arXiv:2301.07353v2 [math.ST] 8 Jan 2024

大样本中的矩阵主序

MUHAMMAD USMAN FAROOQ[1], TOBIAS FRITZ[2],
ERKKAHAAPASALO[3], MARCO TOMAMICHEL[3,4]

Abstract. 当存在一个将第一组中的概率向量转换为另一组中概率向量的单个随机矩阵时，前者的一组概率向量比后者的另一组概率向量包含更多信息。这称为矩阵主序。我们解决了 Mu 等人提出的一个开放问题：如果两个组之间的某些单调量——即 R´enyi 散度的多变量扩展——严格有序，那么对于充分大的 n，存在一个随机矩阵，将每个输入分布的 n 重克罗内克幂变换为相应输出分布的 n 重克罗内克幂。对于大样本中的这种矩阵主序，若单调量采用非严格排序，则相同的条件也是必要的。

我们的结果还给出了存在一列统计映射的条件，这些映射可以渐近地（误差趋于零）将每个输入分布的单个副本转换为相应的输出分布，同时借助一个保持原样返还的催化剂。允许以任意小的误差进行转换时，我们发现，这种催化矩阵主序成立的条件既是必要的，也是充分的。

我们在一位作者近期提出的预序半环一般代数理论基础上推导出这些结果。这也使我们能够以统一的方式重新得到关于大样本主序、催化体制下的主序以及相对主序的若干现有结果。

1. Introduction

统计实验是许多经典统计学及其应用的基础，例如假设检验[15, 26, 29]：每个假设都对应于可能观测结果上的一个概率分布，因此，统计实验不过是定义在同一结果空间上的一族概率分布。给定两个统计实验，自然会引出一个问题：哪一个更具“信息性”？例如，可以比较哪一个在可能采用的各种检验中产生

[1]Department of Mathematics, City University of Hong Kong, Hong Kong
[2]Department of Mathematics, University of Innsbruck, Austria
[3]Centre for Quantum Technologies, National University of Singapore, Singapore[4]Department of Electrical and Computer Engineering,
National Univer-sity of Singapore, Singapore

更高的期望收益。因此，比较统计实验便成为一个非常重要的问题，其研究可追溯至 Blackwell 的工作[3, 4]，此后 Stein [25] 和 Torgersen [28, 29]等人也继续开展了相关研究。

Blackwell 将以下观点形式化：若统计实验 $P=\{p^\vartheta\}_{\vartheta\in\Theta}$ 在某个（标准 Borel）可测空间上具有概率测度 p^ϑ，则它比另一个统计实验 $Q=\{q^\vartheta\}_{\vartheta\in\Theta}$ 包含更多信息；其形式化方式是要求存在一个统计映射 T（以马尔可夫核的形式），使得对所有 $\vartheta\in\Theta$，$Tp^\vartheta=q^\vartheta$ 成立。可以将这个方程理解为：实验 Q 可以通过先执行 P，再对其结果应用一个（可能是随机的）函数来模拟。根据 Blackwell 著名定理，这等价于：在任何包含 P 的检验中，所获得的期望收益至少与包含 Q的检验一样高。在这种情况下，我们称 P 主序于 Q，记作 $P\succeq Q$。

我们尤其关注概率测度的结果空间和参数空间 Θ 均为有限集的情形。在此框架下，统计实验的主序与矩阵主序的概念相一致[8]：若矩阵 P 的元素均非负，则当存在一个随机矩阵 T使得 $TP=Q$时，称它主序于另一个元素均非负的矩阵 Q。若对于 $k=1$，矩阵 P的列 $p^{(k)}$和矩阵 Q 的列 $q^{(k)}$分别为概率向量（即 P和 Q 是随机的），这就意味着实验 $(p^{(1)},\ldots,p^{(d)})$在上述意义下主序于 $(q^{(1)},\ldots,q^{(d)})$。我们注意到，对于 $d=2$，这一问题也称为相对主序。

除了研究单个统计实验之间的主序，我们还可以探讨在每个实验都独立地重复进行时，一个实验比另一个实验信息量更大的条件。用矩阵主序的术语来说，我们关心条件

$$\left((p^{(1)})^{\otimes n},\ldots,(p^{(d)})^{\otimes n}\right)\succeq\left((q^{(1)})^{\otimes n},\ldots,(q^{(d)})^{\otimes n}\right) \tag{1}$$

是否在重复次数 $n\in\mathbb{N}$ 足够大时成立。我们将其称为大样本中的矩阵主序。众所周知，存在这样的矩阵对（或一般而言，统计实验对）：两者互不主序，但在大样本中仍成立主序 [25, 28]。此外，若存在概率向量 $r^{(1)}$、……、$r^{(d)}$，使得 P催化主序于 Q，我们就称之为

$$(p^{(1)}\otimes r^{(1)},\ldots,p^{(d)}\otimes r^{(d)})\succeq(q^{(1)}\otimes r^{(1)},\ldots,q^{(d)}\otimes r^{(d)}), \tag{2}$$

也就是说，拥有催化矩阵 $R=(r^{(1)}、\ldots\ldots、r^{(d)})$，便能实现从 P 到 Q的变换。最后，若存在一个矩阵序列 $(Q_n)_n$ 收敛于 Q，且对所有 $n\in\mathbb{N}$，P都催化主序于 Q_n，我们就称 P 渐近催化主序于 Q。

正如我们将看到的，大样本中的主序蕴含催化主序，这可由一个已知的一般构造得出（例如参见 [10]）。关于大样本中主序的充分条件及一般情况下的必要条件，在

情形 $d = 2$已由 Mu 等人 在 [21], 中确定，而对于一般 d 情形的类似条件则被提出为猜想。在本文中，我们证明了他们猜想的一个略有变化的版本（区别见注 20 ）。这为一般情形下大样本中的矩阵主序给出了充分条件及一般情况下的必要条件。我们的证明运用了本文一位作者在 [12, 13], 中推导出的实代数工具，即预序半环理论。根据这些结果，某些类型的预序半环在大样本下的序关系，可以用涉及到 $\mathbb{R}_+$ 的单调同态以及若干类似单调映射的不等式来刻画。应用于一个适当的、用于描述矩阵主序的预序半环时，这些单调量与 Mu 等人所考察的量非常接近。

更具体地说，假设矩阵 P 和 Q 满足：所有向量 $p^{(k)}$ 和 $q^{(k)}$ 都具有全支撑。在这种情况下，相关的单调量包括如下形式的量

$$D_{\underline{\alpha}}(P) = \frac{1}{\max_{1\leqslant k\leqslant d}\alpha_k - 1}\log\sum_i\prod_{k=1}^{d}\left(p_i^{(k)}\right)^{\alpha_k}, \tag{3}$$

其中，$\underline{\alpha} = (\alpha_1, \ldots\ldots, \alpha_d)$ 是任意一组 d 元实参数，其和为 1，并且满足以下条件之一：$0 \leqslant \alpha_1, \ldots\ldots, \alpha_d < 1$；或者对于某个 k，有 $\alpha_k > 1$，且对所有 $\ell \neq k$，均有 $\alpha_\ell \leqslant 0$。这些映射的各种逐点极限也与此相关，我们将在正式陈述结果时加以讨论。

对于 $d = 2$，映射 $D_{\underline{\alpha}}(P)$ 可专化为某些 Rényi 散度，即当 $\alpha_1 \in [\frac{1}{2}, 1) \cup (1, \infty)$ 时为 $D_{\alpha_1}(p^{(1)}\|p^{(2)})$；类似地，当 $\alpha_2 \in [\frac{1}{2}, 1) \cup (1, \infty)$ 时为 $D_{\alpha_2}(p^{(2)}\|p^{(1)})$，其中

$$D_\alpha(p\|q) = \frac{1}{\alpha - 1}\log\sum_i p_i^\alpha q_i^{1-\alpha}. \tag{4}$$

这些量已经出现在大样本和催化体制下相对主序的已知必要且充分条件中（例如，参见 [5, 14, 21]）。因此，对于一般的 $d \geqslant 2$，上述映射是 R´ényi 散度的自然推广，这并不令人意外；特别是，它们在张量积下可加，并满足数据处理不等式，即对所有向量应用随机映射时，它们单调不增。

现在可以将我们的两个主要结果非正式地概述如下：

- 我们给出了大样本中矩阵主序的充分条件（以及在一般情况下必要的条件）。在定理19 第3节中，我们证明，若

$$D_{\underline{\alpha}}(P) > D_{\underline{\alpha}}(Q), \tag{5}$$

对上述所有参数元组 $\underline{\alpha}$ 均成立，且某些逐点极限也满足这些严格不等式，那么在大样本中， P 主序于 Q 。此外，这一充分条件几乎也是必要的：若在大样本中， P 主序于 Q ，那么必然有

$$D_{\underline{\alpha}}(P) \geqslant D_{\underline{\alpha}}(Q) \tag{6}$$

对于所有此类 $\underline{\alpha}$。我们还称带有严格不等式的充分条件“在一般情况下是必要的”，以表明在一般情况下，严格不等式通常成立。

- 我们给出了渐近催化矩阵主序的必要且充分条件。如前所述，大样本中的主序意味着催化主序，因此（5）中的条件也足以保证催化主序（并且在一般情况下仍是必要条件）。进一步地，我们在第22节3 的定理中证明，当且仅当满足（6）中的条件时，渐近催化主序才有可能。更确切地说，我们证明以下两个陈述等价：

  a) 对所有有效的参数元组 $\underline{\alpha}$，均有

  $$D_{\underline{\alpha}}(P) \geqslant D_{\underline{\alpha}}(Q)\,. \tag{7}$$

  b) 存在一个随机矩阵序列 Q_n ，它收敛于Q ，使得 P 对所有 n 均催化主序于 Q_n 。此外，我们可以假设 Q_n 的最后一列始终等于 Q 的最后一列。

对于 $d = 2$，后一个陈述恰好恢复了[5], 补充材料中的定理17，至少对于具有全支撑的分布而言是如此。[1]因此，它特别加强了 [14]中的定理20。虽然这两项较早的工作都借助嵌入，将简单主序的结果推广到相对情形，但我们的证明避免了这一技术，且更加直接，因此也能够考虑 $d > 2$ 。另请参见第3.3小节中的讨论。

本文研究的另一个主题是概率向量之间的（简单或绝对）主序。若向量 p 的元素均为非负数，则当存在 q 使得，记均匀 $n \in \mathbb{N}$ 分布的概率向量为 $u_n = (1/n, \ldots, 1/n)$, 我们有 $(p, u_n) \geq (q, u_n)$。
Hardy-Littlewood-P´olya 定理给出了这种主序的一个易于检验的条件，即附录中的不等式（168）。Klimesh [17] 和 Turgut [30]此前给出了催化简单主序的充要条件。Jensen [16], 等人证明了大样本结果；我们将这些结果扩展到具有相同支撑的向量情形。Aubrun 和 Nechita 在 [2]中导出了这些结果的渐近版本。在附录中，我们展示了

[1]论文[14] 对 [5]中证明的完整性提出了一些质疑。

通过这些结果的直接新证明，展示了新颖半环方法的实用性。

本文其余部分的组织如下。我们首先在第 2 节介绍与本文工作相关的定义和核心概念。特别是，我们总结了 [12, 13] 中与我们应用相关的实代数结果。我们在第 3 节正式引入矩阵主序的概念，以及与这一设定相关的预序半环，并由此得到上文总结的结果。在第 4 节，我们简要讨论本文结果的应用和未来展望，同时也提及其与量子信息论之间的一些联系及可能的拓展。附录对本文进行了补充，并重新推导了关于简单主序的一些既有结果：在第 A 节，我们建立用于研究大样本和催化体制下简单主序的代数框架。在第 B节，我们展示了预序半环方法的适用性，重新推导了此前在 [2, 17, 16] 中通过更为 特设的方法证明的、关于大样本和催化体制下主序的核心结果。

2. Preliminaries

鉴于我们即将开展的应用，本节将回顾与概率向量和预序半环相关的数学基础，并回顾已有的分离定理（Vergleichsstellens¨atze），这些定理刻画了在足够良好性质的预序半环上所诱导的大样本预序。

2.1. 基本符号。我们介绍一些关于有限结果概率分布的基本符号和定义，这些内容将在全文中使用。我们采用约定 $\mathbb{N} := \{1, 2, \ldots\}$。我们分别用 log 和 exp 表示以一个固定但未另行指定的、大于 1 的底为基的对数和指数函数。对于每个 $n \in \mathbb{N}$，我们定义从 $\|\cdot\|_1$ 到 $\mathbb{R}^n$ 的映射如下

$$\|x\|_1 := |x_1| + \cdots + |x_n| \tag{8}$$

对所有 $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ 均成立。对于 $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ 以及 $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ $(n, m \in \mathbb{N})$，我们通过下式定义直和 $x \oplus y \in \mathbb{R}^{m+n}$ 和克罗内克积 $x \otimes y \in \mathbb{R}^{mn}$

$$x \oplus y = (x_1, \ldots, x_n, y_1, \ldots, y_m), \tag{9}$$

$$x \otimes y = (x_i y_j)_{i,j}, \tag{10}$$

其中，后者被视为一个长度为 mn 的向量；本文始终认为元素的排列顺序无关紧要，因此不再进一步说明。

回顾一下，对于某个 $n \in \mathbb{N}$，$p \in \mathbb{R}^n$ 是一个概率向量，当且仅当其各项均为非负数且 $\|p\|_1 = 1$。我们将所有 n 项概率向量的集合记为

$\mathcal{P}_n$，并通过以下方式定义所有有限概率向量的集合：

$$\mathcal{P}_{<\infty} := \bigcup_{n\in\mathbb{N}} \mathcal{P}_n. \tag{11}$$

对于每个 $s \in \mathbb{N}$，我们将 s 个结果上的均匀概率分布（$\mathbb{R}^s$ 中的概率向量）记为

$$u_s := \big(\underbrace{1/s, \ldots, 1/s}_{s \text{ entries}}\big). \tag{12}$$

给定有限集 $I \subset \mathbb{N}$，我们类似地将满足以下条件的分布记为 u_I：对于 $i \in I$，其概率为 $(u_I)_i = 1/|I|$；其他情况下概率为 $(u_I)_i = 0$。并对所有 $n \geqslant \max I$ 考虑 $u_I \in \mathbb{R}^n$。注意，稍微滥用一下记号，u_I 并不显式依赖于大小 n；末尾的零在我们的语境中并非关键。

2.2. 预序半环与分离定理。

定义 1（预序交换半环）。预序交换半环是一个元组（S，$+$，$\cdot$，0，1，$\preccurlyeq$），其中 S 是一个集合，配备了两个二元运算 $+$、$\cdot$，它们均满足交换律和结合律，分别以 0 和 1 为单位元，并且满足

$$0 \cdot x = 0, \qquad (x+y)\cdot z = x\cdot z + y\cdot z \tag{13}$$

对于所有 x、y、$z \in S$。要求预序 $\preccurlyeq$（一种自反且传递的二元关系）满足：$a \preccurlyeq b$ 蕴含 $a + c \preccurlyeq b + c$ 且 $ac \preccurlyeq bc$，对所有 a、b、$c \in S$均成立。

我们将交替使用记号 $a \preccurlyeq b$ 和 $b \succcurlyeq a$。请注意，与环不同，半环中通常无法进行减法。从现在起及全文中，我们都用“半环”作为“交换半环”的简称。

定义 2 (预序半整域 [12])。预序半环 S 是预序半整域，若其没有零因子，也就是说，$xy = 0$ 对于 x，$y \in S$蕴含 $x = 0$ 或 $y = 0$，并且 $x = 0$ 是唯一的 $x \in S$ 满足 $0 \preccurlyeq x \preccurlyeq 0$。

此外，若半环 S 满足：对于 x，$x + y = 0$ 蕴含$x = 0 = y$，则称其为零和自由。进一步地，在预序半环 S中，我们通过如下方式定义等价关系 $\sim$：若存在有限序列 z_0、.....、$z_n \in S$，使得对所有 i，都有 $z_i \preccurlyeq z_{i+1}$ 或 $z_i \succcurlyeq z_{i+1}$，且 $z_0 = x$ 和 $z_n = y$，则称 $\sim$ 成立；等价地说，若存在锯齿链 $x \preccurlyeq z_1 \succcurlyeq z_2 \preccurlyeq \cdots \succcurlyeq z_n \preccurlyeq y$，则称 $\sim$ 成立。换言之，$\sim$是由 $\preccurlyeq$生成的等价关系。

我们用 $\mathbb{R}_+$ 表示配备通常的加法与乘法、并以通常的 $\leqslant$关系作为预序的非负实数预序半环。我们进一步定义热带实数半环 $\mathbb{TR}_+$，使其作为预序集与 $\mathbb{R}_+$ 相同，且$\mathbb{TR}_+$ 的乘法与序关系均为通常定义，但其加法 $+$ 改为 $x + y = \max\{x, y\}$

。我们还使用 $\mathbb{R}_+^{\mathrm{op}}$ 和 $\mathbb{TR}_+^{\mathrm{op}}$ ，其中上标 $_{\mathrm{op}}$ 表示反转预序。（反转预序半环上的预序会得到另一个预序半环。）

预序半环 S 可能具有、也可能不具有一项重要性质，即多项式增长，其定义如下。

定义 3 (多项式增长 [12])。对于预序半环 S，我们称 S 具有多项式增长，如果它存在一个幂通用对 $(u_-,\ u_+)$，即非零的 $u_\pm \in S$ ，满足：

- $u_- \preccurlyeq u_+$,
- 对于所有满足 $x \preccurlyeq y$ 的非零 $x, y \in S$ ，都存在 $k \in \mathbb{N}$ ，使得$u_-^k y \preccurlyeq u_+^k x$。

如果 u_- 有乘法逆元，则这对元素可以替换为单个幂通用 $u \trianglerighteq 1$ ，它具有如下性质：对于非零x，若 $x \preccurlyeq y$ ，$y \in S$，则存在 $k \in \mathbb{N}$ ，使得 $y \preccurlyeq x u^k$。

定义 4（单调同态）。给定预序半环 S，令$\mathbb{K} \in \{\mathbb{R}_+, \mathbb{R}_+^{\mathrm{op}}, \mathbb{TR}_+, \mathbb{TR}_+^{\mathrm{op}}\}$。那么，到 $\mathbb{K}$ 的单调同态是映射$\Phi: S \to \mathbb{K}$ ，满足对所有 $a, b \in S$，

- Φ 是可加的，即 $\Phi(a+b) = \Phi(a) + \Phi(b)$；
- Φ 是乘法的，即 $\Phi(ab) = \Phi(a) \cdot \Phi(b)$；
- Φ 是归一化的，即 $\Phi(0) = 0$ ，且 $\Phi(1) = 1$、
- 和 Φ 是单调的，即 $a \preccurlyeq b \implies \Phi(a) \preccurlyeq \Phi(b)$。

正如我们将看到的，在主序半环的情形下，最后一个条件特化为舒尔凸性的条件。

最后，当 S 是一个预序半环时，如果满足以下条件，我们称 x和 $y \in S$ 在大样本下有序：

$$x^n \preccurlyeq y^n \tag{14}$$

对所有充分大的 $n \in \mathbb{N}$都成立。此外，如果存在非零的 $a \in S$ 使得以下条件成立，我们称 x 和 y 是催化有序的：

$$xa \preccurlyeq ya. \tag{15}$$

现在，我们已经具备了该理论的基本工具，接下来正式陈述预序半环的分离定理，这些定理在 [12, 13]中称为 Vergleichsstellens ¨atze。粗略地说，这些分离定理为我们提供了判定给定的两个元素是否在大样本下有序或催化有序的标准。我们以稍短的形式陈述主要定理，只保留与当前应用相关的结果。

定理5（ [12]中的定理 7.15）。设 S 是一个具有多项式增长和幂泛元 u 的零和自由预序半环，并且满足 $0 \preccurlyeq 1$。设 x、$y \in S$ 满足 $x \neq 0$ ，且 y 是幂通用的。若对于每个

$\mathbb{K} \in \{\mathbb{R}_+$、 $\mathbb{TR}_+\}$ 以及每个单调同态 Φ ： $S \to \mathbb{K}$ ，只要满足$\Phi(u) > 1$，就有 $\Phi(x) < \Phi(y)$，那么

(a) 对所有充分大的 $n \in \mathbb{N}$， $x^n \preccurlyeq y^n$ ；以及 (b) 对于形如 $a = \sum_{\ell=0}^{n} x^\ell y^{n-\ell}$ 的催化剂， $ax \preccurlyeq ay$ 对所有充分大的 $n \in \mathbb{N}$ 均成立。

注6. 对于定理中的热带同态 Φ ： $S \to \mathbb{TR}_+$ ，任意幂 $\Phi(\cdot)^\lambda$ （ $\lambda > 0$ ）仍是这样的热带同态；不等式 $\Phi(x)^\lambda \leqslant \Phi(y)^\lambda$ 成立，当且仅当原不等式 $\Phi(x) \leqslant \Phi(y)$ 成立。这表明可以只考虑满足归一化条件 $\log \Phi(u) = 1$ 的热带同态。

若 $1 \preccurlyeq 0$ 在 S 中成立，则将定理应用于反预序S^{op} ，可知大样本序和催化序可以用到 $\mathbb{K} \in \{\mathbb{R}_+^{\mathrm{op}}$ 的单调同态来刻画， $\mathbb{TR}_+^{\mathrm{op}}\}$。这将有助于我们在附录中推导次主序半环的相关结论。尽管后文讨论的次主序半环满足 $0 \preccurlyeq 1$，但它没有幂通用元。反向排序的次主序半环具有定理 5 所要求的幂通用元，但其中 0 主序于 1。

定理5 要求 0 和 1 可比较，这限制了它在例如信息论情境中的适用性。在某些情况下，例如附录中讨论的次主序，我们确实有 $0 \preccurlyeq 1$。因此，我们可以在第5节中以定理B.1为起点，重新推导概率向量的近似大样本主序和催化主序结果。然而，在概率论和信息论中，通常存在一些自然出现的预序半环，其中 $1 \succcurlyeq 0$ 和 $1 \preccurlyeq 0$ 都不成立。这通常是因为概率测度按定义被归一化为 1，而我们只希望归一化相同的测度可以相互比较。尤其是，零测度无法与任何归一化测度比较，因此会导致 $1 \not\preccurlyeq 0$ 和$1 \not\succcurlyeq 0$。鉴于此，我们陈述一个在[13] 中证明的附加分离定理。为考虑 1 和 0 不可比较的情况，该定理还要求考虑到取值于 $\mathbb{R}_+^{\mathrm{op}}$ 和 $\mathbb{TR}_+^{\mathrm{op}}$ 的单调同态，以及单调导子，以刻画某些无穷小信息。

在陈述定理之前，我们先简要介绍一些相关的补充概念。设 S 和 T 是预序半环。若单调同态 Φ： $S \to T$ 可通过 $S/\!\sim$, 分解，或者等价地，若对所有 x、 $y \in S$ 均有，则称其为退化的，

$$x \preccurlyeq y \quad \Longrightarrow \quad \Phi(x) = \Phi(y). \tag{16}$$

否则， Φ 是非退化的。为了陈述定理，我们还需要一个相关概念。设 S 是一个半环， $\Phi : S \to \mathbb{R}_+$ 是一个同态。那么， Φ-导子是一个加法映射 $\Delta : S \to \mathbb{R}$ ，满足莱布尼茨法则

$$\Delta(xy) = \Phi(x)\Delta(y) + \Delta(x)\Phi(y) \tag{17}$$

对所有 x、$y \in S$ 都成立。

接下来的分离定理适用范围相当广，但只能让我们得出催化排序。先给出几个技术性定义。若（预序）半环 S 对每个 $x \in S$ 都存在 $n \in \mathbb{N}$ 和 $y \in S$，使得 $x + y = n$（其中 n 是单位元的 n 重和），则称它具有准补元。若 S 还是半整环，则当对每个非零 $x \in S$ 都存在 $n \in \mathbb{N}$ 和 $y \in S$，使得 $xy = n$ 时，称它具有准逆元。设 S 是预序半环，$\sim$ 是由预序通过前文所述的锯齿链生成的等价关系。我们用 $\mathsf{Frac}(S/\sim)$ 表示 $S/\sim$ 上的形式分式集合。向该集合引入加法逆元，相当于考察张量积 $\mathsf{Frac}(S/\sim) \otimes \mathbb{Z}$。

定理 7 (定理 7.1 见[13])。 设 S 是一个零和自由、具有多项式增长且含有幂通用对 (u_-, u_+)的预序半整环。另假设 $S/\sim$ 具有准补元和准逆元，并且 $\mathsf{Frac}(S/\sim) \otimes \mathbb{Z}$是有限个域的乘积。给定非零的 x，$y \in S$ 满足 $x \sim y$。假设对每个 $\mathbb{K} \in \{\mathbb{R}_+, \mathbb{R}_+^{\mathrm{op}}, \mathbb{TR}_+, \mathbb{TR}_+^{\mathrm{op}}\}$ 以及每个具有平凡核的单调同态$\Phi : S \to \mathbb{K}$，都有：

(i) 若 Φ 非退化，则 $\Phi(x) < \Phi(y)$。**(ii)** 若 $\mathbb{K} = \mathbb{R}_+$ 和 Φ 退化，则对于每个满足 $\Delta(u_+) = \Delta(u_-) + 1$的非零单调 Φ-导子 Δ，都有 $\Delta(x) < \Delta(y)$。

那么存在非零的 $a \in S$，使得 $ax \preccurlyeq ay$。反之，若存在这样的 a，则上述不等式以非严格形式成立。

为了刻画大样本中的序关系，我们还需要对所涉及的预序半环施加一些进一步的条件。可以证明，[13] 下文定理中列出的条件蕴含定理 7中所要求的条件。

定理 8（[13]中的定理 8.6）。 设 S 是一个带有幂泛元 u的零和自由预序半域。假设对于某个 $d \in \mathbb{N}$，存在一个满同态 $\|\cdot\| : S \to \mathbb{R}_{>0}^d \cup \{(0, \ldots, 0)\}$，其核平凡，并且满足

$$a \preccurlyeq b \;\Rightarrow\; \|a\| = \|b\| \quad \text{and} \quad \|a\| = \|b\| \;\Rightarrow\; a \sim b. \tag{18}$$

