arXiv:2303.08197v1 [math.ST] 14 Mar 2023

高维数据的自适应检验

BY YANGFAN ZHANG[1,a], RUNMIN WANG[2,c] AND XIAOFENG SHAO[1,b]

[1]伊利诺伊大学厄巴纳-香槟分校统计系，[a]yangfan3@illinois.edu；[b]xshao@illinois.edu

[2]德克萨斯农工大学统计系，[c]runminw@tamu.edu

本文提出了一类基于 L_q范数的 U 统计量，用于解决一系列与高维数据相关的全局检验问题。其中包括均值向量及其空间符号检验、线性模型系数的同时检验，以及高维观测的分量独立性检验等。在原假设下，我们在温和的矩条件和累积量条件下，推导出若干种 L_q范数 U 统计量的渐近正态性，以及它们之间的渐近独立性 q。我们提出了一种简单的组合方法，通过两个学生化的 L_q检验统计量的 p值进行组合，并证明该方法对于不同稀疏程度的备择假设具有很强的检验功效。我们的工作大幅拓展了He 等人(2021 年）的研究。该研究主要关注均值和协方差检验，而我们则对广泛一类核函数所对应的 L_q范数 U 统计量的渐近独立性进行了较为一般的研究。为减轻计算负担，我们利用求和中的单调指标，引入了所提出 U 统计量的一种变体，从而得到一个具有非对称核的 U 统计量。我们还引入了一种动态规划方法，将计算成本从$O(n^{qr})$（计算完整 U 统计量所需的成本）降至 $O(n^r)$，其中 r 表示核的阶数。数值研究进一步证实，与一些现有方法相比，所提出的自适应检验具有优势。

1. 引言。高维数据的维度 p 可能与样本量 n相当，甚至超过样本量，这类数据无处不在，许多科学领域都经常会遇到。在实践中，人们常常关注检验高维数据中低维特征的某些整体模式。这包括检验均值向量（Chen 和 Qin，2010；Cai 等人，2014；Wang 等人，2015；He 等人，2021）、协方差矩阵（Li 和 Chen，2012；He 等人，2021）、线性模型中的回归系数（Zhong 和 Chen，2011），以及高维观测中的分量独立性（Han 等人，2017；Leung 和 Drton，2018）。

这些问题可以表述为 H_0： $\Theta = \mathbf{0}$，其中 0 是全零向量，$\Theta = \{\theta_l, l \in \mathcal{L}\}$ 是一个高维向量， $\mathcal{L}$ 为索引集， $\{\theta_l\}$ 为相应的关注参数。对于这类全局检验问题，文献中已提出许多方法，主要分为两大类。一类是基于L_2 范数的检验，采用平方和型统计量，广泛应用于均值检验（Bai and Saranadasa, 1996; Chen 和 Qin, 2010; Goeman et al., 2006; Gregory et al., 2015; Srivastava and Du, 2008; Srivastava et al., 2016）、协方差检验（Bai et al., 2009; Chen et al., 2010; Ledoit and Wolf, 2002; Li and Chen, 2012）以及其他检验问题，例如分量独立性检验（Leung 和 Drton, 2018）。另一类是基于 L_∞ 范数的检验，采用最大型统计量。均值检验（Cai et al., 2014; Hall and Jin, 2010）、协方差检验（Cai and Jiang, 2011; Cai et al., 2013; Jiang, 2004; Liu et al., 2008; Shao and Zhou, 2014）以及分量独立性检验（Han 等人, 2017; Drton et al., 2020）中均有相关例子。众所周知，平方和型

关键词与短语：独立性检验，同时检验，空间符号，U统计量。

统计量对稠密备择具有较高的检验功效，此时 Θ 中有较大比例的非零元素，且其 $\|\Theta\|_2$较大；而最大型统计量针对的是稀疏备择，此时Θ 中只有少数非零元素，且其 $\|\Theta\|_\infty$较大。在实践中，假设备择假设具有某种特定类型往往并不现实，而且人们通常对备择假设的类型知之甚少。因此，有必要开发能够适应不同类型备择假设的新检验方法，并使其对广泛的备择假设（包括中度稀疏和中度稠密的备择假设）都具有良好的检验功效。

除了 L_2 和 L_∞ 型统计量外，Xu 等人（2016 年）和 Wu 等人（2019 年）还构造了一类 L_q型 von Mises V 统计量，用于检验 $q=2$ 的 $\|\Theta\|_q^q$，3,$\cdots$；这类统计量通常是有偏估计量。在一篇近期论文中，He 等人（2021）提出了 U 统计量框架，并构造了 $\|\Theta\|_q^q$ 的无偏且渐近独立的估计量。随后，他们通过聚合各个 p 值，组合了基于 L_q 范数的检验统计量，包括 $q=2$，$6,\infty$ 等情形。由此得到的检验具有自适应性，能够捕捉广泛的备择假设。然而，He 等人（2021）所研究的 U 统计量主要聚焦于均值与协方差检验问题，其对应的核函数阶数为 1。一个自然的问题是：基于 L_q 范数的 U 统计量的渐近正态性和独立性，是否适用于更一般的核函数，并能否推广到均值与协方差检验问题之外?

在本文中，我们旨在将基于 L_q-范数的检验及其组合推广到更广泛的检验问题。具体而言，我们的框架提供了更一般的处理方式，并涵盖了阶数大于 1 的 U 统计量；这类统计量自然出现在诸如检验线性回归系数的零假设 [Zhong 和 Chen（2011 年）], 以及基于 Kendall 的 τ 和 Spearman 的ρ （Leung 和 Drton，2018 年）进行分量独立性检验的问题中。此外，我们的框架使我们能够将 Wang 等人（2015 年）提出的基于空间符号的检验从 L_2 推广至其 L_q 对应形式，而这似乎未被纳入 He 等人（2021 年）的研究中。我们的渐近理论涵盖了 He 等人（2021 年）提出的部分理论作为特例，并进一步拓宽了基于 L_q 范数的检验及其组合的适用范围。在理论推导中，我们还利用核函数的特殊结构，推导出一种比经典 Hoeffding 分解更精细的新分解。我们指出，这种新分解适用于一类更广泛的具有张量积结构的核函数，因此可能具有广泛的应用价值。

从计算角度来看，计算 $\|\Theta\|_q^q, q=4,6,\cdots$ 的完整 U 统计量极具挑战性。He 等人（2021）利用统计量中的递归结构，开发了一种高效算法，使计算变得可行。然而，他们的算法似乎仅适用于 1 阶核，要推广到高阶情形则十分困难。为此，我们提出一类求和时采用单调指标的 U统计量，以减轻计算完整 U 统计量的计算负担。这类统计量本质上是采用非对称核的 U统计量，可以通过动态规划比完整 U 统计量快得多地计算出来。尽管这会牺牲一定的统计效率，但采用非对称核的新型 U统计量可以在 $O(n^r)$ 时间内高效计算，而蛮力方法则需要 $O(n^{qr})$ 时间，其中 r 等于核函数的阶数。

本文其余部分安排如下。在第 2 节中，我们介绍问题设定，并针对单样本和双样本检验问题提出一类 L_q 型 U 统计量。我们还列举了若干示例，说明所提出的统计量可应用于哪些场景。第 3 节介绍主要理论结果，包括所提出的 U 统计量在原假设下的渐近正态性，以及不同 q 下这些统计量的独立性。此外，还讨论了备择假设下的检验功效。第 4 节提出一种采用非对称核的 U 统计量变体，可通过动态规划高效计算该统计量。第 5 节汇总了所有模拟结果，第 6 节介绍一项真实数据

应用。我们在第 7 节总结全文。所有技术证明均收录于补充材料中。

记号。我们用大写字母（如 X_i）表示随机向量，并用相应的小写字母（如 x_{ij}）表示其分量。我们用 $\sum^*$ 表示对所有互异指标求和。对于两个函数 $f(X_1,\ldots,X_r)$ 和 $g(Y_1,\ldots,Y_s)$，我们将它们的张量积 $f\otimes g$ 定义为函数

$$(f\otimes g)(X_1,\ldots,X_r;Y_1,\ldots,Y_s)=f(X_1,\ldots,X_r)g(Y_1,\ldots,Y_s).$$

我们递归地定义 $(\otimes)^s f=f\otimes[(\otimes)^{s-1}f]$，其中 $s\geq 2$。我们将 $\{1,\ldots,n\}$ 记作 $[n]$，并将 $\{m,m+1,\ldots,n\}$ 记作 $[m,n]$。我们用 I 表示 $[n]$ 中一列（不一定互异的）指标，即 $(i_1,i_2,\ldots,i_m)$，并将 $(X_{i_1},X_{i_2},\ldots,X_{i_m})$ 记作 X_I。特别地，我们有 $X_{[c]}=(X_1,\ldots,X_c)$。我们将 $P_m(n)$定义为 $[n]$ 的所有基数为 m 的子集构成的集合，即

$$P_m(n)=\{\{i_1,i_2,\ldots,i_m\}\subseteq[n]:\ \text{All indices are distinct with each other}\}.$$

对于 $f(n)$和 $g(n)$，若 $f(n)/g(n)\to 1$为 $n\to\infty$，则记作 $f\asymp g$。在全文中，对于任意向量 $X=(x_1,\ldots,x_p)\in\mathbb{R}^p$和 $q\in\mathbb{Z}_+$，我们定义 $\|X\|_q^q=\sum_{l=1}^p x_l^q$。注意，我们没有对 x_l 取绝对值，因此当 q 为奇数时，$\|\cdot\|_q$不等于经典的 l_q范数。我们用 $e_t(r)$ 表示 $\mathbb{R}^r$ 中的一个向量，其第 t 个元素为 1，其余元素为 0。

2. 背景。

2.1. 单样本检验。假设我们从某个底层分布中观测到一些独立同分布（independent and identically distributed，独立同分布）样本 $X_1,\ldots,X_n\in\mathbb{R}^p$，且 p 与样本量 n 相当或更大。我们的目标是检验某些高维参数$\Theta=\Theta(X_i)=\{\theta_l:\ l\in\mathcal{L}\}=\mathbf{0}$。我们可以为$\|\Theta\|_q^q=\sum_{l\in\mathcal{L}}\theta_l^q$ 构造一族无偏 U 统计量。对于每个 l，我们从满足$\mathbb{E}[h_l(X_{[r]})]=\theta_l$的某些 r 阶核函数 h_l 入手。随后可以引入一个新的 qr 阶核 $\otimes^q h_l$，其期望等于 θ_l^q。回想一下，

$$\otimes^q h_l(X_{[qr]})=h_l(X_{[r]})h_l(X_{[r+1,2r]})\cdots h_l(X_{[(q-1)r+1,qr]}).$$

令 $P_m^n=\frac{n!}{(n-m)!}$。于是我们可以定义相应的 U 统计量

$$U_{n,q}=\sum_{l\in\mathcal{L}}(P_{qr}^n)^{-1}\sum_{1\leq i_1,\ldots,i_{qr}\leq n}^{*}\prod_{c=1}^{q}h_l(X_{i_{(c-1)r+1}},\ldots,X_{i_{cr}}).$$

在上述 L_q范数的定义中，我们没有取绝对值（这与经典定义不同），这是由于 U 统计量的构造方式所致。在极端情况下，当正分量与负分量相互抵消时，对于奇数 q，即使 Θ 不是零向量，L_q 范数 $\|\Theta\|_q^q$ 也可能（几乎）为零。因此，本文主要考虑偶数 q，即 $q\in 2\mathbb{Z}_+$，不过我们的理论结果也可以推广到一般的 $q\in\mathbb{Z}_+$。

请注意，He 等人（2021）研究的单样本均值检验问题包含在所提出的框架中。具体而言，$\mathcal{H}_0$: $\mathbb{E}[X_i]=\mu_0=(\mu_{0,1},\ldots,\mu_{0,p})$，通过取 $\mathcal{L}=[p]$、阶数为 $r=1$ 且 $h_l(X_i)=x_{i,l}-\mu_{0,l}$，即可得到 He 等人（2021）提出的 U统计量族。我们有 $\mathbb{E}[U_{n,q}]=\|\mathbb{E}[X_1]-\mu_0\|_q^q$。对于均值检验统计量，$\{U_{n,q}\}_{q\in Z^+}$ 的渐近独立性是根据渐近联合正态性以及 $\mathrm{cov}(U_{n,q_1},U_{n,q_2})=0$ for $q_1\neq q_2$ 推导得出的。作为说明，考虑 $q_1>q_2$ 的例子。对于U_n中的任一被加项 $X_{i_1}\cdots X_{i_{q_1}}$，以及$U_n$ 中的任一 $X_{j_1}\cdots X_{j_{q_2}}$，$_{q_2}$，必定存在至少一个 c，使得 i_c不属于集合 $\{j_1,\ldots,j_{q_2}\}$，因此它们的协方差为 0。这一论证高度依赖于核函数的阶数为 1 这一事实。对于高阶核，我们不能指望协方差恰好为 0，但通过施加一些正则性假设，仍可能证明协方差的渐近可忽略性（经过适当缩放后），从而证明不同 q之间这些 U统计量的渐近独立性。

2.2. 双样本检验。假设我们观测到两个独立同分布样本 $X_1, \ldots\ldots, X_n \in \mathbb{R}^p$ 和 $Y_1, \ldots\ldots, Y_m \in \mathbb{R}^p$。我们希望检验某个高维参数 $\Theta = \Theta(X_i, Y_i) = \{\theta_l : l \in \mathcal{L}\} = \mathbf{0}$。该方法与单样本检验类似，因为我们现在关注的是 $\|\Theta\|_q^q$，并希望检验 $\|\Theta\|_q^q = 0$是否成立。

对于每个 $l \in \mathcal{L}$，我们可以基于某个核函数 $h_l(X_{[r]}; Y_{[r]})$ 构造一族阶数为 (r, r) 且满足 $\mathbb{E}[h_l(X_{[r]}; Y_{[r]})] = \theta_l$ 的双样本 U 统计量。与单样本检验类似，我们可以考虑 $\otimes^q h_l$，其阶数为 (qr, qr)，且其期望等于 θ_l^q，即

$$\otimes^q h_l(X_{[qr]}; Y_{[qr]}) = h_l(X_{[r]}; Y_{[r]}) \cdots h_l(X_{[(q-1)r+1,qr]}; Y_{[(q-1)r+1,qr]}).$$

接下来可以定义相应的 U 统计量

$$U_{n,m,q} = \sum_{l \in \mathcal{L}} (P_{qr}^n P_{qr}^m)^{-1} \sum_{1 \le i_1, \ldots, i_{qr} \le n}^{*} \sum_{1 \le j_1, \ldots, j_{qr} \le m}^{*} \prod_{c=1}^{q} h_l(X_{i_{(c-1)r+1}}, \ldots, X_{i_{cr}}; Y_{i_{(c-1)r+1}}, \ldots, Y_{i_{cr}}).$$

与单样本检验问题的情形相同，He 等人（2021 年）中用于双样本均值检验的统计量族，也是上述双样本 U 统计量的一个特例。

2.3. 示例。本节给出上述表述中包含的一些示例。我们主要关注单样本检验问题中的示例，因为可以直接构造相应的双样本检验示例。

- 均值（零）检验，$\mathcal{H}_0: \mu := \mathbb{E}[X_i] = \mathbf{0}$（He 等人，2021 年）。在此情形下，$\mathcal{L} = [p]$、$r = 1$ 和 $h_l(X_i) = x_{i,l}$。
- 基于空间符号的检验（Wang 等人，2015 年），$\mathcal{H}_0: \mathbb{E}[X_i / \|X_i\|] = \mathbf{0}$。我们可以取 $\mathcal{L} = [p]$、$r = 1$ 和 $h_l(X_i) = x_{i,l} / \|X_i\|$。
- 零均值数据的协方差检验，$\mathcal{H}_0: \Sigma := \mathrm{var}(X_i) = \Sigma_0$ 对于某个给定的 $\Sigma_0 = (\sigma_{p_1 p_2})_{p_1, p_2 = 1}^{p}$。在这种情况下，$\mathcal{L} = \{l = (p_1, p_2) \; 1 \le p_1 \le p_2 \le p\}$，$r = 1$ 以及 $h_l(X_i) = x_{ip_1} x_{ip_2} - \sigma_{p_1 p_2}$。特别地，为了检验协方差是否具有带状结构，即 $\sigma_{p_1 p_2} = 0 \; |p_1 - p_2| \ge d$，我们可以取 $\mathcal{L} = \{l = (p_1, p_2) \in [p]^2 \; p_2 \ge p_1 + d\}$ $r = 1$ 和 $h_l(X_i) = x_{ip_1} x_{ip_2}$。He 等人（2021 年）也研究了这个例子。
- 分量独立性检验。$\mathcal{H}_0$: x_{i,p_1} 与 x_{i,p_2} 对于任意 $p_1 \neq p_2$ 均相互独立。对于基于 Kendall's τ 的检验（Leung 和 Drton，2018 年），我们实质上检验 $\mathcal{H}_0^\tau: P(x_{1,p_1} < x_{2,p_1}, x_{1,p_2} < x_{2,p_2}) = 1/4$ 对任意 $p_1 \neq p_2$ 成立，与 $\mathcal{H}_1^\tau: \mathcal{H}_0^\tau$ 的否定相对立。令 $\mathcal{L} = \{l = (p_1, p_2) \; 1 \le p_1 < p_2 \le p\}, r = 2$ 和 $h_l(X_i, X_j) = \mathrm{sgn}(x_{i,p_1} - x_{j,p_1})\mathrm{sgn}(x_{i,p_2} - x_{j,p_2})$。
- 基于 Spearman's ρ 的分量独立性检验（Leung 和 Drton，2018 年）。在这种情况下，$\mathcal{H}_0^\rho: P(x_{1,p_1} < x_{2,p_1}, x_{1,p_2} < x_{3,p_2}) = 1/4$ 对任意 $p_1 \neq p_2$ 成立，而 $\mathcal{H}_1^\rho$ 是 $\mathcal{H}_0^\rho$ 的否定。我们将 $\mathcal{L} = \{l = (p_1, p_2) \; 1 \le p_1 < p_2 \le p\}, r = 3$ 和 $h_l(X_i, X_j, X_k)$ 视为 $\mathrm{sgn}(x_{i,p_1} - x_{j,p_1})\mathrm{sgn}(x_{i,p_2} - x_{k,p_2})$ 的对称版本，即

$$\begin{aligned} h_l(X_i, X_j, X_k) = \frac{1}{6} \big[& \mathrm{sgn}(x_{i,p_1} - x_{j,p_1})\mathrm{sgn}(x_{i,p_2} - x_{k,p_2}) + \mathrm{sgn}(x_{i,p_1} - x_{k,p_1})\mathrm{sgn}(x_{i,p_2} - x_{j,p_2}) \\ & + \mathrm{sgn}(x_{j,p_1} - x_{i,p_1})\mathrm{sgn}(x_{j,p_2} - x_{k,p_2}) + \mathrm{sgn}(x_{j,p_1} - x_{k,p_1})\mathrm{sgn}(x_{j,p_2} - x_{i,p_2}) \\ & + \mathrm{sgn}(x_{k,p_1} - x_{i,p_1})\mathrm{sgn}(x_{k,p_2} - x_{j,p_2}) + \mathrm{sgn}(x_{k,p_1} - x_{j,p_1})\mathrm{sgn}(x_{k,p_2} - x_{i,p_2}) \big]. \end{aligned}$$

