon \asymp (m-1)\log(N)/\sqrt{N}$，并将用于估计$\mathbb{P}(\hat{\phi}_{mle} \neq \phi_0)$的重复实验次数设为 5,000。



(a) $\epsilon = \frac{\log(N)(m-1)}{\sqrt{N}}$　　(b) $\epsilon = \frac{m-1}{\sqrt{N}}$

图6：在采用自适应 ϵ 且排序长度不同的情况下，事件 $\hat{\phi}_{mle} = \phi_0$ 在 5,000 次重复实验中的经验概率 $m = 3$，4，5。

图 6 验证了定理2中的理论结果，表明当$\mathbb{P}(\hat{\phi}_{mle} \neq \phi_0)$ 随合成排名的数量以 ϵ 速率递减时，仍然可以实现收敛：$\epsilon \asymp (m-1)\log(N)/\sqrt{N}$。相比之下，若没有对数项，$\hat{\phi}_{mle}$ 的一致性将失效，且经验概率会与某些固定常数保持正距离，正如定理2所述。这两种情况表明，为调和 ϵ 的不一致性与合成排名的效用，$\hat{\phi}_{mle}$ 的最慢速率为 $\epsilon \asymp (m-1)/\sqrt{N}$。

6.3 差分隐私个性化排序

在此模拟中，我们旨在验证与拉普拉斯机制相比，所提出的合成排序算法在下游排序学习任务的排序成对比较中保留更多信息的优势。首先，我们通过 4 维均匀分布生成用户和项目的特征，即 $x_{ul}, y_{il} \sim \text{Unif}(-3,3)$，其中 $l = 1, \ldots, 4$。然后，使用模型 $r_{ui} = \boldsymbol{\alpha}^T \boldsymbol{x}_u + \boldsymbol{\beta}^T \boldsymbol{y}_i$生成用户 u对项目的偏好，其中项目集固定为 $\Omega = \{\mathcal{I}_1, \ldots, \mathcal{I}_m\}$。通过对相关偏好 $\{r_{ui}\}_{i=1}^m$进行排序，得到用户 u的排名。随后，我们通过合成排序算法和拉普拉斯机制生成合成排名，使其满足 ϵ-排名~~差分隐私，分别记为 $\phi = \mathcal{A}_\epsilon(\phi_u)$ 和 $\phi = \mathcal{M}_{2(m-1)\epsilon^{-1}}(\phi_u)$。最后，我们通过最小化以下正则化成对学习任务来估计排序函数。

$$\widetilde{f} = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \sum_{i \neq j} I\big(\widetilde{\phi}_u(\mathcal{I}_i) > \widetilde{\phi}_u(\mathcal{I}_j)\big) \upsilon(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)), \tag{9}$$

$$\widetilde{f}_{lap} = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} \sum_{i \neq j} I\big(\widetilde{\phi}_u^{lap}(\mathcal{I}_i) > \widetilde{\phi}_u^{lap}(\mathcal{I}_j)\big) \upsilon(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)), \tag{10}$$

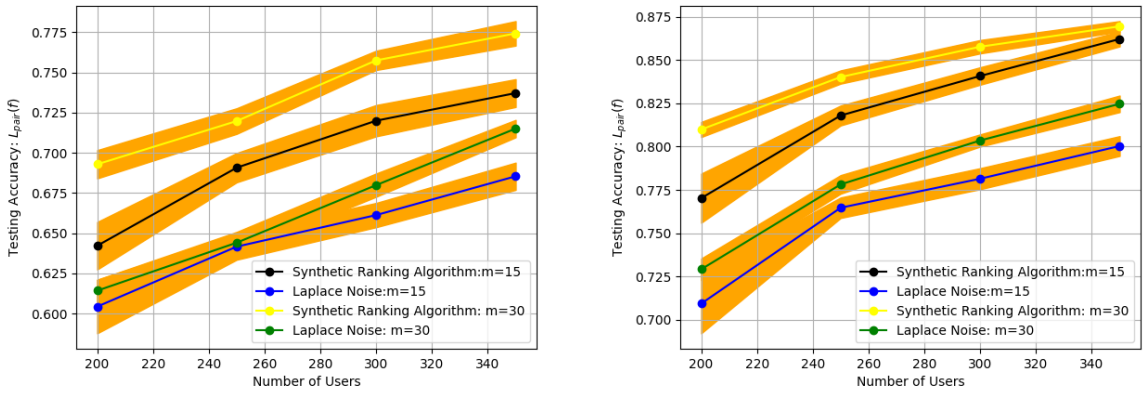
其中 $\mathcal{F}$ 是双塔模型（Wang 等人，2021；Yang 等人，2020）这一类模型，它使用两个并行的神经网络来学习用户和项目特征的表示，并通过对这些表示进行点积来计算偏好分数。

~~此外，为了比较 f_{lap} 和 f的性能，我们按照上述相同的设置，生成一组具有相应特征的新用户。具体来说，我们令 $\{\boldsymbol{x}'_u\}_{u=1}^N$ 表示新用户的特征集合，并将其偏好生成为 $r'_{ui} = \boldsymbol{\alpha}^T \boldsymbol{x}'_u + \boldsymbol{\beta}^T \boldsymbol{y}_i$。对于排序函数 f，我们采用以下度量指标来评估其性能：

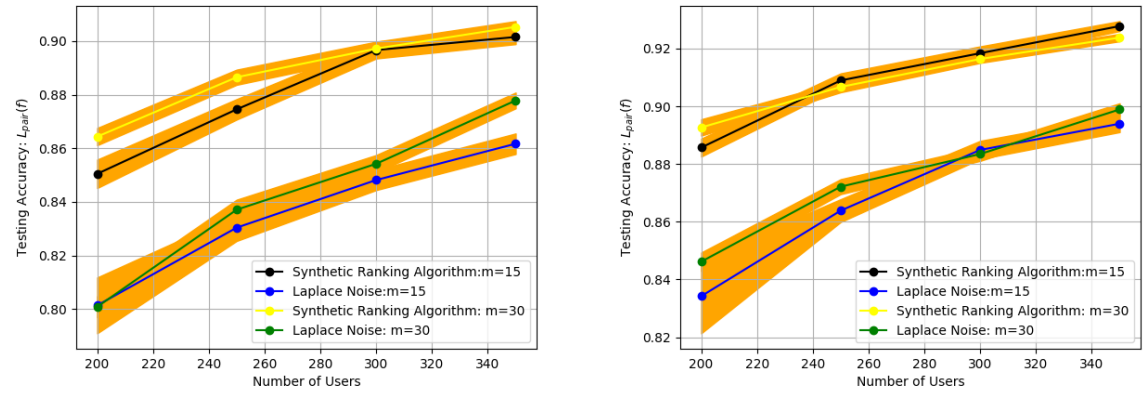
$$L_{pair}(f) = \frac{1}{Nm(m-1)} \sum_{u=1}^{N} \sum_{i \neq j} I\big(r'_{ui} > r'_{uj}\big) I\big(f(\boldsymbol{x}'_u, i) > f(\boldsymbol{x}'_u, j)\big). \tag{11}$$

对于超参数选择，我们将双塔模型中的两个神经网络设为 3 层多层感知机，每层包含 10 个隐藏单元。我们设置 $N = 1,000$ 和

考虑以下情形 $(m,\ n,\ \epsilon) \in \{15, 30\} \times \{200, 250, 300, 350\} \times \{1, 2, 3, 4\}$。每种情形均使用 100 个用户进行验证，并采用监控验证误差的早停方法来调整神经网络的参数。每种情形重复运行 50 次，并在图 7 中报告平均成对测试准确率。



(a) $\epsilon = 1$　　(b) $\epsilon = 2$



(c) $\epsilon = 3$　　(d) $\epsilon = 4$

图7：在不同隐私保证$\epsilon = 1$、2、3、4 和排名长度为 $m = 15$、30 的条件下，进行 50 次重复实验所得的平均成对准确率。

如图 7所示，在所有场景中，所提出方法的平均测试准确率均显著高于拉普拉斯噪声，表明所提出的合成排序算法在为成对学习任务生成效用保持排名方面优于拉普拉斯噪声。

6.4 实际应用

本节中，我们在寿司数据集上评估所开发的合成排序算法的性能，旨在表明，与在相同隐私保证下添加拉普拉斯噪声相比，我们的合成排序算法能为下游学习任务保留更多信息，从而获得更好的泛化性能。

寿司数据集可在此处公开获取 https://www.kamishima.net/sushi/。该数据集包含通过问卷调查收集的 5,000 位用户对 10 种寿司的偏好排序，包括“虾”“海鳗”“金枪鱼”“鱿鱼”“海胆”“鲑鱼卵”“玉子烧”“肥金枪鱼”“金枪鱼卷”和“黄瓜卷”。在此应用中，我们只考虑“金枪鱼”“鲑鱼卵”“玉子烧”“肥金枪鱼”和“黄瓜卷”的偏好排名；与其他寿司相比，用户对这些寿司的偏好排名更为一致。

我们通过第6.3节中的成对学习任务，比较合成排序算法生成的排序与添加拉普拉斯噪声后生成的排序的效用。具体来说，我们首先将数据集划分为包含 3,000 位用户的训练数据集和包含 2,000 位用户的测试数据集。其次，使用合成排序算法对训练数据集中的排序进行置换，并添加满足相同 ϵ-排序差分隐私$\tilde{}$ $\tilde{}$的拉普拉斯噪声。随后，我们按照式（ f 9）和（10），分别估计排序函数$\tilde{}$和待估计的排序函数 $f(\boldsymbol{x}_u,\ \boldsymbol{y}_i) = \boldsymbol{\alpha}\boldsymbol{x}_u + \boldsymbol{\beta}\boldsymbol{y}_i$。我们使用测试数据集中的用户，根据式（ f11$_{\sim}f_{lap}$ ）定义的成对测试准确率，评估和的表现。

图 8 展示了与第 6.3 节类似的结果：在相同的隐私保证下，与拉普拉斯噪声相比，所提出的方法能够为成对学习任务生成更多保留效用的合成排序。

7 总结

本文提出了一种针对排序数据的新型隐私度量，称为 ϵ-排序差分隐私。其基本思想是生成合成排序，以替代真实排序

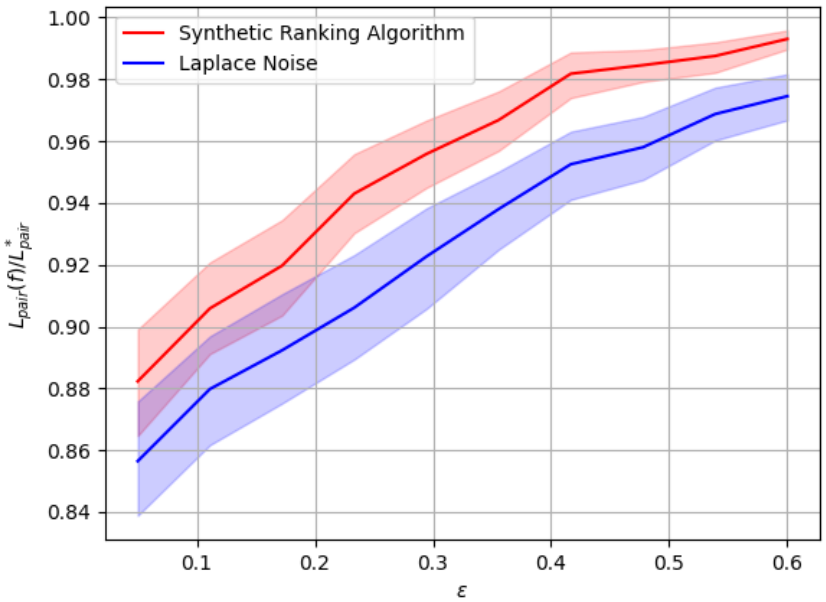


图 8：在 50 次重复实验中，不同隐私保证下的平均相对效用ϵ，隐私保证取 0.06 至 0.5 范围内均匀间隔的值。 L^*_{pair} 表示在 50 次重复实验中，于未进行隐私处理的训练数据集上训练的排序函数的平均成对准确率。

用于共享，同时抵御针对排序列表中特定名次的推断攻击。在此概念的基础上，我们进一步基于多阶段排名模型开发了一种合成排序算法，用于生成满足所提出的 ϵ-排序差分隐私的合成排序；该算法以概率方式对项目的序数排名进行内在置换。此外，我们分析了所提出算法生成的合成排序在下游推断攻击和个性化排序任务中的效用，并建立了刻画相应效用-隐私权衡的统计一致性。值得注意的是，与采用噪声添加方法生成差分隐私排序的现有方法不同，所提出的方法在满足严格隐私保护要求的同时，以隐式方式对排序进行非线性扰动。

参考资料

Alabi, D., Ghazi, B., Kumar, R. 和 Manurangsi, P.（2022）。中心化和本地模型中的隐私排名聚合。载于*AAAI*人工智能大会论文集，第 36 卷，第 5984–5991 页。

Arratia, R. 和 Gordon, L.（1989）。二项分布大偏差教程。《数学生物学通报》，51(1)：125–131。

Balakrishnan, S. 和 Chopra, S.（2012）。协同排序。载于第五届 *ACM* 国际网络搜索与数据挖掘会议论文集，第 143–152 页。

Bartlett, P. L., Jordan, M. I. 和 McAuliffe, J. D.（2006）。凸性、分类与风险界。《美国统计学会会刊》，101(473)：138–156。

Busa-Fekete, R., Fotakis, D. 和 Zampetakis, E.（2021）。排序数据的隐私与非隐私均匀性检验。神经信息处理系统进展，34：9480–9492。

Busa-Fekete, R., H¨ullermeier, E., 和 Sz¨or´enyi, B.（2014）。使用统计模型进行基于偏好的排名 elicitation：以 Mallows 模型为例。载于国际机器学习会议，第 1071–1079 页。PMLR。

Chen, J. 和 Stallaert, J.（2014）。使用行为定向的在线广告经济分析。《*MIS*季刊》，38(2):429–A7。

Choirat, C. 和 Seri, R.（2012）。离散参数模型中的估计。《统计科学》，27(2):278–293。

Critchlow, D. E., Fligner, M. A., 和 Verducci, J. S.（1991）。关于排序的概率模型。《数学心理学杂志》，35(3):294–318。

Dai, B., Shen, X., Wang, J. 和 Qu, A.（2021）。用于个性化预测的可扩展协同排序。《美国统计学会会刊》，116(535):1215–1223。

D´esir, A., Goyal, V., Jagabathula, S. 和 Segev, D.（2016）。Mallows 模型下的品类优化。神经信息处理系统进展，29。

Duchi, J. C. 和 Ruan, F.（2018）。局部隐私估计中恰当的复杂度度量：它不是费 Fisher 信息。*arXiv* 预印本 *arXiv:1806.05756*。

Dwork, C.（2006）。差分隐私。载于自动机、语言与编程国际研讨会，第 1–12 页。Springer。

Dwork, C.、Kenthapadi, K.、McSherry, F.、Mironov, I. 和 Naor, M.（2006）。我们的数据，我们自己：通过分布式噪声生成实现隐私。载于 密码技术理论与应用年度国际会议，第 486–503 页。Springer。