将 $\|\cdot\|$ 的分量同态记为 $\|\cdot\|_{(j)}$、$j = 1$、……、d。设 x、$y \in S\backslash\{0\}$ 满足 $\|x\| = \|y\|$，且其中 y 是幂通用的。如果，

(i) 对每个 $\mathbb{K} \in \{\mathbb{R}_+, \mathbb{R}_+^{\mathrm{op}}, \mathbb{TR}_+, \mathbb{TR}_+^{\mathrm{op}}\}$ 以及每个具有平凡核的非退化单调同态 $\Phi : S \to \mathbb{K}$，都有 $\Phi(x) < \Phi(y)$；**(ii)** 对每个单调 $\|\cdot\|_{(j)}$-导子 $\Delta : S \to \mathbb{R}$，若对所有分量指标 $j = 1, \ldots, d$ 均有 $\Delta(u) = 1$，则 $\Delta(x) < \Delta(y)$，

then

(a) 对所有充分大的 $n \in \mathbb{N}$， $x^n \preccurlyeq y^n$；(b) 对所有充分大的 $n \in \mathbb{N}$，对于形如 $a = \sum_{\ell=0}^{n} x^\ell y^{n-\ell}$ 的催化剂， $ax \preccurlyeq ay$。

反过来，如果对于任意一个 n 或 a，这两个性质中的任意一个成立，那么上述不等式均以非严格形式成立。

注9。 实际上，我们不必考虑上述所有导子，只需关注那些差异不能通过 $\sim$-等价分解的代表性导子，即这些导子并非可互换[13]。关于热带同态归一化的注6 仍然适用；同样，同态 $S \to \mathbb{TR}_+^{\mathrm{op}}$ 也可以归一化为 $\log \Phi(u) = -1$。

这里所述理论的预期应用涉及这样的问题：在大样本或催化条件下，如何将一个有限统计实验（通常由矩阵描述）转化为另一个实验。对实验进行这种比较，便得到本文大部分内容所研究的预序。在这一框架中，为上述分离定理中的大样本可转化性和催化可转化性给出条件的单调同态和导子，可以解释为量化实验信息量的单调量。稍后我们将简要讨论其在量子理论中的应用；在这一领域，这些映射应被视为量子态资源的资源单调量。

3. Matrix majorization

现在我们转而关注概率向量组的比较。给定两个概率向量组 $(p^{(1)}, \ldots, p^{(d)})$ 和 $(q^{(1)}, \ldots, q^{(d)})$，我们的主要问题是确定：当 $n \in \mathbb{N}$ 足够大时，是否存在一个随机映射 T，使得

$$T\big(p^{(k)}\big)^{\otimes n} = \big(q^{(k)}\big)^{\otimes n} \qquad \forall k = 1, \ldots, d, \tag{19}$$

这就是大样本中的矩阵主序。我们还讨论与之相关的催化矩阵主序问题。我们通过构造一个合适的预序半环，研究其到 $\mathbb{K} \in \{\mathbb{R}_+$、$\mathbb{R}_+^{\mathrm{op}}$、$\mathbb{TR}_+$和$\mathbb{TR}_+^{\mathrm{op}}\}$ 的单调同态，以及其上的单调导子来研究这些问题，随后应用分离定理8。

首先介绍一些本节将始终使用的记号。固定 $d \in \mathbb{N}$，并对所有 $n \in \mathbb{N}$，将具有非负元素且各列具有公共支撑的 $(n \times d)$ 矩阵集合记为 $\mathcal{V}_n^d$。我们将矩阵 $P \in \mathcal{V}_n^d$ 与其对应的列元组 $P = (p^{(1)}, \ldots, p^{(d)})$ 等同起来，并通常用相应的小写符号表示这些列。据此，每个 $P = (p^{(1)}, \ldots, p^{(d)}) \in \mathcal{V}_n^d$ 都满足

$$\operatorname{supp} p^{(1)} = \cdots = \operatorname{supp} p^{(d)}. \tag{20}$$

我们还定义

$$\mathcal{V}_{<\infty}^{d} := \bigcup_{n=1}^{\infty} \mathcal{V}_{n}^{d}. \tag{21}$$

我们来定义 $\mathcal{V}_{<\infty}^{d}$ 上的两个二元运算：对于 P、$Q \in \mathcal{V}_{<\infty}^{d}$，我们定义一个新矩阵，将 P 和 Q 依次上下堆叠，

$$P \boxplus Q = \begin{pmatrix} P \\ Q \end{pmatrix}. \tag{22}$$

用上文的元组表示法，这等价于

$$P \boxplus Q = (p^{(1)} \oplus q^{(1)}, \ldots, p^{(d)} \oplus q^{(d)}). \tag{23}$$

类似地，我们将 $P \boxtimes Q$ 定义为一个矩阵，其列是由 P 和 Q 的各列作克罗内克积得到的，

$$P \boxtimes Q = (p^{(1)} \otimes q^{(1)}, \ldots, p^{(d)} \otimes q^{(d)}), \tag{24}$$

其中，行的排列顺序在下文引入的等价关系 $\approx$ 下无关紧要。

我们还约定，对每个 $n \in \mathbb{N}$，用 $0_{n\times d}$ 表示其元素全为零的 $(n \times d)$ 矩阵。对于 P、$Q \in \mathcal{V}_{<\infty}^{d}$，其中 $P \in \mathcal{V}_{m}^{d}$ 且 $Q \in \mathcal{V}_{n}^{d}$，若存在 $t \geqslant m$、n 和一个 $(t \times t)$ 阶置换矩阵 Π，使得 $P \approx Q$，则记作 $P \approx Q$

$$\Pi(P \boxplus 0_{(t-m)\times d}) = Q \boxplus 0_{(t-n)\times d}. \tag{25}$$

这意味着，除去补入全零行的差异后，Q 可由 P 经过行置换得到。对于所有 $P \in \mathcal{V}_{<\infty}^{d}$，我们用 $[P]$ 表示 $\mathcal{V}_{<\infty}^{d}/\approx$ 中的等价类，即 $[P] := \{Q \in \mathcal{V}_{<\infty}^{d} \mid Q \approx P\}$。下文中，我们只关注所讨论矩阵的 $\approx$-等价类；通常，提到矩阵 $P \in \mathcal{V}_{<\infty}^{d}$时，实际上指的是等价类$[P]$。这也使我们可以假设，给定的 P、$Q \in \mathcal{V}_{<\infty}^{d}$、$P$和 Q 具有相同的高度，即对于某个足够大的 $n \in \mathbb{N}$，有 P、$Q \in \mathcal{V}_{n}^{d}$。

定义 $0 = [(0 \cdots 0)]$ 和 $1 = [(1 \cdots 1)]$ 为$(1 \times d)$行矩阵的等价类，并注意，在 $\mathcal{V}_{<\infty}^{d}/\approx$ 中，加法和乘法的形式为

$$[P] + [Q] = [P \boxplus Q], \qquad [P] \cdot [Q] = [P \boxtimes Q] \tag{26}$$

定义良好。我们通过规定 $[Q] \preccurlyeq [P]$（若存在列随机矩阵 T 使得 $Q = TP$）来定义 $\preccurlyeq$ 在 $\mathcal{V}_{<\infty}^{d}/\approx$ 上的预序。在这种情况下，我们也记作 $P \succeq Q$ 或 $Q \preceq P$，并称 P 主序于 Q。就各列而言，这意味着对所有 $k = 1$，.....，d，都有 $q^{(k)} = Tp^{(k)}$。容易验证，$(\mathcal{V}_{<\infty}^{d}/\approx, 0, 1, +, \cdot, \succeq)$ 是一个预序半环，我们作如下定义：

定义 10. 我们将预序半环 $(\mathcal{V}_{<\infty}^{d}/\approx, 0, 1, +, \cdot, \succeq)$ 称为S^{d}，并称其为（长度为 d 的）矩阵主序半环。

注记 11. 对于熟悉代数的读者，注记 27 的以下推广或许会有所帮助：作为一个普通半环，S^d 与群半环 $\mathbb{N}[\mathbb{R}^d_{>0}]$相同，因为将每个矩阵写成其各行的 $\boxplus$ 和，可知半环元素可以视为乘法群 $\mathbb{R}^d_{>0}$中元素的形式和；这种对应使得形式和的通常加法和乘法分别对应于上文引入的运算。

回顾第 2.2 节，其中 $\sim$ 表示由以下关系生成的等价关系$\trianglelefteq$，即 $[P] \sim [Q]$（或者简单地说，$P \sim Q$），如果存在 $n \in \mathbb{N}$和 $R_1, \ldots, R_n \in \mathcal{V}^d_{<\infty}$使得

$$P \trianglelefteq R_1 \trianglerighteq R_2 \trianglelefteq \cdots \trianglerighteq R_n \trianglelefteq Q. \tag{27}$$

事实上，上述序列 $R_1, \ldots, R_n$ 可以取为长度 1。为此，注意到这样一个不等式链意味着对于所有列指标 $k = 1, \ldots, d$，都有 $\|p^{(k)}\|_1 = \|q^{(k)}\|_1$，因为乘以随机矩阵不会改变任何一列的归一化。但我们还可得到

$$P \trianglerighteq R \trianglelefteq Q, \tag{28}$$

其中 $R := \big(\|p^{(1)}\|_1, \ldots, \|p^{(d)}\|_1\big)$ 只有一行。由此还可看出

$$S^d/\!\sim \;\cong \mathbb{R}^d_{>0} \cup \{(0, \ldots, 0)\}, \tag{29}$$

它是一个具有准补元和准逆元的半环，并且 $\mathbf{Frac}(S^d/\sim) \otimes \mathbb{Z} \cong \mathbb{R}^d$ 是有限个域的乘积。定义退化同态 $\|\cdot\| \colon S^d \to \mathbb{R}^d_{>0} \cup \{(0, \ldots, 0)\}$ 为

$$\|P\| := \big(\|p^{(1)}\|_1, \ldots, \|p^{(d)}\|_1\big), \tag{30}$$

由此我们得出

$$P \trianglerighteq Q \;\Rightarrow\; \|P\| = \|Q\| \quad \text{and} \quad \|P\| = \|Q\| \;\Rightarrow\; P \sim Q. \tag{31}$$

因此，我们处于这样一种情形：定理 8的辅助假设成立。

引理 12. 矩阵主序半环具有多项式增长：任意随机矩阵 $P = (p^{(1)}, \ldots, p^{(d)})$，若满足 $p^{(k)} \neq p^{(\ell)}$（对 $k \neq \ell$），则是幂通用的。

证明。为证明这样的 P 是幂通用的，我们需要从满足 $Q \trianglerighteq R$ 的 $R, Q \in \mathcal{V}^d_{<\infty}\backslash\{0\}$ 出发，并证明存在 $n \in \mathbb{N}$ 使得

$$P^{\boxtimes n} \boxtimes R \trianglerighteq Q. \tag{32}$$

通过逐列重新归一化，我们可以假设 R 和 Q 都是随机的，并且不失一般性地假设 Q 的所有元素均严格为正。若我们证明存在 $n \in \mathbb{N}$ 使得 $P^{\boxtimes n} \trianglerighteq Q$，那么便有

$$P^{\boxtimes n} \boxtimes R \trianglerighteq P^{\boxtimes n} \boxtimes (1 \cdots 1) = P^{\boxtimes n} \trianglerighteq Q, \tag{33}$$

于是结论得证。因此，我们将证明存在 $n \in \mathbb{N}$ 使得 $P^{\boxtimes n} \geq Q$。

利用多重假设检验中的一个结果， [19], 存在 $(\varepsilon_k^{(n)}, \ell)_{k,\ \ell=1,\ \ldots\ldots d}$使得

$$P^{\boxtimes n} \geq U^{(n)} = \begin{pmatrix} 1-\varepsilon_{1,1}^{(n)} & \varepsilon_{1,2}^{(n)} & \cdots & \varepsilon_{1,d}^{(n)} \\ \varepsilon_{2,1}^{(n)} & 1-\varepsilon_{2,2}^{(n)} & \cdots & \varepsilon_{2,d}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{d,1}^{(n)} & \varepsilon_{d,2}^{(n)} & \cdots & 1-\varepsilon_{d,d}^{(n)} \end{pmatrix} \tag{34}$$

并且所有误差 $\varepsilon_{k\,\ell}^{(n)}$, 都以指数速度衰减，衰减率由 P 的多重切尔诺夫散度给出， n

$$C(P) := \min_{k,\ell:k\neq\ell} \max_{0\leqslant\alpha\leqslant 1} (1-\alpha) D_\alpha(p^{(k)} \| p^{(\ell)}), \tag{35}$$

其中 $D_\alpha(\cdot\|\cdot)$ 是稍后在（158）中介绍的 Rényi 散度。该量严格为正，例如，由于巴塔查里亚距离 $D_{1/2}$ 满足 $D_{1/2}(p^{(k)}\|p^{(\ell)}) \neq 0$ ，只要 $p^{(k)} \neq p^{(\ell)}$，而我们对所有 $k \neq \ell$ 都作了这一假设，便可看出这一点。因此，当 $n \to \infty$ 时，我们得到 $\varepsilon_k^{(n)}$, $\ell \to 0$ ；于是当 $n \to \infty$ 时，$U^{(n)} \to I_d$ ，其中 I_d 是 $(d \times d)$ 单位矩阵。特别地，当 n 足够大时，$U^{(n)}$ 可逆，我们将其逆记为 $V^{(n)} := (U^{(n)})^{-1}$。由于归一化列矩阵的逆矩阵也具有归一化列[2], ，因此我们知道这适用于 $V^{(n)}$。

为了证明对于充分大的 n， $U^{(n)} \geq Q$ ，现在考虑矩阵

$$T^{(n)} := QV^{(n)}. \tag{36}$$

根据构造， $T^{(n)}U^{(n)} = Q$ 成立，因此只需证明$T^{(n)}$ 是随机的。不过，我们已经知道它的列是归一化的，因为 $V^{(n)}$ 和 Q 都有[3], ，所以现在只需证明其元素非负。事实确实如此，因为 Q 的所有元素都严格为正，并且根据矩阵求逆的连续性和 $U^{(n)} \to I_d$， $V^{(n)} \to I_d$ 如 $n \to \infty$ 。□

我们来定义一族重要的实值函数，其定义域为 S^d。为避免公式中出现多个括号，我们稍微滥用了记号，实际上将这些函数视为定义在 $\mathcal{V}_{<\infty}^d$ 上，其中各个矩阵都与其 $\approx$-等价类相对应。对于每个 $P \in \mathcal{V}_n^d \subset \mathcal{V}_{<\infty}^d$，我们省略零行（因为它们不会对这些函数产生贡献），并记

$$f_{\underline{\alpha}}(P) = \sum_{i=1}^{n} \prod_{k=1}^{d} \left(p_i^{(k)}\right)^{\alpha_k}, \tag{37}$$

[2]取 $e = (1 \cdots 1)$，我们有 $eV^{(n)} = eU^{(n)}V^{(n)} = e$，其中第一步用到 $U^{(n)}$的列是归一化的，第二步用到 $V^{(n)}$ 是 $U^{(n)}$的逆矩阵。[3]使用 $eQV^{(n)} = eV^{(n)} = e$。

$$f^{\mathbb{T}}_{\underline{\beta}}(P) = \max_{1\leqslant i\leqslant n} \prod_{k=1}^{d} \left(p_i^{(k)}\right)^{\beta_k}, \tag{38}$$

$$\begin{aligned}\Delta^{(k)}_{\underline{\gamma}}(P) &= \sum_{i=1}^{n} \sum_{\ell:\,\ell\neq k} \gamma_\ell\, p_i^{(k)} \log \frac{p_i^{(k)}}{p_i^{(\ell)}} \\ &= \sum_{\ell:\,\ell\neq k} \gamma_\ell D_1(p^{(k)} \| p^{(\ell)}),\end{aligned} \tag{39}$$

对所有元组 $\underline{\alpha}$, $\underline{\beta}$, $\underline{\gamma} \in \mathbb{R}^d$ 以及 $k = 1, \ldots, d$。注意，$D_1(\cdot\|\cdot)$ 在（39）中的量是 Kullback-Leibler 散度。不过，我们主要关注更具体的参数范围。因此，我们定义参数集

$$A := \{\underline{\alpha} \in \mathbb{R}^d \,|\, \alpha_1 + \cdots + \alpha_d = 1\}, \tag{40}$$

$$A_- := \bigcup_{k=1}^{d} \{\underline{\alpha} \in A \,|\, \alpha_k \geqslant 1 \ \wedge\ \alpha_\ell \leqslant 0 \ \ \forall \ell \neq k\}, \tag{41}$$

$$A_+ := \{\underline{\alpha} \in A \,|\, \alpha_\ell \geqslant 0 \ \ \forall \ell\}, \tag{42}$$

$$B := \{\underline{\beta} \in \mathbb{R}^d \,|\, \beta_1 + \cdots + \beta_d = 0\}, \tag{43}$$

$$B_- := \bigcup_{k=1}^{d} \{\underline{\beta} \in B \,|\, \beta_k \geqslant 0 \ \wedge\ \beta_\ell \leqslant 0 \ \ \forall \ell \neq k\}. \tag{44}$$

例如，A_- 由恰有一个正项的所有归一化数值元组组成，根据归一化，该正项必然为 $\geqslant 1$；而 A_+ 是 $\mathbb{R}^d$中的标准概率单纯形。将量 $f_{\underline{\alpha}}(P)$ 视为 $\underline{\alpha} \in A_+$ 的函数时，也称为海林格变换 [29, 例 1.4.4]。

3.1. 矩阵主序半环上的单调同态与导子。 接下来，我们将用（S^d 37）和（38）来刻画 e_1, ..., $e_d \in \mathbb{R}^d$ 所表示的标准基向量。

命题 13。非退化单调同态 $S^d \to \mathbb{K}$ 恰好如下：

(i) 对于 $\mathbb{K} = \mathbb{R}_+$，映射 $f_{\underline{\alpha}}$ 适用于所有元组 $\underline{\alpha} \in A_-\backslash\{e_1, \ldots, e_d\}$。(ii) 对于 $\mathbb{K} = \mathbb{R}_+^{\mathrm{op}}$，映射 $f_{\underline{\alpha}}$ 适用于所有元组 $\underline{\alpha} \in A_+\backslash\{e_1, \ldots, e_d\}$。(iii) 对于 $\mathbb{K} = \mathbb{TR}_+$，映射 $f^{\mathbb{T}}_{\underline{\beta}}$ 适用于所有元组 $\underline{\beta} \in B_-\backslash\{0\}$。(iv) 对于 $\mathbb{K} = \mathbb{TR}_+^{\mathrm{op}}$，不存在这样的映射。

证明。设 Φ: $S^d \to \mathbb{K}$ 是这样的单调同态。对于每种情况，我们定义函数 φ_k: $\mathbb{R}_{>0} \to \mathbb{R}_+$，对所有 $k = 1, \ldots, d$，定义如下

$$\varphi_k(x) = \Phi(1 \cdots 1\, x\, 1 \cdots 1), \tag{45}$$

其中 x 出现在第 k 个位置。显然，这些函数是乘法的，即对于所有 x、$y > 0$ 和 $\varphi_k(1) = 1$，都有 $\varphi_k(xy) = \varphi_k(x)\varphi_k(y)$。接下来我们证明，对于某些指数 α_k，它们是幂函数 $x \mapsto x^{\alpha_k}$。根据乘法性以及柯西函数方程的标准理论，只需证明它们在某个区间 $[x, y]$ 上有界，其中 $0 < x < y$。对于所有 $t \in [0, 1]$，定义矩阵

$$T_t = \begin{pmatrix} t & 1-t \\ 1-t & t \end{pmatrix}, \tag{46}$$

$$P_t = \begin{pmatrix} 1 & \cdots & 1 & tx + (1-t)y & 1 & \cdots & 1 \\ 1 & \cdots & 1 & (1-t)x + ty & 1 & \cdots & 1 \end{pmatrix}, \tag{47}$$

我们看到，T_t 是随机的，并且

$$P_{1/2} = T_{1/2}P_t \preceq P_t = T_tP_1 \preceq P_1. \tag{48}$$

将 Φ 应用于这些不等式，并利用 P_t 可表示为其在 S^d 中两行之和这一事实，我们得到

$$\text{case (i): } 2\varphi_k\left(\frac{1}{2}(x+y)\right) \leqslant \varphi_k\big(tx + (1-t)y\big) + \varphi_k\big((1-t)x + ty\big) \tag{49}$$

$$\leqslant \varphi_k(x) + \varphi_k(y), \tag{50}$$

$$\text{case (ii): } 2\varphi_k\left(\frac{1}{2}(x+y)\right) \geqslant \varphi_k\big(tx + (1-t)y\big) + \varphi_k\big((1-t)x + ty\big) \tag{51}$$

$$\geqslant \varphi_k(x) + \varphi_k(y), \tag{52}$$

$$\text{case (iii): } \varphi_k\left(\frac{1}{2}(x+y)\right) \leqslant \max_{s\in\{t,1-t\}} \varphi_k\big(sx + (1-s)y\big) \tag{53}$$

$$\leqslant \max\{\varphi_k(x), \varphi_k(y)\}, \tag{54}$$

$$\text{case (iv): } \varphi_k\left(\frac{1}{2}(x+y)\right) \geqslant \max_{s\in\{t,1-t\}} \varphi_k\big(sx + (1-s)y\big) \tag{55}$$

$$\geqslant \max\{\varphi_k(x), \varphi_k(y)\}. \tag{56}$$

由于 φ_k 取非负值，我们现在看到，对于所有 $z \in [x、y]$，

$$\text{case (i): } 0 \leqslant \varphi_k(z) \leqslant \varphi_k(x) + \varphi_k(y), \tag{57}$$

$$\text{case (ii): } 0 \leqslant \varphi_k(z) \leqslant 2\varphi_k\left(\frac{1}{2}(x+y)\right), \tag{58}$$

$$\text{case (iii): } 0 \leqslant \varphi_k(z) \leqslant \max\{\varphi_k(x), \varphi_k(y)\}, \tag{59}$$

$$\text{case (iv): } 0 \leqslant \varphi_k(z) \leqslant \varphi_k\left(\frac{1}{2}(x+y)\right). \tag{60}$$

特别地，φ_k 在区间 $[x, y]$ 上确实有界。因此，存在 $\alpha_k \in \mathbb{R}$，使得对所有 $x > 0$，都有 $\varphi_k(x) = x^{\alpha_k}$。

在情形 (i) 和 (ii) 中，对于所有具有全支撑的 P，我们现在有：

$$\Phi(P) = \sum_{i=1}^{n} \Phi\big((p_i^{(1)} \cdots p_i^{(d)})\big) \tag{61}$$

$$= \sum_{i=1}^{n} \Phi\big((p_i^{(1)}\, 1 \cdots 1) \boxtimes \cdots \boxtimes (1 \cdots 1\, p_i^{(d)})\big) \tag{62}$$

$$= \sum_{i=1}^{n} \prod_{k=1}^{d} \varphi_k(p_i^{(k)}) = \sum_{i=1}^{n} \prod_{k=1}^{d} (p_i^{(k)})^{\alpha_k}, \tag{63}$$

其中，n 是 P的行数。类似地，在情形 (iii) 和 (iv) 中，我们得到

$$\Phi(P) = \max_{1 \leqslant i \leqslant n} \prod_{k=1}^{d} (p_i^{(k)})^{\alpha_k}. \tag{64}$$

由于这种形式的函数显然确实是同态，因此只需确定它们在哪些参数元组下是单调且非退化的。不难看出

$$\begin{pmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{pmatrix} \geq \begin{pmatrix} 2 \cdots 2 \end{pmatrix} \geq \begin{pmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{pmatrix}, \tag{65}$$

根据单调性可知，

$$\Phi\begin{pmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{pmatrix} = \Phi(2 \cdots 2). \tag{66}$$

在情形 (i) 和 (ii) 中，这意味着 $\alpha_1 + \cdots + \alpha_d = 1$；在情形 (iii) 和 (iv) 中，这意味着 $\alpha_1 + \cdots + \alpha_d = 0$。

暂且聚焦于 (i) 和 (ii)。在这两种情形下，我们现在知道，Φ必定具有形式 $f_{\underline{\alpha}}$，其中参数为 $\underline{\alpha} \in A$。现在来证明，在情形 (i) 中，$\underline{\alpha} \in A_{-}$ 是必要条件；在情形 (ii) 中，$\underline{\alpha} \in A_{+}$ 是必要条件。对于 P、$Q \in \mathcal{V}_n^d$ 和 $t \in [0, 1]$，设 $(n \times 2n)$列随机矩阵 T定义为：对所有 a、$b \in \mathbb{R}^n$，$T(a \oplus b) = a + b$。注意，在情形 (i) 中，我们有

$$\Phi\big(tP + (1-t)Q\big) = \Phi\big(T\big[tP \boxplus (1-t)Q\big]\big) \tag{67}$$

$$\leqslant \Phi\big(tP \boxplus (1-t)Q\big) \tag{68}$$

$$= \Phi(tP) + \Phi\big((1-t)Q\big) \tag{69}$$

$$= t\Phi(P) + (1-t)\Phi(Q), \tag{70}$$

即，Φ是凸的。最后一个等式成立是因为 $\alpha_1 + \cdots + \alpha_d = 1$。类似地，在情形 (ii) 中，我们发现 Φ 必定是凹的。定义单行函数φ：$\mathbb{R}_{>0}^d \to \mathbb{R}_+$ 如下：

$$\varphi(x_1, \ldots, x_d) := \Phi\big((x_1 \cdots x_d)\big) = x_1^{\alpha_1} \cdots x_d^{\alpha_d} \tag{71}$$