- 线性回归系数的同时检验（Zhong 和 Chen，2011 年）。对于线性模型 $Y = X\boldsymbol{\beta} + \varepsilon$，我们检验 $\mathcal{H}_0: \boldsymbol{\beta} = \boldsymbol{\beta}_0$，以对照 $\mathcal{H}_a: \boldsymbol{\beta} \neq \boldsymbol{\beta}_0$。我们取 $\mathcal{L} = [p]$，$r = 2$，并

$$h_l\big((X_1, Y_1), (X_2, Y_2)\big) = \big[(X_1 - X_2)(Y_1 - Y_2 - (X_1 - X_2)^T \boldsymbol{\beta}_0) \big]_l / 2, \tag{1}$$

且感兴趣的参数等于 $\Theta = \Sigma(\boldsymbol{\beta} - \boldsymbol{\beta}_0)$，其中 $\Sigma := \text{方差}(X)$。

• 双样本空间符号检验 (Chakraborty 和 Chaudhuri，2017年)，$\mathcal{H}_0: \Theta = \mathbf{0}$ 其中 $\Theta = \mathbb{E}[(X_i - Y_j)/\|X_i - Y_j\|]$。我们取 $\mathcal{L} = [p]$、$r = 1$ 和双样本核

$$h_l(X_i, Y_j) = (x_{i,l} - y_{j,l})/\|X_i - Y_j\|.$$

3. 理论分析。本节介绍所提出统计量的渐近理论。我们先讨论一阶核函数，再讨论高阶核函数。

3.1. 一阶核的渐近理论。对于阶数为 $r = 1$ 的核函数，我们有

$$U_{n,q} = \sum_{l\in\mathcal{L}} (P_q^n)^{-1} \sum_{1\le i_1,\ldots,i_q\le n}^{*} \prod_{c=1}^{q} h_l(X_{i_c}).$$

记 $\mathcal{L} = [L]$，并定义 $Y_i = (y_{i,1}, \ldots, y_{i,L})^T = (h_1(X_i) \ldots, h_L(X_i))^T \triangleq H(X_i)$，其中 $\Sigma = (\sigma_{l_1,l_2})_{l_1,l_2\in\mathcal{L}} = \mathrm{var}(Y_i)$。这样，我们可以将 $U_{n,q}$ 视为用于检验 $\mathbb{E}[Y_i] \equiv \mathbf{0}$ 的 U 统计量。均值检验的分析可直接应用。例如，我们可以计算方差为

$$\begin{aligned}
\mathrm{var}(U_{n,q}) &= (P_q^n)^{-2} \sum_{l_1,l_2\in\mathcal{L}} \sum_{1\le i_1,\ldots,i_q\le n}^{*} \sum_{1\le j_1,\ldots,j_q\le n}^{*} \mathbb{E}\left[\prod_{c=1}^{q} y_{i_c,l_1} y_{j_c,l_2}\right] \\
&= (P_q^n)^{-2} q! \sum_{l_1,l_2\in\mathcal{L}} \sum_{1\le i_1,\ldots,i_q\le n}^{*} \prod_{c=1}^{q} \mathbb{E}[y_{i_c,l_1} y_{i_c,l_2}] \\
&\asymp q! n^{-q} \sum_{l_1,l_2\in\mathcal{L}} \sigma_{l_1,l_2}^q = q! n^{-q} \|\Sigma\|_q^q.
\end{aligned}$$

对于 $q_1 \neq q_2$，我们有

$$\mathrm{cov}(U_{n,q_1}, U_{n,q_2}) = (P_{q_1}^n)^{-1} (P_{q_2}^n)^{-1} \sum_{l_1,l_2\in\mathcal{L}} \sum_{1\le i_1,\ldots,i_{q_1}\le n}^{*} \sum_{1\le j_1,\ldots,j_{q_2}\le n}^{*} \prod_{c_1=1}^{q_1} \prod_{c_2=1}^{q_2} \mathbb{E}[y_{i_{c_1},l_1} y_{j_{c_2},l_2}] = 0,$$

因为当 $q_1 \neq q_2$ 时，i_{c_1} 或 j_{c_2} 中至少有一个与其他任一指标不同。

可以直接应用 He 等人(2021 年的分析，但我们将改用 Zhang 等人(2021 年) 使用的论证，并施加以下较弱的矩条件和一些累积量假设。

假设 1（$r = 1$）。设 $Y_1, \ldots, Y_n$ 是 Y_0 的独立同分布副本，其方差矩阵为Σ，并满足以下条件。

1. 存在依赖于 p 的常数 $C_p > 0$，使得 $\mathrm{tr}(\Sigma) \ge C_p L$.2. Y_0 至多具有四阶矩，且 $\sup_{l\in\mathcal{L}} \mathbb{E}[y_{0,l}^4] < C$，对于 $u = 2,3,4$，有

$$|\mathrm{cum}(y_{0,l_1}, \ldots, y_{0,l_u})| \le C_p^{u/2} \left(1 \vee \max_{1\le i,j\le u} |l_i - l_j|\right)^{-\kappa}, \quad \text{for some } \kappa > 2.$$

上述假设略弱于 Zhang 等人 (2021 年) 中的假设，但足以推导以下定理所述的渐近分布，因为我们无需像 Zhang 等人 (2021 年) 那样证明过程收敛。上述累积量假设可由几何矩收缩推出 [参见 Wu and Shao (2004年) 中的命题2]，也可由 Wu (2005年) 提出的物理依赖度量推出 [参见第

4节，见 Shao 和 Wu (2007)]，；或者 α–混合（混合系数以多项式速率衰减）[Andrews (1991)，Zhurbenko 和 Zuev (1975)]，适用于时间序列情形。它基本上要求数据中的 p 分量之间满足弱依赖。我们的理论不要求分量具有顺序，只要分量的某种置换 p 满足累积量假设即可，因为我们的检验统计量对分量的置换不变。

定理 3.1 ($r=1$)。 假设*1* 成立。在原假设下，对于任意$q \in 2\mathbb{Z}_+$,

$$(q!)^{-1/2} n^{q/2} \|\Sigma\|_q^{-q/2} U_{n,q} \xrightarrow{\mathcal{D}} N(0,1).$$

此外，对于任意有限集 $I \subseteq 2\mathbb{Z}_+$，$(U_{n,\ q})_{q\in I}$ 渐近联合独立。

对于均值和协方差检验，假设 1 通常在 C_p 为不依赖于 p 的常数时成立。对于基于空间符号的检验，我们有 $H(X_i) = X_i/\|X_i\|_2$。在这种情况下，我们有 $\sum \mathbb{E}[y_{i,l}^2] = \mathbb{E}[\|Y_i\|_2^2] \equiv 1$，因此，如果 C_p 是固定常数$l\in\mathcal{L}$，如 Zhang 等人（2021）所述，假设 1 不成立。为此，我们单独给出证明，以说明基于空间符号的检验统计量的中心极限定理；详情见附录 A.2。总之，对于一阶核，所提出的 U 统计量将 He 等人（2021）研究的统计量作为特例，并且其渐近理论在较弱的假设下得到验证（例如，我们只要求四阶矩存在）。此外，该理论可推广至基于空间符号的检验，而 He 等人（2021）未涵盖这一情形。另一方面，He 等人（2021）得出了基于 L_q 的 U 统计量与用于均值和协方差检验问题的最大型统计量之间的渐近独立性，而本文未涵盖这一结果。

3.2. 高阶核的 Hoeffding 分解。对于阶数为 $r>1$ 的核，其渐近理论更加复杂。我们从一个新的 Hoeffding 分解开始，该分解考虑了 $U_{n,q}$ 的特殊结构。不失一般性，这里假设h_l 对所有 $l \in \mathcal{L}$ 的参数都是对称的。由于我们关注原假设，还假设 $\mathbb{E}[h_l(X_{[r]})] = 0$。

对于 $1 \le t \le r$，定义投影核

$$\begin{aligned} h_{l,t}(x_1,\dots,x_t) =& \mathbb{E}[h_l(x_1,\dots,x_t,X_{t+1},\dots,X_r)] \\ =& \mathbb{E}[h_l(X_1,\dots,X_t,X_{t+1},\dots,X_r) \mid X_1 = x_1,\dots,X_t = x_t]. \end{aligned}$$

如果对于任意 $1 \le t < s$，$h_{l,t} \equiv 0$，而 $h_{l,s} \neq 0$，则称 h_l 具有 $s \in [1,r]$ 阶退化性。特别地，若 $s=1<r$，则称 h_l 为非退化；若 $s=r$，则称其为完全退化。注意，按照惯例，我们将满足 $r=1$ 的 h_l 视为完全退化。

我们进一步定义如下（对于 $1 \le t \le r$，阶数为 t 的）Hajék 投影

$$h_l^{(t)}(x_1,\dots,x_t) = \sum_{s=1}^{t} (-1)^{t-s} \times \sum_{1\le i_1<\cdots<i_s\le t} h_{l,s}(x_{i_1},\dots,x_{i_s}).$$

特别地，我们有 $h_l^{(1)} = h_{l,\ 1}$。当 $r=2$ 时，我们有

$$h_l^{(2)}(X_1,X_2) = h_l(X_1,X_2) - h_l^{(1)}(X_1) - h_l^{(1)}(X_2),$$

whichh 是一个完全退化核。一般而言，$h_l^{(t)}$ 始终是一个 t 阶的完全退化核。

Hajék 投影是推导Hoeffding分解的基石。事实上，经典Hoeffding分解是基于这样一个事实推导出来的：核函数等于不同阶数的所有Hajék投影核之和。然而，我们并不

直接对 $\otimes^q h_l$应用投影，而是将 h_l 写成投影核 $h_l^{(t)}$ 之和，并应用分配律将 $\otimes^q h_l$ 分解为若干个 $h_l^{(t)}$ 的乘积。

作为示例，考虑一个阶数为 $r=2$ 的核 h_l。我们可以写作

$$h_l(X_1,X_2)=h_l^{(1)}(X_1)+h_l^{(1)}(X_2)+h_l^{(2)}(X_1,X_2).$$

应用分配律可得，例如，对于 $q=2$,

$$\begin{aligned}
&(\otimes^2 h_l)(X_1,X_2,X_3,X_4)\\
=&[h_l^{(1)}(X_1)+h_l^{(1)}(X_2)+h_l^{(2)}(X_1,X_2)][h_l^{(1)}(X_3)+h_l^{(1)}(X_4)+h_l^{(2)}(X_3,X_4)]\\
=&(\otimes^2 h_l^{(1)})(X_1,X_3)+(\otimes^2 h_l^{(1)})(X_1,X_4)+(\otimes^2 h_l^{(1)})(X_2,X_3)+(\otimes^2 h_l^{(1)})(X_2,X_4)\\
&+(h_l^{(1)}\otimes h_l^{(2)})(X_1,X_3,X_4)+(h_l^{(1)}\otimes h_l^{(2)})(X_2,X_3,X_4)\\
&+(h_l^{(1)}\otimes h_l^{(2)})(X_3,X_1,X_2)+(h_l^{(1)}\otimes h_l^{(2)})(X_4,X_1,X_2)\\
&+(\otimes^2 h_l^{(2)})(X_1,X_2,X_3,X_4).
\end{aligned}\tag{2}$$

因此，我们可以将 $\otimes^2 h_l$ 分解为 $\otimes^2 h_l^{(1)}$、$h_l^{(1)}\otimes h_l^{(2)}$ 和 $\otimes^2 h_l^{(2)}$。

基于分配律的论证适用于一般情形，我们可以将 $\otimes^q h_l$ 写成许多 $h_l^{(t)}$ 的张量积之和。为刻画这些求和项，我们定义

$$h_l^{\boldsymbol{\alpha}}=[\otimes^{\alpha_1}h_l^{(1)}]\otimes[\otimes^{\alpha_2}h_l^{(2)}]\otimes\cdots\otimes[\otimes^{\alpha_r}h_l^{(r)}],\tag{3}$$

对于 $\boldsymbol{\alpha}\triangleq(\alpha_1,\ .....,\ \alpha_r)$。特别地，$\sum_{t=1}^r t\alpha_t$是 $h_l^{\boldsymbol{\alpha}}$的阶数，而 α_t是公式（$h_l^{(t)}$3）中2。显然，$\otimes^q h_l$ 等于对所有满足 $\sum_{t=1}^r \alpha_t=q$ 的向量 $\boldsymbol{\alpha}\in\mathbb{N}^r$ 求和 $h_l^{\boldsymbol{\alpha}}$。特别地，（2）中的 $\otimes^2 h_l^{(1)}$、$h_l^{(1)}\otimes h_l^{(2)}$ 和 $\otimes^2 h_l^{(2)}$ 分别与 $\boldsymbol{\alpha}=(2,\ 0)$、(1,1) 和 (0,2) 相关联。

尽管 h_l 是一个对称核，但 $h_l^{\boldsymbol{\alpha}}$一般并不对称。将 $h_l^{\boldsymbol{\alpha}}$的对称对应形式定义为 $\tilde{h}_l^{\boldsymbol{\alpha}}$。注意，$\tilde{h}_l^{\boldsymbol{\alpha}}$是一个阶数为 $\sum_{t=1}^r t\alpha_t$的完全退化核，因为 $h_l^{(t)}$ 是完全退化的。例如，与 $\boldsymbol{\alpha}=(1,\ 1)$ 相关联的(2)中的核 $h_l^{(1)}\otimes h_l^{(2)}$ 是一个阶数为 $r=3$的完全退化核。它是不对称的，因为$h_l^{\boldsymbol{\alpha}}(X_1,\ X_2,\ X_3)=h_l^{(1)}(X_1)h_l^{(2)}(X_2,\ X_3)$，其对称对应形式定义如下

$$\tilde{h}_l^{\boldsymbol{\alpha}}(X_1,X_2,X_3)=\frac{1}{3}\Big[h_l^{(1)}(X_1)h_l^{(2)}(X_2,X_3)+h_l^{(1)}(X_2)h_l^{(2)}(X_1,X_3)+h_l^{(1)}(X_3)h_l^{(2)}(X_1,X_2)\Big].$$

我们定义以下与 $h_l^{\boldsymbol{\alpha}}$相关联的 U 统计量，它出现在$U_{n,\ q}$的 Hoeffding 分解中。注意，$c=\sum_{t=1}^r t\alpha_t$ 是 $h_l^{\boldsymbol{\alpha}}$的阶数。

$$\begin{aligned}
U_{n,\boldsymbol{\alpha}}=&(P_c^n)^{-1}\sum_{l\in\mathcal{L}}\sum_{1\le i_1,i_2,\cdots,i_c\le n}^{*}h_l^{\boldsymbol{\alpha}}(X_1,\ldots,X_c)\\
=&\binom{n}{c}^{-1}\sum_{l\in\mathcal{L}}\sum_{1\le i_1<i_2<\cdots<i_c\le n}\tilde{h}_l^{\boldsymbol{\alpha}}(X_1,\ldots,X_c).
\end{aligned}$$

由$U_{n,\ \boldsymbol{\alpha}}$，我们得到以下分解。

引理3.2. 存在一些常数 $C_{\boldsymbol{\alpha}}>0$ 使得

$$U_{n,q}=\sum_{c=qs}^{qr}\sum_{\boldsymbol{\alpha}:\sum_t t\alpha_t=c}C_{\boldsymbol{\alpha}}U_{n,\boldsymbol{\alpha}},\tag{4}$$

其中 s 是 h_l的退化性。

注. 一般而言，Hoeffding 分解（4）比经典分解更为精细。经典分解仅基于 $\otimes^q h_l$的 Hajék 投影，而（4）则考虑了 $\otimes^q h_l$ 的张量积结构，并应用分配律。这种更精细的分解使我们能够直接对 h_l 的 Hajék 投影（以及一些相关量）施加假设；否则，这些假设会过于抽象而难以理解。我们可以用核 $\otimes^q h_l$ 将经典Hoeffding分解写为$U_{nq}=\sum_{c=qs}^{qr} C^{[c]}U_n^{[c]}{}_q$，其中 $C^{[c]}$是一些常数，$U_n^{[c]}{}_q$是与 $\otimes^q h_l$ 的 Hajék 投影相关的 U 统计量，阶数为 c，并包含所有核为 $h_l^{\boldsymbol{\alpha}}$且阶数为$c$的 $U_{n\boldsymbol{\alpha}}$。

例如，考虑 $r=3$ 和 $q=2$，按照与（2）相同的论证，我们可以将 $\otimes^3 h_l$ 分解为六个核函数：$\otimes^2 h_l^{(1)}$，$h_l^{(1)}\otimes h_l^{(2)}$，$h_l^{(1)}\otimes h_l^{(3)}$，$\otimes^2 h_l^{(2)}$，$h_l^{(2)}\otimes h_l^{(3)}$ 和 $\otimes^2 h_l^{(3)}$。而经典Hoeffding分解会将阶数相同的核函数合并，例如 $h_l^{(1)}\otimes h_l^{(3)}$ 和 $\otimes^2 h_l^{(2)}$，因此只包含5项。不难看出，当 $r\le 2$时，所提出的分解与经典分解一致。

一般而言，对于任何可写成若干函数张量积的核函数，都可以利用它们的 Hajék 投影导出更精细的Hoeffding 分解，其方式与我们应用分配律得到（4）相同。

值得注意的是，有一种特殊类型的 $\boldsymbol{\alpha}$ 在上述分解中发挥着重要作用，其中 $\boldsymbol{\alpha}$ 只有第 t个分量等于 q，其余分量均为零，即 $\boldsymbol{\alpha}=qe_t(r)$。对于 $t=1$，……，r。与之相关的 $h_l^{\boldsymbol{\alpha}}$等于 $\otimes^q h_l^{(t)}$。我们将其对称版本记为 $H_l^{(t)}{}_{,\,q}$，因此

$$H_{l,q}^{(t)}(X_{[qt]})=[(qt)!]^{-1}\sum_{I\in P_{qt}(qt)}(\otimes)^q h_l^{(t)}(X_I),$$

相应的 U 统计量记为 $U_n^{(qt)}{}_{,\,q}$。在前述示例中，对于$r=2$，$U_n^{(q)}{}_{,\,q}$和 $U_n^{(2q)}{}_{,\,q}$分别是与 $\otimes^q h_l^{(1)}$ 和 $\otimes^q h_l^{(2)}$ 相关联的 U 统计量。