Dwork, C.、Kumar, R.、Naor, M. 和 Sivakumar, D.（2001）。面向 Web 的排序聚合方法。载于第十届万维网国际会议论文集，第 613–622 页。

Erlingsson, ´U., Pihur, V. 和 Korolova, A. (2014)。RAPPOR：随机化可聚合隐私保护序数响应。载于*2014*年*ACM SIGSAC*计算机与通信安全会议论文集，第 1054–1067 页。

Fligner, M. A. 和 Verducci, J. S. (1986)。基于距离的排序模型。皇家统计学会会刊：*B*辑（方法论），48(3):359–369。

Fligner, M.A. 和 Verducci, J. S. (1988)。多阶段排序模型。《美国统计学会会刊》，83(403):892–901。

Gao, W. 和 Zhou, Z.-H. (2015)。论AUC成对优化的一致性。载于第二十四届国际人工智能联合会议。

Hay, M.、Elagina, L. 和 Miklau, G.（2017）。差分隐私排序聚合。载于*2017*年*SIAM*国际数据挖掘会议论文集，第669–677页。SIAM。

Jeckmans, A. J.、Beye, M.、Erkin, Z.、Hartel, P.、Lagendijk, R. L. 和 Tang, Q.（2013）。推荐系统中的隐私。载于社交媒体检索，第263–281页。Springer。

Jeong, M.、Dytso, A. 和 Cardone, M.（2022）。隐私考量下的排序恢复。《机器学习研究汇刊》。

Kairouz, P.、Oh, S. 和 Viswanath, P.（2015）。差分隐私的组合定理。载于国际机器学习会议，第 1376–1385 页。PMLR。

Karatzoglou, A.、Baltrunas, L. 和 Shi, Y.（2013）。面向推荐系统的学习排序。载于第七届 *ACM* 推荐系统会议论文集，第 493–494 页。

Koltchinskii, V.（2011）。经验风险最小化与稀疏恢复中的预言机不等式´ ´问题：圣弗卢尔概率论暑期学校第三十八届（*2008*），第 2033 卷。Springer 科学与商业媒体。

Lantz, E.、Boyd, K. 和 Page, D.（2015）。子采样指数机制：大输出空间中的差分隐私。载于第八届 *ACM* 人工智能与安全研讨会论文集，第 25–33 页。

Lee, D. T. (2015). 高效、私密且具 ε-策略无操纵性的锦标赛投票规则引出。载于第二十四届国际人工智能联合会议。

Li, Z., Liu, A., Xia, L., Cao, Y., and Wang, H. (2022). 差分隐私孔多塞投票。*arXiv* 预印本 *arXiv:2206.13081*。

Liu, Y.-T., Liu, T.-Y., Qin, T., Ma, Z.-M., and Li, H. (2007). 监督式排序聚合。载于第十六届国际万维网会议论文集，第 481–490 页。Mallows, C. L. (1957年). 非零排序模型。I.《生物计量学》，44(1/2):114–130。

Mandhani, B. 和 Meila, M.（2009）。用于学习排序指数模型的可处理搜索。载于《人工智能与统计学》，第392–399页。PMLR。

Mayer, J. R. 和 Mitchell, J. C.（2012年）。第三方网络跟踪：政策与技术。载于*2012*年*IEEE*安全与隐私研讨会，第413–427页。IEEE。

McSherry, F. 和 Talwar, K.（2007）。通过差分隐私进行机制设计。载于第*48*届*IEEE*计算机科学基础年会（*FOCS'07*），第94–103页。IEEE。

Meil˘a, M. 和 Bao, L.（2010）。无限排序的指数模型。机器学习研究杂志，11(113)：3481–3518。

Nadarajah, S.（2007）。拉普拉斯随机变量的线性组合、乘积与比值。《统计学》，41(6)：535–545。

Oliveira, S. E.、Diniz, V.、Lacerda, A.、Merschmanm, L. 和 Pappa, G. L.（2020）。排名聚合在推荐系统中是否有效？一项实验分析。*ACM*智能系统与技术汇刊（*TIST*），11(2)：1–26。

Pearce, M. 和 Erosheva, E. A.（2022）。一种统一的排序与分数统计学习模型及其在资助评审小组评审中的应用。机器学习研究杂志，23(210)：1–33。

Rendle, S., Freudenthaler, C., Gantner, Z., 和 Schmidt-Thieme, L.（2012年）。贝叶斯个性化排序：基于隐式反馈的贝叶斯个性化排序。*arXiv* 预印本 *arXiv:1205.2618*。

S´anchez, D., Domingo-Ferrer, J., Mart´ınez, S., 和 Soria-Comas, J.（2016）。通过个体排序微聚合实现效用保留的差分隐私数据发布。《信息融合》，30：1–14。

Sason, I.（2015）。关于反向平克尔不等式。*arXiv* 预印本 *arXiv:1503.07118*。

Schapire, R. E.（2003）。机器学习的提升方法：综述。《非线性估计与分类》，第149–171页。

Shang, S.、Wang, T.、Cuff, P. 和 Kulkarni, S.（2014）。差分隐私在排名聚合中的应用：隐私与准确性。载于第*17*届国际信息融合会议（*FUSION*），第1–7页。IEEE。

Shen, X., Tseng, G. C., Zhang, X., and Wong, W. H. (2003). 关于 ψ学习。《美国统计学会会刊》，98(463):724–734。

Song, B., Lan, Q., Li, Y., and Li, G. (2022). 分布式差分隐私排名聚合。载于《亚太知识发现与数据挖掘会议》，第236–248页。Springer。

Soufiani, H. A., Parkes, D., and Xia, L. (2014). 通过排名拆分计算参数化排名模型。载于《国际机器学习会议》，第360–368页。PMLR。

Su, W. (2021). 你是自己论文的最佳审稿人：一种所有者辅助评分机制。《神经信息处理系统进展》，34:27929–27939。

Sun, H., Dong, B., Wang, H., Yu, T., and Qin, Z. (2018). 在本地差分隐私下对稀疏众包数据进行真值推断。载于《2018 IEEE国际大数据会议（大数据）》，第488–497页。IEEE。

Tang, J., Korolova, A., Bai, X., Wang, X., 和 Wang, X. (2017). 苹果在 macOS 10.12 中实现差分隐私所造成的隐私损失。*arXiv* 预印本 *arXiv:1709.02753*。

Tang, W. (2019). Mallows 排名模型：最大似然估计与再生。发表于国际机器学习会议，第 6125–6134 页。PMLR。

Walker, J. 和 Ben-Akiva, M. (2002). 广义随机效用模型。数学社会科学，43(3):303–343。

Wang, J.、Zhu, J. 和 He, X.（2021）。用于训练双塔推荐系统的跨批次负采样。载于第*44*届国际*ACM SIGIR*信息检索研究与发展会议，第 1632–1636 页。

Wang, T.、Blocki、J.、Li、N. 和 Jha、S.（2017）。用于频数估计的局部差分隐私协议。载于第*26*届 *USENIX* 安全研讨会（*USENIX Security 17*），第 729–745 页。

Watson, J.、Lipford, H.R. 和 Besmer, A.（2015）。将用户偏好映射到隐私默认设置。*ACM*《计算机与人机交互汇刊》（*TOCHI*），22(6):1–20。

Williams, O. 和 McSherry, F.（2010）。概率推断与差分隐私。神经信息处理系统进展，23。

Wu, R.、Zhou, J. P.、Weinberger, K. Q. 和 Guo, C.（2022）。标签差分隐私能否防止标签推断攻击？*arXiv* 预印本 *arXiv:2202.12968*。

Yan, Z.、Li, G. 和 Liu, J.（2020）。本地差分隐私下的隐私排名聚合。《国际智能系统杂志》，35(10):1492–1519。

Yang, J.、Cheng, X.、Su, S.、Chen, R.、Ren, Q. 和 Liu, Y.（2019）。在本地差分隐私下收集偏好排序。载于*2019*年 *IEEE* 第*35*届国际数据工程会议（*ICDE*），第1598–1601页。IEEE。

Yang, J., Yi, X., Zhiyuan Cheng, D., Hong, L., Li, Y., Xiaoming Wang, S., Xu, T., 和 Chi, E. H.（2020）。用于推荐的双塔神经网络学习中的混合负采样。载于 *2020*年 *Web*会议伴随论文集，第441–447页。

Young, H. P.（1986年）。基于成对比较的最优排序与选择。《信息汇集与群体决策》，第113–122页。

Zhang, T.（2004）。基于凸风险最小化的分类方法的统计行为与一致性。《统计年鉴》，32(1)：56–85。

Zhu, J. 和 Hastie, T. (2005). 核逻辑回归与导入向量机。《计算与图形统计学杂志》，14(1):185–205。

补充材料
“排序差分隐私”

Shirong Xu、Will Wei Sun 和 Guang Cheng

引理 1的证明：设 ϕ 和 ϕ' 是满足以下条件的两个相邻排名

$$\big(\phi(\mathcal{I}_i)-\phi(\mathcal{I}_j)\big)\big(\phi'(\mathcal{I}_i)-\phi'(\mathcal{I}_j)\big)>0,$$

~对于任意 $\mathcal{I}_i,\ \mathcal{I}_j\in\Omega\setminus\{I_k\}$。因此，对于任意合成排序 ϕ 和 $\mathcal{I}_i,\ \mathcal{I}_j\in\Omega\setminus\{\mathcal{I}_k\}$，有

$$I\Big((\phi(\mathcal{I}_i)-\phi(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i)-\widetilde{\phi}(\mathcal{I}_j))>0\Big)-I\Big((\phi'(\mathcal{I}_i)-\phi'(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i)-\widetilde{\phi}(\mathcal{I}_j))>0\Big)=0.$$

~设 $M_\Omega=\{(\mathcal{I}_i,\ \mathcal{I}_j):\ \mathcal{I}_i,\ \mathcal{I}_j\in\Omega\ \mathcal{I}_i\neq\mathcal{I}_j\}$。对于任意合成排序 ϕ，可得

$$\begin{aligned}\Big|T(\phi,\widetilde{\phi})-T(\phi',\widetilde{\phi})\Big|&=\Big|\frac{1}{|M_\Omega|}\sum_{(i,j)\in M_\Omega}I\Big((\phi(\mathcal{I}_i)-\phi(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i)-\widetilde{\phi}(\mathcal{I}_j))>0\Big)\\&\quad-\frac{1}{|M_\Omega|}\sum_{(i,j)\in M_\Omega}I\Big((\phi'(\mathcal{I}_i)-\phi'(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i)-\widetilde{\phi}(\mathcal{I}_j))>0\Big)\Big|\\&=\Big|\frac{2}{|M_\Omega|}\sum_{\mathcal{I}_j\in\Omega\setminus\{\mathcal{I}_k\}}I\Big((\phi(\mathcal{I}_k)-\phi(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_k)-\widetilde{\phi}(\mathcal{I}_j))>0\Big)\\&\quad-\frac{2}{|M_\Omega|}\sum_{I_j\in\Omega\setminus\{\mathcal{I}_k\}}I\Big((\phi'(\mathcal{I}_k)-\phi'(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_k)-\widetilde{\phi}(\mathcal{I}_j))>0\Big)\Big|\\&\leq\frac{2}{|\Omega|}.\end{aligned}$$

因此，对于任意相邻排名 ϕ 和 ϕ'，有

$$\max_{\widetilde{\phi}}\Big|T(\phi,\widetilde{\phi})-T(\phi',\widetilde{\phi})\Big|\leq\frac{2}{|\Omega|}.$$

令 $\theta=2^{-1}|\Omega|\epsilon$，即可立即得到所需结果。 □

~**引理的证明 2：** 设 $\mathcal{A}$ 表示算法 1，而 $\phi=\mathcal{A}(\phi)$ 表示输出的合成~排序。设 $\Omega=\{\mathcal{I}_1,\mathcal{I}_2,\ldots,\mathcal{I}_L\}$ 表示项目集。假设 ϕ 满足

$\sim\phi(\mathcal{I}_i) = s_{\mathcal{I}_i}$ 其中 $i \in [L]$ 表示输出的合成排序，其中 $\boldsymbol{s} = (s_{\mathcal{I}_i})_{i\in[L]}$ 是 $(i)_{i\in[L]}$的一个排列。

根据算法 1，在第 $t \in \{2, \ldots, L\}$ 次迭代中，所选项目 $\mathcal{I}_j$的排名为 t，在这些项目中的排名，也就是 $\mathcal{I}_j = \phi^{-1}(t)$。注意，先前选出的所有项目的序数排名都更小。因此不难看出

$$\mathbb{P}(V_\epsilon^{(t)} = k) \propto \exp\big(\epsilon(L-1)^{-1}k\big),$$

对于 $k = 0, \ldots, t-1$。由此，我们还可以得到

$$\mathbb{P}(V_\epsilon^{(t)} = k) = \frac{\exp\big(\epsilon k(L-1)^{-1}\big)}{\sum_{k=0}^{t-1}\exp\big(\epsilon k(L-1)^{-1}\big)}, k = 0, \ldots, t-1.$$

定义 $H_t(\boldsymbol{s}) = \{i\colon \phi(\mathcal{I}_i) < t,\ s_{\mathcal{I}_i} < s_{\phi^{-1}(t)}\ i \in [L]\}$ 对于 $t = 2, \ldots, L$。容易验证，$|H_t(\boldsymbol{s})| = k$ 意味着 $V^{(t)} = k$。注意到 $|H_1(\boldsymbol{s})| = 0$，由此可得

$$\mathbb{P}\Big(\mathcal{A}(\phi) = \widetilde{\phi}\Big) = \prod_{t=2}^{L}\frac{\exp(\epsilon(L-1)^{-1}|H_t(\boldsymbol{s})|)}{\sum_{k=0}^{t-1}\exp(\epsilon(L-1)^{-1}k)} \propto \exp\Big(\epsilon(L-1)^{-1}\sum_{t=1}^{L}|H_t(\boldsymbol{s})|\Big). \tag{12}$$

接下来，还需证明 $\sum_{t=1}^{L}|H_t(\boldsymbol{s})| = 2^{-1}|\Omega|(|\Omega|-1)T(\phi, \widetilde{\phi})$。根据$H_t(\boldsymbol{s})$的定义，我们有

$$\begin{aligned}|H_t(\boldsymbol{s})| &= \sum_{i:\phi(\mathcal{I}_i)<t} I\Big((\phi(\mathcal{I}_i) - t)(s_{\mathcal{I}_i} - s_{\phi^{-1}(t)}) > 0\Big)\\ &= \sum_{i:\phi(\mathcal{I}_i)<t} I\Big((\phi(\mathcal{I}_i) - \phi(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i) - \widetilde{\phi}(\mathcal{I}_j)) > 0\Big),\end{aligned}$$