对所有 $(x_1、\ldots\ldots、x_d) \in \mathbb{R}^d_{>0}$均成立。显然，在情形 (i) 中，$\varphi$ 也必定是凸的；在情形 (ii) 中则必定是凹的。为了分析 φ的凸性，我们来计算其海森矩阵：

$$H_\varphi(x_1,\ldots,x_d) = \left(\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}(x_1,\ldots,x_d)\right)_{i,j=1}^{d} \tag{72}$$

$$= \varphi(x_1,\ldots,x_d)\left(\frac{1}{x_i x_j}\right)_{i,j=1}^{d} \star \underbrace{(\alpha_i\alpha_j - \delta_{i,j}\alpha_i)_{i,j=1}^{d}}_{=:\mathcal{A}} \tag{73}$$

其中 $\star$ 表示舒尔积，即形状相同的矩阵逐元素相乘。由于与半正定矩阵作舒尔积会保持矩阵的半定性，且 φ 取正值，因此，将其与半正定矩阵 $(x_i x_j)_{i相j=1}^{d}$乘后，H_φ 为正（分别为负）半定，当且仅当 $\mathcal{A}$为正（分别为负）半定。根据 [21],的引理 8，我们有：

- $\mathcal{A}$ 为半正定，当且仅当 $\underline{\alpha} \in A_-$。
- $\mathcal{A}$ 为负半定，当且仅当 $\underline{\alpha} \in A_+$。

φ 的凸性（凹性）意味着其满足结论中所述的参数 $\underline{\alpha}$ 的条件；其中，$\underline{\alpha}$ 为标准基向量 e_k 的情形可以排除，因为此时 $f_{\underline{\alpha}}$ 由 $f_{\underline{\alpha}}(P) = \|p^{(k)}\|_1$给出，而它是退化的。

接下来我们证明，所有这些同态确实都是单调且非退化的。对于情形 (i)，设 $\underline{\alpha} \in A_-$；此时根据上述内容，我们知道 φ 是凸的。设 $P \in \mathcal{V}_n^d$ 的各行为 $p_1、\ldots\ldots、p_n$，并令 $T = (T_{i,\,j})_{i,\,j}$ 为一个 $(m \times n)$-随机矩阵。不失一般性，假设该矩阵没有零行，并记为 $T_i := \sum_{j=1}^n T_{i,\,j}$。记 $Q := TP \in \mathcal{V}_m^d$，其各行为 $q_1、\ldots\ldots、q_m$。由于 φ 也因 $\alpha_1 + \cdots + \alpha_d = 1$ 而保持数乘运算，我们有

$$f_{\underline{\alpha}}(Q) = \sum_{i=1}^m \varphi(q_i) = \sum_{i=1}^m \varphi\left(\sum_{j=1}^n T_{i,j}p_j\right) = \sum_{i=1}^m T_i\,\varphi\left(\sum_{j=1}^n \frac{T_{i,j}}{T_i}p_j\right) \tag{74}$$

$$\leqslant \sum_{i=1}^m T_i \sum_{j=1}^n \frac{T_{i,j}}{T_i}\varphi(p_j) = \sum_{i=1}^m\sum_{j=1}^n T_{i,j}\varphi(p_j) = \sum_{j=1}^n \varphi(p_j) = f_{\underline{\alpha}}(P), \tag{75}$$

其中第二步利用了标量乘法的保运算性，而该不等式由 φ 的凸性得出。因此，$f_{\underline{\alpha}} : S^d \to \mathbb{R}_+$ 确实关于 $\alpha \in A_-$ 单调。对于情形 (ii)，用方向相反的不等式作同样的论证，可知 $f_{\underline{\alpha}} : S^d \to \mathbb{R}_+^{\mathrm{op}}$ 对所有 $\alpha \in A_+$ 都是单调的。为了完成情形 (i) 和 (ii) 的证明，还需证明当 $\underline{\alpha}$ 不是标准基向量时的非退化性。我们来证明，只要存在一个指标 k

满足 $\alpha_k \neq 0$、1，则 $f_{\underline{\alpha}}$ 是非退化的。定义矩阵

$$P_t := \begin{pmatrix} \frac{1}{2} & \cdots & \frac{1}{2} & 1-t & \frac{1}{2} & \cdots & \frac{1}{2} \\ \frac{1}{2} & \cdots & \frac{1}{2} & t & \frac{1}{2} & \cdots & \frac{1}{2} \end{pmatrix} \in \mathcal{V}^d_{<\infty}, \tag{76}$$

其中非平凡列是第 k 列。注意，每当 $0 < s \leqslant t \leqslant \frac{1}{2}$ 时，$P_s \geq P_t$。容易看出，这些值

$$f_{\underline{\alpha}}(P_t) = 2^{\alpha_k - 1}\big((1-t)^{\alpha_k} + t^{\alpha_k}\big) \tag{77}$$

只要 $\alpha_k \neq 0, 1$，就不会在 t 中保持常数，这表明当 $\underline{\alpha}$ 不是标准基向量时，$f_{\underline{\alpha}}$ 是非退化的。因此，对于情形 (i) 和 (ii)，我们已经证明了该结论。

现在转而讨论热带情形 (iii) 和 (iv)。在情形 (iii) 中，我们可以采用与情形 (i) 类似的推理，证明 Φ 是拟凸的，即

$$\Phi\big(tP + (1-t)Q\big) \leqslant \max\{\Phi(P), \Phi(Q)\} \tag{78}$$

对于所有 P、$Q \in \mathcal{V}^d_{<\infty}$ 以及 $t \in (0, 1)$。特别地，单行函数 φ 必须是定义在 $\mathbb{R}^d_{>0}$ 上的拟凸函数。这种拟凸性等价于次水平集的凸性

$$S_c := \{(x_1, \ldots, x_d) \in \mathbb{R}^d_{>0} \,|\, x_1^{\alpha_1} \cdots x_d^{\alpha_d} \leqslant c\} \tag{79}$$

对所有 $c > 0$。不失一般性，我们假设 $\alpha_1 > 0$ （必要时可通过置换列来实现）。于是，将第一个坐标 x_1 视为因变量，我们注意到 S_c 恰好包含函数 $w\colon \mathbb{R}^{d-1}_{>0} \to \mathbb{R}$ 的图像下方（含边界）的所有点，其中该函数定义为

$$w(x_2, \ldots, x_d) = \frac{x_2^{\mu_2} \cdots x_d^{\mu_d}}{c}, \tag{80}$$

其中 $\mu_k := -\frac{\alpha_k}{\alpha_1}$，对于 $k = 2, \ldots, d$，这些系数现在又满足和为 1。正如我们先前所见，w 的海森矩阵很容易求得：

$$H_w(x_2, \ldots, x_d) = w(x_1, \ldots, x_d)\left(\frac{1}{x_i x_j}\right)^d_{i,j=2} \star \underbrace{(\mu_i\mu_j - \delta_{i,j}\mu_i)^d_{i,j=2}}_{=:\,\mathcal{B}}. \tag{81}$$

因此，要使集合 S_c 为凸集，w 必须为凹函数，或者等价地，H_w 必须为负半定。这与上文相同的论证等价于 $\mu_2, \ldots, \mu_d \geqslant 0$，或者等价地，对于 $k \geqslant 2$，$\alpha_k \leqslant 0$。因此，只有当 $\alpha \in B_-$时，$\Phi = f^{\mathbb{T}}_{\underline{\alpha}}$才是单调的。由于 $f^{\mathbb{T}}_0$ 是将所有非零 P 简单映射为 1 的同态，因此该同态是退化的，我们可以得出结论 $\underline{\alpha} \in B_-\backslash\{0\}$。这就证明了 (iii) 的一个方向。

对于情形 (iv) 中相同且唯一的方向，现在令 $\Phi : S^d \to \mathbb{TR}^{\mathrm{op}}_+$ 为一个非退化单调同态。与情形 (iii) 类似，我们现在有

$$\Phi\big(tP + (1-t)Q\big) \geqslant \max\{\Phi(P), \Phi(Q)\}. \tag{82}$$

对所有 P、$Q \in \mathcal{V}_{<\infty}^{d}$ 以及 $t \in (0, 1)$ 均成立。接下来我们证明，若 $P \neq Q$ 具有相同的支撑，则 $\Phi(P) = \Phi(Q)$。在这种情况下，我们可以找到 s、$t \in (0, 1)$ 和 P'、$Q' \in \mathcal{V}_{<\infty}^{d}$，使得

$$sP + (1-s)P' = Q, \qquad tQ + (1-t)Q' = P. \tag{83}$$

因此，根据（82），我们有

$$\Phi(P) \geqslant \max\{\Phi(Q), \Phi(Q')\} \geqslant \Phi(Q) \geqslant \max\{\Phi(P), \Phi(P')\} \geqslant \Phi(P), \tag{84}$$

即，$\Phi(P) = \Phi(Q)$。这意味着，Φ 仅仅是其自变量支撑的函数。但若 Q 是一个秩 1 随机矩阵，则

$$Q \geq (1 \cdots 1) \geq Q, \tag{85}$$

因此 $\Phi(Q) = 1$。但若 $P \in \mathcal{V}_{<\infty}^{d}$ 非零，我们可以找到任意一个具有相同支撑的秩 1 随机矩阵，从而得出 $\Phi(P) = 1$。因此，不存在到$\mathbb{TR}_{+}^{\mathrm{op}}$ 的非退化同态，情形 (iv) 至此完成。

最后，我们需要证明，满足$\underline{\alpha} \in B_{-}\backslash\{0\}$ 的映射$f_{\underline{\alpha}}^{\mathbb{T}}$是非退化的单调同态 $S^d \to \mathbb{TR}_{+}$。验证它们是同态很简单，因此我们重点讨论单调性。根据上文，我们知道单行函数 φ 是拟凸的。根据 $\alpha_1 + \cdots + \alpha_d = 0$，我们还知道，对所有正标量 t，$f_{\underline{\alpha}}^{\mathbb{T}}(tP) = f_{\underline{\alpha}}^{\mathbb{T}}(P)$。然后，取与情形 (i) 对应证明中相同的 P 和 $Q = TP$，我们可以采用与那里类似的论证，

$$f_{\underline{\alpha}}^{\mathbb{T}}(Q) = \max_{1\leqslant i\leqslant m} \varphi(q_i) = \max_{1\leqslant i\leqslant m} \varphi\left(\sum_{j=1}^{n} T_{i,j} p_j\right) \tag{86}$$

$$= \max_{1\leqslant i\leqslant m} \varphi\left(T_i \sum_{j=1}^{n} \frac{T_{i,j}}{T_i} p_j\right) = \max_{1\leqslant i\leqslant m} \varphi\left(\sum_{j=1}^{n} \frac{T_{i,j}}{T_i} p_j\right) \tag{87}$$

$$\leqslant \max_{1\leqslant i\leqslant m} \max_{1\leqslant j\leqslant 1} \varphi(p_j) = \max_{1\leqslant j\leqslant n} \varphi(p_j) = f_{\underline{\alpha}}^{\mathbb{T}}(P). \tag{88}$$

我们来证明，当 $\underline{\alpha} \neq 0$时，所有这些同态都是非退化的；或者等价地，当存在 k 满足 $\alpha_k > 0$时。取 P_t 如（76）所示，可得

$$f_{\underline{\alpha}}^{\mathbb{T}}(P_t) = (2(1-t))^{\alpha_k} \tag{89}$$

对所有 $t \in (0, \frac{1}{2}]$均成立。这显然是一个非常值函数，也就是说，$f_{\underline{\alpha}}^{\mathbb{T}}$是非退化的。□

现在让我们继续研究退化同态 f_{e_k}处的单调导子：对于 $k = 1$，$S^d \to \mathbb{R}_{+}$ 满足 $f_{e_k}(P) = \|p^{(k)}\|$，……，d。对于这些导子，莱布尼茨法则为

$$\Delta(P \boxtimes Q) = \Delta(P)\|q^{(k)}\|_1 + \|p^{(k)}\|_1 \Delta(Q). \tag{90}$$

现在转而考察 $\Delta_{\underline{\gamma}}^{(k)}$（39）中的映射，我们很容易看出，这些映射是可加的（在 $\boxplus$下），并满足（90）。接下来我们证明，满足$\underline{\gamma} \in \mathbb{R}_+^d$ 的这些映射恰好是 f_{e_k}处的单调导子。

命题 **14**。设 $k \in \{1, \ldots, d\}$。对于满足 $\|p^{(k)}\|_1 \in \mathbb{Q}$ 的矩阵 $P \in \mathcal{V}_{<\infty}^d$，在 f_{e_k} 处作用于 S^d 的单调导子恰好是满足 $\underline{\gamma} \in \mathbb{R}_+^d$ 的 $\Delta_{\underline{\gamma}}^{(k)}$。

注意，$\Delta_{\underline{\gamma}}^{(k)}$ 实际上并不依赖于 γ_k，因此这一分量的值完全是任意的。

证明。我们固定 $k = 1$；其他情形类似。设 $\Delta\colon\ S^d \to \mathbb{R}$ 是在 f_{e_1} 处的单调导子。定义 $\delta\colon\ \mathbb{R}_{>0}^d \cup \{(0, \ldots, 0)\} \to \mathbb{R}$ 为单行函数

$$\delta(x_1, \ldots, x_d) = \Delta\big(x_1 \ \cdots \ x_d\big). \tag{91}$$

我们首先证明，对于所有 $x \in \mathbb{Q}_+$，都有 $\delta(x, \ldots, x) = 0$。由莱布尼茨法则立即可得 $\delta(0, \ldots, 0) = 0 = \delta(1, \ldots, 1)$。对于每个 $n \in \mathbb{N}$，容易看出

$$\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \geq (n \ \cdots \ n) \geq \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \tag{92}$$

其中全 **1** 矩阵为 $(n \times d)$。利用 Δ 的单调性和可加性，我们有

$$\delta(n, \ldots, n) = \Delta\left(\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}\right) = n\delta(1, \ldots, 1) = 0. \tag{93}$$

由此也可轻易得出莱布尼茨法则 $\delta(1/n, \ldots, 1/n) = -n^{-2}\delta(n, \ldots, n) = 0$。最后，对于所有 m、$n \in \mathbb{N}$,

$$\delta\Big(\frac{m}{n}, \ldots, \frac{m}{n}\Big) = m \underbrace{\delta\Big(\frac{1}{n}, \ldots, \frac{1}{n}\Big)}_{=0} + \underbrace{\delta(m, \ldots, m)}_{=0} \frac{1}{n} = 0, \tag{94}$$

这表明对于所有 $x \in \mathbb{Q}_+$，$\delta(x, \ldots\ldots, x) = 0$。

现在令 $p \in \mathbb{R}_+^n$ 为任意长度 $n \in \mathbb{N}$ 的列向量，且满足$\|p\|_1 \in \mathbb{Q}$。利用显而易见的事实:

$$(p, \ldots, p) \geq (\|p\|_1 \ \cdots \ \|p\|_1) \geq (p, \ldots, p), \tag{95}$$

Δ的单调性，以及上文证明的关于 δ的消失性质，可得

$$\Delta(p, \ldots, p) = \delta(\|p\|_1, \ldots, \|p\|_1) = 0. \tag{96}$$

设 $P \in \mathcal{V}_n^d$ 满足 $\|p^{(1)}\|_1 \in \mathbb{Q}$，并且不失一般性地假设其具有全支撑。多次使用莱布尼茨法则以及上述消失性质，可得

$$\Delta(P) = \sum_{i=1}^{n} \delta(p_i^{(1)}, \dots, p_i^{(d)}) = \sum_{i=1}^{n} \delta\left(p_i^{(1)} \cdot 1, p_i^{(1)} \cdot \frac{p_i^{(2)}}{p_i^{(1)}}, \dots, p_i^{(1)} \cdot \frac{p_i^{(d)}}{p_i^{(1)}}\right) \tag{97}$$

$$= \sum_{i=1}^{n} \left[\delta(p_i^{(1)}, \dots, p_i^{(1)}) + p_i^{(1)} \delta\left(1, \frac{p_i^{(2)}}{p_i^{(1)}}, \dots, \frac{p_i^{(d)}}{p_i^{(1)}}\right) \right] \tag{98}$$

$$= \underbrace{\Delta(p^{(1)}, \dots, p^{(1)})}_{=0} + \sum_{i=1}^{n} p_i^{(1)} \delta\left(1, \frac{p_i^{(2)}}{p_i^{(1)}}, \dots, \frac{p_i^{(d)}}{p_i^{(1)}}\right) \tag{99}$$

$$= \sum_{i=1}^{n} p_i^{(1)} \delta\left(1, \frac{p_i^{(2)}}{p_i^{(1)}} \cdot 1, 1 \cdot \frac{p_i^{(3)}}{p_i^{(1)}}, \dots, 1 \cdot \frac{p_i^{(d)}}{p_i^{(1)}}\right) \tag{100}$$

$$= \sum_{i=1}^{n} p_i^{(1)} \left[\delta\left(1, \frac{p_i^{(2)}}{p_i^{(1)}}, 1, \dots, 1\right) + \delta\left(1, 1, \frac{p_i^{(3)}}{p_i^{(1)}}, \dots, \frac{p_i^{(d)}}{p_i^{(1)}}\right) \right] \tag{101}$$

$$= \dots = \sum_{i=1}^{n} \sum_{\ell=2}^{d} p_i^{(1)} g_\ell\left(\frac{p_i^{(\ell)}}{p_i^{(1)}}\right), \tag{102}$$

其中，对于所有 $\ell = 2$、……、d，我们使用了

$$g_\ell(x) := \delta(1, \dots, 1, x, 1, \dots, 1), \tag{103}$$

其中 x 出现在第 ℓ 个位置。利用莱布尼茨法则，很容易看出，对于所有 x、$y > 0$ 以及所有 ℓ，

$$g_\ell(xy) = g_\ell(x) + g_\ell(y), \tag{104}$$

因此，g_ℓ 满足与对数相关的函数方程。我们通过证明 g_ℓ 在 $\mathbb{R}_{>0}$ 上连续，来说明它是对数的倍数。令 x、$y > 0$ 和 $t \in (0, 1)$。利用以下事实：

$$\begin{pmatrix} t & \cdots & t & tx & t & \cdots & t \\ 1-t & \cdots & 1-t & (1-t)y & 1-t & \cdots & 1-t \end{pmatrix} \tag{105}$$

$$\geq \big(1 \cdots 1\, tx + (1-t)y\, 1 \cdots 1\big), \tag{106}$$

其中 x 和 y 出现在第 ℓ 列，应用 Δ，根据 δ 的单调性和导子性质可得，

$$t g_\ell(x) + (1-t) g_\ell(y) \geqslant g_\ell\big(tx + (1-t)y\big), \tag{107}$$

其中左侧由公式（102）得到。

因此， g_ℓ 在其定义域 $\mathbb{R}_{>0}$上是凸的，因而连续。这意味着存在 γ_2，……，$\gamma_d \in \mathbb{R}$ 使得

$$g_\ell = -\gamma_\ell \log \tag{108}$$

对每个 $\ell = 2, \ldots, d$。现在我们旨在证明 $\gamma_\ell \geqslant 0$。对所有 $t \in (0, \frac{1}{2}]$，定义矩阵

$$P_t = \begin{pmatrix} \frac{1}{2} & \cdots & \frac{1}{2} & 1-t & \frac{1}{2} & \cdots & \frac{1}{2} \\ \frac{1}{2} & \cdots & \frac{1}{2} & t & \frac{1}{2} & \cdots & \frac{1}{2} \end{pmatrix}, \tag{109}$$

其中包含 t 的元素位于第 ℓ 列，我们注意到函数 n

$$\Delta(P_t) \overset{(102)}{=} \frac{1}{2}\gamma_\ell \log \frac{1}{4t(1-t)} \tag{110}$$

应当在 $t \in (0, \frac{1}{2}]$ 上非递增，这意味着 $\gamma_\ell \geqslant 0$。再次由（102），可知每个单调导子在第一列具有有理数 1-范数的所有矩阵上都必定具有所声称的形式。

最后，我们来证明映射 $\Delta^{(1)}_{\underline{\gamma}}$ 在 f_{e_1} 处是单调导子，当且仅当 $\gamma_2, \ldots, \gamma_d \geqslant 0$。如前所述，直接计算即可得出它是导子。为证明单调性，任取 n，令 $P \in \mathcal{V}_n^d$，不失一般性地假设其具有全支撑，并令 $T = (T_{i,j})$ 为一个没有零行的 $(m \times n)$-随机矩阵。令 $Q := TP \in \mathcal{V}_m^d$，由函数 $x \mapsto \log \frac{1}{x}$ 的凸性可得，

$$\Delta^{(1)}_{\underline{\gamma}}(P) = \sum_{j=1}^{n}\sum_{\ell=2}^{d} \gamma_\ell p_j^{(1)} \log \frac{p_j^{(1)}}{p_j^{(\ell)}} = \sum_{i=1}^{m}\sum_{\ell=2}^{d}\sum_{j=1}^{n} \gamma_\ell q_i^{(1)} \frac{T_{i,j}p_j^{(1)}}{q_i^{(1)}} \log \frac{p_j^{(1)}}{p_j^{(\ell)}} \tag{111}$$

$$\geqslant \sum_{i=1}^{m}\sum_{\ell=2}^{d} \gamma_\ell q_i^{(1)} \log \frac{q_i^{(1)}}{\sum_{j=1}^{n} T_{i,j}p_j^{(\ell)}} = \sum_{i=1}^{m}\sum_{\ell=1}^{d} \gamma_\ell q_i^{(1)} \log \frac{q_i^{(1)}}{q_i^{(\ell)}} \tag{112}$$

$$= \Delta^{(1)}_{\underline{\gamma}}(Q), \tag{113}$$

这表明 $\Delta^{(1)}_{\underline{\gamma}}\colon\ S^d \to \mathbb{R}_+$ 确实是单调的。 □

3.2. 大样本和催化矩阵主序的充分条件。 我们的目标现在是应用定理 8 来 S^d，以获得大样本和催化矩阵主序的充分条件以及一般必要条件。在本小节中以及下文中，我们记

$$\alpha_{\max} := \max_{k=1,\ldots,d} \alpha_k, \qquad \alpha_{\min} := \min_{k=1,\ldots,d} \alpha_k.$$

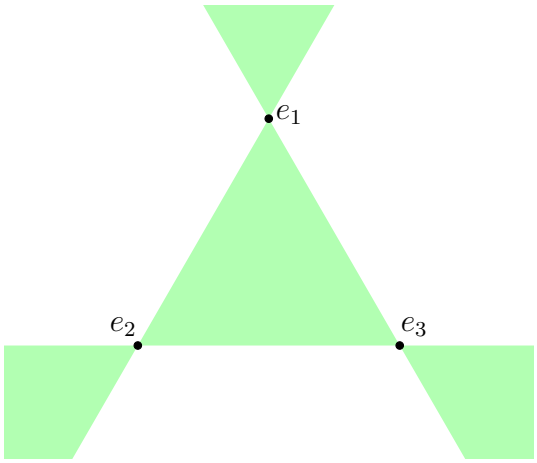

FIGURE 1. 参数范围 $(A_- \cup A_+)\backslash\{e_1, \ldots, e_d\}$ 是定义在 S^3 上的非退化单调同态 $f_{\underline{\alpha}}$ 的参数范围；图中以绿色标出，它是二维仿射空间 A 的一个子集。集合 A_+ 是三个结果上的概率单纯形（中心的三角形区域）。构成 A_- 的其他区域是从各顶点延伸出的锥；这些顶点对应于退化同态 f_{e_k}。

那么，对于每个 $\underline{\alpha} \in (A_+ \cup A_-)\backslash\{e_1、\ldots\ldots、e_d\}$ 和 $\underline{\beta} \in B_-\backslash\{0\}$，我们定义任意 $P \in \mathcal{V}^d_{<\infty}$ 上的相应矩阵$\underline{\alpha}$-散度如下：

$$\begin{aligned} D_{\underline{\alpha}}(P) &:= \frac{1}{\alpha_{\max} - 1} \log f_{\underline{\alpha}}(P) = \frac{1}{\alpha_{\max} - 1} \log \sum_i \prod_{k=1}^{d} \left(p_i^{(k)}\right)^{\alpha_k}, \\ D^{\mathbb{T}}_{\underline{\beta}}(P) &:= \frac{1}{\beta_{\max}} \log f^{\mathbb{T}}_{\underline{\beta}}(P) = \frac{1}{\beta_{\max}} \log \max_i \prod_{k=1}^{d} \left(p_i^{(k)}\right)^{\beta_k}, \end{aligned} \tag{114}$$

其中 i 遍历 P 的非零行。我们作出这些定义，主要是因为与 $f_{\underline{\alpha}}$ 不同，$D_{\underline{\alpha}}$ 作为映射 $S^d \to \mathbb{R}$ 是单调的，其中 $\mathbb{R}$ 配备了标准序。我们很快就会看到，这些定义还使我们能够将热带映射和导子写成 $D_{\underline{\alpha}}$ 的极限。

注15。 （114）中的归一化方式可以有不同选择，也值得考虑。例如，一个自然的选择是采用 [13], 中的对数求值映射；它适用于任何具有固定幂泛元的预序半环。在当前情形下，根据引理 12这相当于选择一个 U 具有全支撑的矩阵，并使所有列都经过归一化且彼此不同。于是，对数求值映射定义为

$$P \longmapsto \frac{\log \Phi(P)}{\log \Phi(U)} \tag{115}$$

对每个非退化单调同态 Φ 均成立。

例如，一个不错的选择是令 U 为对角线元素均为 $\frac{2}{d+1}$、且 $\frac{1}{d+1}$ 非对角线元素的矩阵。采用这一选择，可得到与（114）中相同的 $D_{\underline{\alpha}}$

$$\log\left(\frac{1}{d+1}\sum_k 2^{\alpha_k}\right) \tag{116}$$

以 $\alpha_{\max}-1$ 代替。尽管这样做的优点是它是 $\underline{\alpha}$ 的光滑函数，且在后续证明中无需进行任何分类讨论，但与此同时，计算起来更加繁琐，因此我们仍采用（114）作为首选的归一化方式。

注 16. 对于每个 $\lambda>0$，归一化（114）满足

$$D^{\mathbb{T}}_{\lambda\underline{\beta}} = D^{\mathbb{T}}_{\underline{\beta}}. \tag{117}$$

因此，只需对 $\underline{\beta}$ 进行归一化，例如使得 $\beta_{\max}=1$，由此即可实现热带同态的归一化（注 6）。

命题 17. 令 $\underline{\beta}\in B_-$ 满足 $\beta_k=1$，因而对所有 $\ell\neq k$ 都有 $\beta_\ell\leqslant 0$。对于 $\lambda\geqslant 0$，定义

$$\underline{\alpha}^\lambda := e_k+(\lambda-1)\underline{\beta} \tag{118}$$

$$= \big((\lambda-1)\beta_1,\ldots,(\lambda-1)\beta_{k-1},\lambda,(\lambda-1)\beta_{k+1},\ldots,(\lambda-1)\beta_d\big). \tag{119}$$

因此：

(i) $\underline{\alpha}^\lambda\in(A_+\cup A_-)\backslash\{e_1,\ldots,e_d\}$ *for all* $\lambda\in(0,1)\cup(1,\infty)$.