在第3.3节中，我们证明渐近方差可由以下形式刻画

$$\tilde{\Sigma}_t(q)\triangleq\sum_{l_1,l_2\in\mathcal{L}}\operatorname{cov}(H_{l_1,q}^{(t)}(X_{[qt]}),H_{l_2,q}^{(t)}(X_{[qt]}))=\operatorname{var}(H_q^{(t)}(X_{[qt]})),\tag{5}$$

其中，我们定义了 $H_q^{(t)}\triangleq\sum_{l\in\mathcal{L}}H_l^{(t)}{}_{,\,q}$。

接下来，我们还定义

$$\sigma_t(l_1,l_2)\triangleq\operatorname{cov}(h_{l_1}^{(t)}(X_{[t]}),h_{l_2}^{(t)}(X_{[t]})),\quad l_1,l_2\in\mathcal{L},$$

and

$$\Sigma_t(q)\triangleq\sum_{l_1,l_2\in\mathcal{L}}\sigma_t^q(l_1,l_2)=\sum_{l_1,l_2\in\mathcal{L}}\operatorname{cov}\big((\otimes^q h_{l_1}^{(t)})(X_{[qt]}),(\otimes^q h_{l_2}^{(t)})(X_{[qt]})\big)=\operatorname{var}\Big(\sum_{l\in\mathcal{L}}(\otimes^q h_l^{(t)})(X_{[qt]})\Big),$$

$\tilde{\Sigma}_t(q)$ 对于 $q\in 2\mathbb{Z}_+, 1\le t\le r$。其定义与 $\Sigma_t(q)$ 类似，特别地，我们有 $\tilde{\Sigma}_1(q)=\Sigma_1(q)$，因为 $\otimes^q h_l^{(1)}$ 是对称的，所以 $H_l^{(1)}{}_{,q}=\otimes^q h_l^{(1)}$。一般而言，由柯西-施瓦茨不等式可得 $\tilde{\Sigma}_t(q)\le\Sigma_t(q)$。举例来说，考虑 $t=2$ 和 $q=2$。我们有

$$H_2^{(2)}(X_1,X_2,X_3,X_4)=\sum_{l\in\mathcal{L}}\frac{1}{3}\Big[h_l^{(2)}(X_1,X_2)h_l^{(2)}(X_3,X_4)+h_l^{(2)}(X_1,X_3)h_l^{(2)}(X_2,X_4)$$
$$+h_l^{(2)}(X_1,X_4)h_l^{(2)}(X_2,X_3)\Big].$$

柯西-施瓦茨不等式由此可推出

$$\operatorname{cov}\left(H_2^{(2)}, H_2^{(2)}\right) \leq \operatorname{var}\left(\sum_{l\in\mathcal{L}} \otimes^2 h_l^{(2)}\right) = \Sigma_2(2),$$

因此 $\tilde{\Sigma}_2(2) \leq \Sigma_2(2)$。我们指出，$\Sigma_t(q)$ 更便于进行阶数分析，尽管它并不能刻画渐近方差。

3.3. 一阶核的渐近理论。由于上一节所述的分解是传统 Hoeffding 分解的一个更精细版本，因此它继承了 Hoeffding分解的大多数性质，例如不同阶数投影之间的正交性。因此，它可用于分析 $U_{n,q}$ 的渐近分布。为推导该分解中的主导项，我们作出以下假设。

假设 2. 设 h_l 具有 s 阶退化性。对于任意 $s < t \leq r$ 和 $q \in 2\mathbb{Z}_+$，有 $\Sigma_t(q) = o\left(n^{q(t-s)}\Sigma_s(q)\right)$，且 $\Sigma_s(q) = O\left(\tilde{\Sigma}_s(q)\right)$。

当 $s = 1$ 时，该假设等价于 $\Sigma_t(q) = o\left(n^{q(t-1)}\Sigma_1(q)\right)$，因为 $\Sigma_1(q) = \tilde{\Sigma}_1(q)$。举例来说，假设 h_l 的阶数为 2。若其退化，则假设 2 中的第一条自动成立，因为在这种情况下 $s = r = 2$，因此只需验证第二条。另一方面，若 h_l 不退化，即 $s = 1$，则假设中的第二条成立，因为 $\Sigma_1(q) = \tilde{\Sigma}_1(q)$，因此只需验证 $\Sigma_2(q) = o\left(n^q\Sigma_1(q)\right)$，即

$$\sum_{l_1,l_2\in\mathcal{L}} \mathbb{E}^q[h_{l_1}^{(2)}(X_{[2]})h_{l_2}^{(2)}(X_{[2]})] = o\left(n^q \sum_{l_1,l_2\in\mathcal{L}} \mathbb{E}^q[h_{l_1}^{(1)}(X_1)h_{l_2}^{(1)}(X_1)]\right).$$

若 $\mathbb{E}[h_{l_1}^{(2)}(X_{[2]})h_{l_2}^{(2)}(X_{[2]})]$ 且 $\mathbb{E}[h_{l_1}^{(1)}(X_1)h_{l_2}^{(1)}(X_1)]$（因此求和项也如此）的阶数相同，则该条件显然成立；在大多数情形下都是如此，例如第 2.3 节讨论的线性模型系数同时检验和分量独立性检验。举例而言，对于 Kendall 的 τ，我们有 $\Sigma_2(q) = 7^q\Sigma_1(q) = (7/9)^q p(p-1)/2$（更多细节见附录 A.4）。

一般而言，当 h_l 的退化性阶数为 s时，对于任意 $1 \leq t < s$，Hajék 投影 $h_l^{(t)} \equiv 0$，这意味着 h_l, $\boldsymbol{\alpha} = 0$ 若其阶数 $\sum_t t\alpha_t$ 小于 qs，且最低阶的非零核为 $H_{q\,l}^{(s)}$，（对应于 $\boldsymbol{\alpha} = q \cdot e_t$（$r$））。请注意，这是唯一一个阶数为 qs的非零核。下述引理指出，核为 $H_{q\,l}^{(s)}$，的 U 统计量 $U_n^{(qs)}$, $_q$ 确实是主导项。

引理3.3. 假设核 h_l的退化性阶数为 $s \in [1, r]$。在假设2下，在原假设下，对任意 $q \in 2\mathbb{Z}_+$，均有 $U_{n,\ q} = \binom{r}{s}^q U_n^{(qs)}$, $_q$ $[1 + o_p(1)]$

鉴于引理 3.3，Hoeffding 分解中的主导项 $U_n^{(qs)}$, $_q$ 是一个 U 统计量，其退化核函数 $\otimes^q h_l^{(s)}$ 由 $h_l^{(s)}$ 诱导。因此，不失一般性，只需研究与完全退化（$r = s$）核 h_l 相关的 U 统计量的渐近理论，然后将结果应用于 $h_l^{(s)}$。我们对任意完全退化核作如下假设，该假设适用于 $h_l^{(s)}$。回顾一下，$\sigma_t(l_1, l_2) \triangleq$ 协方差$(h_{l_1}^{(t)}(X_{[t]})\ h_{l_2}^{(t)}(X_{[t]}))$。

假设 3. 设 h_l 是一个阶数为 $r \geq 1$的完全退化核函数。对任意 $q \in 2\mathbb{Z}_+$，以下结论均成立。
(a) $\sum_{l_1,l_2,l_3,l_4 \in \mathcal{L}} [\sigma_r(l_1,l_2)\sigma_r(l_3,l_4)\sigma_r(l_1,l_4)\sigma_r(l_2,l_3)]^{q/2} = o\left(\tilde{\Sigma}_r^2(q)\right)$.
(b1) h_l 具有至多4阶矩，且$\sup_{l \in \mathcal{L}} \mathbb{E}[h_l^4(X_{[r]})] < C$，并且对于任意 q，存在常数 C_q，使得

$$\sum_{l_1,\ldots,l_u \in \mathcal{L}} |\mathrm{cum}^q(h_{l_1}(X_{i_1^{(1)}},\ldots,X_{i_r^{(1)}}),\ldots,h_{l_u}(X_{i_1^{(u)}},\cdots,X_{i_r^{(u)}}))| \leq C_q \tilde{\Sigma}_r^{u/2}(q),$$

对于 $u=2$，3，4 以及任意 u 组互异指标 $(i_1^{(u)}$，...，$i_r^{(u)})$。
此外，若 $r>1$，则下述结论成立。
(b2) h_l 具有至多 q 阶矩，且$\sup_{l \in \mathcal{L}} \mathbb{E}[h_l^q(X_{[r]})] < C$，并且

$$\sum_{l_1,\ldots,l_u \in \mathcal{L}} |\mathrm{cum}^{q/c}(h_{l_1}(X_{I_1}),\ldots,h_{l_u}(X_{I_c}))| = o(\tilde{\Sigma}_r^2(q)),$$

对于 $u=3$，4；$1 \leq c \leq q$ 和 I_1，...，$I_c \in P_{cr}(n)$，满足任一元素仅出现在两个 I_t 中，且对于任意 $t \neq s$，$I_t \neq I_s$。

备注。总之，假设 3 要求 $H_{l_1\,q}^{(s)}$、$H_{l_2\,q}^{(s)}$、$H_{l_3\,q}^{(s)}$、$H_{l_4}^{(s)}$、$_q$ 的联合累积量对于所有 l_1、...、$l_4 \in \mathcal{L}$ 满足某种可求和性。特别地，与 $r=1$ 的情形相比，假设3（b2）是针对 $r>1$ 提出的一项新假设。这是由于具有 $r>1$ 的核函数的四阶矩较为复杂。这些可求和性假设保证渐近方差 $U_n^{(qr)}$、$_q$ 确实由 $\tilde{\Sigma}_r(q)$ 刻画，且鞅中心极限定理成立。事实上，假设 1 蕴含假设3(a)和(b1)，它们是对 Zhang 等人（2021）所作假设的自然推广；其论证侧重于 $r=1$。

假设 3 描述了 h_l 的抽象性质，因为我们的框架允许使用一般核，而验证这些核需要逐一进行分析。为说明其可行性，我们通过几个常见的 U 统计量展示假设的可验证性；这些假设都可以直接验证。结果见附录A。

命题 3.4 （退化核，$r=s$）。假设 *3* 成立。在原假设下，对任意 $q \in 2\mathbb{Z}_+$，

$$[(qr)!]^{-1/2} n^{qr/2} \tilde{\Sigma}_r^{-1/2}(q) U_{n,q} \xrightarrow{\mathcal{D}} N(0,1).$$

*Furt*此外，对任意有限集 $I \subseteq 2\mathbb{Z}_+$，$(U_{n,q})_{q \in I}$ 渐近联合独立。

结合命题 3.4 和引理 3.3，我们立即得到以下结果 lt.

定理 3.5 （一般情形）。设 h_l 是一个阶数为 r、退化阶数为 $s \leq r$的核函数。此外，假设 *2* 和*3* 对 $h_l^{(s)}$ 成立。那么在原假设下，对任意 $q \in 2\mathbb{Z}_+$，

$$[(qs)!]^{-1/2} \binom{r}{s}^{-q} n^{qs/2} \tilde{\Sigma}_s^{-1/2}(q) U_{n,q} \xrightarrow{\mathcal{D}} N(0,1).$$

此外，对于任意有限集 $I \subseteq 2\mathbb{Z}_+$，$(U_{n,\ q})_{\ q \in I}$ 渐近联合独立。

请注意，按照我们的约定，任意阶数为 $r=1$ 的核函数均称为完全退化。因此，命题 3.4 和定理 3.5 均将定理 3.1 作为特例，其中包括 He 等人 (2021 年的均值和协方差检验，以及 Wang 等人 (2015 年研究的基于空间符号的检验。此外，在附录 A 中，我们证明这些假设在更广泛的情形下也成立，例如线性模型系数的同时检验，以及基于 Kendall's τ 和 Spearman's ρ的分量独立性检验。

最后，我们考虑一个特殊情形来进行说明，即核 h_l 是非退化的（$s=1$）。正如我们所讨论的，只需验证 3 （a）和（b1）中的假设，适用于 $h_l^{(1)}$，而这可由 1 中的假设推出，适用于 $h_l^{(1)}$ （因为 Zhang 等人 （ 2021 年）中的假设 2.1 蕴含其中的假设 6.1 ）。我们将结果总结为以下推论。

C推论 3.6 ($s=1$)。设 h_l 是一个阶数为 r的非退化核，并且满足假设 2，使得对于任意 $1<t\le r$， $\Sigma_t(q)=o(n^{q(t-1)}\Sigma_1(q))$。此外，对于 $Y_i=(h_l^{(1)}(X_i))_{l\in\mathcal{L}}$，假设1 成立。那么在原假设下，对于任意 $q\in 2\mathbb{Z}_+$,

$$(q!)^{-1/2}r^{-q}n^{q/2}\Sigma_1^{-1/2}(q)U_{n,q}\xrightarrow{\mathcal{D}}N(0,1).$$

此外，对于任意有限集 $I\subseteq 2\mathbb{Z}_+$， $(U_{n,q})_{q\in I}$ 在渐近意义下联合独立。

3.4. 功效分析。在本节中，我们推导备择假设下 U 统计量$U_{n,q}$的渐近理论。假设 $\mathbb{E}[h_l(X_1,\ldots\ldots,X_r)]=\theta_l$。定义中心化 U 统计量$h_l=h_l-\theta_l$。我们有

$$\begin{aligned}U_{n,q}=&(P_{qr}^n)^{-1}\sum_{l\in\mathcal{L}}\sum_{I\in P_{qr}(n)}[(\otimes)^q(\tilde{h}_l+\theta_l)](X_I)\\=&\|\Theta\|_q^q+\tilde{U}_{n,q}+\sum_{c=1}^{q}\binom{q}{c}(P_{cr}^n)^{-1}\sum_{l\in\mathcal{L}}\sum_{I\in P_{cr}(n)}\theta_l^{q-c}[(\otimes)^c\tilde{h}_l](X_I),\end{aligned}\tag{6}$$

其中，我们将常数 θ_l 视为一个函数，使得 $\theta_l(X_I)\equiv\theta_l$，而 $U_{n,q}$ 是与 h_l 相关联的所提出的 L_q 型 U 统计量。

由于 $\tilde{h}_l$ 具有零均值，我们得到 $n^{qs/2}\tilde{\Sigma}_s^{-1/2}(q)\tilde{U}_{n,q}$的渐近正态性，因此将信号水平定义为

$$\gamma_{n,q}=n^{qs/2}\tilde{\Sigma}_s^{-1/2}(q)\|\Theta\|_q^q.$$

在局部备择下，其中 $\gamma_{n,q}$ 趋于某个非零常数，我们还作出以下假设，以界定交叉乘积项。

假设 4。设 h_l 的退化性阶数为 s。对于任意 $q\in 2\mathbb{Z}_+$ 和$c=1,\ldots\ldots,q-1$，有

$$\sum_{l_1,l_2\in\mathcal{L}}\theta_{l_1}^{q-c}\theta_{l_2}^{q-c}\sigma_s^c(l_1,l_2)=o\big(n^{-(q-c)s}\tilde{\Sigma}_s(q)\big).$$

在假设4下，对于备择假设，我们有以下定理。

定理 3.7。 在与定理3.5相同的假设下，对于任意 $q\in 2\mathbb{Z}_+$,

- 设 $\gamma_{n,q}\to\infty$。则 $n^{qs/2}\tilde{\Sigma}_s^{-1/2}(q)U_{n,q}\xrightarrow{\mathcal{P}}\infty$。
- 假设 $\gamma_{n,q}\to 0$。那么 $U_{n,q}$ 具有与极限原假设相同的渐近分布。

• 假设 *4* 成立。在局部备择下，其中 $\gamma_{n,q} \to \gamma \in (0,\infty)$，我们有

$$n^{qs/2}\tilde{\Sigma}_s^{-1/2}(q)U_{n,q} \xrightarrow{\mathcal{D}} N\left(\gamma, (qs)!\binom{r}{s}^{2q}\right). \tag{7}$$

根据定理 3.7，当 $\gamma_{n,\ q} \to \infty$ 时，我们有 $U_{n,\ q}/\|\Theta\|_q^q = 1 + o_p(1)$，因此 $U_{n,\ q}$ 是 $\|\Theta\|_q^q$ 的比率相合估计量。

由于 Zhang 等人 (2021 年 的引理 6.1 指出，对于 $s=1$，假设 4 由假设1 （施加于 $h_l^{(1)}$ ）蕴含，因此我们得到以下推论。

推论 3.8（$s=1$）。假设 h_l 是非退化的，且假设 *1* 对$h_l^{(1)}$ 成立。此外，假设 *2* 也成立。则定理 *3.7* 中的前两种情形成立。另外，在局部备择 $\gamma_{n,\ q} \to \gamma \in (0,\ \infty)$下，对于任意 $q \in 2\mathbb{Z}_+$，

$$n^{q/2}\tilde{\Sigma}_1^{-1/2}(q)U_{n,q} \xrightarrow{\mathcal{D}} N\left(\gamma, q!r^{2q}\right).$$

为说明不同 $q \in 2\mathbb{Z}_+$ 下的功效表现，我们考虑 $\Theta = \delta \cdot (\mathbf{1}_d,\ \mathbf{0}_{L-d})^T (L = |\mathcal{L}|)$ 以简化表达。根据（7），要比较有限样本功效，只需考察

$$\frac{\mathbb{E}[U_{n,q}]}{\sqrt{\text{var}(U_{n,q})}} \asymp \frac{n^{qs/2}\tilde{\Sigma}_s^{-1/2}(q)\|\Theta\|_q^q}{[(qs)!]^{1/2}\binom{r}{s}^q},$$

˜对于不同的 q。我们假设，对于某些 $a>0$，有 $\Sigma_s(q) \asymp a^q N$，且 $N = N(L)$ 依赖于L，因此我们有

$$\frac{\mathbb{E}[U_{n,q}]}{\sqrt{\text{var}(U_{n,q})}} \asymp [(qs)!]^{-1/2}\binom{r}{s}^{-q} n^{qs/2}a^{-q/2}N^{-1/2}\delta^q d. \tag{8}$$

为达到（渐近）检验功效 $\Phi(R - z_{1-\alpha})$，对于某些 $R>0$，我们令 $\mathbb{E}[U_{n,q}]/\sqrt{\text{var}(U_{n,q})} = R$ 并求解 $\delta = \delta(q)$，由（8）得到

$$\delta(q) \asymp \binom{r}{s}\sqrt{a}[(qs)!]^{1/2q}(\sqrt{N}R/d)^{1/q}n^{s/2}. \tag{9}$$