其中 $\mathcal{I}_j = \phi^{-1}(t)$。对 t 从 1 到 L 求和可得

$$\begin{aligned}\sum_{t=1}^{L}|H_t(\boldsymbol{s})| =& \sum_{t=1}^{L}\sum_{i:\phi(\mathcal{I}_i)<t} I\Big((\phi(\mathcal{I}_i) - t)(s_{\mathcal{I}_i} - s_{\phi^{-1}(t)}) > 0\Big)\\ =& \sum_{j=1}^{L}\sum_{\{i:\phi(\mathcal{I}_i)<\phi(\mathcal{I}_j)\}} I\Big((\phi(\mathcal{I}_i) - \phi(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i) - \widetilde{\phi}(\mathcal{I}_j)) > 0\Big)\\ =& \frac{1}{2}\sum_{j=1}^{L}\sum_{i\neq j} I\Big((\phi(\mathcal{I}_i) - \phi(\mathcal{I}_j))(\widetilde{\phi}(\mathcal{I}_i) - \widetilde{\phi}(\mathcal{I}_j)) > 0\Big)\\ =& \frac{1}{2}|\Omega|(|\Omega|-1)T(\phi, \widetilde{\phi}).\end{aligned} \tag{13}$$

将（13）代入（12），可得

$$\mathbb{P}\Big(\mathcal{A}(\phi)=\widetilde{\phi}\Big)\propto\exp(2^{-1}\epsilon|\Omega|T(\phi,\widetilde{\phi})).$$

证明至此完成。 □

定理 1 的证明：首先，我们证明性质 (1)。令 $\Omega=\{\mathcal{I}_1,\ldots,\mathcal{I}_L\}$ 为项目集。我们考虑 $\mathbb{P}\big(\Phi(\mathcal{I}_i)>\Phi(\mathcal{I}_j)\big)>\mathbb{P}\big(\Phi(\mathcal{I}_i)<\Phi(\mathcal{I}_j)\big)$ 的情形，这意味着$2\eta_{ij}-1>0$。此外，对于任意 $\mathcal{I}_i,\mathcal{I}_j\in\Omega$，我们定义两个排序集如下

$$Y_{ij}^{g}=\big\{\phi:\big(\phi(\mathcal{I}_i)\big)_{\mathcal{I}_i\in\Omega}\in\Upsilon(L),\phi(\mathcal{I}_i)>\phi(\mathcal{I}_j)\big\},$$
$$Y_{ij}^{l}=\big\{\phi:\big(\phi(\mathcal{I}_i)\big)_{\mathcal{I}_i\in\Omega}\in\Upsilon(L),\phi(\mathcal{I}_i)<\phi(\mathcal{I}_j)\big\},$$

其中，Y_{ij}^{g}和 Y_{ij}^{l}分别表示其元素满足 $\phi(\mathcal{I}_i)>\phi(\mathcal{I}_j)$ 和$\phi(\mathcal{I}_j)>\phi(\mathcal{I}_i)$ 的排序集。

令 $\widetilde{\Phi}=\mathcal{A}(\Phi)$ 表示合成排序的随机变量。接下来，我们打算证明$\mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i)>\widetilde{\Phi}(\mathcal{I}_j)\big)>\mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i)<\widetilde{\Phi}(\mathcal{I}_j)\big)$，其定义如下

$$\mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i)>\widetilde{\Phi}(\mathcal{I}_j)\big)=\mathbb{P}\big(\widetilde{\Phi}\in Y_{ij}^{g}\big)=\sum_{\widetilde{\phi}\in Y_{ij}^{g}}\mathbb{P}\Big(\mathcal{A}(\Phi)=\widetilde{\phi}\Big),$$
$$\mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i)<\widetilde{\Phi}(\mathcal{I}_j)\big)=\mathbb{P}\big(\widetilde{\Phi}\in Y_{ij}^{l}\big)=\sum_{\widetilde{\phi}\in Y_{ij}^{l}}\mathbb{P}\Big(\mathcal{A}(\Phi)=\widetilde{\phi}\Big),$$

分别。

考虑两个合成排序 $\phi_1\in Y_{ij}^{g}$和 $\phi_2\in Y_{ij}^{l}$，它们满足 $\phi_1(\mathcal{I}_i)=\phi_2(\mathcal{I}_j)=k_1$，$\phi_1(\mathcal{I}_j)=\phi_2(\mathcal{I}_i)=k_2$，其中 $k_1>k_2$，且 $\phi_1(\mathcal{I}_l)=\phi_2(\mathcal{I}_l)$，对所有 $l\in[L]\setminus\{i,j\}$ 均成立。注意，由 $\phi_1\in Y_{ij}^{g}$ 与 $\phi_2\in Y_{ij}^{l}$ 之间的一一对应关系可知，ϕ_1 和 ϕ_2 是一对序数排序，因此不难看出$|Y_{ij}^{l}|=|Y_{ij}^{g}|$。根据引理 2，我们有

$$\frac{\mathbb{P}\Big(\mathcal{A}(\phi)=\widetilde{\phi}_1\Big)}{\mathbb{P}\Big(\mathcal{A}(\phi)=\widetilde{\phi}_2\Big)}=\exp\Big(2^{-1}\epsilon L\big(T(\phi,\widetilde{\phi}_1)-T(\phi,\widetilde{\phi}_2)\big)\Big),$$

因此，只需比较质量函数 T 的值，即可确定 $\mathbb{P}\big(\mathcal{A}(\phi)=\widetilde{\phi}_1\big)$ 和 $\mathbb{P}\big(\mathcal{A}(\phi)=\widetilde{\phi}_2\big)$的相对顺序。

对于任意排序 ϕ，定义 $S(\phi,\ k_2,\ k_1) = \{i:\ k_2 < \phi(\mathcal{I}_i) < k_1,\ i \in [L]\}$ 为 Ω 的子集，其中的序数排名小于 k_1 且大于 k_2。根据 ϕ_1 和 ϕ_2的定义，不难验证

$$S(\widetilde{\phi}_1, k_2, k_1) = S(\widetilde{\phi}_2, k_2, k_1).$$

此外，我们记 $\Delta_{ij}(\phi,\ \widetilde{\phi}_1) = I\big((\phi(\mathcal{I}_i)-\phi(\mathcal{I}_j))(\widetilde{\phi}_1(\mathcal{I}_i)-\widetilde{\phi}_1(\mathcal{I}_j)) > 0\big)$。根据 T 的定义，我们得到

$$\begin{aligned}
T(\phi,\widetilde{\phi}_1) - T(\phi,\widetilde{\phi}_2) &= \frac{1}{L(L-1)} \sum_{\mathcal{I}_i,\mathcal{I}_j\in\Omega} \Big(\Delta_{ij}(\phi,\widetilde{\phi}_1) - \Delta_{ij}(\phi,\widetilde{\phi}_2)\Big) \\
=&\frac{2}{L(L-1)} \sum_{l\in S(\widetilde{\phi}_1,k_2,k_1)} \Big(\Delta_{il}(\phi,\widetilde{\phi}_1) + \Delta_{jl}(\phi,\widetilde{\phi}_1) - \Delta_{jl}(\phi,\widetilde{\phi}_2) - \Delta_{il}(\phi,\widetilde{\phi}_2)\Big) \\
+&\frac{2}{L(L-1)}\Big(\Delta_{ij}(\phi,\widetilde{\phi}_1) - \Delta_{ij}(\phi,\widetilde{\phi}_2)\Big). \qquad (14)
\end{aligned}$$

接下来，我们考虑 $\phi \in Y_{ij}^g$ 和 $\phi \in Y_{ij}^l$ 两种情况。注意，$\phi \in Y_{ij}^g$ 蕴含 $\phi(\mathcal{I}_i) > \phi(\mathcal{I}_j)$，且 $\Delta_{ij}(\phi,\ \widetilde{\phi}_1) - \Delta_{ij}(\phi,\ \widetilde{\phi}_2) = 1$。对于任意 $l \in S(\phi_1,\ k_2,\ k_1)$，可以验证

$$\Delta_{il}(\phi,\widetilde{\phi}_1) + \Delta_{jl}(\phi,\widetilde{\phi}_1) - \Delta_{jl}(\phi,\widetilde{\phi}_2) - \Delta_{il}(\phi,\widetilde{\phi}_2) = \begin{cases} 2, \phi(\mathcal{I}_j) < \phi(\mathcal{I}_l) < \phi(\mathcal{I}_i), \\ 0, \text{otherwise.} \end{cases}$$

因此，如果 $\phi \in Y_{ij}^g$，（14）可以写成

$$T(\phi,\widetilde{\phi}_1) - T(\phi,\widetilde{\phi}_2) = \frac{4|S(\widetilde{\phi}_1, k_2, k_1) \cap S(\phi, \phi(\mathcal{I}_j), \phi(\mathcal{I}_i))| + 2}{L(L-1)}.$$

因此可得

$$\frac{2}{L(L-1)} \le T(\phi,\widetilde{\phi}_1) - T(\phi,\widetilde{\phi}_2) \le \frac{4(L-2)+2}{L(L-1)}.$$

因此，对于任意 $\phi \in Y_{ij}^g$，均有

$$\begin{aligned}
\frac{\mathbb{P}\big(\mathcal{A}(\phi) \in Y_{ij}^g\big)}{\mathbb{P}\big(\mathcal{A}(\phi) \in Y_{ij}^l\big)} &= \frac{\sum_{\widetilde{\phi}_1\in Y_{ij}^g} \mathbb{P}\big(\mathcal{A}(\phi) = \widetilde{\phi}_1\big)}{\sum_{\widetilde{\phi}_2\in Y_{ij}^l} \mathbb{P}\big(\mathcal{A}(\phi) = \widetilde{\phi}_2\big)} \ge \exp\Big((L-1)^{-1}\epsilon\Big), \\
\frac{\mathbb{P}\big(\mathcal{A}(\phi) \in Y_{ij}^g\big)}{\mathbb{P}\big(\mathcal{A}(\phi) \in Y_{ij}^l\big)} &= \frac{\sum_{\widetilde{\phi}_1\in Y_{ij}^g} \mathbb{P}\big(\mathcal{A}(\phi) = \widetilde{\phi}_1\big)}{\sum_{\widetilde{\phi}_2\in Y_{ij}^l} \mathbb{P}\big(\mathcal{A}(\phi) = \widetilde{\phi}_2\big)} \le \exp\Big(\frac{2L-3}{L-1}\epsilon\Big).
\end{aligned}$$

对 $\phi \in Y_{ij}^l$ 的情况应用类似的论证，可得

$$\exp\Big((L-1)^{-1}\epsilon\Big) \leq \frac{\mathbb{P}\big(\mathcal{A}(\phi) \in Y_{ij}^l\big)}{\mathbb{P}\big(\mathcal{A}(\phi) \in Y_{ij}^g\big)} \leq \exp\Big(\frac{2L-3}{L-1}\epsilon\Big).$$

注意到 Y_{ij}^g 和 Y_{ij}^l 互斥且 $\Upsilon = Y_{ij}^g \cup Y_{ij}^l$，我们得到

$$\begin{aligned}
&\mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i) > \widetilde{\Phi}(\mathcal{I}_j)\big) - \mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_j) > \widetilde{\Phi}(\mathcal{I}_i)\big) \\
=&\Big(\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^g\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^g\big)\Big)\mathbb{P}\big(\Phi \in Y_{ij}^g\big) \\
+&\Big(\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^l\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^l\big)\Big)\mathbb{P}\big(\Phi \in Y_{ij}^l\big). \qquad (15)
\end{aligned}$$

根据 Y_{ij}^l 和 Y_{ij}^g 之间的对称性，可得

$$\begin{aligned}
&\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^g\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^g\big) \\
=&\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^l\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^l\big). \qquad (16)
\end{aligned}$$

将（16）代入（15）可得

$$\begin{aligned}
&\mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i) > \widetilde{\Phi}(\mathcal{I}_j)\big) - \mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_j) > \widetilde{\Phi}(\mathcal{I}_i)\big) \\
=&\Big(\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^g\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^g\big)\Big)\Big(\mathbb{P}\big(\Phi \in Y_{ij}^g\big) - \mathbb{P}\big(\Phi \in Y_{ij}^l\big)\Big) \\
=&\Big(\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^g\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^g\big)\Big)\Big(\mathbb{P}\big(\Phi(\mathcal{I}_i) > \Phi(\mathcal{I}_j)\big) - \mathbb{P}\big(\Phi(\mathcal{I}_j) > \Phi(\mathcal{I}_i)\big)\Big) \\
=&\Big(\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^g\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^g\big)\Big)(2\eta_{ij} - 1) > 0. \qquad (17)
\end{aligned}$$

由 $\mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^g\big) - \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^g\big) > 0$这一事实可知

$$\Big(\mathbb{P}\big(\Phi(\mathcal{I}_i) > \Phi(I_j)\big) - \mathbb{P}\big(\Phi(\mathcal{I}_i) < \Phi(\mathcal{I}_j)\big)\Big)\Big(\mathbb{P}\big(\widetilde{\Phi}(I_i) > \widetilde{\Phi}(\mathcal{I}_j)\big) - \mathbb{P}\big(\widetilde{\Phi}(\mathcal{I}_i) < \widetilde{\Phi}(\mathcal{I}_j)\big)\Big) > 0,$$

这就完成了性质（1）的证明。

接下来，我们来证明性质 (2)。为方便记号，我们令 $A_1 = \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^g | \Phi \in Y_{ij}^g\big)$和 $A_2 = \mathbb{P}\big(\widetilde{\Phi} \in Y_{ij}^l | \Phi \in Y_{ij}^g\big)$。根据 $A_1 + A_2 = 1$和 $\exp\big((L-1)^{-1}\epsilon\big) \leq A_1/A_2 \leq \exp(\frac{2L-3}{L-1}\epsilon)$, 可得

$$A_2 \leq \exp\big(-(L-1)^{-1}\epsilon\big)A_1 \text{ and } A_1 \leq \exp\big(\frac{2L-3}{L-1}\epsilon\big)A_2.$$

结合这些结果与 $A_1 + A_2 = 1$，可得

$$\frac{\exp\left((L-1)^{-1}\epsilon\right)-1}{\exp\left((L-1)^{-1}\epsilon\right)+1} \leq A_1 - A_2 \leq \frac{\exp\left(\frac{2L-3}{L-1}\epsilon\right)-1}{\exp\left(\frac{2L-3}{L-1}\epsilon\right)+1}.$$

结合式（17），可得

$$\frac{\exp\left((L-1)^{-1}\epsilon\right)-1}{\exp\left((L-1)^{-1}\epsilon\right)+1} \leq \frac{|2\widetilde{\eta}_{ij}-1|}{|2\eta_{ij}-1|} \leq \frac{\exp\left(\frac{2L-3}{L-1}\epsilon\right)-1}{\exp\left(\frac{2L-3}{L-1}\epsilon\right)+1},$$

其中 $\eta_{ij} = \mathbb{P}\big(\Phi(I_i) > \Phi(I_j)\big)$ 且 $\widetilde{\eta}_{ij} = \mathbb{P}\big(\widetilde{\Phi}(I_i) > \widetilde{\Phi}(I_j)\big)$。由此完成了性质（2）的证明。