(ii)对任意具有全支撑且列已归一化的 P，有

$$D^{\mathbb{T}}_{\underline{\beta}}(P) = \lim_{\lambda\to\infty} D_{\underline{\alpha}^\lambda}(P), \tag{120}$$

$$\Delta^{(k)}_{-\underline{\beta}}(P) = \lim_{\lambda\to 1} D_{\underline{\alpha}^\lambda}(P). \tag{121}$$

(iii) 对于这样的 P，函数 $\lambda\mapsto D_{\underline{\alpha}^\lambda}$ 可连续延拓至 $[0、\infty]$，且当 $\lambda\geqslant\frac{\beta_{\min}}{\beta_{\min}-1}$ 时，该延拓单调不减。

从图 1来看，轨迹 $\lambda\mapsto\underline{\alpha}^\lambda$ 从概率单纯形 A_+ 的面上 $\lambda=0$ 出发，该面由 $\alpha_k=0$定义；轨迹经过顶点 e_k，该顶点位于$\lambda=1$，随后沿方向 $\underline{\beta}$ 继续在 A_- 中前进。作为 $\lambda\geqslant 0$的函数，$D_{\underline{\alpha}^\lambda}(P)$ 在 $\underline{\alpha}^\lambda$ 越过从 A_+的中心连至 A_+ 上离轨迹最近的边中点的直线后，单调不减。

证明。对于第 (i) 项，容易验证 $\underline{\alpha}^\lambda\in A_+$ 当 $\lambda\in[0,1]$ 时，以及 $\underline{\alpha}^\lambda\in A_-$ 当 $\lambda\in[1,\infty)$ 时。

接下来证明第 (ii) 项中关于极限的结论。两个公式都可通过如下直接计算得到。对于 (120)，回想一下，对于所有实数$r_1,\dots,r_m>0$和 $s_1,\dots,s_m>0$，我们有[4]

$$\lim_{\lambda\to\infty}\frac{1}{\lambda-1}\log\sum_i r_i s_i^{\lambda}=\log\max_i s_i. \tag{122}$$

这直接意味着

$$\lim_{\lambda\to\infty}D_{\underline{\alpha}^{\lambda}}(P)=\lim_{\lambda\to\infty}\frac{1}{\lambda-1}\log\sum_i\left[\prod_{\ell\neq k}\left(p_i^{(\ell)}\right)^{-\beta_\ell}\right]\left[p_i^{(k)}\prod_{\ell\neq k}\left(p_i^{(\ell)}\right)^{\beta_\ell}\right]^{\lambda} \tag{123}$$

$$=\log\max_i p_i^{(k)}\prod_{\ell\neq k}\left(p_i^{(\ell)}\right)^{\beta_\ell}=D^{\mathbb{T}}_{\underline{\beta}}(P), \tag{124}$$

正如要证明的那样。对于（121），使用与（122）中相同的数，并假设 $\sum_i r_i=1$ 此外，我们类似地注意到

$$\lim_{\lambda\to1}\frac{1}{\lambda-1}\log\sum_i r_i s_i^{\lambda-1}=\sum_i r_i\log s_i, \tag{125}$$

由洛必达法则直接应用可得。代入后得到

$$\lim_{\lambda\to1}D_{\underline{\alpha}^{\lambda}}(P)=\lim_{\lambda\to1}\frac{1}{\lambda-1}\log\sum_i p_i^{(k)}\prod_{\ell=1}^{d}\left(p_i^{(\ell)}\right)^{\beta_\ell(\lambda-1)} \tag{126}$$

$$=\sum_i p_i^{(k)}\log\prod_{\ell=1}^{d}\left(p_i^{(\ell)}\right)^{\beta_\ell} \tag{127}$$

$$=-\sum_i\sum_{\ell\neq k}\beta_\ell p_i^{(k)}\log\frac{p_i^{(k)}}{p_i^{(\ell)}}=\Delta^{(k)}_{-\underline{\beta}}(P), \tag{128}$$

正如要证明的那样。

对于第（iii）项，刚刚证明的极限存在性表明，我们可以作连续延拓，使其在 $\lambda=1$ 处取值 $\Delta^{(k)}_{-\underline{\beta}}(P)$，在 $\lambda=\infty$ 处取值 $D^{\mathbb{T}}_{\underline{\beta}}$。关于单调性，令 $\frac{\beta_{\min}}{\beta_{\min}-1}\leqslant\lambda\leqslant\lambda'\leqslant\infty$。不妨假设$\lambda$、$\lambda'\neq1$、$\infty$，并证明单调性

$$D_{\underline{\alpha}^{\lambda}}(P)\leqslant D_{\underline{\alpha}^{\lambda'}}(P) \tag{129}$$

然后，由于其余情况可通过连续性得出。对于 $z:=\frac{\lambda-1}{\lambda'-1}$，我们区分三种情况：

- $\lambda\leqslant\lambda'<1$：现在 $z\geqslant1$，因此 $x\mapsto x^z$ 是凸的。
- $\lambda<1<\lambda'$：现在 $z<0$，因此 $x\mapsto x^z$ 又是凸的。

[4]这类公式在热带数学和马斯洛夫去量子化中经常出现，读者可以轻松推导出这个直接的证明。

- $1 < \lambda \leqslant \lambda'$：现在 $0 < z \leqslant 1$，因此 $x \mapsto x^z$ 是凹的。

考虑以上所有情形并将分母的符号考虑在内，根据詹森不等式，

$$\frac{1}{\lambda-1}\log\sum_i r_i s_i^z \leqslant \frac{z}{\lambda-1}\log\sum_i r_i s_i = \frac{1}{\lambda'-1}\log\sum_i r_i s_i \tag{130}$$

对所有概率向量 r 和所有正实数 s_i 均成立。因此，我们有

$$D_{\underline{\alpha}^\lambda}(P) = \frac{1}{\lambda-1}\log\sum_i\left(p_i^{(k)}\right)^\lambda\prod_{\ell\neq k}\left(p_i^{(\ell)}\right)^{(\lambda-1)\beta_\ell} \tag{131}$$

$$= \frac{1}{\lambda-1}\log\sum_i p_i^{(k)}\prod_{\ell\neq k}\left(\frac{p_i^{(\ell)}}{p_i^{(k)}}\right)^{(\lambda-1)\beta_\ell} \tag{132}$$

$$= \frac{1}{\lambda-1}\log\sum_i p_i^{(k)}\left[\prod_{\ell\neq k}\left(\frac{p_i^{(\ell)}}{p_i^{(k)}}\right)^{(\lambda'-1)\beta_\ell}\right]^z \tag{133}$$

$$\leqslant \frac{1}{\lambda'-1}\log\sum_i p_i^{(k)}\prod_{\ell\neq k}\left(\frac{p_i^{(\ell)}}{p_i^{(k)}}\right)^{(\lambda'-1)\beta_\ell} = D_{\underline{\alpha}^{\lambda'}}(P), \tag{134}$$

这正是要证明的结论。 □

因此，所有相关单调量的集合可如图 2 所示。

命题 18。设 $P \in \mathcal{V}_n^d$ 为随机的。那么：

(i) 对于 $\underline{\alpha} \in (A_+ \cup A_-)\backslash\{e_1,\ldots,e_d\}$，有 $D_{\underline{\alpha}}(P) \geqslant 0$。当且仅当所有满足 $\alpha_k \neq 0$ 的列 $p^{(k)}$ 都相等时，等号成立。(ii) 对于 $\underline{\beta} \in B_-\backslash\{0\}$，有 $D^{\mathbb{T}}_{\underline{\beta}}(P) \geqslant 0$。当且仅当所有满足 $\beta_k \neq 0$ 的列 $p^{(k)}$ 都相等时，等号成立。

证明。我们针对情形 (i) 给出证明，并指出情形 (ii) 完全类似。等价地，我们证明 $f_{\underline{\alpha}}(P) \geqslant 1$，当且仅当所有满足 $\alpha_k \neq 0$ 的列 $p^{(k)}$ 都相等时，等号成立。为此，注意到由单调性以及

$$P \geq (1\ldots1), \tag{135}$$

根据 P 是随机的这一假设，该结论成立。为了确定等号成立的条件，先假设 P 的任意两列都不相等。那么 P 在 S^d 中是幂通用的，由引理 12 可知。因此，如果 $f_{\underline{\alpha}}(P) = 1$ 成立，那么 $f_{\underline{\alpha}}$ 就会是退化的，而根据命题 13，我们知道事实并非如此。至此，P 的各列互不相同的情形证明完毕。

如果 P 有若干列相同，那么不失一般性，我们假设 P 的最后两列相同。我们将问题归结为规模更小的 d情形，从而可以通过对

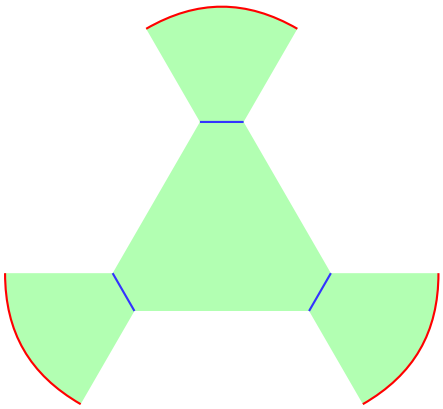

FIGURE 2. 预序半环上的相关单调量 S^3按照测试谱的[13] 示意图如下。绿色部分是矩阵$\underline{\alpha}$-散度 $D_{\underline{\alpha}}$ ，对应 $\underline{\alpha} \in (A_+ \cup A_-)\backslash\{e_1$、……、$e_d\}$，见（114）。红色部分是热带量$D^{\mathbb{T}}_{\underline{\beta}}$，对应 $\underline{\beta} \in B_-\backslash\{0\}$，按备注16进行缩放。蓝色部分是导出量 $\Delta^{(k)}_{\underline{\gamma}}$，对应非零的 $\underline{\gamma}$，同样按比例缩放。拓扑结构对应命题17中的极限。与图1相比，红色的热带部分可以看作无穷远处的点，由方向$\underline{\beta} \in B_-\backslash\{0\}$指定，可作为一种紧化，类似于射影平面中的无穷远点。蓝色的导子部分则将图1 中的每个顶点“膨胀”为一条直线。

d的归纳法得出整体证明。将所讨论矩阵的最后一列复制一份，定义了一个单调半环同态 $S^{d-1} \to S^d$，因此我们可以将 $S^{d-1} \subseteq S^d$ 视为具有诱导预序的子半环。将矩阵散度 $D_{\underline{\alpha}}$ 限制到这个子半环上，便定义了一个类型相同、但参数元组为新的矩阵散度

$$\underline{\alpha}' := (\alpha_1, \ldots, \alpha_{d-2}, \alpha_{d-1} + \alpha_d)$$

代替 $\underline{\alpha}$。若 $\underline{\alpha}$ 的支撑位于最后两列上，则限制后的同态是退化的，待证的 $f_\alpha(P) = 1$ 由（135）得出。否则，限制后的同态仍然是非退化的，因为 $\underline{\alpha}'$ 属于 $A_+ \cup A_-$，但仍不是标准基向量。由于 P 属于子半环 S^{d-1}，我们已完成将问题归约到 $d-1$。 □

现在我们将这些结果综合起来，并对 8 应用定理 S^d。除了备注 20中讨论的一些细节外，这为[21]第 6 节和附录 K 中提出的一个猜想给出了肯定回答。

定理 19. 设 $P = (p^{(1)}, \ldots, p^{(d)})$ 和 $Q = (q^{(1)}, \ldots, q^{(d)})$ 是由概率向量组成的 d 元组，且每个元组中的向量具有一致支撑。就

来自 (114) 的矩阵散度和 Kullback-Leibler 散度 $D_1(\cdot\|\cdot)$而言，假设

$$D_{\underline{\alpha}}(P) > D_{\underline{\alpha}}(Q), \tag{136}$$

$$D^{\mathbb{T}}_{\underline{\beta}}(P) > D^{\mathbb{T}}_{\underline{\beta}}(Q), \tag{137}$$

$$D_1\big(p^{(k)}\|p^{(\ell)}\big) > D_1\big(q^{(k)}\|q^{(\ell)}\big) \tag{138}$$

对所有 $\underline{\alpha}\in(A_+\cup A_-)\backslash\{e_1,\ \ldots e_d\}$，$\underline{\beta}\in B_-\backslash\{0\}$ 以及 $k\neq\ell$，其中参数集如 (41)–(44)所示。于是，对所有足够大的 $n\in\mathbb{N}$，存在一个随机矩阵 T_n 使得

$$(q^{(k)})^{\otimes n} = T_n(p^{(k)})^{\otimes n} \qquad \forall k=1,\ldots,d. \tag{139}$$

反之，如果对于某个 $n\geqslant 1$ 存在这样的 T_n，那么上述不等式以非严格形式成立。

根据备注16，在（137）中还可以施加诸如 $\beta_{\max}=1$ 这样的归一化条件。另请注意，条件（136）–（138）实际上构成一个连续的不等式族，如图2所示。

证明。我们已经指出，定理 8 的辅助假设均已满足。因此，尚需证明假设中的不等式与由此得到的不等式相符，并且 $P=(p^{(1)},\ \ldots\ldots,\ p^{(d)})$是幂通用的。后者成立，是因为所假设的（138）蕴含$p^{(k)}\neq p^{(\ell)}$ 对 $k\neq\ell$成立，因此根据引理 12，P 是幂通用的。

根据命题 13，非退化单调同态 $S^d\to\mathbb{R}_+$恰好是 $f_{\underline{\alpha}}$对于 $\alpha\in A_-\backslash\{e_1,\ \ldots\ldots,\ e_d\}$。但由于

$$f_{\underline{\alpha}}(P) > f_{\underline{\alpha}}(Q) \qquad\Longleftrightarrow\qquad D_{\underline{\alpha}}(P) > D_{\underline{\alpha}}(Q) \tag{140}$$

对于这样的 $\underline{\alpha}$，由 $\alpha_{\max}>1$ 可知，当 $\mathbb{K}=\mathbb{R}_+$ 时，不等式确实吻合。类似地，对于 $\mathbb{K}=\mathbb{R}_+^{\mathrm{op}}$，相关量是带有 $\underline{\alpha}\in A_+$ 的 $f_{\underline{\alpha}}$。因此$\alpha_{\max}<1$，我们有

$$f_{\underline{\alpha}}(P) < f_{\underline{\alpha}}(Q) \qquad\Longleftrightarrow\qquad D_{\underline{\alpha}}(P) > D_{\underline{\alpha}}(Q), \tag{141}$$

并且不等式再次吻合。对于 $\mathbb{K}=\mathbb{TR}_+$，显然有

$$f^{\mathbb{T}}_{\underline{\beta}}(P) > f^{\mathbb{T}}_{\underline{\beta}}(Q) \qquad\Longleftrightarrow\qquad D^{\mathbb{T}}_{\underline{\beta}}(P) > D^{\mathbb{T}}_{\underline{\beta}}(Q), \tag{142}$$

这也就完成了这一情形 . 情形 $\mathbb{K}=\mathbb{TR}_+^{\mathrm{op}}$ 根据命题1，为空3。接下来考虑在 f_{e_k}对每个 $k=1,\ldots,d$ 上的单调导子。根据命题 14，这些导子都具有如下形式：$\Delta^{(k)}_{\underline{\gamma}}$对于 $\gamma\in\mathbb{R}^d_+$。由于$\Delta^{(k)}_{\underline{\gamma}}$关于 $\underline{\gamma}$可加，因此只需假设极端 $\underline{\gamma}$满足不等式；这些极端值恰好是标准基向量。此外，导子 $\Delta^{(k)}_{\underline{\gamma}}$非零，当且仅当 $\gamma_\ell\neq 0$ 对某个 $\ell\neq k$成立。综合以上陈述，我们得出只需考虑满足以下条件的导子：$\Delta^{(k)}_{e_\ell}$其中 $\ell\neq k$，这恰好得到（138）。□

备注 20。将 Mu 等人在 [21] 的在线附录中为其猜想所作的假设转化到我们的框架下，可以发现充分条件存在以下差异：对于热带矩阵散度 $D^{\mathbb{T}}_{\underline{\beta}}$，他们仅对 $\underline{\beta} = e_k - e_\ell$（当 $k, \ell = 1, \ldots, d$时） $k \neq \ell$假设严格不等式（这是他们的一般性条件），而在其他情况下则假设非严格不等式（根据我们的命题 17，这一点是隐含的）。换言之，我们关于大样本中矩阵主序的充分条件（137）要求对所有 $D^{\mathbb{T}}_{\underline{\beta}}$ 和所有 $\underline{\beta} \in B_-\backslash\{0\}$,

$$D_\infty\big(p^{(k)} \| p^{(\ell)}\big) > D_\infty\big(q^{(k)} \| q^{(\ell)}\big) \tag{143}$$

对于 k、 $\ell \in \{1, \ldots, d\}$、 $k \neq \ell$，在 [21] 中均成立。我们不知道是否存在满足他们的假设但不满足我们的假设的任何 P、 $Q \in \mathcal{V}^d_{<\infty}$。

注 21。还请注意，在定理 19 由（136）、（137）和（138）给出的假设下，根据定理 8 的（b）项，我们还得到以下催化结果：对每个 $n \in \mathbb{N}$ 足够大的值，定义概率向量

$$r^{(k)} := \frac{1}{n} \bigoplus_{s=0}^{n-1} \big(p^{(k)}\big)^{\otimes(n-1-s)} \otimes \big(q^{(k)}\big)^{\otimes s}, \tag{144}$$

对于 $k = 1, \ldots, d$，存在一个随机映射 T_n，使得 $T_n(p^{(k)} \otimes r^{(k)}) = q^{(k)} \otimes r^{(k)}$ 对所有 $k = 1, \ldots, d$ 均成立。

现在，我们已经给出了催化矩阵主序的充分条件和一般而言必要的条件。我们注意到，作为推论，使用相同的预序，我们还可以得到渐近催化矩阵主序的充分条件。

定理22. 令 $n \in \mathbb{N}$，并令 $\mathcal{P}_n$ 为具有 n个分量的概率向量集合。令 $P = (p^{(1)}, \ldots, p^{(d)}) \in \mathcal{P}_n^d$和 $Q = (q^{(1)}, \ldots, q^{(d)}) \in \mathcal{P}_n^d$为具有全支撑的 d元概率向量组，且对所有 $k \neq \ell$，均有 $p^{(k)} \neq p^{(\ell)}$。那么以下条件等价：

(i) 对每个 $\varepsilon > 0$，都存在 $(q_\varepsilon^{(1)}, \ldots, q_\varepsilon^{(d)}) \in \mathcal{P}_n^d$和 $(r_\varepsilon^{(1)}, \ldots, r_\varepsilon^{(d)}) \in \mathcal{P}_n^d$，以及一个随机矩阵 T_ε，使得

$$\|q^{(k)} - q_\varepsilon^{(k)}\|_1 \leqslant \varepsilon \qquad \forall k = 1, \ldots, d, \tag{145}$$

and

$$T_\varepsilon(p^{(k)} \otimes r_\varepsilon^{(k)}) = q_\varepsilon^{(k)} \otimes r_\varepsilon^{(k)}, \qquad \forall k = 1, \ldots, d. \tag{146}$$

(ii) 对所有 $\underline{\alpha} \in (A_- \cup A_+)\backslash\{e_1, \ldots, e_d\}$，都有

$$D_{\underline{\alpha}}(P) \geqslant D_{\underline{\alpha}}(Q). \tag{147}$$

此外，对于一个 $k \in \{1, \ldots, d\}$，对于所有 $\varepsilon > 0$，我们都可以在 (i) 中取 $q_\varepsilon^{(k)} = q^{(k)}$。

证明。在整个证明过程中，我们将 $P = (p^{(1)}, \dots, p^{(d)})$ 记作其他元组的类似形式。首先假设该命题的条件 (ii) 成立。对于任意概率向量 $w \in \mathcal{P}_n$，我们定义 $n \times d$ 噪声矩阵

$$W = (w, \dots, w). \tag{148}$$

定义 $Q_\varepsilon := (1 - \frac{\varepsilon}{2})Q + \frac{\varepsilon}{2}W$对所有 $\varepsilon \in (0, 2]$均成立。将 Q_ε的列记为$q_\varepsilon^{(k)}$，于是有

$$\|q^{(k)} - q_\varepsilon^{(k)}\|_1 = \left\|q^{(k)} - \left(1 - \frac{\varepsilon}{2}\right)q^{(k)} - \frac{\varepsilon}{2}w\right\|_1 \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \tag{149}$$

对每个 k均成立。取 $w := q^{(k)}$，其中固定 $k \in \{1,$ 在以下证明中始终令 $\dots, d\}$，即可得到最后的结论。

我们旨在证明 Q_ε 和 P 严格满足所有相关不等式（136）–（138）。给定任意 $\underline{\alpha} \in A_+ \backslash \{e_1, \dots, e_d\}$，我们计算

$$\begin{aligned}
D_{\underline{\alpha}}(Q_\varepsilon) &= \frac{1}{\alpha_{\max} - 1} \log \underbrace{f_{\underline{\alpha}}\left(\left(1 - \frac{\varepsilon}{2}\right)Q + \frac{\varepsilon}{2}W\right)}_{\geqslant \left(1 - \frac{\varepsilon}{2}\right) f_{\underline{\alpha}}(Q) + \frac{\varepsilon}{2} f_{\underline{\alpha}}(W)} && (150)\\
&\leqslant \frac{1}{\alpha_{\max} - 1} \log \left[\left(1 - \frac{\varepsilon}{2}\right) f_{\underline{\alpha}}(Q) + \frac{\varepsilon}{2} f_{\underline{\alpha}}(W)\right] && (151)\\
&\leqslant \frac{1}{\alpha_{\max} - 1} \left[\left(1 - \frac{\varepsilon}{2}\right) \log f_{\underline{\alpha}}(Q) + \frac{\varepsilon}{2} \log f_{\underline{\alpha}}(W)\right] && (152)\\
&= \left(1 - \frac{\varepsilon}{2}\right) D_{\underline{\alpha}}(Q) + \frac{\varepsilon}{2} \underbrace{D_{\underline{\alpha}}(W)}_{=0} \leqslant D_{\underline{\alpha}}(Q) \leqslant D_{\underline{\alpha}}(P), && (153)
\end{aligned}$$

这里用到了 $f_{\underline{\alpha}}$ 的凹性（参见命题 13的证明）以及 $(\alpha_{\max} - 1)^{-1} \log$ 在第一个不等式中单调不增这一事实，还用到了 $(\alpha_{\max} - 1)^{-1}\log$ 在第二个不等式中的凸性，以及最后一个不等式中的假设。如果 $D_{\underline{\alpha}}(Q) = 0$，则根据$D_{\underline{\alpha}}(P) > 0$，由关于 P 的假设和命题 18，最后一个不等式为严格不等式。如果 $D_{\underline{\alpha}}(Q) > 0$，则倒数第二个不等式为严格不等式。因此，$D_{\underline{\alpha}}(Q_\varepsilon) < D_{\underline{\alpha}}(P)$ 在所有情况下均成立。同样，对于 $\underline{\alpha} \in A_- \backslash \{e_1, \dots, e_d\}$，我们可以用基本相同的论证证明 $D_{\underline{\alpha}}(Q_\varepsilon) < D_{\underline{\alpha}}(P)$ 也成立。

对于热带情形，令 $\underline{\beta} \in B_- \backslash \{0\}$，我们将证明 $D^{\mathbb{T}}_{\underline{\beta}}(Q_\varepsilon) < D^{\mathbb{T}}_{\underline{\beta}}(P)$。根据备注 16，我们可以假设 $\beta_{\max} = 1$。于是，例如由前一种情形和命题 17，可得

$$D^{\mathbb{T}}_{\underline{\beta}}(Q_\varepsilon) \leqslant D^{\mathbb{T}}_{\underline{\beta}}(Q), \tag{154}$$

或者等价地

$$\max_i \prod_{k=1}^{d} \left(\left(1 - \frac{\varepsilon}{2}\right) q_i^{(k)} + \frac{\varepsilon}{2} w_i\right)^{\beta_k} \leqslant \max_i \prod_{k=1}^{d} \left(q_i^{(k)}\right)^{\beta_k} \tag{155}$$

对每个 $\varepsilon \in (0,2]$ 都成立。左侧关于 ε 非递增，这是因为当 $\varepsilon \leqslant \varepsilon'$ 时有 $Q_\varepsilon \geq Q_{\varepsilon'}$，并且当 $\varepsilon \to 0$ 时与右侧相等。现在，如果对于某些 $\varepsilon \in (0,2]$，这个不等式是非严格的，那么左侧必须关于 ε 一直保持常数，直至 $\varepsilon \to 0$。由于左侧是有限个实解析函数的逐点最大值，因此在所有这些 ε 上，最大值必定由同一个行索引 i 取得，并且乘法对于这个 i 恒为常数。再由实解析性可知，这意味着左侧在整个范围 $[0,2]$ 内都是常数。因此，如果这个不等式是非严格的，那么我们必定已经有 $D_{\underline{\beta}}^{\mathbb{T}}(Q) = D_{\underline{\beta}}^{\mathbb{T}}(W) = 0$。