给定 $\delta = \delta(q)$，U 统计量$U_{n,\ q}$具有固定的渐近检验功效 $\Phi(R - z_{1-\alpha})$。因此，我们希望找到 $q \in 2\mathbb{Z}_+$，使其对应于最小的 $\delta(q)$，因为这样往往能获得最高的检验功效。注意，我们考虑的是单侧检验，因为即使 也意味着该 L_q 参数的范数始终非负。我们得到以下结果，其形式类似于 He 等人（2021 年）的命题 2.3。

命题 3.9。最优的 q 使（ $\delta(q)$ 中（*9*）的最小值在以下位置取得：

• 2，如果 $d \geq \sqrt{N}R$;

• 某个 $q \in 2\mathbb{Z}_+$，且在 $\sqrt{N}R/d$ 中递增，如果 $d > \sqrt{N}R$。

备注。命题 3.9 是在假设 $\Theta = \delta(\mathbf{1}_d, \mathbf{0}_{L-d})$ 的条件下推导得出的，并且可以直接推广到更一般的 Θ，例如，对于某些常数r （或 a）和 $j \leq d$，$\theta_j = \delta r^j$ （或 δj^a），只要能够显式计算 $\|\Theta\|_q$ 即可。这基本表明，当 Θ 稠密时，最优 q 为 2。随着 Θ 变得更稀疏，最优 q 会增大。这与我们的直觉一致。

4. 检验统计量与计算。

4.1. 使用动态规划计算一阶核函数。当 h_l 的阶数为 1，即 $r=1$ 时，我们有

$$U_{n,q}=\sum_{l\in\mathcal{L}}(P_q^n)^{-1}\sum_{1\le i_1,\dots,i_q\le n}^{*}\prod_{c=1}^{q}h_l(X_{i_c}).$$

在这种情况下，我们可以运用动态规划加快计算，并精确计算统计量。具体而言，我们对 $q\le m\le n$ 定义

$$D_{q,l}(m)=\sum_{1\le i_1<\cdots<i_q\le m}\prod_{c=1}^{q}h_l(X_{i_c}).$$

我们有 $U_{n,\,q}=\binom{n}{q}^{-1}\sum_{l\in\mathcal{L}}D_{q,\,l}(n)$。固定一个分量 $l\in\mathcal{L}$。注意，$D_{1,\,l}(m)=\sum_{i=1}^{m}h_l(X_i)$ 只是序列 $\{h_l(X_i)\}_{i=1}^{n}$ 的累积和，可以在 $O(n)$ 时间内高效计算。之后，我们可以递归地使用以下方法，从 $D_{c-1,\,l}$ 计算 $D_{c,\,l}$，

$$D_{c,l}(m)=D_{c,l}(m-1)+D_{c-1,l}(m-1)h_l(X_m),\quad m\ge c,$$

其中 $D_{c,\,l}(m)=0$ 对于 $1\le m<c$。我们对 $c=2,\ \dots,\ q$ 执行这一步，每一步需要 $O(n)$ 时间。

总之，计算 $D_{q,\,l}(n)$ 和 $U_{n,\,q}$ 的总开销分别为 $O(qn)$ 和 $O(qn|\mathcal{L}|)$，相比需要 $O(qn^q|\mathcal{L}|)$ 计算量的暴力计算，大幅减少了计算开销。

4.2. 高阶核的非对称 U 统计量。上一节的方法仅适用于阶数为 1 的 h_l。不过，这启发我们考虑以下针对 $r>1$ 的 U 统计量，并可对其应用动态规划。

$$U_{n,q}^{M}=\sum_{l\in\mathcal{L}}\binom{n}{qr}^{-1}\sum_{1\le i_1<\cdots<i_{qr}\le n}\prod_{c=1}^{q}h_l(X_{i_{(c-1)r+1}},\dots,X_{i_{cr}}).$$

这里，我们仅考虑每个被加项中的所有单调指标，因此可以应用下文将讨论的动态规划。我们用上标 M 强调其单调性，并将其与 $U_{n,\,q}$区分开来。注意，$U_{n}^{M,\,q}$ 本质上是一个具有非对称核的 U 统计量。

要看出 $U_{n,\,q}$ 与 $U_{n}^{M,\,q}$ 之间的区别，考虑 $r=2$ 和 $q=2$ 的情形。此时，$U_{n,\,q}$ 本质上是一个 阶 U 统计量。在这种情况下，$U_{n,\,q}$ 的（对称）核为

$$\frac{1}{3}\big[h_l(X_1,X_2)h_l(X_3,X_4)+h_l(X_1,X_3)h_l(X_2,X_4)+h_l(X_1,X_4)h_l(X_2,X_3)\big],$$

而 $U_{n}^{M,\,q}$ 的核为 $h_l(X_1,\ X_2)h_l(X_3,\ X_4)$。尽管 $U_{n}^{M,\,q}$也是一个 4 阶 U 统计量，但 $U_{n,\,q}$ 中不同被加项的数量是 $U_{n}^{M,\,q}$ 的 3 倍。

为应用动态规划，我们针对 $qr\le m\le n$ 定义，

$$D_{q,l}^{M}(m)=\sum_{1\le i_1<\cdots<i_{qr}\le m}\prod_{c=1}^{q}h_l(X_{i_{(c-1)r+1}},\dots,X_{i_{cr}}).$$

我们首先初始化 $D^M_{0\,l,}\ (m)\ \equiv 1$，其中 $1 \le m \le n$。然后对 $c = 1$，……，q 和 $m \ge cr$ 使用以下递归更新公式，

$$D^M_{c,l}(m) = D^M_{c,l}(m-1) + \sum_{(c-1)r < i_1 < \cdots < i_{r-1} < m} D^M_{c-1,l}(i_1 - 1) h_l(X_{i_1}, \cdots, X_{i_{r-1}}, X_m),$$

对于 $D^M_{c\,l,}\ (m) = 0$，计算 $1 \le m < cr$。注意，每个 c，m 的计算量为 $O(m^{r-1})$。因此，求得 $D^M_{q\,l,}\ (n)$ 和 $U^M_{n\,,\ q}$ 的总计算量分别为 $O(qn^r)$ 和 $O(qn^r|\mathcal{L}|)$，远优于计算量为 $O(qn^{qr}|\mathcal{L}|)$ 的蛮力方法。

为了说明递归公式的思路，再次考虑 $r = 2, q = 2$ 这一情形。假设我们已经计算出 $D^M_{2\,l,}(m)$（对于 $m < n$）以及完整序列 $D^M_{1\,l,}$。为了得到 $D^M_{2\,l,}(n)$，我们只需计算 $D^M_{2\,l,}(n)$ 与 $D^M_{2\,l,}(n-1)$ 之间的差，其中包含形如 $h_l(X_i, X_j)h_l(X_k, X_n)$（对于 $1 \le i < j < k < n$）的项。若固定 k，则可将 $h_l(X_k, X_n)$ 提取出来，对 $h_l(X_i, X_j)$ 求和恰好为 $D^M_{1\,l,}(k-1)$，于是得到

$$D^M_{2,l}(n) = D^M_{2,l}(n-1) + \sum_{k=3}^{n-1} D^M_{1,l}(k-1) h_l(X_k, X_n).$$

不难看出，$U^M_{n\,,q}$ 的渐近方差的阶数为 $n^{qs/2}\Sigma_r(q)$，它与 Σ_s 成正比，而非 Σ_s，因为后者描述的是对称核 $H^{(s)}_{l\ \ ,q}$ 的方差，因此 $U_{n,q}$。然而，由于 $U^M_{n\,,q}$ 的 Hoeffding 型分解中的主导项是一个加权 U 统计量，精确推导其渐近方差的形式在技术上较为困难，但我们仍预期渐近正态性和独立性成立，并猜想存在某个常数 $C^M_{q\,,s}$ 使得

$$n^{qs/2}\Sigma_s^{-1/2}(q) U^{M,(qs)}_{n,q} \xrightarrow{\mathcal{D}} N(0, C^M_{q,s}),$$

严格证明留待未来研究。如果 h_l 完全退化，即 $s = r$，那么不难看出 $C^M_{q\,,s} = (qr)!$ 当 $U_{n,q}$ 且 $U^M_{n\,,q}$ 在此情形下相同。在第4.3节，我们提出一种基于置换的方差估计程序，无需计算精确方差。

4.3. 方差估计量。 在第 3.3 节中，我们已证明

$$[(qs)!]^{-1/2} \binom{r}{s}^{-q} n^{qs/2} \tilde{\Sigma}_s^{-1/2}(q) U_{n,q} \xrightarrow{\mathcal{D}} N(0,1).$$

~ ˆ~然而，由于 $\Sigma_s(q)$ 未知，我们将提出一个无偏估计量 $\Sigma_s(q)$ 来估计 $\Sigma_s(q)$，并据此定义如下学生化统计量，

$$T_{n,q} = [(qs)!]^{-1/2} \binom{r}{s}^{-q} n^{qs/2} \hat{\Sigma}_s^{-1/2}(q) U_{n,q} \tag{10}$$

ˆ有趣的是，估计量 $\Sigma_s(q)$ 仍属于本文提出的框架，即它是所提出的 U 统计量 $U_{n,q}$，其对应的核函数由 h_l 推导而来。因此，第 3.3 节中的理论保证了比率一致性。

例如，考虑阶数为 1 的 h_l。在这种情况下，我们有

$$\tilde{\Sigma}_1(q) = \sum_{l_1, l_2 \in \mathcal{L}} \mathrm{cov}^q\big(h_{l_1}(X_1), h_{l_2}(X_1)\big).$$

考虑 $\mathcal{L}^2=\{\boldsymbol{l}=(l_1,l_2):\ l_1,l_2\in\mathcal{L}\}$，以及 $\Theta=\left(\theta_{\boldsymbol{l}}\right)_{\boldsymbol{l}\in\mathcal{L}^2}$，其中 $\theta_{\boldsymbol{l}}=\mathsf{cov}\big(h_{l_1}(X_1)\ h_{l_2}(X_1)\big)$，核 $g_{\boldsymbol{l}}(X_1)=h_{l_1}(X_1)h_{l_2}(X_1)$。

我们可以估计 $\tilde{\Sigma}_1(q)=\|\tilde{\Theta}\|_q^q$ 利用与 $g_{\boldsymbol{l}}$相关的 U 统计量，即

$$\sum_{\boldsymbol{l}\in\mathcal{L}^2}(P_q^n)^{-1}\prod_{c=1}^{q}g_{\boldsymbol{l}}(X_c)=\sum_{l_1,l_2\in\mathcal{L}}(P_q^n)^{-1}\prod_{c=1}^{q}h_{l_1}(X_c)h_{l_2}(X_c).$$

这是原假设 $\mathbb{E}[h_l]\equiv 0$下的无偏估计量。为处理备择假设下的非零均值，我们只需将 h_l 中心化，并定义$\tilde{h}_l(X_1)=h_l(X_1)-\frac{1}{n}\sum_{i=1}^{n}h_l(X_i)$。我们相应地修改核函数 $g_{\boldsymbol{l}}$，并定义 $\tilde{g}_{\boldsymbol{l}}(X_1)=h_{l_1}(X_1)h_{l_2}(X_1)$。然后，我们可以用相关统计量估计 $\Sigma_1(q)$

$$\hat{\Sigma}_1(q)\triangleq\sum_{\boldsymbol{l}\in\mathcal{L}^2}(P_q^n)^{-1}\prod_{c=1}^{q}\tilde{g}_{\boldsymbol{l}}(X_c)=\sum_{l_1,l_2\in\mathcal{L}}(P_q^n)^{-1}\prod_{c=1}^{q}\tilde{h}_{l_1}(X_c)\tilde{h}_{l_2}(X_c).$$

$\hat{}$ $\tilde{}$预期 $\Sigma_1(q)$ 是 $\Sigma_1(q)$的比率相合估计量，尽管正式论证超出了本文的范围。

一般而言，对于任意具有阶数 $r>1$ 的非退化（$s=1$）核，我们都可以采用类似的步骤。在原假设下，我们可以定义核 $g_{\boldsymbol{l}}(X_1)=h_{l_1}^{(1)}(X_1)h_{l_2}^{(1)}(X_1)$。当 $h_l^{(1)}$ 的具体形式未知时，一种方法是根据样本对其进行估计，并推导出一个代入估计量。另一种方法是利用以下事实:

$$\mathbb{E}[h_{l_1}^{(1)}(X_1)h_{l_2}^{(1)}(X_1)]=\mathbb{E}[h_{l_1}(X_1,\ldots,X_r)h_{l_2}(X_r,\ldots,X_{2r-1})],$$

并将 $g_{\boldsymbol{l}}=h_{l_1}(X_1,\ldots,X_r)h_{l_2}(X_r,\ldots,X_{2r-1})$定义为一个阶数为 $2r-1$ 的核。为保证在备择假设下的一致性，我们必须定义 $\tilde{g}_{\boldsymbol{l}}$，即通过对 $h_l^{(1)}$ 或 h_l 进行中心化而得到的 $g_{\boldsymbol{l}}$ 的对应量。此外，我们还可以考虑具有单调指标的 U 统计量，即 $U_{n,q}^M$，以节省计算成本，因为我们只需得到 $\Sigma_1(q)$ 的比率相合估计量 $\tilde{}$。

当 $s>1$ 时，其形式会变得更加复杂。由于对于 $r\geq 2$，通常无法在 q 较大时计算 $U_{n,q}$，因此我们建议在实际应用中使用 $U_{n,q}^M$ 来构造检验，因此不讨论 $s\geq 2$ 情况下 $U_{n,q}$ 的方差估计。

对于 $U_{n,q}^M$，如第 4 节所述，预期其渐近正态，渐近方差与 $\Sigma_s(q)$ 成正比，我们采用一些非参数方法来估计渐近方差。一种方法是使用基于置换的方差估计量。这适用于所有在原假设下检验统计量的分布在某些置换下保持不变的检验问题，因此，在置换数据上计算的所提出的 U 统计量与原假设分布具有相同的分布。例如，任何分量独立性检验都满足这一性质。基于置换的方法也适用于第 2.3 节中介绍的线性模型系数检验示例（在某些模型假设下）。

举例来说，令 $\sigma:\ [n]\to[n]$ 表示 $[n]$ 上的一个置换。为了模拟下列 U 统计量的分布，该统计量是为检验线性模型系数而提出的。

$$\begin{aligned}U_{n,q}^M=&(P_{2q}^n)^{-1}\sum_{l\in\mathcal{L}}\sum_{1\leq i_1<\cdots<i_{2q}\leq n}2^{-q}\prod_{c=1}^{q}[(X_{i_{2c-1}}-X_{i_{2c}})(Y_{i_{2c-1}}-Y_{i_{2c}}-(X_{i_{2c-1}}-X_{i_{2c}})^T\boldsymbol{\beta}_0)]_l\\=&(P_{2q}^n)^{-1}\sum_{l\in\mathcal{L}}\sum_{1\leq i_1<\cdots<i_{2q}\leq n}2^{-q}\prod_{c=1}^{q}[(X_{i_{2c-1}}-X_{i_{2c}})(E_{i_{2c-1}}-E_{i_{2c}})]_l,\end{aligned}$$

其中，我们定义了 $E_i \triangleq Y_i - X_i\boldsymbol{\beta}_0$，对于 $i=1$，……，n。那么我们可以考虑

$$U_{n,q}^{M,\sigma} = (P_{2q}^n)^{-1} \sum_{l\in\mathcal{L}} \sum_{1\le i_1<\cdots<i_{2q}\le n} \prod_{c=1}^{q} h_l\big((X_{i_{2c-1}}, Y_{\sigma(i_{2c-1})}), (X_{i_{2c}}, Y_{\sigma(i_{2c})})\big),$$

对于任意随机置换 σ。在原假设 $\boldsymbol{\beta}=\boldsymbol{\beta}_0$ 下，$E_i=\varepsilon_i$ 是独立同分布的，且与 X_i 独立。因此，在原假设下，$U_{n,\,q}^{M,\,\sigma}$的分布与$U_{n,\,q}^{M}$的分布一致，可用于估计渐近方差。更准确地说，考虑 B 个随机置换 σ_1，…，σ_B。我们用基于置换的样本方差估计 $\mathrm{var}(U_{n,\,q}^{M})$，即

$$V_{n,q}^{M} = \frac{1}{B-1}\sum_{b=1}^{B}\left(U_{n,q}^{M,\sigma_b} - \frac{1}{B}\sum_{b'=1}^{B} U_{n,q}^{M,\sigma_{b'}}\right)^2.$$

学生化统计量定义为 $T_{n,\,q}^{M} = \frac{U_{n,\,q}^{M}}{\sqrt{V_{n,\,q}^{M}}}$。我们也可以通过直接从 $\{U_{n,q}^{M,\sigma_b}\}_{b=1}^{B}$ 的经验分布中获取临界值，进行基于置换的检验。根据经验，当 $B=100$ 时，这两种方法的表现相似，因此我们仅在模拟中报告 $T_{n,\,q}^{M}$ 的结果。

我们指出，如果计算资源足以计算完整的 U 统计量 $U_{n,\,q}$，这种基于置换的方法仍适用于估计 $U_{n,\,q}$ 的方差；只需将 $U_{n,\,q}^{M,\,\sigma}$替换为 $U_{n,\,q}^{\sigma}$，也就是在置换样本上计算的完整 U 统计量。

另一种可行的方法是（乘子）自助法。然而，对高阶自助法 U 统计量进行理论分析颇具挑战性（Arcones 和 Giné，1992年），因此本文主要采用基于置换的方法。

4.4. 自适应检验程序。设 I 为一组 $q\in 2\mathbb{Z}_+$（例如{2,6}）。在原假设$T_{n\,\text{下},\,q}$对于不同的 $q\in I$渐近独立，因此我们可以通过组合 $q\in I$直接构造自适应检验。将与$T_{n,\,q}$ 对应的 p值记为p_q，则可将与 $I\subseteq 2\mathbb{Z}_+$ 对应的自适应检验统计量定义为 $p_{ada}=$最小值$_{q\in I}\,p_q$，其对应的 p值为 $1-(1-p_{ada})^{|I|}$。

等价地，如果我们要进行水平为α 的检验，可以将带有 I 的自适应检验函数定义为

$$\phi_{I,\alpha} = \max_{q\in I}\phi_{q,1-(1-\alpha)^{1/|I|}},$$

其中 $\phi_{q,\,\alpha}$ 是基于 $T_{n,\,q}$的水平为 α 的检验，即 $\phi_{q,\,\alpha}=\mathbb{I}(T_{n,\,q}>\Phi(1-\alpha))$，其中 $\Phi(x)$ 表示标准正态分布的累积分布函数。因此，如果 $T_{n,\,q}$ 中任一项超过其临界值 $\Phi((1-\alpha)^{1/|I|})$，自适应检验就拒绝原假设。类似的步骤也可直接应用于 $T_{n,\,q}^{M}$。