接下来，我们证明性质（3）。在第 t 次迭代中，$V_\epsilon^{(t)}$服从如下分布

$$\mathbb{P}(V_\epsilon^{(t)} = k) = \frac{\exp\left(\epsilon(|\Omega|-1)^{-1}\tau(k,\chi^{(t-1)})\right)}{\sum_{k\in\rho^{(t)}}\exp\left(\epsilon(|\Omega|-1)^{-1}\tau(k,\chi^{(t-1)})\right)}, k = 0,\ldots,t-1.$$

注意，项目是按其序数排名依次选出的，因此 $t > \sigma(l)$ 对于 $l \in \chi^{(t)}$和 $V_\epsilon^{(t)} = k$ 意味着 $\tau(k, \chi^{(t)}) = k$。因此，我们有

$$\mathbb{E}\big[V_\epsilon^{(t)}\big] = \sum_{k=0}^{t-1}\mathbb{P}(V_\epsilon^{(t)} = k)k = \sum_{k=0}^{t-1}\frac{k\exp\left(\epsilon(|\Omega|-1)^{-1}\tau(k,\chi^{(t)})\right)}{\sum_{k\in\rho^{(t)}}\exp\left(\epsilon(|\Omega|-1)^{-1}\tau(k,\chi^{(t)})\right)}.$$

为简化记号，我们记 $Q = \sum_{k\in\rho^{(t)}}\exp\left(\epsilon(|\Omega|-1)^{-1}\tau(k,\chi^{(t)})\right)$。我们有

$$\mathbb{E}\big[V_\epsilon^{(t)}\big] = Q^{-1}\sum_{k=1}^{t-1}k\exp\left(\epsilon(|\Omega|-1)^{-1}k\right).$$

直接代数运算表明

$$\begin{aligned}\mathbb{E}\big[V_\epsilon^{(t)}\big] =& \frac{Q^{-1}(t-1)\exp\left(\epsilon(|\Omega|-1)^{-1}t\right) - Q^{-1}\sum_{k=1}^{t-1}\exp\left(\epsilon(|\Omega|-1)^{-1}k\right)}{\exp\left(\epsilon(|\Omega|-1)^{-1}\right)-1} \\ =& \frac{(t-1)q_\epsilon^t}{q_\epsilon^t-1} - \frac{q_\epsilon^t - q_\epsilon}{(q_\epsilon-1)(q_\epsilon^t-1)},\end{aligned}$$

其中 $q_\epsilon = \exp\left(\epsilon(|\Omega|-1)^{-1}\right)$。当 ϵ 趋于无穷大时，可以验证 $\mathbb{E}\big[V_{\left(\frac{t}{\epsilon}\right)}\big]$收敛到 $t-1$。

注意，$\mathrm{Var}(V_\epsilon^{(t)}) = \mathbb{E}(V_{(t)\epsilon})^2 - \big(\mathbb{E}(V_{(t)\epsilon})\big)^2$，因此还需计算 $\mathbb{E}(V_{(t)\epsilon})^2$。

$$\mathbb{E}(V_\epsilon^{(t)})^2 = Q^{-1}\sum_{k=1}^{t-1}k^2\exp\left(\epsilon(|\Omega|-1)^{-1}k\right).$$

应用类似的论证，我们得到

$$\begin{aligned}\mathbb{E}(V_\epsilon^{(t)})^2 &= \frac{Q^{-1}(t-1)^2\exp\left(\epsilon(|\Omega|-1)^{-1}t\right)}{\exp\left(\epsilon(|\Omega|-1)^{-1}\right)-1} - \frac{Q^{-1}\sum_{k=1}^{t-1}(2k-1)\exp\left(\epsilon(|\Omega|-1)^{-1}k\right)}{\exp\left(\epsilon(|\Omega|-1)^{-1}\right)-1} \\ &= \frac{(t-1)^2\exp\left(\epsilon(|\Omega|-1)^{-1}t\right)}{\exp\left(\epsilon(|\Omega|-1)^{-1}t\right)-1} - \frac{Q^{-1}\sum_{k=1}^{t-1}(2k-1)\exp\left(\epsilon(|\Omega|-1)^{-1}k\right)}{\exp\left(\epsilon(|\Omega|-1)^{-1}\right)-1} \\ &= \frac{(t-1)^2\exp\left(\epsilon(|\Omega|-1)^{-1}t\right)}{\exp\left(\epsilon(|\Omega|-1)^{-1}t\right)-1} - \frac{2\mathbb{E}(V_\epsilon^{(t)})}{\exp\left(\epsilon(|\Omega|-1)^{-1}\right)-1} \\ &+ \frac{\exp\left(\epsilon(|\Omega|-1)^{-1}t\right)-\exp\left(\epsilon(|\Omega|-1)^{-1}\right)}{\exp\left(\epsilon(|\Omega|-1)^{-1}t\right)-1}\frac{1}{\exp\left(\epsilon(|\Omega|-1)^{-1}\right)-1}.\end{aligned}$$

容易验证，$\lim_{\epsilon\to+\infty}\mathbb{E}(V_{(t)\epsilon})^2=(t-1)^2$。至此，整个证明就完成了。 □

定理 2的证明：注意， $\mathcal{L}(\phi)$ 的最小化解位于一个离散空间中。因此，^无法推导出 ϕ的解析形式。不失一般性，我们假设真实排名 ϕ 满足 $\phi(I_i)=i$，其中 $i=1,\ldots,L$。

我们在此给出的证明与 Tang（2019）中定理 3.2 的证明相似。令 $\mathcal{Q}_1=\{\phi:\phi(I_1)=2,\phi(I_2)=1\}$ 表示项目 I_1 和 I_2 交换名次的排序集合，令 $\mathcal{Q}_2=\{\phi:\phi(I_1)=1,\phi(I_2)=2\}$ 表示项目 I_1 和I_2 名次正确的排序集合。这里需要注意，真实排名是 $\mathcal{Q}_2$ 的一个元素。对于任意 $\phi_1\in\mathcal{Q}_1$，存在一个 $\phi_2\in\mathcal{Q}_2$，使得对于 $l\in\{3,\ldots,L\}$，$\phi_1(I_l)=\phi_2(I_l)$。令$\Pi_{ij}(\mathcal{S})=\sum_{l=1}^N I\left(\widetilde{\phi}_l(I_i)>\widetilde{\phi}_l(I_j)\right)$ 表示项目 I_i 的排名高于项目 I_j的频数。如果 $\Pi_{12}(\mathcal{S})>N/2$，则容易验证

$$\mathcal{L}(\phi_1)>\mathcal{L}(\phi_2).$$

因此，有

$$\mathcal{L}(\phi)\le\max_{\phi_2\in\mathcal{Q}_2}\mathcal{L}(\phi)<\max_{\phi_1\in\mathcal{Q}_1}\mathcal{L}(\phi).$$

^因此， $\Pi_{12}(\mathcal{S})>N/2$ 意味着 $\phi\neq\phi$。

$$\mathbb{P}_{\mathcal{S}}\left(\widehat{\phi}_{mle}\neq\phi\right)\ge\mathbb{P}_{\mathcal{S}}\left(\Pi_{12}(\mathcal{S})>N/2\right).$$

回想一下，在算法 1中， I_1 和 I_2 的名次由概率决定

$$\mathbb{P}\big(\widetilde{\phi}_l(I_2) < \widetilde{\phi}_l(I_1)\big) = \frac{1}{1+\exp(\epsilon(|\Omega|-1)^{-1})}, \text{ for } l = 1, \ldots, N.$$

为便于记号表示，我们记作 $p_{\epsilon,\Omega} = \frac{\exp(\epsilon(|\Omega|-1)^{-1})}{1+\exp(\epsilon(|\Omega|-1)^{-1})}$。于是可得

$$\mathbb{P}_{\mathcal{S}}\Big(\Pi_{12}(\mathcal{S}) > N/2\Big) \geq \mathbb{P}_{\mathcal{S}}\Big(\Pi_{21}(\mathcal{S}) \leq \lfloor N/2 \rfloor\Big) = \sum_{k=0}^{\lfloor N/2 \rfloor} \binom{N}{k} p_{\epsilon,\Omega}^{k} (1-p_{\epsilon,\Omega})^{N-k}.$$

根据Arratia 和 Gordon (1989的定理 2，存在一些正的常数 C_0，使得通过用斯特林公式近似二项式系数，我们得到

$$\mathbb{P}_{\mathcal{S}}\Big(\Pi_{12}(\mathcal{S}) > N/2\Big) \geq C_0 \frac{\exp(\epsilon(|\Omega|-1)^{-1})+1}{\exp(\epsilon(|\Omega|-1)^{-1})-1} \sqrt{\frac{2}{\pi N}} \exp\Big(-N D_{KL}(1/2 \| p_{\epsilon,\Omega})\Big), \tag{18}$$

其中 $D_{KL}(1/2\|p_{\epsilon,\Omega})$ 表示参数分别为 $1/2$ 和$p_{\epsilon,\Omega}$的两个伯努利随机变量之间的 KL 散度。对右侧应用反向 Pinsker 不等式（Sason，2015）于（18），进一步可得

$$\begin{aligned}
\mathbb{P}_{\mathcal{S}}\Big(\Pi_{12}(\mathcal{S}) > N/2\Big) \geq & C_0 \frac{\exp(\epsilon(|\Omega|-1)^{-1})+1}{\exp(\epsilon(|\Omega|-1)^{-1})-1} \sqrt{\frac{2}{\pi N}} \exp\Big\{-2N \frac{\Big(\exp(\epsilon(|\Omega|-1)^{-1})-1\Big)^2}{1+\exp(\epsilon(|\Omega|-1)^{-1})}\Big\} \\
\geq & C_0 \frac{\exp(\epsilon(|\Omega|-1)^{-1})+1}{\exp(\epsilon(|\Omega|-1)^{-1})-1} \sqrt{\frac{2}{\pi N}} \exp\Big\{-2N \frac{\Big(\exp(\epsilon(|\Omega|-1)^{-1})-1\Big)^2}{2}\Big\} \\
\geq & C_0 \frac{\exp(\epsilon(|\Omega|-1)^{-1})+1}{\exp(\epsilon(|\Omega|-1)^{-1})-1} \sqrt{\frac{2}{\pi N}} \exp\Big\{-N\Big(\exp(\epsilon(|\Omega|-1)^{-1})-1\Big)^2\Big\}.
\end{aligned}$$

记 $C(\epsilon) = \exp(\epsilon(|\Omega|-1)^{-1}) - 1$，则下界可改写为

$$\mathbb{P}_{\mathcal{S}}\Big(\Pi_{12}(\mathcal{S}) > N/2\Big) \geq C_0 \sqrt{\frac{8}{C^2(\epsilon)\pi N}} \exp\Big\{-N C^2(\epsilon)\Big\}. \tag{19}$$

此外，令 $C(\epsilon) = O\big(N^{-1/2}\big)$，则对于任意 $N \geq 1$，式（19）的右侧都与 0 保持正距离。注意到 $C(\epsilon) = O\big(N^{-1/2}\big)$ 蕴含$\epsilon = O\big((|\Omega|-1)N^{-1/2}\big)$，所需结果立即可得。

□

引理的证明3：不失一般性，我们假设原始排序 ϕ 为

$$\phi(\mathcal{I}_i) = i, i \in [m].$$

设 $\widetilde{\phi}_{lap} = \mathcal{M}_{\lambda}^{lap}(\phi)$ 和 $\widetilde{\phi}'_{lap} = \mathcal{M}_{\lambda}^{lap}(\phi')$ 是拉普拉斯机制的输出。对于任意 $\boldsymbol{r} \in \mathbb{R}^m$，$\phi_{lap} = \boldsymbol{r}$ 和 $\phi'_{lap} = \boldsymbol{r}$ 的联合概率密度函数可写为

$$\mathbb{P}(\widetilde{\phi}_{lap} = \boldsymbol{r}) = \frac{1}{2^m \lambda^m} \exp\Big(- \frac{\sum_{i=1}^m |r_i - \phi(\mathcal{I}_i)|}{\lambda} \Big),$$
$$\mathbb{P}(\widetilde{\phi}'_{lap} = \boldsymbol{r}) = \frac{1}{2^m \lambda^m} \exp\Big(- \frac{\sum_{i=1}^m |r_i - \phi'(\mathcal{I}_i)|}{\lambda} \Big).$$

根据相邻排名的定义，我们可以很容易地验证

$$\Big| \sum_{i=1}^m |r_i - \phi(\mathcal{I}_i)| - \sum_{i=1}^m |r_i - \phi'(\mathcal{I}_i)| \Big| \leq 2(m-1),$$

其中，当 ϕ' 满足 $\phi'(\mathcal{I}_i) = i+1$ 对 $i = 1$、……、$m-1$ 以及 $\phi'(\mathcal{I}_1) = m$ 且 $r_i \geq m$ 对 $i = 1$、……、m 时，等号成立。因此，令 $\lambda = 2(m-1)/\epsilon$，我们有

$$\Big| \log \frac{\mathbb{P}(\widetilde{\phi}_{lap} = \boldsymbol{r})}{\mathbb{P}(\widetilde{\phi}'_{lap} = \boldsymbol{r})} \Big| \leq \epsilon.$$

证明至此完成。 □

引理的证明 4：不失一般性，我们假设原始排序 ϕ 满足 $\phi(\mathcal{I}_i) = i$ 对 $i = 1$、……、m。

记 $\widetilde{\phi}_{lap} = \mathcal{M}_2^{lap}{}_{(m-1)\epsilon^{-1}}(\phi)$。对于任意 $i > j$，我们定义一个新的随机变量：$Z_{ij} = \phi_{lap}(\mathcal{I}_i) - \phi_{lap}(\mathcal{I}_j)$。根据 Nadarajah (2007) 的推论 2，Z_{ij} 的累积分布函数形式如下：

$$F_{Z_{ij}}(z) = \begin{cases} \frac{1}{2}\exp\Big(\frac{\epsilon(z-i+j)}{2(m-1)}\Big) - \frac{(z-i+j)\epsilon}{8(m-1)}\exp\Big(\frac{\epsilon(z-i+j)}{2(m-1)}\Big), z < i-j, \\ 1 - \frac{1}{2}\exp\Big(-\frac{\epsilon(z-i+j)}{2(m-1)}\Big) + \frac{(z-i+j)\epsilon}{8(m-1)}\exp\Big(-\frac{\epsilon(z-i+j)}{2(m-1)}\Big) z > i-j. \end{cases}$$

因此，我们有

$$\mathbb{P}\Big(\widetilde{\phi}_{lap}(\mathcal{I}_i) > \widetilde{\phi}_{lap}(\mathcal{I}_j)\Big) = \mathbb{P}\Big(Z_{ij} > 0\Big) = 1 - F_{Z_{ij}}(0)$$
$$= 1 - \frac{1}{2}\exp\Big(-\frac{\epsilon(i-j)}{2(m-1)}\Big) - \frac{(i-j)\epsilon}{8(m-1)}\exp\Big(-\frac{\epsilon(i-j)}{2(m-1)}\Big).$$