因此，如果 $D_{\underline{\beta}}^{\mathbb{T}}(Q) > 0$，那么我们得到 $D_{\underline{\beta}}^{\mathbb{T}}(Q_\varepsilon) < D_{\underline{\beta}}^{\mathbb{T}}(Q)$。而如果 $D_{\underline{\beta}}^{\mathbb{T}}(Q) = 0$，那么根据命题 $D_{\underline{\beta}}^{\mathbb{T}}(Q) < D_{\underline{\beta}}^{\mathbb{T}}(P)$，我们有 18。在这两种情况下，我们都得到

$$D_{\underline{\beta}}^{\mathbb{T}}(Q_\varepsilon) \leqslant D_{\underline{\beta}}^{\mathbb{T}}(Q) \leqslant D_{\underline{\beta}}^{\mathbb{T}}(P), \tag{156}$$

其中至少一个不等式是严格的，因此 $D_{\underline{\beta}}^{\mathbb{T}}(Q_\varepsilon) < D_{\underline{\beta}}^{\mathbb{T}}(P)$。对于 Kullback-Leibler 散度（138），固定互不相同的列指标 k，$\ell \in \{1, \ldots\ldots, d\}$。由 D_1 的凸性，我们得到[5]

$$\begin{aligned} D_1\big(q_\varepsilon^{(k)} \| q_\varepsilon^{(\ell)}\big) &\leqslant \Big(1 - \frac{\varepsilon}{2}\Big) D_1(q^{(k)} \| q^{(\ell)}) + \frac{\varepsilon}{2} \underbrace{D_1(w \| w)}_{=0} \\ &\leqslant D_1(q^{(k)} \| q^{(\ell)}) \leqslant D_1(p^{(k)} \| p^{(\ell)}). \end{aligned} \tag{157}$$

和前面一样，如果 $D_1(q^{(k)} \| q^{(\ell)}) > 0$，那么上面的倒数第二个不等式是严格的；而如果 $D_1(q^{(k)} \| q^{(\ell)}) = 0$，那么最后一个不等式是严格的。总之，我们也得到 $D_1(q_\varepsilon^{(k)} \| q_\varepsilon^{(\ell)}) < D_1(p^{(k)} \| p^{(\ell)})$。

因此，定理 19 的条件对 Q_ε（任意 $\varepsilon \in (0、 1]$）和 P均成立，因此我们得出，存在一个随机映射 T_ε 以及一个 d元概率向量组 $r_\varepsilon^{(k)} \in \mathcal{P}_n$，使得 $T_\varepsilon(p^{(k)} \otimes r_\varepsilon^{(k)}) = q_\varepsilon^{(k)} \otimes r_\varepsilon^{(k)}$对所有 $k = 1, \ldots\ldots, d$（注 21）。

˜对于逆向推导，假设我们有如陈述所述的 Q 。接着，将 $D_{\underline{\alpha}}$ 应用于（146），并结合单调性与可加性˜，得到 $D_{\underline{\alpha}}(P) \geqslant D_{\underline{\alpha}}(Q)$ 对所有 $\underline{\alpha} \in A_- \cup A_+$均成立。但由 $D_{\underline{\alpha}}(P) \geqslant D_{\underline{\alpha}}(Q)$在 $D_{\underline{\alpha}}$ 上的连续性，还可得 $\mathbb{R}_{>0}^{n \times d}$。 □

3.3. 情形 $d = 2$：相对主序。我们来快速看一下 $d = 2$ 的情形，也就是说，矩阵退化为二分布 $(p^{(1)}、 p^{(2)}) \in \mathcal{V}_{<\infty}^2$ 的情形。在此设定下，矩阵主序也称为相对主序。

在二分法 $(p, q) \in \mathcal{V}_{<\infty}^2$ 中，要求向量 p 和 q 的支撑集相同。不过，我们暂且去掉这一假设，只考虑一对具有非负元素的向量 (p, q)。每当共同考察两个这样的向量 $p = (p_1, \ldots, p_n)$ 和 $q = (q_1, \ldots, q_n)$ 时，

[5]这种凸性也可以由 $D_{\underline{\alpha}}$的凸性得出，它是（70）的一个实例，并将$\Delta_{e_k - e_\ell}$ 写成 $D_{\underline{\alpha}}$ 的逐点极限，如命题17所述。

我们都假设已通过补零使它们具有相同的长度 n。鉴于此，并定义 $0\log 0:=0$，我们来为所有有限概率向量 p 和 q 定义 Rényi 散度：

$$D_\alpha(p\|q)=\begin{cases}-\log\sum_{i\in\operatorname{supp}p}q_i & \text{if } \alpha=0 \text{ and } \operatorname{supp}p\cap\operatorname{supp}q\neq\varnothing,\\ \frac{1}{\alpha-1}\log\sum_{i=1}^n p_i^\alpha q_i^{1-\alpha} & \text{if } \alpha\in(0,1) \text{ and } \operatorname{supp}p\cap\operatorname{supp}q\neq\varnothing,\\ \sum_{i=1}^n p_i\log\frac{p_i}{q_i} & \text{if } \alpha=1 \text{ and } \operatorname{supp}p\subseteq\operatorname{supp}q,\\ \frac{1}{\alpha-1}\log\sum_{i=1}^n p_i^\alpha q_i^{1-\alpha} & \text{if } \alpha\in(1,\infty) \text{ and } \operatorname{supp}p\subseteq\operatorname{supp}q,\\ \max_{i\in\operatorname{supp}q}\log\frac{p_i}{q_i} & \text{if } \alpha=\infty \text{ and } \operatorname{supp}p\subseteq\operatorname{supp}q,\\ \infty & \text{otherwise,}\end{cases} \tag{158}$$

对所有 R´恩尼参数值 $\alpha\geqslant 0$。注意，D_1 是 Kullback-Leibler 散度。由定理 19，我们得到主结果的一种变体，见 [21]，即如下结论：

推论23. 设 $p^{(1)}$ 和 $p^{(2)}$ 为支撑集相同的概率向量，$q^{(1)}$ 和 $q^{(2)}$也同样如此。若

$$D_\alpha\big(p^{(1)}\|p^{(2)}\big)>D_\alpha\big(q^{(1)}\|q^{(2)}\big), \tag{159}$$

$$D_\alpha\big(p^{(2)}\|p^{(1)}\big)>D_\alpha\big(q^{(2)}\|q^{(1)}\big) \tag{160}$$

对所有 $\alpha\in[1/2,\infty]$, 均成立，则对每个足够大的 $n\in\mathbb{N}$，有

$$\Big(\big(p^{(1)}\big)^{\otimes n},\big(p^{(2)}\big)^{\otimes n}\Big)\succeq\Big(\big(q^{(1)}\big)^{\otimes n},\big(q^{(2)}\big)^{\otimes n}\Big). \tag{161}$$

证明。假设（159）和（160）对所有 $\alpha\in[1/2,\infty]$ 均成立，并记$p^{(k)}=(p_1^{(k)},\dots,p_m^{(k)})$和 $q^{(k)}=(q_1^{(k)},\dots,q_n^{(k)})$ 分别为 $k=1$，2，其中所有元素均为正数。首先，令 $\alpha\in\mathbb{R}\backslash\{0,\ 1\}$。容易验证

$$D_{(\alpha,1-\alpha)}(p^{(1)},p^{(2)})=\begin{cases}D_\alpha(p^{(1)}\|p^{(2)}), & \alpha\in[1/2,1)\cup(1,\infty),\\ D_{1-\alpha}(p^{(2)}\|p^{(1)}), & \alpha\in(-\infty,0)\cup(0,1/2).\end{cases} \tag{162}$$

因此，(159) 和 (160) 在 $\alpha\in(0,\ 1)\cup(1,\ \infty)$ 的情况下对应于条件 (136)，其中 $d=2$。接下来，设 $\beta>0$。回顾备注 16，很容易看出

$$D^{\mathbb{T}}_{(\beta,-\beta)}(p^{(1)},p^{(2)})=D^{\mathbb{T}}_{(1,-1)}(p^{(1)},p^{(2)})=\max_{1\leqslant i\leqslant m}\log\frac{p_i^{(1)}}{p_i^{(2)}}=D_\infty(p^{(1)}\|p^{(2)}), \tag{163}$$

$$D^{\mathbb{T}}_{(-\beta,\beta)}(p^{(1)},p^{(2)})=D^{\mathbb{T}}_{(-1,1)}(p^{(1)},p^{(2)})=\max_{1\leqslant i\leqslant m}\log\frac{p_i^{(2)}}{p_i^{(1)}}=D_\infty(p^{(2)}\|p^{(1)}). \tag{164}$$

因此，(159) 和 (160) 在 $\alpha=\infty$ 的情况下对应于条件 (137)。最后，(159) 和 (160) 在 $\alpha=1$ 的情况下显然对应于

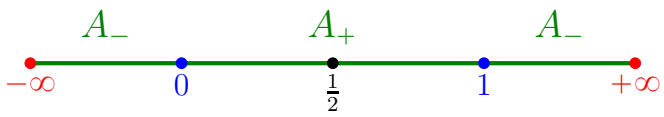

FIGURE 3. 这里我们展示相对主序半环的测试谱 S^2。它对应于 R´ényi 散度的集合，如推论 23 的证明所示；此处将其绘制为 $\mathbb{R}$ 的两点紧化。在中心点$\alpha = 1/2$ （黑色）的右侧，我们有散度 $(p^{(1)}, p^{(2)}) \mapsto D_\alpha(p^{(1)}\|p^{(2)})$；在中心点的左侧，我们有散度$(p^{(1)}, p^{(2)}) \mapsto D_\alpha(p^{(2)}\|p^{(1)})$。点 $\alpha = 0$ 和 1（蓝色）对应于两个本质导子，即参数顺序不同的两种 Kullback-Leibler 散度，而点 $\alpha = \pm\infty$ （红色）对应于散度 D_∞ 的两种不同参数顺序。

条件 (138) 在 $d = 2$的情况下。总而言之，矩阵散度与 Rényi 散度有简单的联系，另请参见图 3。因此，该结论由定理 19得出。□

我们还要指出，定理22 蕴含了关于催化渐近相对主序的以下结果；这可以看作是对 [14]中定理 20 的加强。

推论 24。对于具有公共支撑的概率向量二分布 $(p^{(1)}、 p^{(2)})$，以及具有公共支撑的概率向量二分布 $(q^{(1)}、 q^{(2)})$，且二者均属于 $\mathcal{P}_n^2$，以下条件等价：

（i）存在 $\mathcal{P}_n$ 中的概率向量序列 $(q_k^{(1)})_{k=1}^\infty$，使 $\|q^{(1)} - q_k^{(1)}\|_1 \to 0$ 为 $k \to \infty$，并且存在由概率向量二分布构成的序列 $(r_k^{(1)} 、 r_k^{(2)})_{k=1}^\infty$，使得

$$\left(p^{(1)} \otimes r_k^{(1)}, p^{(2)} \otimes r_k^{(2)}\right) \succeq \left(q_k^{(1)} \otimes r_k^{(1)}, q^{(2)} \otimes r_k^{(2)}\right) \tag{165}$$

对所有 $k \in \mathbb{N}$ 均成立。

（ii）对所有 $\alpha \geqslant 1/2$,

$$D_\alpha\left(p^{(1)}\middle\|p^{(2)}\right) \geqslant D_\alpha\left(q^{(1)}\middle\|q^{(2)}\right), \tag{166}$$

$$D_\alpha\left(p^{(2)}\middle\|p^{(1)}\right) \geqslant D_\alpha\left(q^{(2)}\middle\|q^{(1)}\right). \tag{167}$$

证明可直接由定理 22 以及前面的证明中对这些散度 $D_{(\alpha,\ 1-\alpha)}$ 与 R´ényi 散度的对应关系得出。请注意，我们只对 $q^{(1)}$ 进行近似，而不对 $q^{(2)}$ 进行近似；换言之，我们要求二分布的第二个分量满足精确催化主序，这可由定理 22实现。当然，也可以反过来进行

。定理20的 [14] 结论相同，但对 $q^{(1)}$ 和 $q^{(2)}$都采用近似；从这个意义上说，我们当前的结果更强。

4. Conclusion

在这项工作中，我们将主序的若干问题与预序半环理论联系起来。借助这一框架，我们推导出了关于大样本矩阵主序和催化矩阵主序的新充分条件，以及一般情况下必要的条件；也就是说，这些条件适用于有限结果情形下的 Blackwell 优势。在后一种情形中，我们特别将相关的单调量确定为矩阵 $\underline{\alpha}$-散度，可将其视为标准 R´enyi 散度单参数族的推广。

我们想指出，这一理论不仅与经典统计学中的问题有着众所周知的联系，也与资源理论在量子信息处理背景下的问题有关。所有基本的量子设备——量子态、信道和可观测量——也都可以描述为统计实验的适当推广（例如参见[18]）。具体而言，在纯量子态的纠缠资源理论中，如果存在一个局域操作与经典通信（LOCC）量子信道 φ，使得 AB，那么我们可以说复合系统的纯（向量）态ψ ‘高于’另一个纯态 Φ。众所周知 $\Phi(|\varphi\rangle\langle\varphi|) = |\psi\rangle\langle\psi|$，当且仅当[22]的施密特向量主序于 ψ 的施密特向量时，这一条件才成立。因此，根据定理41，可以得到纯态催化 LOCC 排序的条件，正如 [17]中指出的那样。纯态的非相干变换也有类似结果[9]。

量子态的多重主序也是一个相关问题。这只是矩阵主序的量子对应形式：给定 d元组$(\rho_1$、……、 $\rho_d)$ 和 $(\sigma_1$、……、 $\sigma_d)$（它们是量子态，可能属于不同的系统），我们可以问是否存在一个量子信道 Λ，使得 $\Lambda(\rho_k) = \sigma_k$ 对 k =1，……， d均成立。这个问题的单次条件例如可由[7, 24]得到。由 [12, 13]中的比较定理 ¨构成的代数工具也适用于量子情形：我们可以按照构造矩阵主序半环的相同方式，构造一个具有多项式增长的量子主序半环。由此，大样本下或通过催化实现的量子统计实验主序，可以用量子主序半环上的单调同态和单调导子所涉及的不等式来刻画（它们共同构成该半环的测试谱）。

然而，这种方法的问题在于，目前仍无法刻画量子主序半环的测试谱。首先，我们无法像在命题13 和14中那样，对量子情形下的加法映射进行类似的分解，因为 d元组中的状态通常具有不同的特征基。相对情形$d = 2$ [27]已经出现了问题：目前已定义了无数种量子散度

推广经典 R´ényi 散度的量子散度已有定义，但我们尚不清楚其中哪些与主序问题真正相关。因此，即使在 $d = 2$情形下，我们目前也无法给出大样本中量子主序或催化量子主序的具体、明确条件。可以运用本文提出的方法研究一些特殊情形，但这将留待未来研究。此外，采用较宽松的预序（例如量子版次主序）或许会有帮助，因为这样有时能够将问题缩小到一组可处理的充分条件[23, 6]。

Appendix A. The majorization and submajorization semirings

在这些附录中，我们将重点讨论概率向量的（简单）主序，并在此基础上重新推导一些此前已确立的关于大样本主序和催化主序的结果。我们将看到，与矩阵主序的情况一样，我们所掌握的新代数工具在简单主序问题上也十分强大。

若存在一个双随机矩阵 T 使得 $Tp = q$，则称概率向量 p 主序于 q。这里，必要时我们通过补零，将 p 和 q 视为大小相同的向量 n；双随机意味着该矩阵是一个元素非负的 $(n \times n)$ 矩阵，且其各行和各列的元素之和均为 1。根据 Hardy-Littlewood-P´olya 定理，p 主序于 q 当且仅当

$$\sum_{i=1}^{k} p_i^{\downarrow} \geqslant \sum_{i=1}^{k} q_i^{\downarrow} \tag{168}$$

对所有 $k = 1, \ldots, n$，其中 $p^{\downarrow} = (p_i^{\downarrow})_{i=1}^{n}$和 $q^{\downarrow} = (q_i^{\downarrow})_{i=1}^{n}$是将 p 和 q 的各元素按非递增顺序排列后得到的向量。该定义可推广到具有非负元素的向量 p和 q，也就是说，我们不必假设 $\|p\|_1 = 1 = \|q\|_1$。鉴于主序及相关预序的重要性，以下各小节旨在通过配备主序预序或其某种变体的非负向量交换半环，从代数角度分析主序，并说明如何运用第 2 节介绍的分离定理一般理论，重新得到一些关于催化变换和大样本变换的经典结果。这些结果最初是通过更具体的方法证明的，而那些方法并不适用于其他情况。

A.1. 主序半环。我们聚焦于支撑有限且具有非负元素的向量。这些向量构成集合 $\mathcal{V}_{<\infty} := \bigcup_{n=1}^{\infty} \mathbb{R}_{+}^{n}$。对于每个$p \in \mathcal{V}_{<\infty}$，记 $p^{\downarrow}$ 为 p的递减重排，因此向量 $p^{\downarrow}$和 p 的分量仅排列顺序不同。我们认为 p、 $q \in \mathcal{V}_{<\infty}$等价，并在 $p^{\downarrow}$ 和 $q^{\downarrow}$ 补零后相同时记作 $p \approx q$

。我们用 $\mathcal{V}_{<\infty}/\approx$ 表示 $\approx$-等价类 $[p]$的集合。我们稍微滥用一下记号，将 $\mathbf{1} := [(1)]$记作单元素向量（其元素为 1）的等价类，并将 $\mathbf{0} := [(0)]$记作相应的等价类。 $[p]$、 $[q] \in \mathcal{V}_{<\infty}/\approx$的加法和乘法定义如下：

$$[p] + [q] := [p \oplus q], \qquad [p] \cdot [q] := [p \otimes q], \tag{169}$$

其中，克罗内克积中各分量的顺序在 $\approx$ 意义下无关。

定义 25。我们在 $\mathcal{V}_{<\infty}/\approx$ 上定义三个预序如下。给定两个一般元素 $p = (p_1、\ \ldots\ldots、\ p_n) \in \mathcal{V}_{<\infty}$ 和 $q = (q_1、\ \ldots\ldots、\ q_n) \in \mathcal{V}_{<\infty}$，假设向量 p 和 q 的长度相同，为 n；通过补零，总可以在 $\approx$ 意义下做到这一点。

(a) 次主序预序：若对所有 $k = 1、\ \ldots\ldots、\ n$， $[q] \trianglelefteq [p]$，

$$\sum_{i=1}^{k} q_i^{\downarrow} \leqslant \sum_{i=1}^{k} p_i^{\downarrow}, \tag{170}$$

或者等价地，若存在双随机矩阵$T = (T_{i,\ j})_{i,\ j=1}^{n} \in \mathcal{M}_n(\mathbb{R}_+)$，使得 $q \leqslant Tp$（逐项不等式）。赋予次主序后， $\mathcal{V}_{<\infty}/\approx$ 成为一个预序半环（见下文）。我们将其记为 S，并称之为次主序半环。对于单个向量 p 和 q，今后当 $[q] \trianglelefteq [p]$时，我们记作 $p \succeq q$ 或 $q \preceq p$。

(b) 主序预序： $[q] \trianglelefteq_1 [p]$ 当且仅当 $[q] \trianglelefteq [p]$ 和 $\|p\|_1 = \|q\|_1$。赋予主序后， $\mathcal{V}_{<\infty}/\approx$ 成为一个预序半环，我们称之为主序半环，并记作 S_1。对于单个向量 p 和 q，从现在起，当 $[q] \trianglelefteq_1 [p]$时，我们记作 $p \succeq_1 q$ 或$q \preceq_1 p$。

(c) 修正主序：$[q] \trianglelefteq_{0,\ 1} [p]$ 当且仅当 $[q] \trianglelefteq_1 [p]$ 和 $\|p\|_0 = \|q\|_0$，其中 $\|\cdot\|_0$表示支撑 $p|$的大小 $|$，即满足 $p_i > 0$的结果 i 的数量。赋予修正主序后， $\mathcal{V}_{<\infty}/\approx$ 成为一个预序半环，我们称之为修正主序半环，并记作 $S_{0,\ 1}$。对于单个向量 p 和 q，从现在起，当 $[q] \trianglelefteq_{0,\ 1} [p]$时，我们记作 $p \succeq_{0,\ 1} q$ 或 $q \preceq_{0,\ 1} p$。

命题 26. S、 S_1 和$S_{0,\ 1}$ 都是预序半环。

关于 S 的证明。对于 p， $q \in \mathcal{V}_{<\infty}$，假设 $q \preceq p$ 由一个双随机矩阵 T 见证，且 $Tp \geqslant q$，[6] 不失一般性，假设 p 和q 长度相同 n。我们需要证明，这种主序关系在与任意 $r \in \mathcal{V}_{<\infty}$相加和相乘时保持稳定。定义

[6]请注意，此处符号 $\geqslant$ 表示向量之间的逐元素不等式。

分块矩阵 $S := T \oplus I_m$，其中 m 是 r 的大小， I_m 是 $(m \times m)$-单位矩阵。因此

$$S(p \oplus r) = Tp \oplus r \geqslant q \oplus r. \tag{171}$$

显然， S 是双随机的，因此 $[q] + [r] \preccurlyeq [p] + [r]$。

同理，将映射 T 与单位矩阵 I_m作张量积，可得

$$(T \otimes I_m)(p \otimes r) = Tp \otimes r \geqslant q \otimes r, \tag{172}$$

因此，所需的 $[p] \cdot [r] \trianglerighteq [q] \cdot [r]$成立。 □

S_1 和$S_{0,\ 1}$的证明。根据定义， S_1 上的预序是 S上的预序与具有相等 $\|\cdot\|_1$ 这一关系的交集。由于后一个映射是一个同态 $S \to \mathbb{R}_+$，具有相等 $\|\cdot\|_1$ 这一关系本身就是一个（对称的）半环预序。由此可知 S_1 的结论成立，因为两个半环预序的交集仍是半环预序。$S_{0,\ 1}$ 的证明类似。□

备注27。对于熟悉代数的读者，不妨注意，我们的主序半环的纯半环结构（即不考虑预序时）恰好就是群半环 $\mathbb{N}[\mathbb{R}_{>0}]$。这里所说的群半环，是指由群 $(\mathbb{R}_{>0}, \cdot)$的所有元素的形式和构成的半环；两个这样的形式和相乘时，先应用分配律，再利用 $\mathbb{R}_{>0}$ 中的乘法，得到一个新的形式和。这个半环与我们的 S一致，因为 S 中的每个元素都可以看作正实数的形式和，而这些形式和按照上述方式相乘。

这个群半环由如下泛性质刻画：到任意其他半环 T 的同态 $S \to T$ 与群同态 $\mathbb{R}_{>0} \to T^\times$ 一一对应，其中 $T^\times$ 是 T 的单位群。正是这一泛性质使我们能够在下文找到主序半环上单调同态的简单刻画。

为了利用主序半环上的分离定理，我们需要证明这些预序半环都具有多项式增长。

命题28. 预序半环 S、 S_1 和 $S_{0,1}$ 都具有多项式增长。此外， S 和 S_1^{op} 具有幂泛元。

证明。我们先考虑次主序半环 S。单项元素 $[(a)] =: u$ 与任意 $a > 1$ 都是幂泛元。为此，设$p = (p_1, \ldots, p_n) \in \mathcal{V}_{<\infty}$和 $q = (q_1, \ldots, q_n) \in \mathcal{V}_{<\infty}$ 均非零。显然，对于 $a > 1$，存在 $k \in \mathbb{N}$ 使得

$$\sum_{i=1}^{\ell} q_i^{\downarrow} \leqslant a^k \sum_{i=1}^{\ell} p_i^{\downarrow} \tag{173}$$

对所有 $\ell = 1, \ldots, n$ 均成立，即 $a^k p \succeq q$。这意味着 u 是 S中的幂泛元。类似地，我们也可以证明 S^{op} 同样有一个幂泛元，即$[(a)]$，其中 $0 < a < 1$。

接下来讨论主序半环 S_1。我们证明， S_1^{op} 具有一个幂泛元，因此 S_1^{op}，进而 S_1 具有多项式增长。具体而言，我们证明 $[r]$，其中 $r = (t, 1-t)$，对任意 $t \in (0, 1)$，都是 S_1^{op} 的幂泛元。我们不妨假设 $t \geqslant 1/2$。设 p、$q \in \mathcal{V}_{<\infty}$ 满足 $\|p\|_1 = \|q\|_1$。我们需要证明存在 $k \in \mathbb{N}$ 使得 $p \succeq_1 r^{\otimes k} \otimes q$。记 p 的最小非零元素为 $p_{\min}$，并记 q 的最大元素为 $q_{\max}$。显然，对于每个充分大的 $k \in \mathbb{N}$， $r^{\otimes k} \otimes q$的最大元素$t^k q_{\max}$ 有上界 $p_{\min}$。这意味着 $p \succeq_1 r^{\otimes k} \otimes q$。更具体地说，我们可以任选

$$k \geqslant \frac{\log p_{\min} - \log q_{\max}}{\log t} \tag{174}$$

以获得所需的 $p \succeq_1 r^{\otimes k} \otimes q$。

最后，我们证明修正主序半环 $S_{0,\,1}$ 具有多项式增长。对于每个 $s \in \mathbb{N}$，回顾一下，我们记

$$u_s := (1/s, \ldots, 1/s) \tag{175}$$

为 s 个结果上的均匀分布。定义 $t_- := u_2$ 以及 $t_+ := (1-\varepsilon, \varepsilon)$，对每个固定的 $\varepsilon \in (0, 1/2)$均如此。接下来我们证明， $([t_-], [t_+])$ 是 $S_{0,\,1}$的幂通用对。事实上，$t_- \preceq_{0,\,1} t_+$ 显然成立。对于主要条件，只需证明：对于每个满足 $\|p\|_1 = \|q\|_1$ 和 $\|p\|_0 = \|q\|_0$的 $p, q \in \mathcal{V}_{<\infty}$，存在 $k \in \mathbb{N}$ 使得 $t_-^{\otimes k} \otimes p \preceq_{0,\,1} t_+^{\otimes k} \otimes q$。显然，我们可以假设$\|p\|_1 = 1 = \|q\|_1$。记 $n := \|p\|_0 = \|q\|_0$，并将 p 中最小的非零元素记为 $p_{\min}$。假设 $n \geqslant 2$，现在定义

$$\tilde{p} := \Big(1 - p_{\min}, \underbrace{\frac{p_{\min}}{n-1}, \ldots, \frac{p_{\min}}{n-1}}_{n-1 \text{ copies}}\Big), \tag{176}$$