一般而言，较小的 q（例如 $q=2$）对稠密备择假设的检验功效更高，而较大的 q 则更适用于稀疏备择假设。通过自适应检验，我们可以结合不同 q的优势，在稠密和稀疏备择假设下都获得较高的检验功效；只要某个检验统计量 $T_{n,\,q}$ 的渐近功效为 1，自适应检验的检验功效就会趋于 1。由于校正后的显著性水平 $1-(1-\alpha)^{1/|I|}$ 随 I增加而降低，即便每个基于单个 q 的检验可能都针对某种特定备择假设达到最优，将过多 q 纳入 I也可能损害自适应检验的整体检验功效。为获得更好的有限样本表现，我们建议在实践中结合两个基于单个 q 的检验，这样对于稀疏性不同的备择假设都能保持相对较高的检验功效。

4.5. 双样本非对称 U 统计量。本节介绍带有单调指标的双样本 U 统计量，动态规划也可应用于此。为作说明，我们主要考虑阶数为 (1,1) 的核函数，因为主要的动机示例是双样本空间符号检验（Chakraborty 和 Chaudhuri，2017年）。假设我们观测到两个独立同分布样本 $X_1, \ldots\ldots, X_n \in \mathbb{R}^p$ 和 $Y_1, \ldots\ldots, Y_m \in \mathbb{R}^p$。我们希望检验某个高维参数 $\Theta = \Theta(X_i, Y_i) = \{\theta_l : l \in \mathcal{L}\} = \mathbf{0}$。回顾一下，完整的 U 统计量定义为

$$U_{n,m,q} = \sum_{l\in\mathcal{L}} (P_q^n P_q^m)^{-1} \sum_{1\le i_1,\ldots,i_q\le n}^{*} \sum_{1\le j_1,\ldots,j_q\le m}^{*} \prod_{c=1}^{q} h_l(X_{i_c}, Y_{j_c}).$$

特别地，对于空间符号检验，我们取 $\mathcal{L} = \{1, \ldots, p\}$，并令 $h_l(X_i, Y_j) = (x_{i,l} - y_{j,l}) / \|X_i - Y_j\|$。

由于计算 $U_{n,m,q}$ 的复杂度为 $O(n^q m^q)$，我们考虑如下单调索引对应形式，

$$U_{n,m,q}^M = \sum_{l\in\mathcal{L}} \binom{n}{q}^{-1} \binom{m}{q}^{-1} \sum_{1\le i_1<\cdots<i_q\le n} \sum_{1\le j_1<\cdots<j_q\le m} \prod_{c=1}^{q} h_l(X_{i_c}, Y_{j_c}).$$

利用动态规划，我们可以将计算 $U_{n,m,q}^M$ 的复杂度加速至 $O(mn)$。对于 $q \le i \le n$ 和 $q \le j \le m$，定义

$$S_{q,l}^M(i,j) = \sum_{1\le i_1<\cdots<i_{q-1}<i} \sum_{1\le j_1<\cdots<j_{q-1}<j} h_l(X_i, Y_j) \prod_{c=1}^{q-1} h_l(X_{i_c}, Y_{j_c}),$$

其中包含 $U_{n,m,q}^M$ 中的所有求和项，使得 X 和 Y 的最大索引分别为 i 和 j。我们还定义

$$D_{q,l}^M(i,j) = \sum_{1\le i_1<\cdots<i_q\le i} \sum_{1\le j_1<\cdots<j_q\le j} \prod_{c=1}^{q} h_l(X_{i_c}, Y_{j_c}).$$

如果将 $S_{q,l}(i,j)$ 视为以 $i \in [1,n]$，$j \in [1,m]$ 为索引的矩阵，那么 $D_{q,l}(i,j)$ 将位于(i,j)元素上方和左侧的所有元素相加。特别地，我们有

$$U_{n,m,q}^M = \binom{n}{q}^{-1} \binom{m}{q}^{-1} \sum_{l\in\mathcal{L}} D_{q,l}^M(n,m),$$

以及 $S_{q,l}^M(i,j) = D_{q-1,l}^M(i-1, j-1) h_l(X_i, Y_j)$。我们首先为所有 i，j 初始化 $D_{0,l}^M(i,j) \equiv 1$。然后，我们使用以下递归更新公式来计算 $c = 1, \ldots, q$ 和 $i, j \ge c$，

$$\begin{aligned} D_{c,l}^M(i,j) &= S_{c,l}^M(i,j) + D_{c,l}^M(i-1,j) + D_{c,l}^M(i,j-1) - D_{c,l}^M(i-1,j-1) \\ &= D_{c-1,l}^M(i-1,j-1) h_l(X_i, Y_j) + D_{c,l}^M(i-1,j) + D_{c,l}^M(i,j-1) - D_{c,l}^M(i-1,j-1). \end{aligned}$$

我们推测，统计量 $U_{n,m,q}$ 和 $U_{n,m,q}^M$ 都是渐近正态的。它们的方差可以通过置换进行估计（即在两个样本之间随机置换数据），与单样本情形类似。

5. 模拟研究。本节通过模拟考察所提出的单样本检验q和自适应检验的有限样本表现。本节中，对于具有$r = 1$ 的核（例如空间符号检验），我们采用$T_{n,q}$其基于第4.3节介绍的方差估计量；对于 $r > 1$，我们采用T_{n}^{M}构造检验，q并使用基于置换的方差估计量。第5.1 节展示空间符号单样本检验的模拟结果，第5.4 节则介绍其双样本对应检验。第5.2 节给出分量独立性检验的结果，第5.3 节则对应于线性模型系数为零的检验问题。对于所有检验问题，我们均进行 1000 次蒙特卡洛重复。

5.1. 空间符号检验。本小节中，我们希望检验

$$\mathcal{H}_0 : \mathbb{E}[X_i/\|X_i\|] = \mathbf{0} \quad v.s. \quad \mathcal{H}_a : \mathbb{E}[X_i/\|X_i\|] \neq \mathbf{0}.$$

这与标准的均值检验问题类似，$\mathcal{H}_0$： $\mathbb{E}[X_i] = \mathbf{0}$ 和 $\mathcal{H}_a$： $\mathbb{E}[X_i] \neq \mathbf{0}$，并满足某种椭圆对称假设，但不要求一阶矩存在。Wang 等人(2015 年扩展了Chen 和 Qin (2010)提出的基于 U 统计量的双样本检验，并提出了 L_2-范数型空间符号检验统计量。该统计量实际上是我们采用核函数 $h_l(X_i) = x_{i,\ l}/\|X_i\|$ 和 $q = 2$的 U 统计量的一个特例，尽管他们使用了不同的方差估计量。通过理论分析和模拟，Wang 等人(2015 年)发现，对于高斯数据，他们的检验功效略低于Chen 和 Qin (2010)提出的检验，但对于重尾数据，检验功效则高得多。然而，基于 L_2-范数的空间符号检验主要针对稠密备择，其在稀疏备择下的表现尚未得到考察。

下面我们根据模型生成模拟数据 $X_i = Z_i + \mu$，其中 $Z_i \in \mathbb{R}^p$ 具有独立同分布的标准正态分布或 t_3 分量。我们纳入稀疏和稠密备择假设（分别对应 $r = 2$ 和 $r = p$），并同时考虑轻尾和重尾分布（即标准正态分布与 t_3）。表 1 汇总了检验水平和检验功效的结果。在所有比较的检验中，我们指出，参数为 $q = 2$ 的均值检验对应于 Chen 和 Qin (2010) 提出的检验（方差估计量有所不同）；参数为$q = 2$ 的空间符号检验对应于 Wang 等人 (2015) 提出的检验（方差估计量有所不同）。我们的检验在参数为 $q = 2$ 时，与 Chen 和 Qin (2010) 提出的检验（均值检验）以及 Wang 等人 (2015) 提出的检验（空间符号检验）表现非常相似，因此我们省略了这些检验的结果。

DGP	δ, r	Test	$q=2$	$q=4$	$q=6$	$q=2,4$	$q=2,6$
高斯	0, NA	mean	6.1	5.9	3.0	6.4	5.4
		空间符号	5.8	6.1	3.2	6.4	5.6
	0.3, 2	mean	74.3	96.3	90.0	96.3	91.4
		空间符号	74.5	96.0	90.0	96.2	92.1
	0.1, p	mean	100	92.2	37.8	100	100
		空间符号	100	92.1	37.3	100	100
t_3	0, NA	mean	6.1	4.5	1.9	6.8	4.8
		空间符号	6.2	4.8	2.1	6.9	4.9
	0.4, 2	mean	42.7	73.1	57.6	72.7	63.6
		空间符号	51.1	79.3	65.4	80.1	71.1
	0.1, p	mean	98.2	30.1	14.5	97.0	97.0
		空间符号	99.8	36.5	15.6	99.3	99.3

表格 1 均值检验和空间符号检验的检验水平与检验功效（百分比）

如表1所示，对于较小的 q（即 $q = 2$，4），其检验水平往往比对于较大的 q（即 $q = 6$）更准确；后者表现出一定程度的检验水平偏低。基于$q = 2$，4 的自适应检验略微偏大，而 $q = 2$，6 的检验水平似乎较为准确。就检验功效而言，基于较大和较小 q 的单一 q 检验，正如预期的那样，分别对稀疏和稠密备择假设具有较高的检验功效。此外，对于两种备择假设和所有 q，比较高斯数据下的空间符号检验与均值检验时，检验功效没有明显损失；但对于服从 t_3 分布的数据，前者的检验功效显著更高。这表明，Wang 等人（2015）首次发现的基于 L_2 范数的空间符号检验在检验功效上的稳健性和优势，也适用于更一般的 L_q 范数情形。最后可以看到，自适应检验对稀疏和稠密备择假设总体上都具有良好的检验功效，而且其检验功效始终与最佳单一 q 检验非常接近。在稀疏情形下，采用 q =(2,4) 的自适应检验似乎略优于采用 q =(2,6) 的对应检验；这可能与前者在两种设定下检验水平略微偏高有关。

5.2. 分量独立性检验。在本小节中，我们比较几种基于 Kendall τ 统计量的分量独立性检验。我们按照以下两种设定生成独立同分布数据 X_1、……、$X_n \in \mathbb{R}^p$，这些设定见 Leung 和 Drton(2018)。

• 具有带状协方差矩阵的高斯向量：$X_i \overset{i.i.d.}{\sim} N(0, \Sigma)$ 对于 $i=1, \ldots\ldots, n$，其中 $\Sigma = (\sigma_{ab})_{a,\, b=1}^{p}$ 且 $\sigma_{ab} = (1-\delta)1\{a=b\} + \delta 1\{a \le r,\, b \le r\}$。• t 具有带状协方差矩阵的 - 分布：$X_i \overset{i.i.d.}{\sim} \Sigma^{1/2} Z$ 对于 $i=1, \ldots\ldots, n$，其中 Σ 如上所述，且 $Z_{ij} \overset{i.i.d.}{\sim} t_3$；$i=1, \ldots\ldots, n$；$j=1, \ldots\ldots, p$。

在原假设下，我们令 $\delta = 0$；对于具有不同稀疏度的备择假设，令 (δ, r) =(0.5, 3)、(0.3, 4)、(0.25, 6)、(0.15, 10)、(p, 0.05)，并令 (n, p) =(100, 50)。

在我们比较的检验中，包括我们提出的检验 $U_{n,\, q}^{M}$ 它采用基于置换的方差估计量；Han 等人(2017年提出了基于 Kendall's τ 的最大型统计量（表 5.2中记为 L_∞）；Leung 和 Drton (2018 年提出了基于 Kendall's τ 的平方和型统计量（表 5.2中记为 L_2）；此外，还有基于 Bonferroni 校正、将 L_2 与 L_∞简单组合得到的检验（表中记为 $L_2\&L_\infty$，见表 5.2）。

注意，Leung 和 Drton（2018）提出的 L_2 检验包含对角项，且仅为渐近无偏。他们还提出了一种无偏的 L_2 型统计量，它是我们提出的、满足 $q=2$ 的完整 U 统计量的一个特例。根据模拟结果，其性能与有偏版本非常相似，因此这里不予报告。Leung 和 Drton（2018）还提出了一种基于容斥原理的高效算法，可将计算复杂度从 $O(n^{2r})$ 降低至 $O(n^r)$（在本例中为 $r=2$）。然而，这一思路只能应用于 $q=2$；一般而言，他们的方法可以将精确 U 统计量的计算复杂度从 $O(n^{qr})$ 降低至$O(n^{(q-1)r})$。我们的单调索引 U 统计量，$_q$ 与 $U_{n,\, q}$ 相比，效率会固定损失一定比例，但对任意 q，都可将计算负担降低至 $O(n^r)$。

DGP	δ, r	$q=2$	$q=4$	$q=6$	$q=2,4$	$q=2,6$	L_2	L_∞	$L_2\&L_\infty$
高斯	0, NA	5.5	5.0	5.8	6.4	7.3	5.8	4.6	5.6
	0.5, 3	38.8	96.2	95.6	96.4	95.2	43.8	98.6	97.2
	0.3, 4	25.4	42.6	34.0	45.0	38.8	29.0	39.4	41.4
	0.25, 6	43.8	47.0	30.0	56.4	48.0	51.8	37.0	53.2
	0.15, 10	48.4	22.2	11.6	48.2	45.4	53.6	14.6	50.0
	0.05, p	95.4	14.4	9.0	93.6	93.8	96.2	13.0	94.6
t_3	0, NA	4.3	4.9	4.7	5.2	5.1	5.2	2.9	4.3
	0.5, 3	54.0	99.8	99.6	99.8	99.8	60.0	100	99.8
	0.3, 4	40.4	72.0	63.2	70.2	64.2	44.2	71.2	68.4
	0.25, 6	71.2	80.0	64.4	83.6	77.4	76.2	66.8	78.4
	0.15, 10	78.6	47.6	22.6	76.0	73.0	83.2	27.8	76.8
	0.05, p	100	30.8	16.0	100	100	100	23.8	100

TABLE 2 Kendall's τ 检验的检验水平和检验功效
（百分比）

在表 5.2 中，我们报告了针对稀疏程度各异的一系列备择假设，所有基于 Kendall τ 的检验的检验水平和检验功效。可以看出，大多数检验的检验水平都很准确，但高斯数据下采用 $q=(2,6)$ 的自适应检验以及 L_∞ 情形下的 t_3 方法除外。就检验功效而言，对于稠密备择，本文提出的统计量以及 Leung 和 Drton（2018 年）中的 $q=2$ 和 L_2 方法表现更佳；而在稀疏备择下，基于较大的 q 或 Han 等人（2017年）中的 L_∞ 方法构造的统计量检验功效更高。L_2 方法略优于我们经学生化处理的 $U_{n,\, q}^{M}$（采用 $q=2$），这在预料之中，因为单调

索引 U 统计量与完整的 U 统计量相比，统计效率较低。基于 $q = (2,4)$ 的自适应检验似乎与 $L_2\&L_\infty$（即基于朴素 Bonferroni 校正、简单结合 L_2 和 L_∞ 的方法）表现相当，尽管文献中似乎尚未确立基于 Kendall τ 的 L_2 与 L_∞ 统计量之间的渐近独立性。在一些稀疏或稠密程度适中的情形下（即 $(\delta, r) = (0.3, 4), (0.25, 6)$），我们发现基于 $q = (2, 4)$ 的自适应检验略优于 $L_2\&L_\infty$，这表明将基于 L_q 范数的检验纳入自适应检验可带来额外优势，其中 $q \neq 2, \infty$。

5.3. 线性模型系数检验。本小节中，我们给出一些用于检验线性回归系数零假设的模拟结果。设数据 $\{(X_i, Y_i)\}_{i=1}^n$来自线性模型 $Y_i = X_i\boldsymbol{\beta} + \varepsilon_i$ 的独立同分布样本，目标是检验 $\mathcal{H}_0: \boldsymbol{\beta} = \mathbf{0}$，对比 $\mathcal{H}_a: \boldsymbol{\beta} \neq \mathbf{0}$。回顾一下，核函数的阶数为 2，如式（1）所示，其中 $\boldsymbol{\beta}_0 = \mathbf{0}$,$h_l(W_1, W_2) = (x_{1,l} - x_{2,l})(Y_1 - Y_2)/2$，待检验参数等于 $\Sigma\boldsymbol{\beta}$。Zhong 和 Chen (2011) 提出的检验统计量实际上是我们方法的一个特例，其中 $q = 2$（因此他们的核函数阶数为 qr=4），但采用了不同的方差估计量。我们设 $(n, p) = (100, 50), (200, 100)$，$\boldsymbol{\beta} = \delta(\mathbf{1}_r, \mathbf{0}_{p-r})$，并生成与$\varepsilon_i \sim N(0, 1)$独立的 $X_i \sim N(0, \Sigma)$。我们考虑来自 AR(1) 结构的 $\Sigma = (\sigma_{a,b})_{a,b=1}^p$，其中 $\sigma_{a,b} = \rho^{|a-b|}$，且$q = 2$为 4、6。

在表 3 中，我们展示了基于$U_{n,2}$, $U_{n,q}^M$ 的检验结果（采用基于置换的方差估计量），适用于 $q = 2, 4, 6, (2, 4)$ 和 $(2, 6)$；还展示了针对 $q = 4, 6$，将$U_{n,2}$ 与$U_{n,q}^M$ 进行 Bonferroni 型组合的结果。从计算角度来看，对于任意 q，单调索引 U 统计量$U_{n,q}^M$ 和学生化统计量$T_{n,q}^M$ 的计算成本为 $O(n^2)$。$U_{n,2}$ 的完整计算通过 Leung 和 Drton（2018）提出的高效算法实现，因此计算成本从 $O(n^4)$ 降低至 $O(n^2)$。在未报告的模拟中，我们还将$U_{n,2}$（采用基于置换的方差估计量）的性能与 Zhong and Chen (2011) 的方法进行了比较，发现二者结果非常相似。请注意，Zhong and Chen (2011) 中学生化检验的计算需要 $O(n^4)$ 计算。

(n, p)	δ, r, ρ	$U_{n,q}$	$U_{n,q}^M$					$U_{n,2}$ & $U_{n,q}^M$	
		$q = 2$	$q = 4$	$q = 6$	$q = 2, 4$	$q = 2, 6$		$q = 4$	$q = 6$
(100,50)	0, NA, 0	6.3	5.7	6.4	3.1	6.3	4.6	7.2	5.7
	0, NA, 0.5	7.0	6.5	4.6	4.1	6.5	5.7	7.0	6.4
	0.25, 2, 0	29.2	21.2	28.4	17.8	31.2	25.4	34.4	29.4
	0.25, 2, 0.5	60.4	45.6	62.6	49.6	65.2	57.0	69.6	62.8
	0.05, p, 0	31.2	24.8	8.0	6.8	22.8	22.2	27.8	26.4
	0.05, p, 0.5	98.0	92.2	45.6	19.6	89.6	89.6	96.8	97.0
(200,100)	0, NA, 0	5.7	5.5	4.9	3.5	5.1	5.6	5.2	5.3
	0, NA, 0.5	6.6	5.9	3.9	2.6	6.3	5.1	6.8	5.5
	0.25, 2, 0	45.2	32.4	56.8	46.8	58.2	52.6	63.8	57.8
	0.25, 2, 0.5	81.6	62.2	93.6	90.4	92.8	92.0	94.6	93.8
	0.035, p, 0	42.4	30.4	10.2	7.6	27.8	28.0	36.2	36.6
	0.035, p, 0.5	100	99.4	57.6	28.6	99.4	99.4	100	100