令 $U_{\epsilon}^{t-1} = \mathbb{E}\big[\sum_{j=1}^{t-1} I(\widetilde{\phi}_{lap}(\mathcal{I}_t) > \widetilde{\phi}_{lap}(\mathcal{I}_j))\big]$ 表示正确偏序的预期数量

项目 t 及排名较低的项目。由此可知，对于任意 $t \geq 2$,

$$\begin{aligned}
\mathbb{E}\big[U_\epsilon^{t-1}\big] =& \sum_{j=1}^{t-1} \mathbb{P}\Big(\widetilde{\phi}_{lap}(\mathcal{I}_t) > \widetilde{\phi}_{lap}(\mathcal{I}_j)\Big) \\
=& t-1-\frac{\exp(-t\lambda_\epsilon)}{2}\sum_{j=1}^{t-1}\exp\left(j\lambda_\epsilon\right) - \sum_{j=1}^{t-1}\frac{j\lambda_\epsilon}{4}\exp\left(-j\lambda_\epsilon\right) \\
=& t-1-\frac{1-\exp(-(t-1)\lambda_\epsilon)}{2(\exp(\lambda_\epsilon)-1)} + \frac{\lambda_\epsilon(t-1)\exp(-(t-1)\lambda_\epsilon)}{4(\exp(\lambda_\epsilon)-1)} - \lambda_\epsilon\frac{\exp(-\lambda_\epsilon)-\exp(-t\lambda_\epsilon)}{4(1-\exp(-\lambda_\epsilon))^2} \\
=& t-1-\frac{1-p_\epsilon^{-(t-1)}}{2(p_\epsilon-1)} + \frac{\lambda_\epsilon(t-1)p_\epsilon^{-(t-1)}}{4(p_\epsilon-1)} - \lambda_\epsilon\frac{p_\epsilon-p_\epsilon^{-(t-2)}}{4(p_\epsilon-1)^2} \\
=& t-1-\frac{p_\epsilon^{t-1}-1}{2(p_\epsilon^t-p_\epsilon^{t-1})} + \frac{\lambda_\epsilon(t-1)}{4(p_\epsilon^t-p_\epsilon^{t-1})} - \lambda_\epsilon\frac{p_\epsilon-p_\epsilon^{-(t-2)}}{4(p_\epsilon-1)^2}
\end{aligned}$$

其中 $p_\epsilon = \exp(\lambda_\epsilon)$ 且 $\lambda_\epsilon = \epsilon 2^{-1}(m-1)^{-1}$。根据定理 1 的性质 (3)，可得

$$\mathbb{E}\big[V_\epsilon^{t-1}\big] = t-1+\frac{t}{q_\epsilon^t-1}-\frac{1}{q_\epsilon-1},$$

其中 $q_\epsilon = p_\epsilon^2$。

N 接下来，我们证明对于任意 $2 \leq t \leq M$ 和 $\epsilon >$， $\mathbb{E}\big[V_\epsilon^{t-1}\big] - \mathbb{E}\big[U_\epsilon^{t-1}\big] > 0$ 0.

$$\mathbb{E}\big[V_\epsilon^{t-1}\big] - \mathbb{E}\big[U_\epsilon^{t-1}\big] = \frac{t}{q_\epsilon^t-1}-\frac{1}{q_\epsilon-1}+\frac{p_\epsilon^{t-1}-1}{2(p_\epsilon^t-p_\epsilon^{t-1})}-\frac{\lambda_\epsilon(t-1)}{4(p_\epsilon^t-p_\epsilon^{t-1})}+\lambda_\epsilon\frac{p_\epsilon-p_\epsilon^{-(t-2)}}{4(p_\epsilon-1)^2}.$$

注意到

$$-\frac{\lambda_\epsilon(t-1)}{4(p_\epsilon^t-p_\epsilon^{t-1})}+\lambda_\epsilon\frac{p_\epsilon-p_\epsilon^{-(t-2)}}{4(p_\epsilon-1)^2} = \frac{\lambda_\epsilon(p_\epsilon^t-tp_\epsilon+t-1)}{4(p_\epsilon^t-p_\epsilon^{t-1})(p_\epsilon-1)} \geq 0,$$

对于任意 $p_\epsilon \geq 1$。那么，当 $p_\epsilon \geq 2$ 时，我们有

$$\mathbb{E}\big[V_\epsilon^{t-1}\big] - \mathbb{E}\big[U_\epsilon^{t-1}\big] \geq \frac{t}{q_\epsilon^t-1}+\frac{p_\epsilon^{t-1}(p_\epsilon-1)-(p_\epsilon+1)}{2(p_\epsilon+1)(p_\epsilon^t-p_\epsilon^{t-1})} > 0.$$

接下来，我们考虑 当 $p_\epsilon < 2$ 时的情况。根据 $3/2\lambda_\epsilon \geq p_\epsilon - 1$ 当 $p_\epsilon < 2$ 时，我们得到

$$\frac{\lambda_\epsilon(p_\epsilon^t-tp_\epsilon+t-1)}{4(p_\epsilon^t-p_\epsilon^{t-1})(p_\epsilon-1)} \geq \frac{(p_\epsilon^t-tp_\epsilon+t-1)}{6(p_\epsilon^t-p_\epsilon^{t-1})}.$$

接下来，我们分析该比值

$$\begin{aligned}
Ra(p_\epsilon,t) =& \frac{\frac{1}{p_\epsilon^2-1}-\frac{t}{p_\epsilon^{2t}-1}}{\frac{p_\epsilon^{t-1}-1}{2(p_\epsilon^t-p_\epsilon^{t-1})}+\frac{(p_\epsilon^t-tp_\epsilon+t-1)}{6(p_\epsilon^t-p_\epsilon^{t-1})}} = \frac{6(p_\epsilon^{2t}-tp_\epsilon^2+t-1)(p_\epsilon^t-p_\epsilon^{t-1})}{(p_\epsilon^t+3p_\epsilon^{t-1}-tp_\epsilon+t-4)(p_\epsilon^2-1)(p_\epsilon^{2t}-1)} \\
=& \frac{6(p_\epsilon^{2t}-tp_\epsilon^2+t-1)p_\epsilon^{t-1}}{(p_\epsilon^t+3p_\epsilon^{t-1}-tp_\epsilon+t-4)(p_\epsilon+1)(p_\epsilon^{2t}-1)} \\
=& \frac{6p_\epsilon^{t-1}}{(1+p_\epsilon)(p_\epsilon^t+1)}\frac{(p_\epsilon^{2t}-tp_\epsilon^2+t-1)}{p_\epsilon^{2t}+3p_\epsilon^{2t-1}-tp_\epsilon^{t+1}+(t-5)p_\epsilon^t-3p_\epsilon^{t-1}+tp_\epsilon-t+4}. \qquad (20)
\end{aligned}$$

根据洛必达法则，我们有 $\lim_{p_\epsilon \to 1} Ra(p_\epsilon, t) = 1$ 对于任意 $t \geq 2$。此外，很容易验证，式（20）的分子和分母在$p_\epsilon \in [1, 2]$ 对于任意 $t \geq 2$ 上均为正且递增，并且对于任意 $p_\epsilon \in [1, 2]$，分母都大于分子，且 $t \geq 2$。因此可得 $Ra(p_\epsilon, t) \leq 1$，这意味着

$$\mathbb{E}\big[V_\epsilon^{t-1}\big] - \mathbb{E}\big[U_\epsilon^{t-1}\big] \geq 0,$$

当且仅当 $p_\epsilon = 1$ 表示 $\epsilon = 0$时，等号成立。

综上可得

$$\mathbb{E}\Big[T\big(\phi, \mathcal{A}_\epsilon(\phi)\big)\Big] = 2\sum_{t=2}^{m} \mathbb{E}[V_\epsilon^{t-1}] > 2\sum_{t=2}^{m} \mathbb{E}[U_\epsilon^{t-1}] = \mathbb{E}\Big[T\big(\phi, \mathcal{M}^{lap}_{2(m-1)\epsilon^{-1}}(\phi)\big)\Big],$$

对任意 $\epsilon > 0$均成立。证明完毕。 □

引理 6的证明：根据定理 1的性质 **(1)**，可得

$$\Big(\mathbb{P}\big(\Phi_u(\mathcal{I}_i) > \Phi(\mathcal{I}_j)\big) - \mathbb{P}\big(\Phi_u(\mathcal{I}_i) < \Phi(\mathcal{I}_j)\big)\Big)\Big(\mathbb{P}\big(\widetilde{\Phi}_u(\mathcal{I}_i) > \widetilde{\Phi}_u(\mathcal{I}_j)\big) - \mathbb{P}\big(\widetilde{\Phi}_u(\mathcal{I}_i) < \widetilde{\Phi}_u(\mathcal{I}_j)\big)\Big) > 0.$$

结合引理 5 可得

$$\Big(f^*(\boldsymbol{x}_u, \boldsymbol{y}_i) - f^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\Big)\Big(\widetilde{f}^*(\boldsymbol{x}_u, \boldsymbol{y}_i) - \widetilde{f}^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\Big) > 0, \tag{21}$$

对任意 $\sim u \in [n]$ 且 $i,\ j \in [m]$。根据 $E_{uij}(f)$ 和 $E_{uij}(f)$ 的定义，我们有

$$E_{uij}(f) - E^*_{uij} = |2\eta_{uij} - 1| I\Big(\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big)\big(f^*(\boldsymbol{x}_u, \boldsymbol{y}_i) - f^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) > 0\Big),$$
$$\widetilde{E}_{uij}(f) - \widetilde{E}^*_{uij} = |2\widetilde{\eta}_{uij} - 1| I\Big(\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big)\big(\widetilde{f}^*(\boldsymbol{x}_u, \boldsymbol{y}_i) - \widetilde{f}^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) > 0\Big).$$

所需结果可直接由定理 1 的性质（2）得出。 □

引理 7 的证明：根据定理1 中的性质（1），我们有

$$\frac{\exp\big((|\Omega_u| - 1)^{-1}\epsilon_u\big) - 1}{\exp\big((|\Omega_u| - 1)^{-1}\epsilon_u\big) + 1} \leq \frac{|2\widetilde{\eta}_{uij} - 1|}{|2\eta_{uij} - 1|} \leq \frac{\exp\big(\epsilon_u\big) - 1}{\exp\big(\epsilon_u\big) + 1}.$$

令 $\Psi_u = \exp\big((|\Omega_u|-1)^{-1}\epsilon_u\big)+1 \exp\big((|\Omega_u|-1)^{-1}\epsilon_u\big)-1$，则不难看出 $|2\widetilde{\eta}_{uij} - 1| \leq \beta$蕴含 $|2\eta_{uij} - 1| \leq \Psi_u\beta$。由假设C可得

$$\mathbb{P}\Big(|2\widetilde{\eta}_{uij} - 1| \leq \beta\Big) \leq \mathbb{P}\Big(|2\eta_{uij} - 1| \leq \Psi_u\beta\Big) \leq C_1(\Psi_u)^\gamma\beta^\gamma.$$

证明至此完成。 □

引理 8. 在假设 4下，存在一些正的常数 $C_4 > 0$ 使得

$$Var\big(G_f(u) - G_{f^*}(u)\big) \le 2C_4 \mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big]\big(e_{\upsilon}(f, f^*)\big)^{\gamma/(\gamma+1)}.$$

引理 8 的证明：根据全方差公式，我们有

$$\mathrm{Var}\big(G_f(u) - G_{f^*}(u)\big) = \mathbb{E}\big[\mathrm{Var}_u\big(G_f(u) - G_{f^*}(u)\big)\big] + \mathrm{Var}\big[\mathbb{E}_u\big(G_f(u) - G_{f^*}(u)\big)\big],$$

其中，$\mathrm{Var}_u\big(G_f(u) - G_{f^*}(u)\big)$ 表示在固定用户 u的条件下，$G_f(u) - G_{f^*}(u)$的条件方差；$\mathbb{E}_u\big(G_f(u) - G_{f^*}(u)\big)$ 则是对给定用户u时的合成排序取期望。

假设项目是在固定规模下均匀生成的，则 $\mathrm{Var}_u\big(G_f(u) - G_{f^*}(u)\big)$ 的上界为

$$\begin{aligned}
&\mathrm{Var}_u\big(G_f(u) - G_{f^*}(u)\big) \\
=&\frac{1}{m^2(m-1)^2}\mathrm{Var}\Big(\sum_{i\neq j} I(\widetilde{\Phi}_u(\mathcal{I}_i) > \widetilde{\Phi}_u(\mathcal{I}_j)))\overline{\upsilon}\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(f(\boldsymbol{x}_u, \boldsymbol{y}_j))\big)\Big) \\
\le&\frac{1}{m(m-1)^2}\sum_{i\neq j}\mathrm{Var}\Big(I(\widetilde{\Phi}_u(\mathcal{I}_i) > \widetilde{\Phi}_u(\mathcal{I}_j)))\overline{\upsilon}\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_i, \boldsymbol{y}_j)\big)\Big),
\end{aligned}$$

其中，$\overline{\upsilon}(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)) = \upsilon(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)) - \upsilon(f^*(\boldsymbol{x}_i, \boldsymbol{y}_i) - f^*(\boldsymbol{x}_i, \boldsymbol{y}_j))$，该不等式源于偏序之间存在部分相关性这一事实。

对于任意 $i, j \in [m]$，我们有

$$\begin{aligned}
&\mathrm{Var}\Big(I(\widetilde{\Phi}_u(\mathcal{I}_i) > \widetilde{\Phi}_u(\mathcal{I}_j))\overline{\upsilon}\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big)\Big) \\
\le&\mathbb{P}\big(\widetilde{\Phi}_u(\mathcal{I}_i) > \widetilde{\Phi}_u(\mathcal{I}_j)\big)\overline{\upsilon}^2\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) \\
&+ \mathbb{P}\big(\widetilde{\Phi}_u(\mathcal{I}_j) > \widetilde{\Phi}_u(\mathcal{I}_i)\big)\overline{\upsilon}^2\big(f(\boldsymbol{x}_u, \boldsymbol{y}_j) - f(\boldsymbol{x}_u, \boldsymbol{y}_i)\big) \\
=&\widetilde{\eta}_{uij}\overline{\upsilon}^2\big(f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)\big) + (1 - \widetilde{\eta}_{uij})\overline{\upsilon}^2\big(f(\boldsymbol{x}_u, \boldsymbol{y}_j) - f(\boldsymbol{x}_u, \boldsymbol{y}_i)\big) \\
\le&K_{\upsilon}^2\Big[f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_i, \boldsymbol{y}_j) - f^*(\boldsymbol{x}_u, \boldsymbol{y}_i) + f^*(\boldsymbol{x}_u, \boldsymbol{y}_j)\Big]^2,
\end{aligned}$$