使得 $\tilde{p} \succeq_{0,\,1} p$。接下来我们证明存在 $k \in \mathbb{N}$，使得 $t_-^{\otimes k} \otimes \tilde{p} \preceq_{0,\,1} t_+^{\otimes k} \otimes u_n$。这将说明我们要证明的结论，因为

$$t_-^{\otimes k} \otimes p \preceq_{0,1} t_-^{\otimes k} \otimes \tilde{p} \preceq_{0,1} t_+^{\otimes k} \otimes u_n \preceq_{0,1} t_+^{\otimes k} \otimes q, \tag{177}$$

其中，最后一个不等式由以下事实得出：$q \succeq_{0,\,1} u_n$。因此，接下来我们证明，当 $k \in \mathbb{N}$ 足够大时，上述中间的不等式成立。

$(t_-^{\otimes k} \otimes \tilde{p})^{\downarrow}$ 的初始部分和线性增长，直到第 2^k 个部分和达到 $1 - p_{\min}$。之后，初始部分和再次线性增长，最终达到 $\mathbf{1}$。由于初始部分和的图像是一个凹函数，且在最终和处达到 $\mathbf{1}$，因此，为了找到 k，使得 $t_-^{\otimes k} \otimes \tilde{p} \preceq_{0,\,1} t_+^{\otimes k} \otimes u_n$，只需确保 $(t_+^{\otimes k} \otimes u_n)^{\downarrow}$ 的第 2^k 个初始部分和至少为 $1 - p_{\min}$。

我们可以将这个初始部分和界定为

$$\sum_{i=1}^{2^k}(t_+^{\otimes k}\otimes u_n)_i^{\downarrow} \geqslant \sum_{j=0}^{a} n\binom{k}{j}\frac{(1-\varepsilon)^{k-j}\varepsilon^j}{n} = \sum_{j=0}^{a}\binom{k}{j}(1-\varepsilon)^{k-j}\varepsilon^j \tag{178}$$

对每个 $a \in [0,\ k]$ 使得$\sum_{j=0}^{a} n\binom{k}{j} \leqslant 2^k$。右侧是在 a 处取值的均值为εk的二项随机变量的累积分布函数。现在我们选择 $\delta > 0$使得 $\varepsilon+2\delta < \frac{1}{2}$且 k满足 $k\delta > 1$。此外，任取 $a \in [k(\varepsilon+\delta),\ k(\varepsilon+2\delta)]$随着我们增大 k，（178）中的累积分布函数将根据大数定律收敛到 1，特别是最终会超过 $1-p_{\min}$。因此，只需检验对于这样的 a，条件$\sum_{j=0}^{a} n\binom{k}{j} \leqslant 2^k$ 在足够大的k时是否成立。为验证这一点，我们使用众所周知的界（例如，参见 [31]中的第 1.4 节），

$$\sum_{j=0}^{a}\binom{k}{j} \leqslant \exp\big(k\,h_1(\varepsilon+2\delta)\big), \tag{179}$$

其中 $h_1:\ [0,1]\to\mathbb{R}$,

$$h_1(x) = -x\log x - (1-x)\log(1-x), \tag{180}$$

$$0\log 0 := 0, \tag{181}$$

是二元熵。注意，由于 $h_1(\varepsilon+2\delta) < 1$按照我们对 δ的选择，显然对于充分大的 k，有 $n\exp\big(k\,h_1(\varepsilon+2\delta)\big) < \exp k$，从而完成证明。□

注 29. 只要 $r \in \mathcal{P}_{<\infty}$ 是任意一个满足 $[r] \neq 1$ （即$\|r\|_0 \geqslant 2$）的概率向量，那么 $[r]$对于 S_1^{op} 已经是幂通用的。为此，将 r 中最小的非零元素记为 $r_{\min}$，并令 $t:=1-r_{\min}\in(0,1)$。由此可得$\tilde{r}:=(t,1-t)$ 主序于 r。利用命题28证明开头针对S_1的论证，其中已证明 $[\tilde{r}]$ 是幂通用的，对于所有满足 $\|p\|_1=\|q\|_1$ 的 $p,q\in\mathcal{V}_{<\infty}$，以及充分大的 $k\in\mathbb{N}$，有

$$r^{\otimes k}\otimes q \preceq_1 \tilde{r}^{\otimes k}\otimes q \preceq_1 p, \tag{182}$$

从而证明该断言。

注30。 我们已经看到， S 具有一个幂泛元，而且很容易看出它也是零和自由的。又已知 $0\preccurlyeq 1$ 在 S中，因此定理5 的条件对 S成立。关于 S_1 和$S_{0,\ 1}$，下面我们论证它们都满足定理7 和8的辅助条件。事实上，记 $\mathcal{V}_{<\infty}$ 上由 $\preccurlyeq_1$ 生成的等价关系为 $\sim_1$，则有 $p\sim_1 q$ 当且仅当 $\|p\|_1=\|q\|_1$。这意味着 $S_1/\sim_1 \simeq \mathbb{R}_{+,}$，显然它具有准补元和准逆元，其意义见 [13], 和

$$\mathsf{Frac}(S_1/\sim_1)\otimes\mathbb{Z}\cong\mathbb{R} \tag{183}$$

是一个域，因此显然是有限个域的乘积。

同理，我们有 $p \sim_{0,\,1} q$ 当且仅当 $\|p\|_1 = \|q\|_1$ 且 $\|p\|_0 = \|q\|_0$。因此

$$S_{0,1}/\sim_{0,1} \simeq (\mathbb{R}_{>0} \times \mathbb{N}) \cup \{(0,0)\}, \tag{184}$$

很容易看出，它具有准补元和准逆元。此外
，
我们有

$$\mathsf{Frac}(S_{0,1}/\sim_{0,1}) \otimes \mathbb{Z} \cong \mathbb{R} \times \mathbb{Q}, \tag{185}$$

它是两个域的乘积。

为了应用分离定理，我们首先聚焦于$S_{0,\,1}$并刻画所有非退化单调同态 $\Phi : S_{0,\,1} \to \mathbb{K}$，其中$\mathbb{K} \in \{\mathbb{R}_+$、$\mathbb{R}_+^{\mathrm{op}}$、$\mathbb{TR}_+$、$\mathbb{TR}_+^{\mathrm{op}}\}$，以及退化单调同态处的单调导子。然后，我们将从$S_{0,\,1}$的相应结果推导出 S 和 S_1 的结果。

应用分离定理，便可得到任意一对向量之间大样本主序和催化主序的充分条件，以及一般情形下的必要条件 p，$q \in \mathcal{V}_{<\infty}$。随后，将定理 5 应用于 S，若所有相关单调量的取值都严格有序，我们便可得出大样本主序和催化主序。由于 $\|\cdot\|_1 \colon S \to \mathbb{R}_+$ 是一个单调同态，这尤其要求 $\|p\|_1$ 与 $\|q\|_1$ 之间严格有序，因此，若 p 和 q 都是概率向量，我们便无法得出结论。然而，此时 S_1 便能发挥作用：将定理 8 应用于 S_1，并令$\|\cdot\| = \|\cdot\|_1$，只需相对于其他单调同态严格有序，因此仍可得到大样本主序和催化主序的充分条件，且这些条件在一般情形下也是必要的。但由于 $\|\cdot\|_0$是相关单调同态之一，这同样不适用于非一般情形 $\|p\|_0 = \|q\|_0$。不过，对于这种情形，我们可以将定理 7 应用于$S_{0,\,1}$，从而得到催化主序的充分条件，以及一般情形下的必要条件。

A.2. 修正主序半环上的单调同态与导子。我们首先讨论修正主序半环的情形。在我们的主序半环中，由于预序 $\geq_{0,1}$ 是最小的主序预序，因此它拥有数量最多的单调同态。在以下各节中，我们会稍微放宽记号的用法，将 $S_{0,1}$ 上的单调同态和导子，以及 S_1和 S，视为定义在 $\mathcal{V}_{<\infty}$ 上的函数，尽管它们实际上定义在 $\mathcal{V}_{<\infty}/\approx$上；这仅意味着，作为定义在 $\mathcal{V}_{<\infty}$ 上的函数，它们在每个 $\approx$-等价类上均为常数。采用这种记号上的简化，是为了避免额外使用表示等价类的括号，从而使记号更加简洁。

定义在任意 $p = (p_1, \ldots, p_n) \in \mathcal{V}_{<\infty}$ 上、支撑为 $I := \mathsf{supp}(p)$ 的单调同态将在以下函数中找到：$S_{0,1}$

$$f_\alpha(p) = \sum_{i \in I} p_i^\alpha, \qquad \alpha \in \mathbb{R}, \tag{186}$$

$$f_\infty(p) = \max_{i\in I} p_i, \tag{187}$$

$$f_{-\infty}(p) = \Big(\min_{i\in I} p_i\Big)^{-1}, \tag{188}$$

$$H_1(p) = -\sum_{i\in I} p_i \log p_i, \tag{189}$$

$$H_0'(p) = \sum_{i\in I} \log p_i. \tag{190}$$

这里，我们约定空和为零，$\max\varnothing := 0, \min\varnothing := \infty$, $1/\infty :=0$ and $0\log 0 := 0$。注意，方程（189）中定义的映射 H_1 是香农熵。另外 $f_1 = \|\cdot\|_1$ 和 $f_0 = \|\cdot\|_0$，其中这两个“范数”与之前相同。

命题 31. 非退化单调同态$S_{0,\ 1} \to \mathbb{K}$ 恰好如下：

(i) 对于 $\mathbb{K} = \mathbb{R}_+$，映射 f_α 适用于 $\alpha \in (-\infty,\ 0) \cup [1,\ \infty)$。(ii) 对于 $\mathbb{K} = \mathbb{R}_+^{\mathrm{op}}$，映射 f_α 适用于 $\alpha \in (0,\ 1)$。(iii) 对于 $\mathbb{K} = \mathbb{TR}_+$，映射 $f_{\mathrm{sign}(\alpha)\infty}(\cdot)^{|\alpha|}$ 适用于 $\alpha \in (-\infty,\ 0) \cup (0,\ \infty)$。(iv) 对于 $\mathbb{K} = \mathbb{TR}_+^{\mathrm{op}}$，不存在这样的映射。

唯一的退化同态 $S_{0,\ 1} \to \mathbb{R}_+$ 是 $f_0 = \|\cdot\|_0$ 和 $f_1 = \|\cdot\|_{1}$.

证明. 给定一个单调同态 $\Phi : S_{0,\ 1} \to \mathbb{K}$，我们将 $\varphi : \mathbb{R}_+ \to \mathbb{K}$记作 Φ 在单项向量（的等价类）集合上的限制，$\varphi(x) := \Phi([(x)])$。因此，$\varphi$ 是乘法的，即

$$\varphi(xy) = \varphi(x)\varphi(y) \qquad \forall x, y \in \mathbb{R}_+, \qquad \varphi(1) = 1, \tag{191}$$

根据 Φ的乘法性。这几乎可以确保 φ 是一个单项式函数$\varphi(x) = x^\alpha$：根据柯西函数方程的标准理论，要得出这一结论，只需另外证明 φ在某个区间 $[x$ 上有界即可，$y]$ 对于任意 $0 \leqslant x < y$ [1, 第 2 章，定理 8][7]，，下面我们将对此进行证明。为此，注意对于所有 $t \in [0,\ 1]$,

$$\Big(\frac{x+y}{2}, \frac{x+y}{2}\Big) \preceq_{0,1} \big(tx + (1-t)y, (1-t)x + ty\big) \preceq_{0,1} (x, y), \tag{192}$$

正如双随机矩阵所见

$$T_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad T_2 = \begin{pmatrix} t & 1-t \\ 1-t & t \end{pmatrix}. \tag{193}$$

因此，根据 Φ 的可加性和单调性，我们有

$$\text{case (i): } 2\varphi\Big(\frac{x+y}{2}\Big) \leqslant \varphi\big(tx + (1-t)y\big) + \varphi\big((1-t)x + ty\big) \tag{194}$$

[7]注意，该定理讨论的是柯西函数方程的线性形式。我们可以改为考虑函数 $\log \circ\varphi\circ \exp$，将问题转化为这种形式。

$$\leqslant \varphi(x)+\varphi(y), \tag{195}$$

$$\text{case (ii): } 2\varphi\Big(\frac{x+y}{2}\Big) \geqslant \varphi\big(tx+(1-t)y\big)+\varphi\big((1-t)x+ty\big) \tag{196}$$

$$\geqslant \varphi(x)+\varphi(y), \tag{197}$$

$$\text{case (iii): } \varphi\Big(\frac{x+y}{2}\Big) \leqslant \max_{s\in\{t,1-t\}} \varphi\big(sx+(1-s)y\big) \tag{198}$$

$$\leqslant \max\{\varphi(x),\varphi(y)\}, \tag{199}$$

$$\text{case (iv): } \varphi\Big(\frac{x+y}{2}\Big) \geqslant \max_{s\in\{t,1-t\}} \varphi\big(sx+(1-s)y\big) \tag{200}$$

$$\geqslant \max\{\varphi(x),\varphi(y)\}, \tag{201}$$

其中，不等式是相对于通常的数值大小顺序而言的。现在只需注意，在每种情况下，这些不等式都意味着 $\varphi\big(tx+(1-t)y\big)$ 存在一个与 t 无关的上界，这基于 φ 仅取非负值这一假设。因此， φ 在 $[x、y]$ 上有界，故存在 $\alpha\in\mathbb{R}$，使得对所有 $x\geqslant 0$，都有 $\varphi(x)=x^{\alpha}$。

转向主要结论，考虑任意 $p\in\mathcal{V}_{<\infty}$ 并记 $I:=\operatorname{supp} p$。现在我们有

$$\Phi(p)=\Phi\left(\bigoplus_{i\in I}(p_i)\right)=\sum_{i\in I}\varphi(p_i)=\sum_{i\in I}p_i^{\alpha}=f_{\alpha}(p) \tag{202}$$

在情形 (i) 和 (ii) 中，以及

$$\Phi(p)=\Phi\left(\bigoplus_{i\in I}(p_i)\right)=\max_{i\in I}\varphi(p_i)=\max_{i\in I}p_i^{\alpha}=\begin{cases} f_{\infty}(p)^{\alpha} & \text{if } \alpha\geqslant 0,\\ f_{-\infty}(p)^{-\alpha} & \text{if } \alpha\leqslant 0\end{cases} \tag{203}$$

在情形 (iii) 和 (iv) 中。因此， Φ 在所有情形下都具有所需的形式；使这些映射成为非退化单调同态的 α 取值范围仍有待确定。

容易看出，映射 f_{α} 确实是同态$S_{0,\ 1}\to\mathbb{R}_{+}$ 对所有 $\alpha\in\mathbb{R}$都成立；同理，映射 $f_{\infty}(\cdot)^{\alpha}$ 和 $f_{-\infty}(\cdot)^{-\alpha}$ 对所有 $\alpha\in\mathbb{R}$都是同态$S_{0,\ 1}\to\mathbb{TR}_{+}$。因此，要完成主要结论的证明，只需证明当且仅当 α 满足所述条件时，这些映射是非退化且单调的。现在先聚焦于情形 (i) 和 (ii)，并首先考虑 $\alpha\in[0、1]$。为分析 f_{α}的单调性，令$p、q\in\mathcal{V}_{<\infty}$ 满足$p\succeq_{0,\ 1}q$。因此，我们可以假设存在 $n\in\mathbb{N}$使得 $p=(p_1、\cdots\cdots、p_n)\in\mathbb{R}^n_{>0}$、 $q=(q_1、\cdots\cdots、q_n)\in\mathbb{R}^n_{>0}$，并且存在双随机$T=(T_{i,\ j})_{i,\ j=1}^{n}$ 使得 $q=Tp$。利用函数$x\mapsto x^{\alpha}$ 在 $\alpha\in[0、1]$时为凹函数这一事实，我们可以计算

$$f_{\alpha}(q)=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}T_{i,j}p_j\right)^{\alpha}\geqslant\sum_{i,j=1}^{n}T_{i,j}p_j^{\alpha}=\sum_{j=1}^{n}p_j^{\alpha}=f_{\alpha}(p). \tag{204}$$

因此， $f_\alpha : S_{0,\,1} \to \mathbb{R}_+^{\mathrm{op}}$ 是关于 $\alpha \in [0、\ 1]$的单调同态。在$\alpha = 0$ 和 $\alpha = 1$处，我们得到同态 $\|\cdot\|_0$ 和 $\|\cdot\|_1$；根据$\succeq_{0,\,1}$的定义，它们是退化的。对于 $\alpha \in (0、\ 1)$，例如可以考虑 $p_t = (1-t、\ t)$（对于 $t \in (0、\ 1/2],$），从而证明其非退化性；这样，每当$0 < s \leqslant t$时，我们都有$p_s \succeq_{0,\,1} p_t$。于是得到 $f_\alpha(p_t) = (1-t)^\alpha + t^\alpha$，不难看出，对于 $\alpha \in (0、\ 1)$，它是关于 t 的严格递增函数。

类似地，利用 $x \mapsto x^\alpha$ 在 $\alpha \leqslant 0$ 或 $\alpha \geqslant 1$时的凸性，我们得到$f_\alpha : S_1 \to \mathbb{R}_+$ 在这些情形下是单调同态。[8] 对于 $\alpha \neq 0,\ 1$ 的非退化性证明同理。因此我们得到结论 (i) 和 (ii)。

最后来看情形 (iii) 和 (iv)。映射 $f_\infty(\cdot)^\alpha : S_{0,\,1} \to \mathbb{TR}_+$ 对任意 $\alpha \geqslant 0$都是单调同态。为此，任取 $p,\ q \in \mathcal{V}_{<\infty}$ 如（204）中所示，并类似地计算

$$f_\infty(q)^\alpha = \left(\max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} T_{i,j} p_j\right)^\alpha \leqslant \left(\max_{1\leqslant i\leqslant n} p_i\right)^\alpha = f_\infty(p)^\alpha, \tag{205}$$

这里我们用到了这样一个事实：凸组合不会增大参与组合的元素的最大值。对于 $f_{-\infty}(\cdot)^\alpha$ 和 $\alpha \geqslant 0$，我们也可以完全照此处理。由于对于 $\alpha > 0$，在上述 p_t 这类一般情形下，这些不等式都是严格的，因此除非 $\alpha = 0$，$f_{\pm\infty}(\cdot)^\alpha$ 作为到 $\mathbb{TR}_+^{\mathrm{op}}$ 的映射并非单调。特别地，对于所有 $\alpha \neq 0$，我们也得到映射$S_{0,\,1} \to \mathbb{TR}_+$ 的非退化性。□

由于我们有两个从$S_{0到1}$ 的退化同态 $\mathbb{R}_+$，因此必须同时考虑在 $\|\cdot\|_1$ 和 $\|\cdot\|_0$处的单调导子。利用关于香农熵公理体系的现有结果，可以完成对 $\|\cdot\|_1$ 处单调导子的分类（见备注34）；不过，我们仍给出完整证明，因为这再次展示了半环论方法的实用性。正如我们稍后将看到的，在 $\|\cdot\|_0$ 处求导的情形比在 $\|\cdot\|_1$处简单得多。

在深入探讨这些导子的分类细节之前，先给出并证明一个有用的引理；在刻画 $\|\cdot\|_1$处的单调导子时，我们将用到它。

引理 32。固定 $n \in \mathbb{N}$，并令 $\Delta : \mathbb{R}_+^n \to \mathbb{R}$ 满足

$$p \succeq_{0,1} q \quad \Longrightarrow \quad \Delta(p) \geqslant \Delta(q) \tag{206}$$

对所有 $p, q \in \mathbb{R}_+^n$均成立。假设存在一个连续映射 $\Lambda : \mathbb{R}_+^n \to \mathbb{R}$，使得$\Delta(p) = \Lambda(p)$ 对所有 $p \in \mathbb{Q}_+^n$均成立。那么，对所有满足 $\|p\|_1 \in \mathbb{Q}$的 $p \in \mathbb{R}_+^n$，均有 $\Delta(p) = \Lambda(p)$。

证明。我们通过证明以下命题来证明该断言：对于每个 $p \in \mathbb{R}_+^n$ with$\|p\|_1 \in \mathbb{Q}$ 和 $\varepsilon > 0$，存在 q、$r \in \mathbb{Q}_+^n$，使得 $q \succeq_{0,1} p \succeq_{0,1} r$ 并且

[8] 请注意，当 $\alpha < 0$ 时，关于相同的支撑大小的假设（即可以假设上述 p 和 q 的所有元素均非零）至关重要；否则，我们最终会得到形如 0^α 的无定义的表达式。

$\|q-r\|_1 \leqslant \varepsilon$ 以及 $\|q-p\|_1 \leqslant \varepsilon$。这是足够的，因为将 Δ 和 Λ 应用于该不等式可得

$$\Lambda(q) = \Delta(q) \geqslant \Delta(p) \geqslant \Delta(r) = \Lambda(r), \tag{207}$$

其中我们使用了 Δ 的 $\geq_{0,1}$-单调性。由于固定归一化的非负向量集合是紧的，我们知道 Λ 在其上是一致连续的。因此取 $\varepsilon \to 0$ 后可得 $\Delta(p) = \Lambda(p)$。所以，只需证明存在如上所述的 q 和 r。

为此，我们不妨自由地假设给定的 $p = (p_1、\ldots、p_m、0、\ldots、0) \in \mathbb{R}_+^n$ 已按非递增顺序排列，且满足 $p_m > 0$，因为 Δ 的单调性意味着 Δ 在置换下保持不变。将范围限制为 $\mathbb{R}_+^m$ 后，我们还可假设 $m = n$，从而 p 具有全支撑。将所有量按有理数 $\|p\|_1^{-1}$ 进行缩放后，我们还可假设 $\|p\|_1 = 1$。对于给定的 $\varepsilon > 0$，令 $\delta := \frac{\varepsilon}{4(n-1)}$。对于每个 $i = 1、\ldots、n-1$，我们取一个有理数 $q_i \in [p_i、\ p_i + \delta]$，使得 $q_1 \geqslant \cdots \geqslant q_{n-1}$。然后定义有理数 $q_n := 1 - \sum_{i=1}^{n-1} q_i$，根据 $\|p\|_1 = 1$，它满足

$$p_n - (n-1)\delta \leqslant q_n \leqslant p_n. \tag{208}$$

对于充分小的 ε，这尤其保证了 $q_n > 0$，因此 $q = (q_1、\ldots\ldots、q_n)$ 是一个概率向量，并且 $\|q\|_0 = \|p\|_0$。q 的各个分量也已经按非递增顺序排列，而且根据构造，对于所有 $k = 1、\ldots\ldots、n-1$，我们都有 $\sum_{i=1}^k q_i \geqslant \sum_{i=1}^k p_i$。因此 $p \preceq_{0,\,1} q$。此外，

$$\|q-p\|_1 = \sum_{i=1}^{n-1}(q_i - p_i) + p_n - q_n \leqslant 2(n-1)\delta = \frac{\varepsilon}{2}. \tag{209}$$

为了构造所需的向量 r，我们还假设 p 不是均匀的，因为在均匀情形下我们只需取 $r = p$ 即可。于是令 ℓ 为满足 $p_\ell > p_{\ell+1} = p_n$ 的唯一索引（即，p 中的最后 $n-\ell$ 个分量与最低概率 p_n 相同）。定义

$$\delta' := \min\left\{\frac{\varepsilon}{4\ell}, \frac{n-\ell}{n}(p_\ell - p_{\ell+1})\right\}. \tag{210}$$

随后，对于所有 $i = 1、\ldots\ldots、\ell$，我们取正有理数 r_i，使得 $\frac{n-\ell-1}{n-\ell}\delta' \leqslant p_i - r_i \leqslant \delta'$ 且 $r_1 \geqslant \cdots \geqslant r_\ell$。我们还令

$$r_{\ell+1} = \cdots = r_n = \frac{1}{n-\ell}\left(1 - \sum_{i=1}^{\ell} r_i\right) \leqslant p_n + \frac{\ell}{n-\ell}\delta'. \tag{211}$$

δ' 的定义保证

$$r_\ell \geqslant p_\ell - \delta' \geqslant p_{\ell+1} + \frac{\ell}{n-\ell}\delta' \geqslant r_{\ell+1} = \cdots = r_n, \tag{212}$$

因此整个向量 $r := (r_1, \dots, r_n)$ 也按非递增顺序排列，且所有分量均非零。我们立即得到，对于所有$k = 1, \dots, \ell$，都有 $\sum_{i=1}^{k} p_i \geqslant \sum_{i=1}^{k} r_i$。如果 $\ell = n - 1$，这就意味着$p \succeq_{0,\,1} r$。现在我们通过证明其余不等式，来说明一般情形下同样成立。对于 $j = 1, \dots, n-\ell-1$,

$$\sum_{i=1}^{\ell+j} r_i = \sum_{i=1}^{\ell} r_i + \sum_{i=\ell+1}^{\ell+j} r_i \tag{213}$$

$$\leqslant \sum_{i=1}^{\ell} p_i - \ell \frac{n-\ell-1}{n-\ell}\delta' + \sum_{i=\ell+1}^{\ell+j} p_i + j\frac{\ell}{n-\ell}\delta' \tag{214}$$

$$= \sum_{i=1}^{\ell+j} p_i - \ell \frac{n-\ell-1-j}{n-\ell}\delta' \leqslant \sum_{i=1}^{\ell+j} p_i, \tag{215}$$