表格 3 线性模型系数检验的检验水平和检验功效百分比

从表3可以看出，所有检验的检验水平均相当合理，但某些情况下存在轻微的尺寸扭曲。正如预期，较大的 q 对稀疏备择具有更高的检验功效，而较小的 q在稠密备择下表现更佳。比较$U_{n,2}$和$U_{n,2}^M$，可以看到$U_{n,2}^M$的检验功效明显降低。这并不意外，因为单调索引 U 统计量往往会损失一些统计效率。如果计算

L_2-范数检验统计量的复杂度能够达到 $O(n^2)$，我们建议计算完整的 U 统计量，即 $U_{n,2}$，并将其与 $U^M_{n,q}$ 结合，用于 $q = 4, 6$。对于这个例子，基于 $U_{n,2}$ 和 $U^M_{n,4}$ 的自适应检验，其检验功效似乎最佳，或接近检验功效最佳的单一-q 检验。

5.4. 空间符号双样本检验。 在本小节中，我们研究由（Chakraborty 和 Chaudhuri, 2017年）提出的双样本空间符号检验的检验水平和检验功效，并将其与基于 L_q-范数的扩展及自适应检验进行比较。第一个样本由 $\{X_i\}_{i=1}^n = \{Z_i + \mu\}_{i=1}^n$ 生成，第二个样本为 $\{Y_j\}_{j=1}^m$，其中我们设定 $\mu = \delta\,(\mathbf{1}_r^T,\ \mathbf{0}_{p-r}^T)^T$，并按以下方式生成 $\{Y_i\}$ 和 $\{Z_j\}$：

- 标准高斯分布： $Y_i, Z_j \overset{i.i.d.}{\sim} N(\mathbf{0}, I_p)$; $i = 1, \ldots, n$; $j = 1, \ldots, m$。
- t-分布： $Y_{il}, Z_{jl} \overset{i.i.d.}{\sim} t_3$; $i = 1, \ldots, n$; $j = 1, \ldots, m$; $l = 1, \ldots, p$。

我们令 $n = m = 50$, $p = 100$。

DGP	δ, r	$q=2$	$q=4$	$q=6$	$q=2,4$	$q=2,6$
高斯	0, NA	2.9	6.2	6.8	4.2	5.2
	1, 2	17.2	94.2	91.0	91.2	90.8
	0.3, p	96.4	69.2	37.2	95.6	96.2
t_3	0, NA	2.4	5.0	5.6	4.2	4.7
	1, 2	3.8	37.8	41.4	34.0	38.6
	0.5, p	98.3	76.6	45.6	97.4	97.5

表4 双样本空间符号检验的检验水平和检验功效（百分比）

根据表4，除 $q = 2$这种情况外，大多数检验的检验水平都较为准确；在这种情况下，双样本空间符号检验的检验水平略显保守。这与单样本检验的发现相反，在单样本检验中，较大的 q 往往较为保守。与单样本空间符号检验的情况相同，较小和较大的 q 分别对应于针对稠密和稀疏备择假设的高检验功效。同样，基于$(2, 4)$ 或$(2, 6)$ 的自适应检验结合了单一 q 检验的优势，其检验功效与最佳单一-q 检验相当。

6. 真实数据应用。本节中，我们将双样本基于空间符号的检验（采用置换法）应用于机器学习 UCI 网站（http://www.cs.ucr.edu/eamonn/time_序列_数据）上公开的 Connectionist Bench 数据集。我们考察的数据集包含来自金属圆柱体的 111 个声呐信号样本，以及来自岩石的 97 个样本。我们遵循（Chakraborty 和 Chaudhuri, 2017年）中的方法，研究所提出的双样本检验在不同 q下的检验水平和检验功效。对于检验水平模拟，我们从岩石样本中随机抽取两个样本量为 40 的子样本，并计算平均拒绝率。对于检验功效，我们同样分别从金属和岩石样本中抽取两个样本量为 40 的子样本，并基于 1000 次重复计算平均拒绝概率。结果汇总于表5。

统计量	$q=2$	$q=4$	$q=6$	$q=2,4$	$q=2,6$
Size	2.1	5.9	6.2	4.1	5.5
检验功效	22.2	48.7	30.1	42.6	33.2

表 5 真实数据应用中的检验水平和检验功效（百分比）

从表中可以看出， $q=2$ 的检验水平较为保守，而其他检验的检验水平都相对准确。就检验功效而言， $q=4$ 的检验功效最高，其次是 $q=6$，它优于 $q=2$，这表明，$q=4$ 比 $q=2$更能捕捉潜在的备择假设。采用结合 $q=2$、4 （或2、6）的自适应检验，其检验功效高于在 q 下使用 $q=2$ 的单一基于统计量的检验，正如Chakraborty 和 Chaudhuri（2017年）的研究所示。这清楚地体现了我们开发的自适应检验的价值。在实践中，确定针对当前数据哪种 q 检验功效更高，以及在自适应组合中应使用哪一对q，都会很有帮助。我们将这一问题留待未来研究。

7. 结论。本文针对高维独立同分布数据提出了一类 L_q范数型 U 统计量，并证明在较弱的矩和累积量假设下，对于不同的 q，这些统计量渐近正态且相互独立。我们的研究受到He 等人(2021 年的启发。他们首次证明了用于检验均值向量和协方差矩阵的 L_q范数型 U 统计量具有渐近正态性和独立性。我们的结果适用于一大类核函数，可视为对He 等人 (2021 年所作工作的重要扩展；他们的研究仅限于阶数为1的核函数。特别是，我们通过理论分析和模拟，展示了这些结果在若干He 等人 (2021 年未涉及的全局检验问题中的实用性，包括单样本空间符号检验、线性模型系数的同时检验、通过 Kendall'sτ进行的分量独立性检验，以及双样本空间符号检验。

为兼顾稀疏和稠密备择，我们基于各自的 p 值，简单地将两个基于学生化 L_q 范数的检验统计量按 Bonferroni 类型进行组合，并证明这种组合可构成自适应检验，对稠密和稀疏备择均具有较高的检验功效。为减轻计算负担，我们提出一种对称 U 统计量的变体，在求和时只考虑单调指标，从而可以应用动态规划的思想。在统计效率仅有适度损失的情况下，我们将计算成本从 $O(n^{qr})$（对应完整的对称 U 统计量）降至 $O(n^r)$，其中 r 是核的阶数。根据模拟结果，将现有的基于 L_2 范数的检验与基于单调指标的 L_q 范数检验（$q=4$ 或 6）相结合，似乎能够以 $O(n^2)$ 阶的计算成本，对稀疏和稠密备择均取得令人满意的检验功效。

最后，我们简要指出几个值得探讨的未来研究方向。将渐近理论推广到高维时间序列将会很有意义，因为我们在本文中研究的许多全局检验问题都可以自然地针对高维时间序列数据提出。我们预计这一推广并非易事，因为时间依赖性可能会对 L_q范数检验统计量所用正态近似的有效性产生不可忽视的影响。此外，将我们的方法和理论推广到变点检测问题也可能很有意义。由于变点位置未知，该问题与双样本检验问题相似，但更为复杂；参见Zhang 等人(2021 年)，其中使用了 L_q范数检验统计量来进行高维数据中的均值变点检验。最后，渐近方差估计是推断中的一个重要问题，但在高维情形下尚未得到解决。我们推测自助法或许可行，但鉴于观测数据的高维性，应用时仍需谨慎，并且还需要严谨的理论支持。这些开放问题留待未来研究。

补充材料

《高维数据的自适应检验》的补充材料

The 补充材料包含论文中所述理论结果的全部证明 r.

参考文献

Andrews, D. W. (1991). 异方差与自相关一致的协方差矩阵估计。《计量经济学会杂志》，第817–858页。Arcones, M. A. and Gine, E. (1992). 论U统计量和V统计量的自助法。《统计年刊》，20(2):655–674。Bai, Z., Jiang, D., Yao, J.-F., and Zheng, S. (2009). 随机矩阵理论对大维协方差矩阵似然比检验的修正。《统计年刊》，37(6B):3822–3840。Bai, Z. and Saranadasa, H. (1996). 高维的影响：以一个双样本问题为例。《中华统计学志》，6(2):311–329。Cai, T., Liu, W., and Xia, Y. (2013). 高维稀疏情形下的双样本协方差矩阵检验与支撑集恢复。《美国统计学会杂志》，108(501):265–277。Cai, T. T. and Jiang, T. (2011). 随机矩阵相干性的极限定律及其在协方差结构检验和构造压缩感知矩阵中的应用。《统计年刊》，39(3):1496–1525。Cai, T. T., Liu, W., and Xia, Y. (2014). 依赖情形下高维均值的双样本检验。英国皇家统计学会会刊：B辑：统计方法学，76(2):349–372。Chakraborty, A. and Chaudhuri, P. (2017). 基于均值、空间符号和空间秩的高维数据检验。《统计年刊》，45(2):771–799。Chen, S. X. and Qin, Y.-L. (2010). 高维数据的双样本检验及其在基因集检验中的应用。《统计年刊》，38(2):808–835。Chen, S. X., Zhang, L.-X., and Zhong, P.-S. (2010). 高维协方差矩阵检验。《美国统计学会杂志》，105(490):810–819。Drton, M., Han, F., and Shi, H. (2020). 通过秩相关的最大值进行高维一致独立性检验。《统计年刊》，48(6):3206–3227。El Karoui, N. (2009). 测度集中与随机矩阵的谱：在相关矩阵、椭圆分布及其他领域中的应用。《应用概率年刊》，19(6):2362–2405。Goeman, J. J., Van De Geer, S. A., and Van Houwelingen, H. C. (2006). 针对高维备择假设的检验。英国皇家统计学会会刊：B辑（统计方法学），68(3):477–493。Gregory, K. B., Carroll, R. J., Baladandayuthapani, V., and Lahiri, S. N. (2015). 高维情形下均值相等的双样本检验。《美国统计学会杂志》，110(510):837–849。Hall, P. and Jin, J. (2010). 用于检测相关噪声中稀疏信号的创新高阶检验。《统计年刊》，38(3):1686–1732。Han, F., Chen, S., and Liu, H. (2017). 高维无分布独立性检验。《生物统计学》，104(4):813–828。He, Y., Xu, G., Wu, C., and Pan, W. (2021). 高维检验中的渐近独立U统计量。《统计年刊》，49(1):154–181。Jiang, T. (2004). 样本相关矩阵最大元素的渐近分布。《应用概率年刊》，14(2):865–880。Ledoit, O. and Wolf, M. (2002). 维数远大于样本量时关于协方差矩阵的一些假设检验。《统计年刊》，30(4):1081–1102。Leung, D. and Drton, M. (2018). 利用秩相关和进行高维独立性检验。《统计年刊》，46(1):280–307。Li, J. and Chen, S. X. (2012). 高维协方差矩阵的双样本检验。《统计年刊》，40(2):908–940。Liu, W.-D., Lin, Z., and Shao, Q.-M. (2008). 高维独立性新检验的渐近分布与Berry–Esseen界及其在随机优化中的应用。《应用概率年刊》，18(6):2337–2366。Shao, Q.-M. and Zhou, W.-X. (2014). 超高维随机矩阵相干性渐近分布的充要条件。《概率年刊》，42(2):623–648。Shao, X. and Wu, W. B. (2007). 非线性过程分数阶积分的局部Whittle估计。《计量经济学理论》，23(5):899–929。Srivastava, M. S. and Du, M. (2008). 一个观测数少于维数的均值向量检验。《多元分析杂志》，99(3):386–402。Srivastava, R., Li, P., and Ruppert, D. (2016). Raptt：一种利用随机投影的高维精确双样本检验。《计算与图形统计学杂志》，25(3):954–970。Wang, L., Peng, B., and Li, R. (2015). 一种针对均值向量的高维非参数多元检验。《美国统计学会杂志》，110(512):1658–1669。Wang, R., Zhu, C., Volgushev, S., and Shao, X. (2022). 通过自归一化推断高维数据中的变点。《统计年刊》，50(2):781–806。

Wu, C., Xu, G. 和 Pan, W.（2019）。广义线性模型中高维参数的自适应检验。《中华统计学志》，29(4)：2163–2186。Wu, W. B.（2005）。非线性系统理论：重新审视依赖性。《美国国家科学院院刊》，102(40)：14150–14154。Wu, W. B. 和 Shao, X.（2004）。迭代随机函数的极限定理。《应用概率学杂志》，41(2)：425–436。Xu, G., Lin, L., Wei, P. 和 Pan, W.（2016）。高维均值的自适应双样本检验。《生物统计学》，103(3)：609–624。Zhang, Y., Wang, R. 和 Shao, X.（2021）。高维数据变点的自适应推断。《美国统计学会杂志》，即将刊出。Zhong, P.-S. 和 Chen, S. X.（2011）。析因设计下高维回归系数的检验。《美国统计学会杂志》，106(493)：260–274。Zhurbenko, I. 和 Zuev, N.（1975）。具有混合性的平稳过程的高阶谱密度。《乌克兰数学杂志》，27(4)：364–373。

《高维数据的自适应检验》的补充材料

作者：Yangfan Zhang、Runmin Wang 和 Xiaofeng Shao

补充材料的组织如下。在附录 A中，我们针对所关注的若干检验问题验证主要理论所需的假设。附录B 包含所有技术证明。

附录 A：假设的验证

本节针对不同的检验问题验证相关假设。我们首先总结需要检验的假设。若核函数的阶数为 1（$r=1$），或更一般地，若核函数完全退化（$r=s$），则假设 3 (b2) 仅对退化核适用（$s>1$）。

A.1. 均值检验与协方差检验。 这两个例子的核函数阶数为$r=1$，因此我们只需检验假设 3 (a) 和 (b1)。它们可由假设1 推出；均值检验的证明见 Wang 等人 (2022)，其中 $C_p=c_0$，且某个常数 c_0。因此，只要 X_i 满足假设 1，我们的理论便适用。对于协方差检验，只需将 $(x_{i,p_1}x_{i,p_2})_{(p_1,p_2)\in\mathcal{L}}$ 视为 $\mathbb{R}^{|\mathcal{L}|}$ 中的独立同分布随机向量，并验证它们满足假设1 即可。

A.2. 空间符号检验。 对于空间符号，回顾 $h_l(X_i)=x_{i,\,l}/\|X_i\|$ 以及 $H_{l,\,q}(X_1,\ \ldots,X_q)=h_l(X_1)\cdots h_l(X_q)$。直接验证 1 对 h_l 的成立在技术上很困难。我们转而证明 $D_{n,\,q}$ 的鞅中心极限定理。由于需要完成附录B中主要定理的证明，我们仅强调空间符号所需的新论证。回顾一下，相关的鞅差序列（经过适当标准化后）为

$$\begin{aligned}\xi_{n,i}=&\alpha_1 q!n^{-q/2}\tilde{\Sigma}_1^{-1/2}(q)\sum_{l\in\mathcal{L}}\sum_{1\le i_1<\cdots<i_{q-1}<i}H_{l,q}^{(t)}(X_{i_1},\ldots,X_{i_{q-1}},X_i)\\&+\alpha_2 Q!n^{-Q/2}\tilde{\Sigma}_1^{-1/2}(Q)\sum_{l\in\mathcal{L}}\sum_{1\le i_1<\cdots<i_{Q-1}<i}H_{l,Q}^{(t)}(X_{i_1},\ldots,X_{i_{Q-1}},X_i)\\\overset{\Delta}{=}&\alpha_1 q!\xi_{n,i}^{(1)}+\alpha_2 Q!\xi_{n,i}^{(2)}.\end{aligned}$$

F遵循 Wang 等人 (2015 年)，我们考虑椭圆分布的随机向量 X_i wi th

$$X_i=\mu+\varepsilon_i,\quad \varepsilon_i=\Gamma R_i U_i, \tag{11}$$

其中 $\Gamma\in\mathbb{R}^{p\times p}$, U_i 在 $\mathbb{R}^p$ 的单位球面上均匀分布，且 $R_i\ge 0$ 是一个与 U_i独立的尺度随机变量。记 $\Omega=\Gamma\Gamma^T$，则协方差矩阵为 $\Sigma=p^{-1}D\Omega$，其中 $D\overset{\Delta}{=}\mathbb{E}[R_i^2]$。记 $W_i=\Gamma U_i=(w_{i,\,1},\ \ldots,\ w_{i,\,p})^T$，并且有 $X_i=R_iW_i$，因此在原假设下 $\mathbb{E}[W_i]=\mathbb{E}[X_i]=\mathbf{0}$。

考虑事件 $A_i=\left\{\|W_i\|^2\ge \mathrm{tr}_{2p}^{(\Omega)}\right\}$。根据 Wang 等人 （2015 年）中的引理 3，我们有

$$P(A_i)\ge 1-c_1\exp\left\{\frac{\mathrm{tr}^2(\Omega)}{-128p\lambda_{\max}^2(\Omega)}\right\},$$

其中 $c_1=2\exp(\pi/2)$ 为一个常数。

我们作出以下假设，该假设类似于Wang 等人（2015 年）中的假设（C2），以便应用（El Karoui，2009）提出的集中不等式。

假设5. 假设随机向量 X_i 服从（11）所示的椭圆分布，并且

$$p^{2q-2}\exp\left\{\frac{\mathrm{tr}^2(\Omega)}{-128p\lambda_{\max}^2(\Omega)}\right\}=O(1). \tag{12}$$

该假设略强于 Wang 等人 (2015 年) 所用的假设，并且可以通过类似的论证加以验证，因为我们预期指数项会以较快的速度衰减至 0。更准确地说，令 $\lambda_1\leq\lambda_2\leq\ldots\leq\lambda_p$ 表示 Σ的特征值。假设有k_1 个特征值以 p^a 的速度衰减至 0，另有 k_2 个特征值以 p^a 的速度发散至 ∞，其中 $a>0$为某个常数。其余$(p-k_1-k_2)$ 个特征值假设有某个 $b>0$作为下界。我们假设 k_1 和 k_2 有界，并回顾一下，我们考虑的是满足$p\to\infty$的高维情形。我们有

$$\frac{\mathrm{tr}(\Sigma)}{\sqrt{p}\lambda_{\max}(\Sigma)}\geq\frac{b(p-k_1-k_2)}{\sqrt{p}\lambda_p}.$$