其中，最后一个不等式由 υ的利普希茨连续性得出。为简化记号，我们令 $g_f(u, i, j) = f(\boldsymbol{x}_u, \boldsymbol{y}_i) - f(\boldsymbol{x}_u, \boldsymbol{y}_j)$。由于 f^*是最优极小化器，因此

对于每个用户 u， $\mathbb{E}_u\big(G_f(u) - G_{f^*}(u)\big)$ 可以下界为

$$
\begin{aligned}
&\mathbb{E}_u\big(G_f(u) - G_{f^*}(u)\big) \qquad (22)\\
=&\frac{1}{m(m-1)}\sum_{i<j}\widetilde{\eta}_{uij}\overline{\phi}\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i) - f(\boldsymbol{x}_u,\boldsymbol{y}_j)\big) + (1-\widetilde{\eta}_{uij})\overline{\phi}\big(f(\boldsymbol{x}_u,\boldsymbol{y}_j) - f(\boldsymbol{x}_u,\boldsymbol{y}_i)\big)\\
\geq&\frac{T_1}{m(m-1)}\sum_{i<j}\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i) - f(\boldsymbol{x}_u,\boldsymbol{y}_j) - f^*(\boldsymbol{x}_u,\boldsymbol{y}_i) + f^*(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)^2|1-2\widetilde{\eta}_{uij}|\\
\geq&T_1\mathbb{E}_u\big[(g_f(u,i,j) - g_{f^*}(u,i,j))^2\beta I(|1-2\widetilde{\eta}_{uij}|\geq\beta)\big], \qquad (23)
\end{aligned}
$$

对于某些正常数 $T_1>0$，第一个不等式由以下事实得出： f^* 是最优极小化器，并且当 $\widetilde{\eta}_{uij}=1/2$ 几乎必然成立或 $f=f^*$ 时，该不等式成立。

结合（22）和引理 7，可得

$$
\begin{aligned}
\mathbb{E}_u\big(G_f(u) - G_{f^*}(u)\big) \geq&T_1\beta\mathbb{E}_u\big[(g_f(u,i,j) - g_{f^*}(u,i,j))^2\big] - C_2\Psi_u^{\gamma}\beta^{\gamma+1}\\
=&T_1\big(\frac{T_2}{2C_2}\big)^{\frac{\gamma+1}{\gamma}}\mathbb{E}_u^{\frac{\gamma+1}{\gamma}}\big[(g_f(u,i,j) - g_{f^*}(u,i,j))^2\big]\Psi_u^{-1},
\end{aligned}
$$

其中，最后一个等式是通过取 $\beta=(2^{-1}T_1C_2^{-1})^{\frac{1}{\gamma}}\Psi_u^{-1}\mathbb{E}^{\frac{1}{\gamma}}\big[(g_f(u,i,j) - g_{f^*}(u,i,j))^2\big]$ 得出的。

因此，对于某些依赖于 γ 的正常数 C_4，有

$$
\mathrm{Var}_u\big(G_f(u) - G_{f^*}(u)\big) \leq \frac{C_4}{|\Omega|-1}\Psi_u^{\frac{\gamma}{\gamma+1}}\mathbb{E}_u^{\frac{\gamma}{\gamma+1}}\big(G_f(u) - G_{f^*}(u)\big).
$$

对两边取期望可得

$$
\begin{aligned}
\mathbb{E}\big[\mathrm{Var}_u\big(G_f(u) - G_{f^*}(u)\big)\big] \leq&C_4\mathbb{E}\big[(|\Omega|-1)^{-1}\Psi_u^{\frac{\gamma}{\gamma+1}}\mathbb{E}_u^{\frac{\gamma}{\gamma+1}}\big(G_f(u) - G_{f^*}(u)\big)\big]\\
\leq&\frac{C_4}{|\Omega|-1}\mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big]\big(e_v(f,f^*)\big)^{\frac{\gamma}{\gamma+1}}, \qquad (24)
\end{aligned}
$$

其中，最后一个不等式由 Hölder 不等式得出。

接下来，我们建立 $\mathrm{Var}\big[\mathbb{E}_u\big(G_f(u) - G_{f^*}(u)\big)\big]$ 与 $e_v(f,f^*)$ 之间的关系。

$$
\begin{aligned}
&\mathrm{Var}\big[\mathbb{E}_u\big(G_f(u) - G_{f^*}(u)\big)\big] = \mathbb{E}\Big(\mathbb{E}_u^2\big(G_f(u) - G_{f^*}(u)\big)\Big) - e_v^2(f,f^*)\\
=&\mathbb{E}\Big(\frac{1}{m(m-1)}\sum_{i<j}\widetilde{\eta}_{uij}\overline{v}\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i) - f(\boldsymbol{x}_u,\boldsymbol{y}_j)\big) + (1-\widetilde{\eta}_{uij})\overline{v}\big(f(\boldsymbol{x}_u,\boldsymbol{y}_j) - f(\boldsymbol{x}_u,\boldsymbol{y}_i)\big)\Big)^2 - e_v^2(f,f^*)\\
\leq&\mathbb{E}\Big(\frac{1}{m(m-1)}\sum_{i<j}\widetilde{\eta}_{uij}\overline{v}^2\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i) - f(\boldsymbol{x}_u,\boldsymbol{y}_j)\big) + (1-\widetilde{\eta}_{uij})\overline{v}^2\big(f(\boldsymbol{x}_u,\boldsymbol{y}_j) - f(\boldsymbol{x}_u,\boldsymbol{y}_i)\big)\Big) - e_v^2(f,f^*)\\
\leq&C_4\mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big]\big(e_v(f,f^*)\big)^{\gamma/(\gamma+1)}, \qquad (25)
\end{aligned}
$$

其中，最后一个不等式是通过应用上述相同步骤得到的。结合（24）和（25），可得

$$\mathrm{Var}\big(G_f(u) - G_{f^*}(u)\big) \leq 2C_4 \mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big]\big(e_v(f, f^*)\big)^{\frac{\gamma}{\gamma+1}}.$$

证明至此完成。 □

引理 9。 若 $\big(n^{-1}|\Theta|\log\big(|\Theta|^{-1}n\big)\big)^{\frac{\gamma+1}{\gamma+2}}\mathbb{E}^{\frac{1}{\gamma+2}}\big(\Psi_u^{\gamma}\big) = O(\delta_n)$，则对于 $0 \leq i \leq i_0$ 和 $0 \leq j \leq j_0$，有

$$\mathbb{E}\Big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\Big) \leq 1/2M(i,j), \tag{26}$$

其中 $U_n(f) = n^{-1}\sum_{u=1}^{n}\big[DG_f(u) - \mathbb{E}\big(DG_f(u)\big)\big]$。

引理 9 的证明：对 $\mathbb{E}(\sup_{f\in\mathcal{H}_{ij}} U_n(f))$ 进行界定的主要思路基于拉德马赫复杂度。首先，根据对称化论证，我们有

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\Big) = \mathbb{E}\Big(\sup_{f\in\mathcal{H}_{ij}} \frac{1}{n}\sum_{u=1}^{n}\big(DG_f(u)\big) - \mathbb{E}\big(DG_f(u)\big)\Big)\\
=&\mathbb{E}_{\Omega}\Big[\sup_{f\in\mathcal{H}_{ij}} \mathbb{E}_{\Omega'}\Big(\frac{1}{n}\sum_{u,u'}\big(DG_f(u) - DG_f(u')\big)\Big|\Omega\Big)\Big] \leq \mathbb{E}_{\mathcal{U},\mathcal{U}'}\Big[\sup_{f\in\mathcal{H}_{ij}} \frac{1}{n}\sum_{u,u'}\big(DG_f(u) - DG_f(u')\big)\Big]\\
=&\mathbb{E}_{\mathcal{U},\mathcal{U}'}\Big[\sup_{f\in\mathcal{H}_{ij}} \mathbb{E}_{\boldsymbol{v}}\Big(\frac{1}{n}\sum_{u,u'}\sigma_u\big(DG_f(u) - DG_f(u')\big)\Big)\Big] \leq 2\mathbb{E}_{\mathcal{U}}\Big[\frac{1}{n}\mathbb{E}_{\boldsymbol{\sigma}}\Big(\sup_{f\in\mathcal{H}_{ij}}\sum_{u=1}^{n}\sigma_u DG_f(u)\Big)\Big]\\
=&2\mathbb{E}_{\mathcal{U}}[\mathcal{R}_n(\mathcal{H}_{ij})],
\end{aligned}$$

其中，第一个不等式由詹森不等式得出，$(\sigma_u)_{u=1}^n$ 是相互独立的拉德马赫随机变量，$\mathcal{U}$表示用户集合，$\mathcal{U}'$是 $\mathcal{U}$的一个独立副本，而 $\mathbb{E}[\mathcal{R}_n(\mathcal{H}_{ij})]$是 $\mathcal{H}_{ij}$的拉德马赫复杂度。

接下来，我们来界定 $\mathbb{E}[\mathcal{R}_n(\mathcal{H}_{ij})]$。根据Koltchinskii (2011中的定理3.11，存在一些常数 $A_1 > 0$ 使得

$$\mathbb{E}[\mathcal{R}_n(\mathcal{H}_{ij})] \leq A_1 n^{-1/2}\mathbb{E}\Big[\int_0^{2\sigma_n(\mathcal{H}_{ij})}\sqrt{\log\mathcal{N}(\mathcal{H}_{ij}, v, \|\cdot\|_{\mathcal{U}})}dv\Big], \tag{27}$$

其中 $\|f\|_{\mathcal{U}} = \sqrt{n^{-1}\sum_{u=1}^{n} DG_f^2(u)}$ 和 $\sigma_n^2(\mathcal{H}_{ij}) = \sup_{f\in\mathcal{H}_{ij}} n^{-1}\sum_{u=1}^{n} DG_f^2(u)$。接下来，我们

推导 $\log \mathcal{N}(\mathcal{H}_{ij},\ v,\ \|\cdot\|_{\mathcal{U}})$ 的一个上界。对于任意 $f_1,\ f_2 \in \mathcal{H}$,

$$\begin{aligned}
\|f_1 - f_2\|_{\mathcal{U}}^2 =& \frac{1}{n}\sum_{u=1}^{n}\Big(DG_{f_1}(u) - DG_{f_2}(u)\Big)^2 = \frac{1}{n}\sum_{u=1}^{n}\Big(G_{f_1}(u) - G_{f_2}(u)\Big)^2 \\
\leq& \frac{1}{n}\sum_{u=1}^{n}\frac{1}{m(m-1)}\sum_{i\neq j}K_v^2\big(g_{f_1}(u,i,j) - g_{f_2}(u,i,j)\big)^2 \\
\leq& \frac{1}{n}\sum_{u=1}^{n}\frac{1}{m(m-1)}\sum_{i\neq j}K_v^2\Big(2\big(f_1(\boldsymbol{x}_u,\boldsymbol{y}_i) - f_2(\boldsymbol{x}_u,\boldsymbol{y}_i)\big)^2 + 2\big(f_1(\boldsymbol{x}_u,\boldsymbol{y}_j) - f_2(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)^2\Big) \\
\leq& \frac{1}{n}\sum_{u=1}^{n}\frac{1}{m(m-1)}\sum_{i\neq j}K_v^2\big(f_1(\boldsymbol{x}_u,\boldsymbol{y}_i) - f_2(\boldsymbol{x}_u,\boldsymbol{y}_i)\big)^2 \\
\leq& \frac{C_1}{n}\sum_{u=1}^{n}\Big(\|\boldsymbol{x}_u\|_2^2 + \frac{1}{m}\sum_{i\in[m]}\|\boldsymbol{y}_i\|_2^2\Big)\|\Theta_{f_1} - \Theta_{f_2}\|_\infty^2,
\end{aligned}$$

其中，第一个不等式源于以下事实： $\phi(\cdot)$ 是一个 K_v-利普希茨函数，而最后一个不等式则由假设3得出。

为便于记号，我们记 $F^2(u) = \|\boldsymbol{x}_u\|_2^2 + \frac{1}{m}\sum_{i\in[m]}\|\boldsymbol{y}_i\|_2^2$。注意到， $f \in \mathcal{H}_{ij}$意味着 $\|\Theta_f\|_\infty \leq 2^j J_0$。令 $B_\Theta(2^j J_0)$ 表示以原点为中心、在 $\mathbb{R}^{|\Theta|}$ 中边长为 $2^{j+1}J_0$ 的立方体。于是可以验证

$$\mathcal{N}(\mathcal{H}_{ij}, v, \|\cdot\|_{\mathcal{U}}) \leq \mathcal{N}\Big(B_\Theta(2^j J_0), \frac{v}{\sqrt{C_1}\|F\|}, \|\cdot\|\Big),$$

其中 $\|F\| = \sqrt{n^{-1}\sum_{u=1}^{n}F^2(u)}$。于是可得

$$\log \mathcal{N}(\mathcal{H}_{ij}, v, \|\cdot\|_{\mathcal{U}}) \leq |\Theta|\log\Big(\max\Big\{\frac{\sqrt{C_1}2^{j+1}J_0\|F\|}{v}, 1\Big\}\Big), \tag{28}$$

接下来，结合（27）和（28）可得

$$\begin{aligned}
&\mathbb{E}[\mathcal{R}_n(\mathcal{H}_{ij})] \leq A_1 n^{-1/2}\mathbb{E}\Big[\int_0^{2\sigma_n(\mathcal{H}_{ij})}\sqrt{|\Theta|\log\Big(\max\Big\{\frac{\sqrt{C_1}2^{j+1}J_0\|F\|}{v}, 1\Big\}\Big)}dv\Big] \\
=&A_1 n^{-1/2}\mathbb{E}\Big[\int_0^{2\sigma_n(\mathcal{H}_{ij})}\sqrt{|\Theta|\log\Big(\max\Big\{\frac{\sqrt{C_1}2^{j+1}J_0\|F\|}{v}, 1\Big\}\Big)}dv I\big(\|F\|^2 > 2\mathbb{E}F^2(u)\big)\Big] \\
+&A_1 n^{-1/2}\mathbb{E}\Big[\int_0^{2\sigma_n(\mathcal{H}_{ij})}\sqrt{|\Theta|\log\Big(\max\Big\{\frac{\sqrt{C_1}2^{j+1}J_0\|F\|}{v}, 1\Big\}\Big)}dv I\big(\|F\|^2 \leq 2\mathbb{E}F^2(u)\big)\Big] \\
=&V_1 + V_2.
\end{aligned}$$