其中我们在最终的不等式中使用了 $j \leqslant n-\ell-1$，这是在归一化方程 $\sum_{i=1}^{n} r_i = \sum_{i=1}^{n} p_i$ 之前需要考虑的最后一种情况。因此，$p \succeq_{0,\,1} r$。此外，

$$\|p - r\|_1 = \sum_{i=1}^{\ell}(p_i - r_i) + \sum_{i=\ell+1}^{n}(r_i - p_i) \leqslant 2\ell\delta' \leqslant \frac{\varepsilon}{2}, \tag{216}$$

其中，最后一个不等式由 δ' 的定义得出。

总之，我们有 $q \succeq_{0,\,1} p \succeq_{0,\,1} r$ 以及 $\|q - r\|_1 \leqslant \|q - p\|_1 + \|p - r\|_1 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$，因此辅助结果得证。□

现在我们继续研究这两类导子。导子 Δ_0、Δ_1 的莱布尼茨法则：在 $\|\cdot\|_1$ 处，分别为 $S_{0,\,1} \to \mathbb{R}$ 和 $\|\cdot\|_0$，现在可写为

$$\Delta_1(p \otimes q) = \Delta_1(p)\|q\|_1 + \|p\|_1\Delta_1(q), \tag{217}$$

$$\Delta_0(p \otimes q) = \Delta_0(p)\|q\|_0 + \|p\|_0\Delta_0(q) \tag{218}$$

对于所有 p，$q \in \mathcal{V}_{<\infty}$。容易验证，香农熵 H_1满足（217），而（190）中的函数 H_0' 满足（218）。接下来我们将证明，除缩放因子外，这些是唯一的单调导子，至少在有理数 1-范数的那些向量上如此。不过，缩放因子为负，以适应 $S_{0,\,1}$ 的排序。

命题 33. 在有理数 1-范数的向量上，单调导子Δ：$S_{0,\,1} \to \mathbb{R}$ 恰为以下几种：

(i) 在 $\|\cdot\|_1$ 处，H_1 的负倍数。(ii) 在 $\|\cdot\|_0$ 处，H_0' 的负倍数。

证明。对于 (i)，众所周知，H_1 在主序下是单调的，即若 $p \succeq_{0,1} q$（甚至只需 $p \succeq_1 q$），则 $H_1(p) \leqslant H_1(q)$。因此，假设

满足 $\Delta_1 : S_{0,1} \to \mathbb{R}$ 是 $\|\cdot\|_1$ 处的任意单调导子。令 δ_1 为 Δ_1的单项限制，即对所有 $x \in \mathbb{R}_+$， $\delta_1(x) \coloneqq \Delta_1\big([(x)]\big)$ 。 δ_1 的莱布尼茨法则可写为

$$\delta_1(xy) = \delta_1(x)y + x\delta_1(y) \qquad \forall x, y \geqslant 0. \tag{219}$$

容易看出，对于所有 $x, y > 0$ 和 $m, n \in \mathbb{N}$，

$$\Big(\underbrace{\frac{m}{m+n}x + \frac{n}{m+n}y, \ldots, \frac{m}{m+n}x + \frac{n}{m+n}y}_{m+n \text{ copies}}\Big) \preceq_{0,1} (\underbrace{x, \ldots, x}_{m \text{ copies}}, \underbrace{y, \ldots, y}_{n \text{ copies}}). \tag{220}$$

将 Δ_1 应用于此不等式，并利用单调性和可加性，我们得到

$$\delta_1\big(tx + (1-t)y\big) \leqslant t\delta_1(x) + (1-t)\delta_1(y) \tag{221}$$

对于所有 $t \in [0,1] \cap \mathbb{Q}$ 和 $x, y > 0$。接下来我们证明 $\delta_1(\frac{1}{2}) \leqslant 0$。由于

$$0 = \delta_1(1) = \delta_1\left(2 \cdot \frac{1}{2}\right) = 2\delta_1\left(\frac{1}{2}\right) + \frac{1}{2}\delta_1(2), \tag{222}$$

我们有 $\delta_1(2) = -4\delta_1(\frac{1}{2})$。利用这一点和（221）可得

$$0 = \delta_1(1) = \delta_1\left(\frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 2\right) \leqslant \frac{2}{3}\delta_1\left(\frac{1}{2}\right) + \frac{1}{3}\delta_1(2) \tag{223}$$

$$= \left(\frac{2}{3} - \frac{4}{3}\right)\delta_1\left(\frac{1}{2}\right) = -\frac{2}{3}\delta_1\left(\frac{1}{2}\right). \tag{224}$$

因此，$\delta_1(\frac{1}{2}) \leqslant 0$。但由 $\delta_1(1) = 0$以及再次应用（221），我们得到 $\delta_1(x) \leqslant 0$对所有有理 $x \in [\frac{1}{2}, 1]$，莱布尼茨法则则将此不等式推广到所有有理 $x \in [0, 1]$。

现在考虑 $\eta(x) \coloneqq \frac{\delta_1(x)}{x}$ for $x > 0$。上述针对 δ_1 的莱布尼茨法则表明，该函数满足柯西函数方程的乘法到加法形式，

$$\eta(xy) = \eta(x) + \eta(y) \qquad \forall x, y > 0. \tag{225}$$

现在我们将论证，对于所有 $x \in \mathbb{Q}_{>0}$，都有 $\eta(x) = c\log x$ for，并且某个固定的 $c \geqslant 0$。由于对于有理数 $x \in (0, 1]$，$\eta(x) \leqslant 0$，函数方程意味着 η 在有理数自变量上单调不减。现在我们可以按照柯西函数方程 [1, p. 14] 的标准解法继续：假设存在 x、$y \in \mathbb{Q}_{>0}$，使得这些点

$$(\log x, \eta(x)), \qquad (\log y, \eta(y)) \tag{226}$$

在 $\mathbb{R}^2$ 中线性无关。因此，它们张成整个 $\mathbb{R}^2$，特别是我们可以找到系数 ℓ_x、$\ell_y \in \mathbb{Z}$，使得

$$\ell_x \log x + \ell_y \log y \geqslant 0, \qquad \ell_x\eta(x) + \ell_y\eta(y) < 0. \tag{227}$$

这意味着正有理数 $z := x^{\ell_x} y^{\ell_y}$ 满足 $z \geqslant 1$ 和 $\eta(z) < 0$，这与 $\eta(1) = 0$ 以及 η的单调性矛盾。因此，点 (226) 线性相关，或者等价地

$$\frac{\eta(x)}{\log x} = \frac{\eta(y)}{\log y} \qquad \forall x, y > 0. \tag{228}$$

固定例如 $y := 2$，可知我们有 $\eta(x) = c \log x$，其中 $c = \frac{\eta(2)}{\log 2} \geqslant 0$。但由此可得

$$\delta_1(x) = cx \log x \qquad \forall x \in \mathbb{Q}_+. \tag{229}$$

由 Δ_1 的可加性，对于所有具有有理数项的 $p \in \mathcal{V}_{<\infty}$，我们因此有，

$$\Delta_1(p) = \sum_i \delta_1(p_i) = c \sum_i p_i \log p_i, \tag{230}$$

或者等价地 $\Delta_1(p) = -cH_1(p)$ 对于所有 $p \in \mathbb{Q}_+^n$，$n \in \mathbb{N}$。注意到 H_1 是连续的，且 Δ_1 是单调的，引理 32 推出 Δ_1 与$-cH_1$ 在所有满足 $p \in \mathcal{V}_{<\infty}$ 的 $\|p\|_1 \in \mathbb{Q}$上重合。

对于 (ii)，我们首先证明 $H_0' : S_0^{\mathrm{op}}{}_{,\,1} \to \mathbb{R}$ 确实是单调的（即，H_0' 的负倍数是单调的）。为此，设 $p, q \in \mathcal{V}_{<\infty}$ 满足$p \succeq_{0,\,1} q$。因此存在 $n \in \mathbb{N}$，使得 $p = (p_1, \ldots, p_n) \in \mathbb{R}^n_{>0}$ 和 $q = (q_1, \ldots, q_n) \in \mathbb{R}^n_{>0}$，以及一个双随机矩阵 $T = (T_{i,\,j})_{i,\,j=1}^n$，满足 $Tp = q$；我们不妨假设 p 和 q 具有全支撑。我们可以计算

$$H_0'(p) = \sum_{j=1}^n \log p_j = \sum_{i,j=1}^n T_{i,j} \log p_j \tag{231}$$

$$\leqslant \sum_{i=1}^n \log \sum_{j=1}^n T_{i,j} p_j = \sum_{i=1}^n \log q_i = H_0'(q), \tag{232}$$

其中，在第二个等式中，我们用到了 T 的列之和为 1 这一事实；在不等式中，我们用到了 T 的行之和为 1 这一事实以及对数函数的凹性。因此，H_0' 确实是单调的。

现在令 $\Delta_0 : S_{0,1} \to \mathbb{R}$ 为 $\|\cdot\|_0$ 处的任意单调导子，并令 δ_0 为 Δ_0 的单项限制。特别地，可得 $\delta_0(0) = 0$ 和 $\delta_0(xy) = \delta_0(x) + \delta_0(y)$ 对所有 x, $y > 0$ 均成立。定义函数 $k\colon (0,1) \to \mathbb{R}$ 如下：

$$k(x) = \Delta_0(x, 1-x) = \delta_0(x) + \delta_0(1-x) = \delta_0(x - x^2). \tag{233}$$

对于每个 $t \in [0,1]$，考虑 $p_t = (1-t, t) \in \mathcal{V}_{<\infty}$。利用 Δ_0的单调性，以及对所有满足 $s \leqslant t$的 $s, t \in (0, \frac{1}{2}]$，$p_s \succeq_{0,1} p_t$这一事实，我们可知$k$ 在 $(0, \frac{1}{2}]$上是非递增的。因此，δ_0 在 $(0, \frac{1}{4}]$上是非递增的。但这样一来，$x \mapsto \delta_0(2^x)$ 就是一个可加函数 $\mathbb{R} \to \mathbb{R}$，它在某个非平凡区间上单调，因此具有 $x \mapsto -cx$ 的形式，其中 $c \in \mathbb{R}$为某个常数。由此可知，对于 $x > 0$，$\delta_0(x) = -c \log x$。因此 $\Delta_0 = -cH_0'$，而 Δ_0 的单调性意味着 $c \geqslant 0$。□

注 34. 下面概述如何利用香农熵的一个已知刻画，得到上述在 $\|\cdot\|_1$ 处单调导子的分类。设 Δ_1 是在 $\|\cdot\|_1$ 处的单调导子，且满足 $\Delta_1(\frac{1}{2},\frac{1}{2})=-1$。这个导子具有以下三个性质：

- 当将 Δ_1 视为定义在 $\mathcal{V}_{<\infty}$ 上的函数时（这是我们通常的做法），它是置换不变的，也就是说，论元向量各分量的排列顺序无关紧要。这仅仅是因为，在我们的框架中，Δ_1 形式上是定义在集合 $\mathcal{V}_{<\infty}/\approx$ 上的函数，而该集合由排列顺序不同的非负向量的等价类组成。函数 $h:\ [0,1]\to\mathbb{R}$，其中 $h(x)=-\Delta_1(x,1-x)$，是可测的。这是因为，作为实值函数，h 在区间 $(0,\frac{1}{2}]$ 上单调不减，在区间 $[\frac{1}{2},1)$ 上单调不增。进而，这是由 Δ_1 关于预序 $\geq_{0,\ 1}$ 的单调性以及以下事实得出的：

$$(1-x,x)\geq_{0,1}(1-y,y) \tag{234}$$

  当 $0<x\leqslant y\leqslant\frac{1}{2}$ 时如此，而当 $\frac{1}{2}\leqslant x\leqslant y<1$ 时顺序相反。此外，根据所假设的归一化，$h(\frac{1}{2})=1$。
- 对于每个 $p=(p_1、\ldots\ldots、p_n)\in\mathcal{V}_{<\infty}$ 和 $t\in[0、1]$，我们都有

$$\Delta_1\big(tp_1,(1-t)p_1,p_2,\ldots,p_n\big)=p_1h(t)+\Delta(p). \tag{235}$$

  这由 Δ_1 的莱布尼茨法则以及可加性得出。

根据 [20]，这意味着 $-\Delta_1$ 在 $\mathcal{P}_{<\infty}$ 上与香农熵 H_1 一致。这就是我们应用所需的全部内容。不过，利用证明命题33（i）末尾所用的相同技巧，我们可以证明 $-\Delta_1$ 与 H_1 在所有 $p\in\mathcal{V}_\infty$ 上 $\|p\|_1\in\mathbb{Q}$ 一致。

注35. 定理7 中的单调导子 Δ 被归一化为 $\Delta(u_+)=\Delta(u_-)+1$，其中 (u_-,u_+) 是选定的幂通用对。但当然，这种归一化是任意的；在将定理7应用于 $S_{0\text{中的概率向量时，}1}$ 我们同样可以将它们归一化，使相关的单调导子恰好为 $-H_1$ 和 $-H_0'$。

A.3. 主序半环上的单调同态与导子。接下来，对于主序半环 S_1，我们将推导出单调同态 Φ 的一个简单刻画：$S_1\to\mathbb{K}$，其中 $\mathbb{K}\in\{\mathbb{R}_+,\ \mathbb{R}_+^{\mathrm{op}},\ \mathbb{TR}_+,\ \mathbb{TR}_+^{\mathrm{op}}\}$，该刻画是通过映射（186）和（187）给出的。

命题36。非退化单调同态 $S_1\to\mathbb{K}$ 恰好如下：

（i）对于 $\mathbb{K}=\mathbb{R}_+$，映射 f_α 适用于 $\alpha\in(1,\infty)$。（ii）对于 $\mathbb{K}=\mathbb{R}_+^{\mathrm{op}}$，映射 f_α 适用于 $\alpha\in[0,1)$。（iii）对于 $\mathbb{K}=\mathbb{TR}_+$，映射 $f_\infty(\cdot)^\alpha$ 适用于 $\alpha\in(0,\infty)$。（iv）对于 $\mathbb{K}=\mathbb{TR}_+^{\mathrm{op}}$，不存在这样的映射。

唯一的退化同态 $S_1\to\mathbb{R}_+$ 是 $f_1=\|\cdot\|_1$。

证明。 由于主序预序 $\preccurlyeq_1$ 是修正序$\preccurlyeq_{0\text{的扩展},\ 1}$ 而我们已在命题 31中讨论过修正序，因此我们已经知道，每种情况下的单调同态都属于其中所考虑的那些同态。此外，非退化性显然得以保留，而退化同态$f_0 = \|\cdot\|_0$ 现在是取值于 $\mathbb{R}_+$的非退化单调同态。

因此，要证明情形 (ii) 中的结论，只需证明 f_α 对于$\alpha \in [0,\ 1)$ 仍然是单调的。但这可由前面相同的不等式（204）推出；该不等式仍然适用，因为这里并未用到所有概率均非零这一假设。

情形 (i) 也同样如此，此外我们还需要证明 f_α 在$\alpha \in (-\infty、\ 0)$ 下现在不再是单调的。为此，考虑以下二元结果向量及其在 f_α 下的像，

$$\begin{array}{ccccc} [(\frac{1}{2},\frac{1}{2})] & \preceq_1 & [(1-\varepsilon,\varepsilon)] & \preceq_1 & 1 \\ \downarrow f_\alpha & & \downarrow f_\alpha & & \downarrow f_\alpha \\ 2^{1+|\alpha|} & < & (1-\varepsilon)^{-|\alpha|}+\varepsilon^{-|\alpha|} & > & 1 \end{array} \tag{236}$$

其中，第二行中的不等式对所有充分小的 $\varepsilon > 0$ 都成立。因此，情形 (i) 恰好对应于函数 f_α 在 $\alpha \in (1、\ \infty)$ 下的情况。

情形 (iii) 中的结论可由命题 31(iii) 立即推出。在情形 (iv) 中，映射 f_∞ 及其正整数次幂仍然是单调的，由（205）可知，因为那里并未用到所有概率均非零这一假设。映射 $f_{-\infty}$ 及其正整数次幂则不再是单调的，将它作用于（236）中的相同向量即可看出。□

由于映射 $f_0 = \|\cdot\|_0$ 在 S_1上不再是退化的，我们只需考虑 $\|\cdot\|_1$处的单调导子。由于香农熵 H_1 仍然是单调的，命题 33立即推出以下结果。

命题 37. 在有理数 1-范数的向量上，单调导子$\Delta:\ S_1 \to \mathbb{R}$ 在 $\|\cdot\|_1$ 处恰好是 H_1 的负倍数。

因此，在定理 8 应用于 S_1中的归一化概率向量时，我们可以仅考虑香农熵 H_1 本身，将其作为唯一相关的导子。

A.4.次主序半环上的单调同态。现在我们将推导出单调同态φ 的简单刻画：$S \to \mathbb{K}$ 对于 $\mathbb{K} \in \{\mathbb{R}_+,\ \mathbb{TR}_+\}$。我们仍然用函数 f_α （见（186）和（187））来表述这一分类。事实证明， $\alpha \in [-\infty,\ 1)$ 现在已不再相关。

命题 38. 单调同态 $S \to \mathbb{K}$ 恰好如下：

(i) 对于 $\mathbb{K} = \mathbb{R}_+$，映射 f_α 适用于 $\alpha \in (1,\infty)$。

(ii) For $\mathbb{K} = \mathbb{TR}_+$*, the maps* $f_\infty(\cdot)^\alpha$ *for* $\alpha \in (0, \infty)$*.*

证明。 由于次主序预序 $\preccurlyeq$ 是主序预序 $\preccurlyeq_1$的扩展，因此每种情况下的单调同态都是命题 36中相应同态的子集，而且它们恰好仍然全部是单调的。在情形 (i) 中，关于 $\alpha \in [1, \infty)$ 的单调性证明基本与之前相同，细微差别在于第一步是使用 $q \leqslant Tp$ ，而不是代入 $q = Tp$。情形 (ii) 也类似。 □

AppendixB. Applications to majorization problems of probability vectors

在本节中，我们将定理5、7 和8的分离定理应用于我们的主序半环，以重新推导有关大样本主序和催化主序的结果。尽管这些结果此前已有所知，但本文的主要原创之处在于，这些结果都可以从同一个总体框架中得出。在探讨这些问题之前，我们先提出一个有用的观察。

注 39. 假设 p 和 q 是有限概率向量，且 p在大样本中主序于 q ，即存在 $n \in \mathbb{N}$ 使得 $p^{\otimes n} \succcurlyeq_1 q^{\otimes n}$。那么$p$ 也催化主序于 q，即存在一个有限概率向量 r 使得 $p \otimes r \succcurlyeq_1 q \otimes r$。为此，我们运用一个著名技巧[10]: [9] 定义

$$r := \bigoplus_{j=1}^{n} \left(p^{\otimes(n-j)} \otimes q^{\otimes(j-1)}\right), \qquad s := \bigoplus_{j=1}^{n-1} \left(p^{\otimes(n-j)} \otimes q^{\otimes j}\right). \tag{237}$$

于是很容易看出 $p \otimes r = p^{\otimes n} \oplus s$ 和 $r \otimes q = s \oplus q^{\otimes n}$，从而

$$p \otimes r = p^{\otimes n} \oplus s \succcurlyeq_1 q^{\otimes n} \oplus s = q \otimes r. \tag{238}$$

这就证明了该断言，因为可以对 r 进行归一化，而不影响不等式的有效性。

B.1. 渐近大样本主序与催化主序。对于给定的有限概率向量 $p \in \mathcal{P}_{<\infty}$，定义如下由 p 催化主序以及在大样本中被其主序的概率向量集合:

$$T_{<\infty}(p) := \{q \in \mathcal{P}_{<\infty} \mid q \otimes r \preceq_1 p \otimes r \text{ for some } r \in \mathcal{P}_{<\infty}\}, \tag{239}$$

$$M_{<\infty}(p) := \{q \in \mathcal{P}_{<\infty} \mid q^{\otimes n} \preceq_1 p^{\otimes n} \text{ for some } n \in \mathbb{N}_{>0}\}. \tag{240}$$

随后，我们将这些集合相对于全变差距离在 $\mathcal{P}_{<\infty}$ 中的闭包分别记为$\overline{T_{<\infty}(p)}$ 和 $\overline{M_{<\infty}(p)}$。

我们为 Aubrun 和 Nechita 的一项结果给出了新的证明[2] 。该结果用 Rényi 熵 H_α、$\alpha \geqslant 1$刻画渐近催化或大样本主序成立的条件，其中 R´enyi 熵定义于 $p = (p_1, \ldots\ldots, p_n) \in$

[9]另请参见[11] 中的引理5.4，了解该陈述的一般表述。

$\mathcal{P}_{<\infty}$ 通过

$$H_\alpha(p) = \log \|p\|_0 - D_\alpha(p\|u_{\operatorname{supp} p}) \tag{241}$$

$$= \begin{cases} \frac{1}{1-\alpha} \log \sum_{i=1}^n p_i^\alpha & \text{if } \alpha \in (0,1) \cup (1,\infty), \\ \log \|p\|_0 & \text{if } \alpha = 0, \\ -\sum_{i \in \operatorname{supp} p} p_i \log p_i & \text{if } \alpha = 1, \\ -\log \max_{1 \leqslant i \leqslant n} p_i & \text{if } \alpha = \infty. \end{cases} \tag{242}$$

不过请注意，我们下面只需要 H_α 和 $\alpha \geqslant 1$ 。

定理 40 （ [2]中的定理 1）。 对于任意一对有限概率向量p， $q \in \mathcal{P}_{<\infty}$以下各项彼此等价:

（i） $q \in \overline{M_{<\infty}(p)}$。 （ii） $q \in \overline{T_{<\infty}(p)}$。 （iii）对所有 $\alpha \geqslant 1$, $H_\alpha(p) \leqslant H_\alpha(q)$。

我们注意到，条件（iii）等价于 $f_\alpha(p) \geqslant f_\alpha(q)$ 对所有 $\alpha \geqslant 1$成立，这在 [2] 中表述为 $\|p\|_\alpha \geqslant \|q\|_\alpha$ ，其中 $\|\cdot\|_p$ 是有限概率向量通常的 ℓ^p-范数。我们对该定理的新证明采用了 [2], 中发展出的相同技术，但关键区别在于， [2] 中的作者在证明蕴含关系（iii）′（i）的关键步骤中使用了克莱默大偏差定理，而对我们而言，这一步可直接由定理 5 应用于S ，再结合我们此前根据命题 $S \to \mathbb{R}_+$ 和 $S \to \mathbb{TR}_+$ 38对单调同态的识别$S \to \mathbb{R}_+$ 以及 $S \to \mathbb{TR}_+$ ，即可得到。由于我们的新技术并未为蕴含关系（i） $\Longrightarrow$ （ii）带来任何新的内容，该关系基本上使用备注 39，而蕴含关系（ii） $\Longrightarrow$（iii）也是如此，因此我们仅关注（iii） $\Longrightarrow$ （i）。

证明。 （iii） $\Longrightarrow$ （i）： $f_\alpha(p) \geqslant f_\alpha(q)$ 对所有$\alpha \in [1, \infty]$。由于定理 5 要求严格不等式才能得出大样本下的排序，因此思路是在给定误差范围内，将 q 下界限定为某个 q'_ε ε ，以构造严格不等式 $f_\alpha(p) > f_\alpha(q'_\varepsilon)$。之后，将下界 q'_ε 补全为一个归一化向量，同时确保它在大样本下仍被 p所支配。

因此，设给定 $\varepsilon \in (0, q_{\min}\|q\|_0)$ ，其中 $q_{\min} = f_{-\infty}(q)^{-1}$ 是 q 中最小的非零元素。我们通过从 q 的所有非零元素中减去 $\frac{\varepsilon}{\|q\|_0}$ 来定义 $q'_\varepsilon \in \mathcal{V}_{<\infty}$ 。这是一个满足 $\|q - q'_\varepsilon\|_1 = \varepsilon$ 的次归一化向量。由于对于所有 $\alpha \in [1, \infty]$, 显然有 $f_\alpha(q'_\varepsilon) < f_\alpha(q)$ ，假设 (iii) 意味着对于所有 $\alpha \in [1, \infty]$, 严格不等式 $f_\alpha(q'_\varepsilon) < f_\alpha(p)$ 成立。将定理5应用于 S ，再结合命题 38，即可推出存在 $n_\varepsilon \in \mathbb{N}$ ，使得 $p^{\otimes n_\varepsilon}$ 主序于 $(q'_\varepsilon)^{\otimes n_\varepsilon}$。

现在，我们将 q'_ε 补全为归一化向量

$$q_\varepsilon := q'_\varepsilon \oplus \Big(\underbrace{\frac{\varepsilon}{m}, \ldots, \frac{\varepsilon}{m}}_{m \text{ copies}}\Big), \tag{243}$$

也就是说，我们向 q'_ε 追加 m 个值为 $\frac{\varepsilon}{m}$ 的项，其中假设 m 足够大，以确保 $p^{\otimes n_\varepsilon}$ 仍主序于 $q_\varepsilon^{\otimes n_\varepsilon}$。一种做法是要求

$$\frac{\varepsilon}{m} \leqslant \min\{(q'_\varepsilon)_{\min}, p_{\min}^{n_\varepsilon}\}. \tag{244}$$

下面证明这样做确实有效。我们断言，只要 $k \geqslant 1$,

$$\sum_{i=1}^{k} (q_\varepsilon^{\otimes n_\varepsilon})_i^{\downarrow} \leqslant \sum_{i=1}^{k} (p^{\otimes n_\varepsilon})_i^{\downarrow}. \tag{245}$$