因此，只要满足 $a<1/2$，指数项就会比 p^{2-2q} 更快地衰减至 0。即使 k_1 和 k_2 也发散至 ∞，只要施加额外限制，这些假设仍然可以成立。第二个陈述实质上假设，经过适当标准化后， R_i 具有有界的四阶矩。

要应用鞅中心极限定理（Billingsley（2008）中的定理 35.12），我们需要验证以下两个条件。

$$(i)\quad \forall\varepsilon>0,\sum_{i=1}^{n}\mathbb{E}\left[\xi_{n,i}^2\mathbf{1}\left\{|\xi_{n,i}|>\varepsilon\right\}\middle|\mathcal{F}_{i-1}\right]\xrightarrow{p}0.$$

$$(ii)\quad V_n=\sum_{i=1}^{n}\mathbb{E}\left[\xi_{n,i}^2|\mathcal{F}_{i-1}\right]\xrightarrow{p}\alpha_1^2(qr)!+\alpha_2^2(Qr)!.$$

根据引理 7.2 (i)，需要满足以下等式，即,(14)

$$\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}\left[\prod_{u=1}^{4}h_{l_u}(X_{i_1^{(u)}})\cdots h_{l_u}(X_{i_q^{(u)}})\right]=\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}\left[\prod_{u=1}^{4}H_{l_u}(X_{i_1^{(u)}},\cdots,X_{i_q^{(u)}})\right]\leq\tilde{\Sigma}_1^2(q).$$

考虑事件 $A_q=\left\{\|W_i\|^2\geq\mathrm{tr}_{2p}^{(\Omega)},\quad i=1,\quad\ldots\ldots,\quad q.\right\}$。根据 Wang 等人（2015 年）(13)中的引理 3，我们有

$$P(A_q)\geq1-qc_1\exp\left\{\frac{\mathrm{tr}^2(\Omega)}{-128p\lambda_{\max}^2(\Omega)}\right\},$$

其中 $c_1=2\exp(\pi/2)$ 是一个常数。

请注意，我们有

$$\begin{aligned}&\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}\left[\prod_{u=1}^{4}H_{l_u}(X_{i_1^{(u)}},\cdots,X_{i_q^{(u)}})\right]\\ \leq&\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}H_l(X_1,\cdots,X_q)\right)^4\right]\\ \leq&\mathbb{E}\left[\left(\frac{\sum_{l\in\mathcal{L}}w_{1,l}\cdots w_{q,l}}{\|W_1\|\cdots\|W_q\|}\right)^4(1_{A_q}+1_{A_q^c})\right]\end{aligned}$$

$$\leq(\mathrm{tr}(\Omega)/2p)^{-2q}\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}w_{1,l}\cdots w_{q,l}\right)^4\right]+P(A_q^c)\quad(q\geq 2).$$

$$\leq(\mathrm{tr}(\Omega)/2p)^{-2q}(\mathbb{E}[R_1^4])^{-q}\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}x_{1,l}\cdots x_{q,l}\right)^4\right]+P(A_q^c)\quad(x_{i,l}=R_iw_{i,l}).$$

$$\lesssim(D\mathrm{tr}(\Omega)/2p)^{-2q}\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}x_{1,l}\cdots x_{q,l}\right)^4\right]+P(A_q^c)\quad(\mathbb{E}[R_i^4]\geq\mathbb{E}^2[R_i^2]=D^2)$$

$$\lesssim\mathrm{tr}^{-2q}(\Sigma)\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}x_{1,l}\cdots x_{q,l}\right)^4\right]+P(A_q^c)\quad(\Sigma=p^{-1}D\Omega)$$

请注意，我们有（由均值检验的验证可知）

$$\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}x_{1,l}\cdots x_{q,l}\right)^4\right]\lesssim\|\Sigma\|_q^{2q},$$

的阶数至多为 p^2，而且由于 $\mathrm{tr}(\Sigma)$ 的阶数至少为 p，因此可推出 s

$$\mathrm{tr}^{-2q}(\Sigma)\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}x_{1,l}\cdots x_{q,l}\right)^4\right]\lesssim p^{2-2q}.$$

˜关于 $\Sigma_1(q)$，我们有

$$\begin{aligned}\tilde{\Sigma}_1^{1/q}(q)=&\left(\sum_{l_1,l_2\in\mathcal{L}}\mathbb{E}^q\left[\frac{x_{1,l_1}x_{1,l_2}}{\|X_1\|^2}\right]\right)^{1/q}\geq\left(\sum_{l\in\mathcal{L}}\mathbb{E}^q\left[\frac{x_{1,l}^2}{\|X_1\|^2}\right]\right)^{1/q}\\=&p^{1/q-1}\left(\sum_{l\in\mathcal{L}}\mathbb{E}^q\left[\frac{x_{1,l}^2}{\|X_1\|^2}\right]\right)^{1/q}\left(\sum_{l\in\mathcal{L}}1^{q/(q-1)}\right)^{1-1/q}\\\geq&p^{1/q-1}\sum_{l\in\mathcal{L}}\mathbb{E}\left[\frac{x_{1,l}^2}{\|X_1\|^2}\right]=p^{1/q-1}.\end{aligned}$$

因此，我们得出结论：$\tilde{\Sigma}_1(q)\geq p^{1-q}$。最后，为了得出结论 (13)，还需注意到，(12) 蕴含 $P(A_q^c)\lesssim p^{2-2q}$。证明至此完成。

A.3. 线性回归系数零假设的检验。记 $W_i=(X_i,Y_i)$。不失一般性，我们假设 $\mathbb{E}[X]=0$，并令 $\Sigma=(\Sigma_{l_1,l_2})_{l_1,l_2=1}^p\overset{\Delta}{=}\mathrm{var}(X)$。在原假设下，我们有

$$h_l(W_{[2]})=(\varepsilon_1-\varepsilon_2)(x_{1,l}-x_{2,l}),\quad l=1,\ldots,p,$$

其中，为了简化记号，我们省略了因子 1/2。因此，

$$h_l^{(1)}(W_1)=\mathbb{E}[h_l(W_{[2]})|W_1]=\varepsilon_1x_{1,l},$$

and

$$h_l^{(2)}(W_{[2]})=h_l(W_{[2]})-h_l^{(1)}(W_1)-h_l^{(1)}(W_2)=-(\varepsilon_1x_{2,l}+\varepsilon_2x_{1,l}).$$

由于核非退化，只需检验假设2 和3 (a,b1)。然而，通过定义独立同分布的零均值随机变量 $Y_i \triangleq \varepsilon_i X_i$ 使得方差$(Y) = \sigma^2 \cdot \Sigma$ 且$\sigma^2 \triangleq$方差(ε_1)，我们可以看出 $h_l^{(1)}(W_i) = y_{i,\ l}$。因此，若独立同分布随机向量3 (a) 和 (b1) 成立，则假设1 对独立同分布随机向量成立 $\varepsilon_i X_i$。

接下来我们检验假设2，即$\Sigma_2(q) = o(n^q \Sigma_1(q))$。注意，

$$\sigma_1(l_1,l_2) = \mathbb{E}[\varepsilon_1 x_{1,l_1} \varepsilon_1 x_{1,l_2}] = \sigma^2 \Sigma_{l_1,l_2},$$

$$\sigma_2(l_1,l_2) = \mathbb{E}[(\varepsilon_1 x_{2,l_1} + \varepsilon_2 x_{1,l_1})(\varepsilon_1 x_{2,l_2} + \varepsilon_2 x_{1,l_2})] = 2\sigma^2 \Sigma_{l_1,l_2} = 2\sigma_1(l_1,l_2).$$

因此，我们有 $\Sigma_2(q) = 2^q \Sigma_1(q)$，因此假设 2 成立。

A.4. 基于 Kendall's τ 的分量独立性检验。我们首先通过证明 $\Sigma_2(q) = o(n^q \Sigma_1(q))$ 来验证假设2。回顾一下，在此例中 $\mathcal{L} = \{l = (d,d')\ 1 \le d < d' \le p\}$，并记 $L = |\mathcal{L}| = p(p-1)/2$。对于 $l = (d, d')$，由于

$$h_l(X_{[2]}) = \text{sgn}(x_{1,d} - x_{2,d})\text{sgn}(x_{1,d} - x_{2,d'}),$$

在分量独立性的原假设下，我们有，

$$h_l^{(1)}(X_1) = \mathbb{E}[h_l(X_{[2]})|X_1] = (2F_d(x_{1,d}) - 1)(2F_{d'}(x_{1,d'}) - 1),$$

$$h_l^{(2)}(X_{[2]}) = h_l(X_{[2]}) - h_l^{(1)}(X_1) - h_l^{(1)}(X_2),$$

其中我们用 F_d 表示 $x_{1,d}$的累积分布函数。我们假设 X_i 具有连续分布。因此，$U_d \triangleq F_d(x_{1,d}) \sim Unif(0,1)$，且它们在原假设下相互独立。我们有

$$\begin{aligned}\sigma_1(l_1,l_2) &= \mathbb{E}[h_{l_1}^{(1)}(X_1) h_{l_2}^{(1)}(X_1)] = \mathbb{E}[(2U_{d_1} - 1)(2U_{d_1'} - 1)(2U_{d_2} - 1)(2U_{d_2'} - 1)] \\ &= I(l_1 = l_2) \cdot \mathbb{E}[(2U_{d_1} - 1)^2]\mathbb{E}[(2U_{d_1'} - 1)^2] = \frac{1}{9} \cdot I(l_1 = l_2),\end{aligned}$$

$$\begin{aligned}\sigma_2(l_1,l_2) &= \mathbb{E}[h_{l_1}(X_{[2]}) h_{l_1}(X_{[2]})] - 2\mathbb{E}[h_{l_1}^{(1)}(X_1) h_{l_2}^{(1)}(X_1)] \\ &= I(l_1 = l_2) - \frac{2}{9} I(l_1 = l_2) = \frac{7}{9} I(l_1 = l_2).\end{aligned}$$

因此，我们有 $\Sigma_2(q) = 7^q \Sigma_1(q) = (7/9)^q L$，因此假设 2 对 $h_l^{(1)}$ 成立。

由于核非退化，因此还需验证假设 3 (a) 和 (b1)。对于 (a)，我们有

$$\sum_{l_1,l_2,l_3,l_4 \in \mathcal{L}} [\sigma_1(l_1,l_2)\sigma_1(l_3,l_4)\sigma_1(l_1,l_4)\sigma_1(l_2,l_3)]^{q/2} = \left(\frac{7}{9}\right)^{2q} L.$$

确实有 $o(\Sigma_2^2(q))$，因为 $p \to \infty$，且因为 $\Sigma_2^2(q) = (7/9)^{2q} L^2$。

为简化记号，我们用 Z_{l_h}表示 $h_{l_h}^{(1)}(X_1)$。对于 (b1)，我们有 cum$(Z_{l_1}, \ldots, Z_{l_4})$ $= \mathbb{E}[Z_{l_1} Z_{l_2} Z_{l_3} Z_{l_4}] - \mathbb{E}[Z_{l_1} Z_{l_2}]\mathbb{E}[Z_{l_3} Z_{l_4}] - \mathbb{E}[Z_{l_1} Z_{l_3}]\mathbb{E}[Z_{l_2} Z_{l_4}] - \mathbb{E}[Z_{l_1} Z_{l_4}]\mathbb{E}[Z_{l_2} Z_{l_3}]$。容易看出，右侧的每一项都被一个常数所界。此外，若至少存在一个指标 l_h，使得 d_h 或 d_h' 与其他任何指标都不相同，则右侧为零，因为U_d相互独立。因此，要使某一项非零，d_h 和 d_h' 都必须在 $\{l_1,l_2,l_3,l_4\} = \{d_1, d_1', \ldots, d_4, d_4'\}$中至少出现两次，这意味着求和 $\sum_{l_1,\ldots,l_4 \in \mathcal{L}} = \sum_{1 \le d_1 < d_1' \le p, \ldots, 1 \le d_4 < d_4' \le p}$至多包含 $O(p^4)$ 项。由此可得

$$\sum_{l_1,l_2,l_3,l_4 \in \mathcal{L}} \text{cum}(h_{l_1}^{(1)}(X_1), \ldots, h_{l_4}^{(1)}(X_1)) = O(p^4) = O(L^2) = O(\Sigma_2^2(q)),$$

且（b1）成立。

附录 B：定理的技术证明

B.1. 引理 3.2 的证明。为方便记号，我们将 $U_{n,q}$ 和 $U_{n,\boldsymbol{\alpha}}$ 的未归一化版本定义为

$$D_{n,q}=\frac{n!}{(n-qr)!}U_{n,q};\quad D_{n,\boldsymbol{\alpha}}=\frac{n!}{(n-c)!}U_{n,\boldsymbol{\alpha}},$$

其中，$c=\sum_t t\alpha_t$ 是 $h_l^{\boldsymbol{\alpha}}$的阶数。

对于 $I=(i_1,i_2,\ldots,i_{qr})\in P_{qr}(n)$，记 $I_c=(i_{(c-1)r+1},\ldots,i_{cr})\in P_r(n)$为 $c=1,\ldots,q$，因此我们也可以写作 $I=(I_1,I_2,\ldots,I_q)$。我们有

$$\begin{aligned}
D_{n,q}&=\sum_{l\in\mathcal{L}}\sum_{I\in P_{qr}(n)}(\otimes^q h_l)(X_I)\\
&=\sum_{l\in\mathcal{L}}\sum_{I\in P_{qr}(n)}\prod_{c=1}^{q}h_l(X_{I_c})\\
&=\sum_{l\in\mathcal{L}}\sum_{I\in P_{qr}(n)}\prod_{c=1}^{q}\left(\sum_{t=1}^{r}\sum_{J\subseteq I_c:|J|=t}h_l^{(t)}(X_J)\right)\\
&=\sum_{l\in\mathcal{L}}\sum_{c=q}^{qr}\frac{(n-c)!}{(n-qr)!}\sum_{\boldsymbol{\alpha}:\sum_t t\alpha_t=c}C_{\boldsymbol{\alpha}}\sum_{I\in P_c(n)}h_l^{\boldsymbol{\alpha}}(X_I)\\
&=\sum_{c=q}^{qr}\frac{(n-c)!}{(n-qr)!}\sum_{\boldsymbol{\alpha}:\sum_t t\alpha_t=c}C_{\boldsymbol{\alpha}}D_{n,\boldsymbol{\alpha}},
\end{aligned}$$

对于某些常数 $C_{\boldsymbol{\alpha}}$，这里我们用到了如下事实：

$$h_l(X_I)=\sum_{t=1}^{r}\sum_{J\subseteq I:|J|=t}h_l^{(t)}(X_J),\tag{14}$$

并将具有相同核 $h_l^{\boldsymbol{\alpha}}$（即具有相同 $\boldsymbol{\alpha}$）的项合并为相应的 U 统计量。因此我们得出

$$U_{n,q}=\sum_{c=q}^{qr}\sum_{\boldsymbol{\alpha}:\sum_t t\alpha_t=c}C_{\boldsymbol{\alpha}}U_{n,\boldsymbol{\alpha}}.$$

˜由于核 $h_{l,\ \boldsymbol{\alpha}}$ 的阶数 $\sum_t t\alpha_t$ 小于 qs 等于 0，我们也可以从 $c=qs$ 而不是 $c=q$ 开始求和指标。

注意，通常可以显式地计算 $C_{\boldsymbol{\alpha}}$。我们只推导主导项所需的 $\boldsymbol{\alpha}=e_t(r)$。在（14）中，$h_l$ 的分解包含 $h_l^{(t)}$ 中的 $\binom{r}{t}$ 项。因此，$\otimes^q h_l^{(t)}$ 中的项总数等于 $P_{n-qr}^n\binom{r}{t}^q$。因此，对于 $\boldsymbol{\alpha}=e_t(r)$，通过匹配项数 $\otimes^q h_l^{(t)}$，有 $P_{n-qr}^n\binom{r}{t}^q=\frac{(n-c)!}{(n-qr)!}P_c^nC_{\boldsymbol{\alpha}}$，这意味着 $C_{\boldsymbol{\alpha}}=\binom{r}{t}^q$。——

B.2. 引理3.3的证明。我们分析$U_{n,\boldsymbol{\alpha}}$的渐近方差。对于 $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_r)$，
我们定义

$$\tilde{\sigma}_{\boldsymbol{\alpha}}(l_1,l_2)=\mathbb{E}[\tilde{h}_{l_1}^{\boldsymbol{\alpha}}(X_I)\tilde{h}_{l_2}^{\boldsymbol{\alpha}}(X_I)],$$

˜和$\Sigma_{\boldsymbol{\alpha}} = \sum_{l_1,l_2} \tilde{\sigma}_{\boldsymbol{\alpha}}(l_1,l_2)$。注意，根据柯西–施瓦茨不等式，我们有

$$\tilde{\Sigma}_{\boldsymbol{\alpha}} = \mathbb{E}\left[\left(\sum_{l\in\mathcal{L}} \tilde{h}_l^{\boldsymbol{\alpha}}(X_I)\right)^2\right] \le \mathbb{E}\left[\left(\sum_{l\in\mathcal{L}} h_l^{\boldsymbol{\alpha}}(X_I)\right)^2\right] = \sum_{l_1,l_2\in\mathcal{L}} \prod_{t=s}^{r} \sigma_t^{\alpha_t}(l_1,l_2),$$

˜因为 $\sum_{l\in\mathcal{L}} h_l^{\boldsymbol{\alpha}}$是 $\sum_{l\in\mathcal{L}} h_l^{\boldsymbol{\alpha}}$的对称化核。此外，应用 Hölder不等式，我们有

$$\sum_{l_1,l_2\in\mathcal{L}} \prod_{t=s}^{r} \sigma_t^{\alpha_t}(l_1,l_2) \le \prod_{t=s}^{r} \left(\sum_{l_1,l_2\in\mathcal{L}} \sigma_t^{q}(l_1,l_2)\right)^{\alpha_t/q} = \prod_{t=s}^{r} \Sigma_t^{\alpha_t/q}(q).$$

回顾一下，Hoeffding 分解可以写为

$$U_{n,q} = \binom{r}{s}^q U_{n,q}^{(qs)} + \sum_{c=qs+1}^{qr} \sum_{\boldsymbol{\alpha}:\sum_t t\alpha_t = c} C_{\boldsymbol{\alpha}} U_{n,\boldsymbol{\alpha}},$$