接下来，只需分别对 V_1 和 V_2 进行界定。对于 V_1，我们首先注意到

$$\begin{aligned}
&\int_0^{2\sigma_n(\mathcal{H}_{ij})} \sqrt{|\Theta| \log\Big(\max\Big\{\frac{\sqrt{C_1}2^{j+1}J_0\|F\|}{v}, 1\Big\}\Big)}dv \\
=&\sqrt{C_1}2^{j+1}J_0\|F\| \int_0^{2\frac{\sigma_n(\mathcal{H}_{ij})}{\sqrt{C_1}2^{j+1}J_0\|F\|}} \sqrt{|\Theta| \log\Big(\max\Big\{\frac{1}{v}, 1\Big\}\Big)}d\epsilon \\
\leq&\sqrt{C_1}2^{j+1}J_0\|F\| \int_0^{1} \sqrt{|\Theta| \log\big(\epsilon^{-1}\big)}d\epsilon \\
\leq&\sqrt{C_1}2^{j+1}J_0\|F\|\sqrt{|\Theta|} \int_1^{\infty} u^{-2}\sqrt{\log(u)}du \\
\leq&A_2 2^{j+1}J_0\|F\|\sqrt{|\Theta|},
\end{aligned}$$

其中 $A_2 = \sqrt{C_1}\int_0^{\infty} u^{-3/2}du$，最后一个不等式是由以下事实得出的：对于 $u \geq 1$, $\sqrt{\log(u)} \leq \sqrt{u-1}$。据此， V_1 可界定为

$$\begin{aligned}
V_1 \leq&A_1 n^{-1/2}\mathbb{E}\Big[A_2 2^{j+1}J_0\|F\|\sqrt{|\Theta|}I\big(\|F\|^2 > 2\mathbb{E}F^2(u)\big)\Big] \\
\leq&A_1A_2 2^{j+1}J_0 n^{-1/2}\mathbb{E}\Big[\|F\|\sqrt{|\Theta|}I\big(\|F\|^2 > 2\mathbb{E}F^2(u)\big)\Big] \\
\leq&A_1A_2 2^{j+1}J_0 n^{-1/2}\sqrt{|\Theta|\mathbb{E}F^2(u)}\mathbb{P}\Big(\|F\|^2 > 2\mathbb{E}F^2(u)\Big) \\
\leq&A_1A_2 2^{j_0+1}J_0 n^{-1/2}\sqrt{|\Theta|\mathbb{E}F^2(u)}\mathbb{P}\Big(\|F\|^2 > 2\mathbb{E}F^2(u)\Big) \qquad (29)
\end{aligned}$$

其中第三个不等式由 H¨older不等式得出。根据 Bernstein不等式，(29) 的右侧可界定为

$$\mathbb{P}\Big(\|F\|^2 > 2\mathbb{E}F^2(u)\Big) \leq \exp\Big(-\frac{1/2n^2(\mathbb{E}F^2(u))^2}{nA_4 + 1/3A_3\mathbb{E}F^2(u)}\Big), \qquad (30)$$

其中 $A_3 \geq F(u)$ 对任意 u 以及 $A_4 \geq \frac{1}{n}\sum_{u=1}^{n}\mathbb{E}[F^4(u)],$，这由以下事实成立：$\mathcal{X}$ 和 $\mathcal{Y}$ 均为紧致空间。结合 (29) 和 (30) 可知，存在一些正常数 $A_5 > 0$ 使得

$$V_1 \leq A_5 n^{-1}.$$

为了界定 V_2，我们首先注意到，对于 $u \geq 1$

$$\frac{du^{-2}\sqrt{\log(\max\{u,1\})}}{du} = \frac{1}{u^3\sqrt{\log(u)}}(-2\log(u)+1). \qquad (31)$$

令（31）为 0，可得 $u=\exp(1/2)$，因此 $u^{-2}\log(\max\{u,1\})$ 在$1\le u\le\exp(1/2)$ 时递增，在 $u>\exp(1/2)$ 时递减。因此，对于任意 $A>0$,

$$\int_A^{+\infty}u^{-2}\sqrt{\log(\max\{u,1\})}du=\int_A^{\exp(1/2)}u^{-2}\sqrt{\log(\max\{u,1\})}du+\int_{\exp(1/2)}^{+\infty}u^{-2}\sqrt{\log(u)}du$$
$$\le 2^{-1/2}A^{-1}-2^{-1/2}\exp(-1/2)+2^{-1/2}\exp(-1/2)+\int_{\exp(1/2)}^{+\infty}\frac{u^{-2}}{\sqrt{\log(u)}}du$$
$$\le 2^{-1/2}A^{-1}+A^{-a}\int_{\exp(1/2)}^{+\infty}\frac{u^{-2+a}}{\sqrt{\log(u)}}du,$$

其中 $0\le a<1$。可以验证，当 $0\le a<1$ 时，$\int_{\exp(1/2)}^{+\infty}\frac{u^{-2+a}}{\sqrt{\log(u)}}du$ 是有限的。因此，存在常数 A_6，使得对于任意 $A>0$

$$\int_A^{+\infty}u^{-2}\sqrt{\log(\max\{u,1\})}du\le A_6A^{-1}. \tag{32}$$

对于 V_2，由 $\sqrt{\log(x)}$的凹性可知

$$\begin{aligned}
V_2=&A_1n^{-1/2}\mathbb{E}\Big[\int_0^{2\sigma_n(\mathcal{H}_{ij})}\sqrt{|\Theta|\log\Big(\max\Big\{\frac{\sqrt{C_1}2^{j+1}J_0\|F\|}{v},1\Big\}\Big)}dvI\big(\|F\|^2\le 2\mathbb{E}F^2(u)\big)\Big]\\
\le&A_1n^{-1/2}\int_0^{2\sqrt{\mathbb{E}\sigma_n^2(\mathcal{H}_{ij})}}\sqrt{|\Theta|\log\Big(\max\Big\{\frac{\sqrt{C_1}2^{j+1}J_0\sqrt{2\mathbb{E}[F^2(u)]}}{v},1\Big\}\Big)}dv\\
=&A_1\sqrt{C_1}n^{-1/2}2^{j+1}J_0\sqrt{2|\Theta|\mathbb{E}[F^2(u)]}\int_0^{\frac{\sqrt{\mathbb{E}\sigma_n^2(\mathcal{H}_{ij})}}{\sqrt{C_1}2^jJ_0\sqrt{2\mathbb{E}[F^2(u)]}}}\sqrt{\log\Big(\max\Big\{\frac{1}{v},1\Big\}\Big)}dv\\
\le&A_1\sqrt{C_1}n^{-1/2}2^{j+1}J_0\sqrt{2|\Theta|\mathbb{E}[F^2(u)]}\int_{\frac{\sqrt{C_1}2^{j+1}J_0\sqrt{\mathbb{E}[F^2(u)]}}{2\sqrt{\mathbb{E}\sigma_n^2(\mathcal{H}_{ij})}}}^{+\infty}u^{-2}\sqrt{\log(\max\{u,1\})}du\\
\le&2A_1A_6\sqrt{C_1}n^{-1/2}\sqrt{|\Theta|\mathbb{E}\sigma_n^2(\mathcal{H}_{ij})},
\end{aligned}\tag{33}$$

其中，最后一个不等式由（32）得出

如定理 2 所证明的，我们有

$$\mathbb{E}\sigma_n^2(\mathcal{H}_{ij})\le 8C_v\mathbb{E}\big[\mathcal{R}(\mathcal{H}_{ij})\big]+C_5\mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^\gamma\big]M^{\frac{\gamma}{\gamma+1}}(i,j), \tag{34}$$

将（34）代入（33）可得，对于某些正常数 $A_7>0$

$$V_2\le A_7\sqrt{n^{-1}|\Theta|}\Big(\sqrt{8C_v\mathbb{E}\big[\mathcal{R}(\mathcal{H}_{ij})\big]}+C_5\mathbb{E}^{\frac{1}{2(1+\gamma)}}\big[\Psi_u^\gamma\big](2^i\delta_n)^{\frac{\gamma}{2(\gamma+1)}}\Big).$$

因此，我们得到

$$\mathbb{E}[\mathcal{R}_n(\mathcal{H}_{ij})] \leq A_6 n^{-1} + A_7\sqrt{n^{-1}|\Theta|}\Big(\sqrt{8C_v\mathbb{E}\big[\mathcal{R}(\mathcal{H}_{ij})\big]} + C_5\mathbb{E}^{\frac{1}{2(1+\gamma)}}\big[\Psi_u^{\gamma}\big](2^i\delta_n)^{\frac{\gamma}{2(\gamma+1)}}\Big). \quad (35)$$

解（35）可得，$\mathbb{E}[\mathcal{R}_n(\mathcal{H}_{ij})] \leq A_8 n^{-1/2}|\Theta|^{-1/2}(2^i\delta_{ij})^{\frac{\gamma}{(\gamma+1)}}\mathbb{E}^{\frac{1}{(1+\gamma)}}\big[\Psi_u^{\gamma}\big] + A_9 n^{-1}|\Theta|$ 对于某些常数 A_8，$A_9 > 0$。

若满足 $\big(n^{-1}|\Theta|\log\big(|\Theta|^{-1}n\big)\big)^{\frac{\gamma+1}{\gamma+2}}\mathbb{E}^{\frac{1}{\gamma+2}}\big(\Psi_u^{\gamma}\big) = O(\delta_n)$，则可得

$$\frac{\mathbb{E}[\mathcal{R}_n(\mathcal{H}_{ij})]}{M(i,j)} \leq \frac{A_8 n^{-1/2}|\Theta|^{1/2}\mathbb{E}^{\frac{1}{2(1+\gamma)}}\big[\Psi_u^{\gamma}\big]}{(2^i\delta_n)^{\frac{\gamma+2}{2(\gamma+1)}}} + \frac{A_9 n^{-1}|\Theta|}{2^i\delta_n} \leq 1/4,$$

其中，当 n 趋于无穷大时，最后一个不等式成立，因此可推出 $\mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big) \leq 1/2M(i,j)$，至此证明完成。

□

定理 3的证明：根据假设2，对于 $\delta_n > 0$，，有

$$\mathbb{P}\big(e(\widetilde{f}, f^*) \geq \delta_n^{\alpha}\big) \leq \mathbb{P}\big(e_v(\widetilde{f}, f^*) \geq \delta_n\big).$$

Next，我们接下来对 $\mathbb{P}\big(e_\phi(\widehat{f}, f_\phi^*) \geq \delta_n\big)$ 进行界定。根据 $\widehat{f}$ 的定义，我们有 e

$$\mathbb{P}\Big(e_v(\widetilde{f}, f^*) \geq \delta_n\Big) \leq \mathbb{P}\Big(\sup_{f\in\mathcal{H}} \widetilde{R}_{n,v}(f_{\mathcal{F}}^*) + \lambda_n J_0 - \widetilde{R}_{n,v}(f) - \lambda_n J(f) \geq 0\Big) \equiv I, \quad (36)$$

其中 $J_0 = \max\{J(f_{\mathcal{F}}^*), 1\}$。令 i_0 表示满足 $2^{i_0}\delta_n \geq \sup_f e_v(f, f^*)$ 的整数，并令 j_0 表示满足以下条件的整数：对于任意满足 $J(f) \geq 2^{j_0}J_0$ 的 $f \in \mathcal{F}$，都必须有 $\|\Theta_f\|_{min} \geq C_{\mathcal{F}}$，其中 $\|\cdot\|_{min}$ 表示 Θ_f 的最小绝对值。

对于 $0 \leq i \leq i_0$ 和 $0 \leq j \leq j_0$，我们定义

$$\mathcal{H}_{ij} = \big\{f \in \mathcal{F} : 2^{i-1}\delta_n < e_v(f, f^*) \leq 2^i\delta_n, 2^{j-1}J_0 < J(f) \leq 2^j J_0\big\}, 1 \leq i \leq i_0, 1 \leq j \leq j_0$$

$$\mathcal{H}_{i0} = \big\{f \in \mathcal{F} : 2^{i-1}\delta_n < e_v(f, f^*) \leq 2^i\delta_n, J(f) \leq J_0\big\}, 1 \leq i \leq i_0.$$

容易验证，$\mathcal{H}$ 可以表示为 $\mathcal{H} = \cup_{i=1}^{i_0}\cup_{j=0}^{j_0}\mathcal{H}_{ij}$。由此，(36) 可以

以上界限定为

$$
\begin{aligned}
I =&\mathbb{P}\Big(\sup_{f\in\cup_{i=1}^{i_0}\cup_{j=0}^{j_0}\mathcal{H}_{ij}} \widetilde{R}_{n,v}(f^*_{\mathcal{F}}) + \lambda_n J_0 - \widetilde{R}_{n,v}(f) - \lambda_n J(f) \geq 0\Big)\\
\leq&\sum_{i=0}^{i_0}\sum_{j=0}^{j_0}\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \widetilde{R}_{n,v}(f^*_{\mathcal{F}}) + \lambda_n J_0 - \widetilde{R}_{n,v}(f) - \lambda_n J(f) \geq 0\Big)\\
=&\sum_{i=1}^{i_0}\sum_{j=1}^{j_0}\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \widetilde{R}_{n,v}(f^*_{\mathcal{F}}) + \lambda_n J_0 - \widetilde{R}_{n,v}(f) - \lambda_n J(f) \geq 0\Big)\\
&+\sum_{i=1}^{i_0}\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{i0}} \widetilde{R}_{n,v}(f^*_{\mathcal{F}}) + \lambda_n J_0 - \widetilde{R}_{n,v}(f) - \lambda_n J(f) \geq 0\Big) \equiv I_1 + I_2.
\end{aligned}
$$

因此，只需分别对 I_1 和 I_2 进行界定。根据假设 2，对于任意 $1 \leq i \leq i_0$ 和 $j \geq 0$，我们得到

$$
\begin{aligned}
&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \widetilde{R}_{n,v}(f^*_{\mathcal{F}}) + \lambda_n J_0 - \widetilde{R}_{n,v}(f) - \lambda_n J(f) \geq 0\Big)\\
\leq&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \big(\widetilde{R}_{n,v}(f^*_{\mathcal{F}}) - \widetilde{R}_{n,v}(f) - \widetilde{R}_{v}(f^*_{\mathcal{F}}) + \widetilde{R}_{v}(f)\big)\\
&\qquad\geq \lambda_n \inf_{f\in\mathcal{H}_{ij}}(J(f) - J_0) + \inf_{f\in\mathcal{H}_{ij}} \widetilde{R}_{\phi}(f) - \widetilde{R}_{v}(f^*_{\mathcal{F}})\Big)\\
\leq&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \big(\widetilde{R}_{n,v}(f^*_{\mathcal{F}}) - \widetilde{R}_{n,v}(f) - \widetilde{R}_{v}(f^*_{\mathcal{F}}) + \widetilde{R}_{v}(f)\big) \geq \lambda_n(2^{j-1}-1)J_0 + 2^{i-1}\delta_n\Big)\\
\leq&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \big(\widetilde{R}_{n,v}(f^*_{\mathcal{F}}) - \widetilde{R}_{n,v}(f) - \widetilde{R}_{v}(f^*_{\mathcal{F}}) + \widetilde{R}_{v}(f)\big) \geq \lambda_n(2^{j-1}-1)J_0 + 2^{i-1}\delta_n\Big)\\
\leq&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \big(\widetilde{R}_{n,v}(f^*_{\mathcal{F}}) - \widetilde{R}_{n,v}(f) - \widetilde{R}_{v}(f^*_{\mathcal{F}}) + \widetilde{R}_{v}(f)\big) \geq M(i,j)\Big),
\end{aligned}
\tag{37}
$$