事实上， m 的选取使得 $q_\varepsilon^{\otimes n_\varepsilon}$ 中最大的 $\|q\|_0^{n_\varepsilon}$ 个项恰好是 $(q'_\varepsilon)^{\otimes n_\varepsilon}$ 中的项，因此当 $1 \leqslant k \leqslant \|q\|_0^{n_\varepsilon}$ 时，不等式 (245) 可由 $p^{\otimes n_\varepsilon}$ 主序于 $(q'_\varepsilon)^{\otimes n_\varepsilon}$ 这一事实推出。如果 $\|q\|_0^{n_\varepsilon} < k \leqslant \|p\|_0^{n_\varepsilon}$，则由归纳法可得该不等式，因为 m 的选取保证了 $(q_\varepsilon^{\otimes n_\varepsilon})_k^{\downarrow} \leqslant \frac{\varepsilon}{m} \leqslant (p^{\otimes n_\varepsilon})_k^{\downarrow}$。最后，如果 $k > \|p\|_0^{n_\varepsilon}$，则 (245) 显然成立，因为右侧等于 1。总之，$p^{\otimes n_\varepsilon}$ 主序于 $q_\varepsilon^{\otimes n_\varepsilon}$，因此 $q_\varepsilon \in M_{<\infty}(y)$。但构造 q_ε 时使其满足 $\|q - q_\varepsilon\|_1 \leqslant \varepsilon$。由于 ε 是任意的，我们得出 $q \in \overline{M_{<\infty}(p)}$，这表明性质 (i) 成立。
□

B.2. 精确催化主序。 现在我们转而讨论有限概率分布的精确催化主序刻画问题。回顾一下式（´瑞❖伊散度 $D_\alpha(\cdot\|\cdot)$， $\alpha \geqslant 0$（158）。注意，只要 $0 < \alpha < 1$ 且 $\operatorname{supp} p \cap \operatorname{supp} q \neq \varnothing$，就有

$$D_\alpha(p\|q) = \frac{\alpha}{1-\alpha} D_{1-\alpha}(q\|p). \tag{246}$$

利用这些映射，我们对 [17] 的主要结果稍作改述：

定理41. 设 $p, q \in \mathcal{P}_{<\infty}$，并令 $u := u_{\operatorname{supp} p \cup \text{支撑集 } q}$ 为 p 和 q 的支撑集之并上的均匀分布。若对所有 $\alpha \in [\frac{1}{2}, \infty)$,

$$D_\alpha(p\|u) > D_\alpha(q\|u), \tag{247}$$

$$D_\alpha(u\|p) > D_\alpha(u\|q), \tag{248}$$

则存在 $r \in \mathcal{P}_{<\infty}$，使得 $p \otimes r$ 主序于 $q \otimes r$。反之，若存在这样的 r，则上述不等式至少以非严格形式成立。

请注意，不等式（248）左边可能取值∞；在这种情况下，只要右边有限，这些条件就成立。值得指出的是，我们可以对定理 41中的向量 p 和 q 进行排序，使其中一个向量的支撑集包含于另一个向量的支撑集，例如，选择 $p = p^{\downarrow}$ 和 $q = q^{\downarrow}$。在

这样， u 就是 $\operatorname{supp} p$ 和 $\operatorname{supp}\ q$中较大支撑集上的均匀分布。如果不等式成立，那么该较大支撑集总是 $\operatorname{supp} q$，因为 $\|p\|_0 > \|q\|_0$ 会推出 $D_\alpha(u\|p) < \infty = D_\alpha(u\|q)$，这与（248）矛盾。如果排序方式不够恰当，导致两个支撑集互不包含，那么（248）两边都会变为无穷大，此时该定理不适用。

注 42。 定理 41中给出的充分条件实际上也基本上是必要的。事实上，对于 $p^{\downarrow} \neq q^{\downarrow}$，催化主序已蕴含 (247) 和 (248) 中的严格不等式，这得益于 Rényi 散度的严格联合凸性 ´[17]。不过，在本文中，我们着重于充分性的证明，并讨论逆命题仅仅是为了说明这些充分条件与必要条件非常接近。

在证明定理 41之前，我们以一种等价的方式重述该命题，使其与函数 f_α、H_1和 H_0' 之间的联系一目了然：

定理 43。设 p、 $q \in \mathcal{P}_{<\infty}$ 是有限概率向量。假设以下条件成立：

(i) 对所有 $\alpha \in (0,\ 1)$，均有 $f_\alpha(p) < f_\alpha(q)$；(ii) $H_1(p) < H_1(q)$; (iii) 对所有 $\alpha \in (1,\ \infty)$，均有 $f_\alpha(p) > f_\alpha(q)$，

并且以下任一条件成立：

(iv) $\|p\|_0 < \|q\|_0$, *or*

(v) $\|p\|_0 = \|q\|_0$ *and* $H_0'(p) < H_0'(q)$ *and* $f_\alpha(p) > f_\alpha(q)$ *for all* $\alpha \in (-\infty, 0)$.

于是存在一个有限概率向量 r，使得 $p \otimes r$ 主序于 $q \otimes r$。

反之，如果存在这样的 r，那么上述严格不等式以非严格形式成立。

下面说明该陈述的正向结论如何等价于定理 41。设 p、 q 和 u 如定理 41中所述，并且不妨设 $\operatorname{supp} p \subseteq \operatorname{supp}\ q$。如果 $\|p\|_0 = \|q\|_0$，则不等式（247）和（248）中不存在无穷值，并且

$$D_\alpha(p\|u) = \frac{1}{\alpha - 1} \log f_\alpha(p) + \log \|p\|_0, \tag{249}$$

$$D_\alpha(u\|p) = \frac{1}{\alpha - 1} \log f_{1-\alpha}(p) + \frac{\alpha}{1 - \alpha} \log \|p\|_0 \tag{250}$$

对所有 $\alpha \neq 1$，并且

$$D_1(p\|u) = \log \|p\|_0 - H_1(p), \tag{251}$$

$$D_1(u\|p) = -\log \|p\|_0 - \frac{1}{\|p\|_0} H_0'(p), \tag{252}$$

对于 q 也同理。考虑到 $\alpha - 1$ 的符号，我们得到定理 43 与定理 41 在 $\|p\|_0 = \|q\|_0$ 情形下等价，因为例如不等式 $D_1(u\|p) > D_1(u\|q)$ 会变为 $H_0'(p) < H_0'(q)$。

若 $\|p\|_0 < \|q\|_0$，则不等式（248）自动成立，因为左侧为无穷大，而右侧为有限值。因此，只有不等式（247）相关，由此得到定理 43 的（i）—（iii）项。最后，如上所述，$\|p\|_0 > \|q\|_0$ 被两个定理的假设排除。

现在我们证明定理 43。

证明。我们先考虑 $\|p\|_0 < \|q\|_0$ 的情形。由于对于 $\|\cdot\|_0$存在严格不等式，相关的半环为 S_1。根据命题 36，此时非退化单调同态的集合由以下元素构成：$f_\alpha : S_1 \to \mathbb{R}_+$ 对于 $\alpha \in (1, \infty)$ 以及 $f_\alpha : S_1 \to \mathbb{R}_+^{\mathrm{op}}$ 对于 $\alpha \in [0, 1)$ 还有$f_\infty : S_1 \to \mathbb{TR}_+$；而在退化同态 $\|\cdot\|_1$ 处唯一相关的导子是 $-H_1$，由命题 37给出。回想一下，

$$f_\infty(r) = \|r\|_\infty = \lim_{\alpha\to\infty} \|r\|_\alpha = \lim_{\alpha\to\infty} f_\alpha(r)^{1/\alpha} \tag{253}$$

对于所有 $r \in \mathcal{V}_{<\infty}$，其中 $\|\cdot\|_\alpha$对于 $\alpha \in [1, \infty]$是通常的 α-范数，由此可得 $f_\infty(p) \geqslant f_\infty(q)$。因此，若 $f_\infty(p) > f_\infty(q)$，则将定理 7 应用于 S_1 并利用命题 36 和 37，便可立即得到结论。于是考虑子情形 $f_\infty(p) = f_\infty(q)$。我们还可以不失一般性地假设 $p = p^\downarrow = (p_1, \ldots, p_n)$以及 $q = q^\downarrow = (q_1, \ldots, q_n)$，并且 $p \neq q$。我们还记 $I := \{1, \ldots, n\}$为不相交并集 $I = I_{\max} \cup I_0$，其中 $I_{\max} = \{1, \ldots, m\}$是满足条件 $p_i = q_i$ 对于 $i \in I_{\max}$的最长初始序列，因此 $p_{m+1} \neq q_{m+1}$。我们采用以下记号

$$p' := (p_i)_{i\in I_0}, \tag{254}$$

$$q' := (q_i)_{i\in I_0}, \tag{255}$$

$$p'' = q'' := (p_i)_{i\in I_{\max}} = (q_i)_{i\in I_{\max}}. \tag{256}$$

现在，对于所有 $\alpha \in [0、 1)$,

$$f_\alpha(p') + f_\alpha(p'') = f_\alpha(p) < f_\alpha(q) = f_\alpha(q') + f_\alpha(q'') = f_\alpha(q') + f_\alpha(p''), \tag{257}$$

其中我们利用了 f_α的可加性。因此，$f_\alpha(p') < f_\alpha(q')$ 对于所有$\alpha \in [0, 1)$。类似地，我们得到 $H_1(p') < H_1(q')$ 以及 $f_\alpha(p') > f_\alpha(q')$ 对于所有$\alpha > 1$。利用式（253）中的极限，我们得出 $f_\infty(p') \geqslant f_\infty(q')$，即 $p_{m+1} \geqslant q_{m+1}$。由于这些项互不相同，必有 $p_{m+1} > q_{m+1}$，等价地$f_\infty(p') > f_\infty(q')$。因此，可将定理7 应用于 p'，$q' \in S_1$，从而存在 $r \in \mathcal{V}_{<\infty}\backslash\{0\}$ 使得 $p' \otimes r \geq_1 q' \otimes r$。显然，通过重新归一化，我们可以假设 $\|r\|_1 = 1$。综上可得

$$p \otimes r = (p' \otimes r) \oplus (p'' \otimes r) = (p' \otimes r) \oplus (q'' \otimes r) \tag{258}$$

$$\geq_1 (q' \otimes r) \oplus (q'' \otimes r) = q \otimes r, \tag{259}$$

正如所要证明的。

还需考虑以下情形 $\|p\|_0 = \|q\|_0$，其处理方式类似。此时相关的半环为$S_{0,1}$，因为 $\|\cdot\|_0$是退化的。除了 S_1上的单调同态之外，根据命题31，现在还需要 $f_\alpha : S_{0,1} \to \mathbb{R}_+$对于 $\alpha \in (-\infty, 0)$ 以及 $f_{-\infty}$，而在现在这个退化的 $\|\cdot\|_0$ 上，唯一相关的单调导子是 $-H_0'$，根据命题33。利用以下事实：对于所有 $r = (r_1, \ldots, r_n) \in \mathcal{V}_{<\infty}$ 满足 $\tilde{r} := (1/r_i)_{i\in\operatorname{supp} r}$ 我们有

$$f_{-\infty}(r) = \|\tilde{r}\|_\infty = \lim_{\alpha\to-\infty} \|\tilde{r}\|_{-\alpha} = \lim_{\alpha\to-\infty} f_\alpha(\tilde{r})^{-1/\alpha}, \tag{260}$$

由此可得，$f_{-\infty}(p) \geqslant f_{-\infty}(q)$ 除了 $f_\infty(p) \geqslant f_\infty(q)$ 之外，还可由（253）得到。若 $f_{\pm\infty}(p) > f_{\pm\infty}(q)$，则再应用一次定理 7即可得出结论。一般而言，我们可以再次假设 $p = p^\downarrow$ 且 $q = q^\downarrow$，并将这些向量的公共支撑 $I = \{1, \ldots, n\}$ 写成不相交并 $I = I_{\max} \cup I_0 \cup I_{\min}$，其中 $I_{\max} = \{1, \ldots, \ell\}$ 是 I的最长初始部分，而 $I_{\min} = \{\ell + m + 1, \ldots, n\}$ 是 I的最长末尾部分，在该部分上，对所有 $i \in I_{\max} \cup I_{\min}$，均有 $p_i = q_i$。特别地，我们有 $p_{\ell+1} \neq q_{\ell+1}$ 和 $p_{\ell+m} \neq q_{\ell+m}$。现在记

$$p' := (p_i)_{i\in I_0}, \tag{261}$$

$$q' := (q_i)_{i\in I_0}, \tag{262}$$

$$p'' = q'' := (p_i)_{i\in I_{\max}\cup I_{\min}} = (q_i)_{i\in I_{\max}\cup I_{\min}}. \tag{263}$$

与第一种情形类似，p' 和 q' 满足相同的不等式 (i)–(iii) 和 (iv)。此外，再次利用 (253) 和 (260) 中的极限，我们有 $f_{\pm\infty}(p') \geqslant f_{\pm\infty}(q')$，或者等价地，$p_{\ell+1} \geqslant q_{\ell+1}$ 和 $p_{\ell+m} \leqslant q_{\ell+m}$。由于这些不等式中没有等号，我们得出它们都严格成立，或者等价地，$f_{\pm\infty}(p') > f_{\pm\infty}(q')$。现在将定理 7 应用于$S_{0,1}$，可得 $r \in \mathcal{V}_{<\infty}$，使得$p' \otimes r \succeq_{0,1} q' \otimes r$。对 r 进行归一化，并按第一种情形中的方法继续推导，即可证明该结论。□

B.3. 精确大样本主序。现在我们转而讨论大样本中的精确主序。以下结果最早见于[16]。

定理 44（[16] 中的命题 3.7）。设 $p, q \in \mathcal{P}_{<\infty}$。若对所有 $\alpha \in [0, \infty]$，

$$H_\alpha(p) < H_\alpha(q), \tag{264}$$

则对于所有充分大的 $n \in \mathbb{N}$，$p^{\otimes n}$ 主序于 $q^{\otimes n}$。反之，若对所有 $n \in \mathbb{N}$ 都成立，则上述不等式以非严格形式成立。

在 [16], 中，上述大样本变换条件表述为

$$\min_{\alpha\in[0,\infty]} \frac{H_\alpha(q)}{H_\alpha(p)} > 1. \tag{265}$$

但是，由于映射族 $\{H_\alpha\}_{\alpha\in(0,\infty]}$ 在使求值映射连续的最弱拓扑下构成一个紧拓扑空间[10]，，上述最小值的存在性已由（264）中的严格不等式所蕴含。利用这一事实，

$$H_\alpha(r) = \frac{1}{1-\alpha}\log f_\alpha(r) \tag{266}$$

对于所有 $r \in \mathcal{P}_{<\infty}$ 和 $\alpha \in (0,\ 1)\cup(1,\ \infty)$，我们还可以将定理44 表述如下：

定理 45. 设 $p, q \in \mathcal{P}_{<\infty}$ 是有限概率向量。若

$$f_\alpha(p) < f_\alpha(q) \qquad \forall \alpha \in [0,1), \tag{267}$$

$$f_\alpha(p) > f_\alpha(q) \qquad \forall \alpha \in (1,\infty], \tag{268}$$

$$H_1(p) < H_1(q), \tag{269}$$

则对于所有充分大的 $n \in \mathbb{N}$，$p^{\otimes n}$ 主序于 $q^{\otimes n}$。反之，若对所有 $n \in \mathbb{N}$ 这种情况都成立，则上述不等式以非严格形式成立。

这就是我们现在要证明的定理版本。注意，导子H_0' 并未直接涉及这些条件；相反，$\alpha = 0$ 处的相关量是 H_0，也就是说，本质上是 $\|\cdot\|_0$。

证明. 设我们假定 (267)、(268) 和 (269) 中的不等式成立；特别是$\|p\|_0 < \|q\|_0$。因此，相关的预序半环是 S_1，其中根据命题 36，非退化单调同态集合 $\varphi : S_1 \to \mathbb{K} \in \{\mathbb{R}_+, \mathbb{R}_+^{\mathrm{op}}, \mathbb{TR}_+, \mathbb{TR}_+^{\mathrm{op}}\}$ 由 f_α（其中 $\alpha \in [0,1)\cup(1,\infty]$）构成，而在退化的 $\|\cdot\|_1$ 处唯一相关的单调导子是 $-H_1$。由于 $\|q\|_0 > \|p\|_0$，我们知道 $[q] \neq 1$，因此根据注 29，q 在 S_1^{op} 中是幂通用的。这就是我们必须考虑 S_1^{op}，以便应用定理 8 的原因。

从 S_1^{op} 转到 S_1，我们仍可仿照命题 36 和 37，对 S_1^{op} 到 $\mathbb{K} \in \{\mathbb{R}_+、\mathbb{R}_+^{\mathrm{op}}、\mathbb{TR}_+、\mathbb{TR}_+^{\mathrm{op}}\}$ 的单调同态及相关导子进行刻画；不过，$\mathbb{R}_+$ 和 $\mathbb{TR}_+$ 分别与 $\mathbb{R}_+^{\mathrm{op}}$ 和 $\mathbb{TR}_+^{\mathrm{op}}$ 互换位置，且导子的符号改变。这意味着，非退化单调同态 $\Phi : S_1^{\mathrm{op}} \to \mathbb{K}$ 由 $\Phi = f_\alpha$ 给出，其中当 $\mathbb{K} = \mathbb{R}_+$时，$\alpha \in [0、1)$；当 $\mathbb{K} = \mathbb{R}_+^{\mathrm{op}}$ 时，$\Phi = f_\alpha$ 给出，其中 $\alpha \in (1、\infty)$；当 $\mathbb{K} = \mathbb{TR}_+^{\mathrm{op}}$ 时，则由 $\Phi = f_\infty$ 给出；而当$\mathbb{K} = \mathbb{TR}_+^{\mathrm{op}}$ 时，不存在这样的同态。唯一退化同态 $f_1 : S_1^{\mathrm{op}} \to \mathbb{R}_+$ 处的相关单调导子为 H_1（此时符号为正）。将定理 8 的 (a) 项应用于 S_1^{op}，即可得到所述结论。

逆向部分可直接由以下事实得出：H_α 在 $\otimes$下可加，在主序下单调。□

[10]这可以看作关于预序半环的另一条一般性事实的一个实例 [13,命题 8.5]。

注意，当 p，$q \in \mathcal{P}_{<\infty}$ 满足 $H_\alpha(p) < H_\alpha(q)$时，根据备注39 和定理8的(b)部分，$p \otimes r$ 主序于 $q \otimes r$ 其中 r 具有以下形式（237），其中 $n \in \mathbb{N}$充分大。

然而，利用我们关于矩阵主序的结果，即定理 19，结合定理 44，我们得到定理 44 的一个扩展，用以刻画大样本简单主序。与定理 41一样，我们假设：若不等式（270）或（271）左侧为无穷大而右侧有限，则该不等式成立。

推论46。设概率向量 p、$q \in \mathcal{P}_{<\infty}$。记 u 为 $\operatorname{supp} p \cup \operatorname{supp} q$ 上的均匀分布。每当

$$D_\alpha(p\|u) > D_\alpha(q\|u), \tag{270}$$
$$D_\alpha(u\|p) > D_\alpha(u\|q) \tag{271}$$

对所有 $\alpha \in [\frac{1}{2}, \infty]$ 均成立时，则对于充分大的 $n \in \mathbb{N}$，$p^{\otimes n}$ 主序于 $q^{\otimes n}$。

证明。我们仍可假设 $p = p^{\downarrow}$ 且 $q = q^{\downarrow}$。正如我们在定理 41之后的讨论中所指出的，现在可知，若条件（270）和（271）成立，则 $\operatorname{supp} p \subseteq$ 支撑集 q，因此我们可以假设 u 是 $\operatorname{supp} q$上的均匀分布。与我们在定理 43之后的讨论类似，现在可知，命题中的条件根据 p和 q的支撑集大小分为两种情况：

$$f_\alpha(p) < f_\alpha(q) \text{ for all } \alpha \in (0,1), \tag{272}$$
$$H_1(p) < H_1(q), \text{ and} \tag{273}$$
$$f_\alpha(p) > f_\alpha(q) \text{ for all } \alpha \in (1,\infty] \tag{274}$$

并且，此外，对于所有 $\alpha \in [-\infty, 0)$，$\|p\|_0 < \|q\|_0$ 或 $\|p\|_0 = \|q\|_0$ 且 $H_0'(p) < H_0'(q)$ 且$f_\alpha(p) > f_\alpha(q)$。

若 $\|p\|_0 < \|q\|_0$，则可直接由定理 44 得出该结论。因此，我们将重点讨论 $\|p\|_0 = \|q\|_0$ 的情况。现在，均匀分布 u 与 p 和 q 具有相同的支撑。回顾矩阵主序半环 S^2 的定义，我们现在看到 $[(p, u)]$、$[(q, u)] \in S^2$。根据推论23，(270) 和 (271) 中的条件意味着，对于充分大的 $n \in \mathbb{N}$，$(p^{\otimes n}$ 主序于 $(q^{\otimes n}$，其中 $u^{\otimes n})$；而由于 $u^{\otimes n}$ 是与 $p^{\otimes n}$ 和 $q^{\otimes n}$ 具有相同支撑的均匀分布，这意味着 $p^{\otimes n}$ 主序于 $q^{\otimes n}$。□

请注意，定理 41与推论 46 中的条件差异在于，端点 $\alpha = \infty$ 处的条件包含在推论 46 中，而定理 41 中则不包含。这与大样本主序蕴含催化主序这一事实相符。

Acknowledgment

T. F. 感谢 Omer Tamuz 和 Luciano Pomatto 富有见地的讨论。E. H. 和 M. T. 获得新加坡国家研究基金会与 A*STAR 通过其 CQT 衔接资助提供的支持。M. T. 还获得新加坡国立大学启动经费的支持（ R-263-000-E32-133 和 R-263-000-E32-731）。

References

[1]J. Aczel 和 J. Dhombres。《多变量函数方程》。剑桥大学出版社（1989）。[2]G. Aubrun 和 I. Nechita。“催化主序与 ℓ_p 范数”。《数学物理通讯》278(1)：133–144（2007）。[3] D. Blackwell。“实验比较”。载于《第二届伯克利数学统计与概率研讨会论文集》，1950 年，第 93–102 页，（1951）。[4] D. Blackwell。“等价实验比较”。《数理统计年刊》24：265–272，（1953）。[5]F. G. S. L. Brandao、M. Horodecki、N. H. Y. Ng、J. Oppenheim 和 S. Wehner。“量子热力学第二定律”。《美国国家科学院院刊》112(11)：3275–3279（2014）。[6] G. Bunth 和 P. Vrana。“多状态盒的渐近相对次主序”。《数学物理快报》111(4)：1–23（2021）。[7] A. Chefles。“量子布莱克威尔定理与最小错误状态判别”，（2009）。DOI：10.48550/arXiv.0907.0866。[8] G. Dahl。“矩阵主序”。《线性代数及其应用》288：53–73（1999）。[9] S. Du、Z. Bai 和 Y. Guo。“非相干操作下相干性变换的条件”。《物理评论 A》91：052120（2015）。[10] R. Duan、Y. Feng、X. Li 和 M. Ying。“多副本纠缠变换与纠缠催化”。《物理评论 A》71：042319（2005）。[11]T. Fritz。“资源可转换性与有序交换幺半群”。《计算机科学中的数学结构》27(6)：850–938（2017）。[12] T. Fritz。“预序半域与半环的抽象比较定理 I”，（2020）。DOI：10.48550/arXiv.2112.05949。[13] T. Fritz。“预序半域与半环的抽象比较定理 II”，（2021）。DOI：10.48550/arXiv.2112.05949。[14]G. Gour 和 M. Tomamichel。“基于信息论原理的熵与相对熵”。IEEE 信息论汇刊 67(10)：6313–6327（2021）。[15] H. Heyer。《有限实验比较》。施普林格纽约（1982）。[16]A. K. Jensen。“有限概率分布的渐近主序”。IEEE 信息论汇刊 65(12)：8131–8139（2019）。[17] M. Klimesh。“共同完全刻画催化主序关系的不等式”，（2007）。[18] Y. Kuramochi。“冯·诺依曼代数上的极小充分统计实验”。《数学物理杂志》58(6)：062203（2017）。[19] C. Leang 和 D. Johnson。“M 假设贝叶斯检测的渐近性质”。IEEE 信息论汇刊 43(1)：280–282（1997）。

[20] P. M. Lee。“信息论公理”。《数学统计年刊》35：415–418（1964）。[21] X. Mu、L. Pomatto、P. Strack 和 O. Tamuz。“从大样本中的 Blackwell 优势到 Rényi 散度，再回到 Blackwell 优势”。《计量经济学》89(1)：475–506（2020）。[22]M. A. Nielsen。“一类纠缠变换的条件”。《物理评论快报》83：436–439（1999）。[23] C. Perry、P. Vrana 和 A. H. Werner。“二分布半环与渐近相对次主序”。IEEE 信息论汇刊 68(1)：311–321（2022）。[24]E. Shmaya。“信息结构与完全正映射的比较”。《物理学杂志 A：数学与通论》38(44)：9717（2005）。[25] C. Stein。“理论统计研讨会笔记。I. 实验比较”，（1951）。[26] H. Strasser。《统计学数学理论：统计实验与渐近决策理论》。德古意特出版社（1985）。[27]M. Tomamichel。《有限资源下的量子信息处理——数学基础》。施普林格数学物理简明丛书第 5 卷，施普林格国际出版公司（2016）。[28] E. Torgersen。“测度族的主序与近似主序：在实验的局部比较及（舒尔凸性）向量主序理论中的应用”。载于 T. Cali´nski 和 W. Klonecki 编，《线性统计推断》，第 259–310 页，纽约州纽约市（1985）。[29] E. Torgersen。《统计实验比较》。剑桥大学出版社（1991）。[30]S. Turgut。“双体纯态的催化变换”。《物理学杂志 A：数学与理论》40(40)：12185（2007）。[31] J. van Lint。《编码理论导论》。施普林格（1999）。