其中 $M(i,\ j) = \lambda_n(2^{j-1}-1)J_0 + 2^{i-1}\delta_n$ 对于 $1 \leq i \leq i_0$ 和 $j \geq 1$ 以及 $M(i,\ 0) = 2^{i_0}\delta_n$ 对于 $i \geq 1$。此外，我们定义

$$
G_f(u) = \frac{1}{m(m-1)}\sum_{i\neq j} I\Big(\widetilde{\phi}_u(\mathcal{I}_i) > \widetilde{\phi}_u(\mathcal{I}_j)\Big) v(g_f(u,i,j)).
$$

那么，(37) 的右侧可以写成

$$
\begin{aligned}
&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \big(\widetilde{R}_{n,v}(f^*_{\mathcal{F}}) - \widetilde{R}_{n,v}(f) - \widetilde{R}_{v}(f^*_{\mathcal{F}}) + \widetilde{R}_{v}(f)\big) \geq M(i,j)\Big)\\
=&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \big[\frac{1}{n}\sum_{u=1}^{n}\big(G_{f^*_{\mathcal{F}}}(u) - G_f(u)\big) - \mathbb{E}\big(G_{f^*_{\mathcal{F}}}(u) - G_f(u)\big)\big] \geq M(i,j)\Big)\\
=&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} \big[\frac{1}{n}\sum_{u=1}^{n}\big(DG_f(u)\big) - \mathbb{E}\big(DG_f(u)\big)\big] \geq M(i,j)\Big),
\end{aligned}
\tag{38}
$$

其中 $DG_f(u) = G_{f^*_{\mathcal{F}}}(u) - G_f(u)$。令 $U_n(f) = n^{-1}\sum_{u=1}^n \big[DG_f(u) - \mathbb{E}\big(DG_f(u)\big)\big]$，则(38)可重写为

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}}\big[\frac{1}{n}\sum_{u=1}^n\big(DG_f(u)\big) - \mathbb{E}\big(DG_f(u)\big)\big] \geq M(i,j)\Big)\\
=&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big) \geq M(i,j) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big)\Big)\\
\leq&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big) \geq 1/2M(i,j)\Big),
\end{aligned} \tag{39}$$

其中最后一个不等式由引理 9 得出。由此，我们有

$$\begin{aligned}
I_1 &\leq \sum_{i=1}^{i_0}\sum_{j=1}^{j_0}\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big) \geq 1/2M(i,j)\Big),\\
I_2 &\leq \sum_{i=1}^{i_0}\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{i0}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{i0}} U_n(f)\big) \geq 1/2M(i,j)\Big).
\end{aligned}$$

显然，只需对任意$1 \leq i \leq i_0$ 和 $0 \leq j \leq j_0$，界定 $\mathbb{P}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big) \geq 1/2M(i,j)\big)$，对此我们在下一步应用 Talagrand 不等式。

记 $\mathbb{E}[\sigma_n^2(\mathcal{H}_{ij})] = \mathbb{E}[\sup_{f\in\mathcal{H}_{ij}} n^{-1}\sum_{i=1}^n DG_f^2(u)]$。然后，我们建立 $\mathbb{E}[\sigma_n^2(\mathcal{H}_{ij})]$ 与 $M(i,j)$ 之间的关系。通过对称化论证，我们有

$$\begin{aligned}
\mathbb{E}[\sigma_n^2(\mathcal{H}_{ij})] \leq&\mathbb{E}\Big[\sup_{f\in\mathcal{H}_{ij}}\frac{1}{n}\sum_{u=1}^n DG_f^2(u) - \mathbb{E}\big(DG_f^2(u)\big)\Big] + \sup_{f\in\mathcal{H}_{ij}}\frac{1}{n}\sum_{u=1}^n\mathbb{E}\big(DG_f^2(u)\big)\\
\leq&8C_v\mathbb{E}\big[\mathcal{R}(\mathcal{H}_{ij})\big] + \sup_{f\in\mathcal{H}_{ij}}\frac{1}{n}\sum_{u=1}^n\mathbb{E}\Big[\big(G_{f^*_{\mathcal{F}}}(u) - G_f(u)\big)^2\Big].
\end{aligned} \tag{40}$$

(40) 的右侧可以界定为

$$\begin{aligned}
&\sup_{f\in\mathcal{H}_{ij}}\frac{1}{n}\sum_{u=1}^n\mathbb{E}\Big[\big(G_{f^*_{\mathcal{F}}}(u) - G_f(u)\big)^2\Big]\\
\leq&\frac{2}{n}\sum_{u=1}^n\mathbb{E}\Big[\big(G_{f^*_{\mathcal{F}}}(u) - G_{f^*}(u)\big)^2\Big] + \sup_{f\in\mathcal{H}_{ij}}\frac{2}{n}\sum_{u=1}^n\mathbb{E}\Big[\big(G_{f^*}(u) - G_f(u)\big)^2\Big]\\
\leq&\frac{2}{n}\sum_{u=1}^n\mathrm{Var}\big(G_{f^*_{\mathcal{F}}}(u) - G_{f^*}(u)\big) + \sup_{f\in\mathcal{H}_{ij}}\frac{2}{n}\sum_{u=1}^n\mathrm{Var}\big(G_{f^*}(u) - G_u(f)\big)\\
&+ 2e_v^2(f^*_{\mathcal{F}}, f^*) + 2\sup_{f\in\mathcal{H}_{ij}} e_v^2(f, f^*_{\mathcal{F}})\\
\leq&4C_4\mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big]\sup_{f\in\mathcal{H}_{ij}}\big(e_v(f,f^*)\big)^{\gamma/(\gamma+1)} + \sup_{f\in\mathcal{H}_{ij}} e_v^2(f, f^*_{\mathcal{F}}).
\end{aligned}$$

假设 $\mathbb{E}\big[\Psi_u^{\gamma}\big] = O(\delta_n)$。由于 $\gamma/(\gamma+1) < 2$，存在一些常数 C_5，使得

$$\sup_{f\in\mathcal{H}_{ij}} \frac{1}{n}\sum_{u=1}^{n} \mathbb{E}\Big[\big(G_{f_{\mathcal{F}}^*}(u) - G_f(u)\big)^2\Big] \le C_5 \mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big] M^{\frac{\gamma}{\gamma+1}}(i,j). \tag{41}$$

结合引理 9，可得

$$\mathbb{E}\big[\sigma_n^2(\mathcal{H}_{ij})\big] \le 2C_v M(i,j) + C_5 \mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big] M^{\frac{\gamma}{\gamma+1}}(i,j),$$

其中 C_v使得 $G_f(u) \le C_v$对任意 $f \in \mathcal{F}$均成立。由于 $M(i,\ j)$ 渐近小于 1，因此存在一些常数 T_4 使得

$$\mathbb{E}\sigma_n^2(\mathcal{H}_{ij}) \le T_4 \mathbb{E}^{\frac{1}{1+\gamma}}\big[\Psi_u^{\gamma}\big] M^{\frac{\gamma}{\gamma+1}}(i,j), \tag{42}$$

对于 $1 \le i \le i_0$ 和 $0 \le j \le j_0$。然后，将（42）代入（39），可得对于某些常数$T_5 > 0$

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big) \ge 1/2M(i,j)\Big) \\
\le& T_1 \exp\Big(-\frac{nM(i,j)}{2BT_1}\log\big(1 + \frac{B\big(M(i,j)\big)}{T_4\mathbb{E}^{\frac{1}{1+\gamma}}(\Psi_u^{\gamma})M^{\gamma/(\gamma+1)}(i,j)}\big)\Big) \\
\le& T_1 \exp\Big(-\frac{T_5}{2T_1T_4} n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^{\gamma})\big(M(i,j)\big)^{\frac{\gamma+2}{\gamma+1}}\Big),
\end{aligned}$$

其中，T_5 满足 $\log(1+x) \ge T_5 x$ 对于 $x \in [0, (T_4\mathbb{E}^{\frac{1}{1+\gamma}}(\Psi_u^{\gamma}))^{-1}B\big(M(i,\ j)\big)^{1/(1+\gamma)}]$，这由 $M(i,\ j)$ 的有界性可知。

令 $T_6 = \frac{T_5}{2T_1T_4}$，则 I_3 可界定为

$$
\begin{aligned}
I_1 =& \sum_{i=1}^{i_0}\sum_{j=1}^{j_0}\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{ij}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{ij}} U_n(f)\big) \geq 1/2M(i,j)\Big) \\
\leq& \sum_{i=1}^{i_0}\sum_{j=1}^{j_0} T_1 \exp\Big\{ -T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\Big(\delta_n^{\frac{\gamma+2}{\gamma+1}}(2^{i-1})^{\frac{\gamma+2}{\gamma+1}} - (\lambda_n J_0)^{\frac{\gamma+2}{\gamma+1}}(2^{j-1}-1)^{\frac{\gamma+2}{\gamma+1}}\Big)\Big\} \\
\leq& \sum_{i=2}^{\infty}\sum_{j=1}^{\infty} T_1 \exp\Big\{ -T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\Big(\delta_n^{\frac{\gamma+2}{\gamma+1}} i - (\lambda_n J_0)^{\frac{\gamma+2}{\gamma+1}}(j-1)\big)\Big)\Big\} \\
&+ \sum_{j=1}^{\infty} T_1 \exp\Big\{\Big(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\Big(\delta_n^{\frac{\gamma+2}{\gamma+1}} - (\lambda_n J_0)^{\frac{\gamma+2}{\gamma+1}}(j-1)\Big)\Big\} \\
\leq& T_1 \frac{\exp(-2T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}})}{1-\exp(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}))}\frac{1}{1-\exp(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma))(\lambda_n J_0)^{\frac{\gamma+2}{\gamma+1}})} \\
&+ T_1\exp(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}})\frac{1}{1-\exp(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma))(\lambda_n J_0)^{\frac{\gamma+2}{\gamma+1}})} \\
\leq& 6T_1 \exp\Big(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}\Big),
\end{aligned}
$$

其中，倒数第二个不等式由柯西-施瓦茨不等式得出，最后一个不等式在 n 趋于无穷大时成立。

类似地，对于 I_2，我们有

$$
\begin{aligned}
I_4 =& \sum_{i=1}^{i_0}\mathbb{P}\Big(\sup_{f\in\mathcal{H}_{i0}} U_n(f) - \mathbb{E}\big(\sup_{f\in\mathcal{H}_{i0}} U_n(f)\big) \geq 1/2M(i,j)\Big) \\
\leq& \sum_{i=1}^{i_0} T_1 \exp\Big(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}(2^{i-1})^{\frac{\gamma+2}{\gamma+1}} + T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)(\lambda_n J_0)^{\frac{\gamma+2}{\gamma+1}}\Big) \\
\leq& \sum_{i=1}^{i_0} T_1 \exp\Big(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}(2^{i-1}-1/2)^{\frac{\gamma+2}{\gamma+1}}\Big) \\
\leq& T_1 \frac{\exp\big(-2^{-1}T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}\big)}{1-\exp\big(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}\big)} \leq 2T_1\exp\big(-T_6 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}\big).
\end{aligned}
$$

最后，令 $C_2 = T_1$ 和 $C_3 = T_6$，得到

$$
I_1 + I_2 \leq 8C_2\exp\big(-C_3 n\mathbb{E}^{-\frac{1}{1+\gamma}}(\Psi_u^\gamma)\delta_n^{\frac{\gamma+2}{\gamma+1}}\big).
$$

至此，证明完毕。 □

推论的证明 1：首先，根据定理 3中的性质 (2)，我们有

$$\frac{|2\eta_{uij}-1|}{|2\widetilde{\eta}_{uij}-1|} \leq \frac{\exp\left((|\Omega|-1)^{-1}\epsilon_u\right)+1}{\exp\left((|\Omega|-1)^{-1}\epsilon_u\right)-1} = \Psi_u.$$

根据超额风险的定义，我们有

$$\widetilde{R}(f)-\widetilde{R}(f^*) = \frac{1}{m(m-1)}\mathbb{E}_u\Big[\sum_{i<j}|2\widetilde{\eta}_{uij}-1|I\Big(\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i)-f(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)\big(f^*(\boldsymbol{x}_u,\boldsymbol{y}_i)-f^*(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)<0\Big)\Big].$$

因此，我们有

$$\begin{aligned}
&R(f)-R(f^*)\\
=&\frac{1}{m(m-1)}\mathbb{E}_u\Big[\sum_{i<j}|2\eta_{uij}-1|I\Big(\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i)-f(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)\big(f^*(\boldsymbol{x}_u,\boldsymbol{y}_i)-f^*(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)<0\Big)\Big]\\
\leq&\frac{1}{m(m-1)}\mathbb{E}_u\Big[\Psi_u\sum_{i<j}|2\widetilde{\eta}_{uij}-1|I\Big(\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i)-f(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)\big(f^*(\boldsymbol{x}_u,\boldsymbol{y}_i)-f^*(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)<0\Big)\Big]\\
\leq&\mathbb{E}_u(\Psi_u)\frac{1}{m(m-1)}\mathbb{E}_u\Big[\sum_{i<j}|2\widetilde{\eta}_{uij}-1|I\Big(\big(f(\boldsymbol{x}_u,\boldsymbol{y}_i)-f(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)\big(f^*(\boldsymbol{x}_u,\boldsymbol{y}_i)-f^*(\boldsymbol{x}_u,\boldsymbol{y}_j)\big)<0\Big)\Big]\\
=&\mathbb{E}_u(\Psi_u)\Big(\widetilde{R}(f)-\widetilde{R}(f^*)\Big).
\end{aligned}$$

由定理 3 可得

$$R(\widetilde{f})-R(f^*) = O_p\Big(\mathbb{E}(\Psi_u)\mathbb{E}^{\frac{\alpha}{\gamma+2}}(\Psi_u^{\gamma})\big(|\Theta|n^{-1}\log(n/|\Theta|)\big)^{\frac{\alpha(\gamma+1)}{\gamma+2}}\Big).$$

如果 $\epsilon_u=\epsilon$ 对任意用户 u、 $\alpha=1$ 和 $\epsilon=o(1)$，我们有

$$\Psi_u = \frac{\exp\left((|\Omega|-1)^{-1}\epsilon_u\right)+1}{\exp\left((|\Omega|-1)^{-1}\epsilon_u\right)-1} \asymp \frac{1}{\epsilon}.$$

因此，

$$\mathbb{E}(\Psi_u)\mathbb{E}^{\frac{\alpha}{\gamma+2}}(\Psi_u^{\gamma})\big(|\Theta|n^{-1}\log(n/|\Theta|)\big)^{\frac{\alpha(\gamma+1)}{\gamma+2}} \asymp (\Psi_u^{\gamma})\big(|\Theta|(\epsilon^2 n)^{-1}\log(n/|\Theta|)\big)^{\frac{\alpha(\gamma+1)}{\gamma+2}}.$$

证明完成。 □