˜ ˜注意到 $h_l^{\boldsymbol{\alpha}}$是完全退化且相互正交的，这意味着 $\mathrm{var}(U_{n,\ \boldsymbol{\alpha}}) = O(n^{-c}\Sigma_{\boldsymbol{\alpha}})$，其中 $c = \sum_t \alpha_t$ 且 $\mathrm{var}(U_{n}^{(qs)},\ _q) = O(n^{-qs}\tilde{\Sigma}_s(q))$。因此，对于任意满足$\sum_t t\alpha_t > qs$的 $\boldsymbol{\alpha}$，我们有：

$$\begin{aligned}
&\mathrm{var}(U_{n,\boldsymbol{\alpha}}) = O(n^{-c}\tilde{\Sigma}_{\boldsymbol{\alpha}}) \\
=&O(n^{-c}\Sigma_{\boldsymbol{\alpha}}) = O\left(n^{-c}\prod_{t=s}^{r}\Sigma_t^{\alpha_t/q}(q)\right) \quad (\tilde{\Sigma}_{\boldsymbol{\alpha}} \le \Sigma_{\boldsymbol{\alpha}} \text{ and Hölder}) \\
=&o\left(n^{-c}\prod_{t=s}^{r} n^{\alpha_t(t-s)}\Sigma_s^{\alpha_t/q}(q)\right) \\
=&o(n^{-qs}\tilde{\Sigma}_s(q)) = o\big(\mathrm{var}(U_{n,q}^{(qs)})\big),
\end{aligned}$$

根据假设 2。因此，除了 $U_{n}^{(qs)},\ _q$ 之外，所有项都可忽略不计，从而完成证明。

B.3. 命题 **3.4** 的证明。 h_l 的退化性意味着 $\otimes^q h_l$ 也完全退化，并且我们有 $D_{n,\ q} = D_{n\ q}^{(qr)},$。注意，本节稍微滥用记号，定义$D_{n}^{(qr)},\ _q = \binom{n}{qr} U_{n\ q}^{(qr)},$（而不是等于 $!\,(n-qr)!$ 的系数），使得

$$D_{n,q}^{(qr)} = \sum_{l\in\mathcal{L}} \sum_{1\le i_1<\cdots<i_{qr}\le n} H_{l,q}^{(r)}(X_{i_1},\ldots,X_{i_{qr}}).$$

。由沃尔德技巧，只需证明对于任意 $q \ne Q \in 2\mathbb{Z}_+$ 和 $\alpha_1,\ \alpha_2 \in \mathbb{R}$，

$$\alpha_1 n^{-qr/2}\tilde{\Sigma}_r^{-1/2}(q) D_{n,q}^{(qr)} + \alpha_2 n^{-Qr/2}\tilde{\Sigma}_r^{-1/2}(Q) D_{n,Q}^{(Qr)} \xrightarrow{\mathcal{D}} N(0, \alpha_1^2(qr)! + \alpha_2^2(Qr)!).$$

核的退化性意味着左侧是一个鞅，其鞅差为

$$\begin{aligned}
\xi_{n,i} =& \alpha_1(qr)! n^{-qr/2}\tilde{\Sigma}_r^{-1/2}(q) \sum_{l\in\mathcal{L}} \sum_{1\le i_1<\cdots<i_{qr-1}<i} H_{l,q}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X_i) \\
&+ \alpha_2(Qr)! n^{-Qr/2}\tilde{\Sigma}_r^{-1/2}(Q) \sum_{l\in\mathcal{L}} \sum_{1\le i_1<\cdots<i_{Qr-1}<i} H_{l,Q}^{(r)}(X_{i_1},\ldots,X_{i_{Qr-1}},X_i) \\
\overset{\Delta}{=}& \alpha_1(qr)!\xi_{n,i}^{(1)} + \alpha_2(Qr)!\xi_{n,i}^{(2)}.
\end{aligned}$$

注意，第一行中的$(qr)!$和$(Qr)!$是由于在 qr 和 Qr指标之间进行置换而出现的。要应用鞅中心极限定理（Billingsley（2008）中的定理 35.12），我们需要验证以下两个条件。

$$(i) \quad \forall \varepsilon>0, \sum_{i=1}^{n} \mathbb{E}\left[\xi_{n,i}^{2} \mathbf{1}\left\{\left|\xi_{n,i}\right|>\varepsilon\right\} \middle| \mathcal{F}_{i-1}\right] \xrightarrow{p} 0.$$

$$(ii) \quad V_{n}=\sum_{i=1}^{n} \mathbb{E}\left[\xi_{n,i}^{2} | \mathcal{F}_{i-1}\right] \xrightarrow{p} \alpha_{1}^{2}(qr)!+\alpha_{2}^{2}(Qr)!.$$

要证明 (i)，只需证明

$$\sum_{i=1}^{n} \mathbb{E}[\xi_{n,i}^{4}] \lesssim \sum_{i=1}^{n} \mathbb{E}[(\xi_{n,i}^{(1)})^{4}+(\xi_{n,i}^{(2)})^{4}] \to 0.$$

注意到

$$\sum_{i=1}^{n} \mathbb{E}[(\xi_{n,i}^{(1)})^{4}]=n^{-2qr} \tilde{\Sigma}_{r}^{-2}(q) \sum_{i=1}^{n} \mathbb{E}\left[\left(\sum_{1 \leq i_{1}<\cdots<i_{qr-1}<i} \sum_{l \in \mathcal{L}} H_{l,q}^{(r)}(X_{i_{1}}, \ldots, X_{i_{qr-1}}, X_{i})\right)^{4}\right]$$

$$\lesssim \sum_{i=1}^{n} n^{-2qr} \tilde{\Sigma}_{r}^{-2}(q) n^{2(qr-1)} \mathbb{E}\left[\prod_{u=1}^{4} H_{l_{u},q}^{(r)}(X_{i_{1}^{(u)}}, \ldots, X_{i_{qr-1}^{(u)}}, X_{i})\right]$$

$$\lesssim n^{-1} \to 0,$$

其中，第一个不等式利用了所有下标都必须至少出现两次才会有非零期望这一事实，第二个不等式则由引理 B.1 得出。类似的论证可得 $\sum_{i=1}^{n} \mathbb{E}[(\xi_{n\, i}^{(2)},)^{4}] \to 0$.

要证明 (ii)，我们将求和拆分为两部分，

$$\sum_{i=qr}^{n} \mathbb{E}[(\xi_{n,i}^{(1)})^{2} | \mathcal{F}_{i-1}]$$

$$=n^{-qr} \widetilde{\Sigma}_{r}^{-1}(q) \sum_{l_{1}, l_{2} \in \mathcal{L}} \sum_{i=qr}^{n} \sum_{i_{t}^{(1)}<i, i_{t}^{(2)}<i} \mathbb{E}\left[H_{q,l_{1}}^{(r)}(X_{i_{1}^{(1)}}, \ldots, X_{i_{qr-1}^{(1)}}, X') H_{q,l_{2}}^{(r)}(X_{i_{1}^{(2)}}, \ldots, X_{i_{qr-1}^{(2)}}, X_{i}) \mid \mathcal{F}_{i-1}\right]$$

$$=n^{-qr} \widetilde{\Sigma}_{r}^{-1}(q) \sum_{l_{1}, l_{2} \in \mathcal{L}} \sum_{i=qr}^{n} \left(\sum_{i_{t}^{(1)}<i, i_{t}^{(2)}<i}^{(1)} \mathbb{E}\left[H_{q,l_{1}}^{(r)}(X_{i_{1}^{(1)}}, \ldots, X_{i_{qr-1}^{(1)}}, X') H_{q,l_{2}}^{(r)}(X_{i_{1}^{(2)}}, \ldots, X_{i_{qr-1}^{(2)}}, X') \mid \mathcal{F}_{i-1}\right] \right.$$

$$\left. + \sum_{i_{t}^{(1)}<i, i_{t}^{(2)}<i}^{(2)} \mathbb{E}\left[H_{q,l_{1}}^{(r)}(X_{i_{1}^{(1)}}, \ldots, X_{i_{qr-1}^{(1)}}, X') H_{q,l_{2}}^{(r)}(X_{i_{1}^{(2)}}, \ldots, X_{i_{qr-1}^{(2)}}, X') \mid \mathcal{F}_{i-1}\right] \right)$$

$$\triangleq V_{n}^{(1)}+V_{n}^{(2)},$$

其中，$\sum^{(1)}$ 遍历满足以下条件的指标：$i_{t}^{(1)}$ 和 $i_{t}^{(2)}$ 均可配对，且配对方式至多相差一个置换；对于 $t=1$，2，....., $qr-1$，而 $\sum^{(2)}$ 包含其余所有项。由于我们要对 X_{i} 求期望（以其余项为条件），因此用哑变量 X' 替换 X_{i}。注意，在上面的求和表达式中，为简化记号，我们将 $i_{1}^{(u)}<i_{2}^{(u)}<\cdots<i_{qr-1}^{(u)}<i$ 简写为 $i_{t}^{(u)}<i$，其中$u=1$，2。

我们首先证明 $V_{(\frac{1}{n}} \xrightarrow{\mathcal{P}} [)(qr)!\ ]^{-1}$。由于

$$\mathbb{E}[V_n^{(1)}] = n^{-qr}\widetilde{\Sigma}_r^{-1}(q) \sum_{l_1,l_2\in\mathcal{L}} \sum_{1\le i_1<\cdots<i_{qr}\le n} \mathbb{E}\left[H_{q,l_1}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')H_{q,l_2}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')\right]$$

$$\to [(qr)!]^{-1},$$

只需证明

$$\mathbb{E}\left[\left(V_n^{(1)}\right)^2\right] \to [(qr)!]^{-2},$$

这 意味着在 L^2 中收敛，从而依概率收敛。事实上，我们有 ve

$$\mathbb{E}[(V_n^{(1)})^2] = n^{-2qr}\widetilde{\Sigma}_r^{-2}(q) \sum_{l_1,l_2,l_3,l_4\in\mathcal{L}} \sum_{i,j=qr}^{n} \sum_{i_t<i} \sum_{j_t<j} \mathbb{E}\left[H_{q,l_1}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')H_{q,l_2}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')\right.$$

$$\left. H_{q,l_3}^{(r)}(X_{j_1},\ldots,X_{j_{qr-1}},X'')H_{q,l_4}^{(r)}(X_{j_1},\ldots,X_{j_{qr-1}},X'')\right].$$

我们将求和分解为 $\sum^{(a)}$ 和 $\sum^{(b)}$，其中 $\sum^{(a)}$ 包含满足以下条件的项： t $\{i_t\}_{t=1}^{qr-1}$ 和 $\{j_t\}_{t=1}^{qr-1}$ 没有交集，而 $\sum^{(b)}$ 则包含其余所有项。注意，

$$n^{-2qr}\tilde{\Sigma}_r^{-2}(q) \sum_{l_1,l_2,l_3,l_4\in\mathcal{L}} \sum_{i,j=qr}^{n} \sum_{i_t<i,j_t<j}^{(a)} \mathbb{E}\left[H_{q,l_1}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')H_{q,l_2}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')\right.$$

$$\left. H_{q,l_3}^{(r)}(X_{j_1},\ldots,X_{j_{qr-1}},X'')H_{q,l_4}^{(r)}(X_{j_1},\ldots,X_{j_{qr-1}},X'')\right]$$

$$\asymp n^{-2qr}\tilde{\Sigma}_r^{-2}(q) \sum_{i,j=qr}^{n} \binom{i}{qr-1}\binom{j}{qr-1} \sum_{l_1,l_2,l_3,l_4\in\mathcal{L}} \mathbb{E}\big[H_{q,l_1}^{(r)}(X_{[qr]})H_{q,l_2}^{(r)}(X_{[qr]})\big]\mathbb{E}\big[H_{q,l_3}^{(r)}(X_{[qr]})H_{q,l_4}^{(r)}(X_{[qr]})\big]$$

$$\asymp n^{-2qr}\tilde{\Sigma}_r^{-2}(q)[(qr)!]^{-2}n^{2qr}\tilde{\Sigma}_r^{2}(q) = [(qr)!]^{-2}.$$

对于 $\sum^{(b)}$，由于 $\{i_t\}$ 和 $\{j_t\}$ 之间的指标匹配，我们至少损失一个自由度。因此，根据引理 B.1 (i)，我们有

$$n^{-2qr}\widetilde{\Sigma}_r^{-2}(q) \sum_{l_1,l_2,l_3,l_4\in\mathcal{L}} \sum_{i,j=qr}^{n} \sum_{i_t<i,j_t<j}^{(b)} \mathbb{E}\left[H_{q,l_1}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')H_{q,l_2}^{(r)}(X_{i_1},\ldots,X_{i_{qr-1}},X')\right.$$

$$\left. H_{q,l_3}^{(r)}(X_{j_1},\ldots,X_{j_{qr-1}},X'')H_{q,l_4}^{(r)}(X_{j_1},\ldots,X_{j_{qr-1}},X'')\right]$$

$$= O(\frac{1}{n})\tilde{\Sigma}_r^{-2}(q)\cdot O(\tilde{\Sigma}_r^{2}(q)) \to 0..$$

因此，我们得出结论： $V_n^{(1)}$收敛于 $[(qr)!\ ]^{-2}$，在 L^2中。

对于 $V_{(2)n}$，只需证明 $\mathbb{E}[(V_{(2)n}\,)^2] \to 0$ 可以忽略不计。注意， $V_{(2)n}$ 中的被加项仅当 $\{i_t^{(u)}, j_t^{(u)}\}$ 中的某个指标至少出现两次时才非零。求和中有指标出现超过两次的项为 $O(1/n)$，这是由于 n中的阶数额外损失。因此，我们关注所有指标恰好出现两次的项，其求和项的数量也具有 $O(n^{2qr})$的阶数，这与 $\sum^{(a)}$ 在分析 $V_n^{(1)}$时类似。由引理 B.1，可得

$$V_n^{(2)} = n^{-2qr}\tilde{\Sigma}_r^{-2}(q)\cdot O(n^{2qr})\cdot o(\tilde{\Sigma}_r^{2}(q)) \to 0.$$

这意味着 $V_{(2)n}$ 在 L^2 中收敛到 0。因此，我们得出 $\sum_{i=qr}^{n}\mathbb{E}[(\xi_{n\,i}^{(1)},\,)^2|\mathcal{F}_{i-1}]\xrightarrow{\mathcal{P}}[(qr)!\,]^{-2}$。类似地，我们有 $\sum_{i=Qr}^{n}\mathbb{E}[(\xi_{n\,i}^{(2)},\,)^2|\mathcal{F}_{i-1}]\xrightarrow{\mathcal{P}}[(Qr)!\,]^{-2}$，且 $\sum_{i=qr}^{n}\mathbb{E}[\xi_{n,i}^{(1)}\xi_{n,i}^{(2)}|\mathcal{F}_{i-1}]\xrightarrow{\mathcal{P}}0$。将所有部分合并起来，我们便完成了证明。

LEMMA B.1. 给定四组索引，且每组内的索引互不相同。 $\{i_1^{(u)},\ldots,i_{qr}^{(u)}\}$， $\{j_1^{(u)},\ldots,j_{qr}^{(u)}\}$对于 $u=1$，2。 (i) 在假设3 (b1)下，我们有

$$\begin{aligned}&\sum_{l_1,l_2,l_3,l_4\in\mathcal{L}}\mathbb{E}\Big[H_{q,l_1}^{(r)}(X_{i_1^{(1)}},\ldots,X_{i_{qr}^{(1)}})H_{q,l_2}^{(r)}(X_{i_1^{(2)}},\ldots,X_{i_{qr}^{(2)}})\\&\quad H_{q,l_3}^{(r)}(X_{j_1^{(1)}},\ldots,X_{j_{qr}^{(1)}})H_{q,l_4}^{(r)}(X_{j_1^{(2)}},\ldots,X_{j_{qr}^{(2)}})\Big]\\=&O(\tilde{\Sigma}_r^2(q)).\end{aligned}$$

(ii) 此外，假设 $\{i_1^{(1)},\ldots,i_{qr-1}^{(1)}\}$与 $\{i_1^{(2)},\ldots,i_{qr-1}^{(2)}\}$不相同，且 $\{i_t^{(1)},i_t^{(2)},j_t^{(1)},j_t^{(2)}\}_{t=1}^{qr-1}$中的每个元素恰好出现两次。在假设 3 (a) 和 (b2)下，我们有

$$\begin{aligned}&\sum_{l_1,l_2,l_3,l_4\in\mathcal{L}}\mathbb{E}\Big[H_{q,l_1}^{(r)}(X_{i_1^{(1)}},\ldots,X_{i_{qr-1}^{(1)}},X')H_{q,l_2}^{(r)}(X_{i_1^{(2)}},\ldots,X_{i_{qr-1}^{(2)}},X')\\&\quad H_{q,l_3}^{(r)}(X_{j_1^{(1)}},\ldots,X_{j_{qr-1}^{(1)}},X'')H_{q,l_4}^{(r)}(X_{j_1^{(2)}},\ldots,X_{j_{qr-1}^{(2)}},X'')\Big]\\=&o(\tilde{\Sigma}_r^2(q)).\end{aligned}$$

引理 B.1 的证明。对于 (i)，由算术-几何平均不等式，因 $H_{q\,l}^{(r)}$，是 $\otimes^q h_l^{(r)}$ 的对称化核，我们有

$$\begin{aligned}&\sum_{l_1,l_2,l_3,l_4\in\mathcal{L}}\mathbb{E}\Big[H_{q,l_1}^{(r)}(X_{i_1^{(1)}},\ldots,X_{i_{qr}^{(1)}})H_{q,l_2}^{(r)}(X_{i_1^{(2)}},\ldots,X_{i_{qr}^{(2)}})\\&\quad H_{q,l_3}^{(r)}(X_{j_1^{(1)}},\ldots,X_{j_{qr}^{(1)}})H_{q,l_4}^{(r)}(X_{j_1^{(2)}},\ldots,X_{j_{qr}^{(2)}})\Big]\\=&\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}H_{q,l}^{(r)}(X_{i_1^{(1)}},\ldots,X_{i_{qr}^{(1)}})\right)\left(\sum_{l\in\mathcal{L}}H_{q,l}^{(r)}(X_{i_1^{(2)}},\ldots,X_{i_{qr}^{(2)}})\right)\right.\\&\quad\left.\left(\sum_{l\in\mathcal{L}}H_{q,l}^{(r)}(X_{j_1^{(1)}},\ldots,X_{j_{qr}^{(1)}})\right)\left(\sum_{l\in\mathcal{L}}H_{q,l}^{(r)}(X_{j_1^{(2)}},\ldots,X_{j_{qr}^{(2)}})\right)\right]\\\le&\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}H_{q,l}^{(r)}(X_{i_1^{(1)}},\ldots,X_{i_{qr}^{(1)}})\right)^4\right]\quad\text{(Hölder's inequality)}\\\le&\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}\otimes^q h_l^{(r)}(X_1,\ldots,X_{qr})\right)^4\right]\quad\text{(AM-GM inequality, as }H_{q,l}^{(r)}\text{ is an average of }\otimes^q h_l\text{)}\\=&\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}[h_{l_1}(X_1,\ldots,X_r)\cdots h_{l_1}(X_{(q-1)r+1},\ldots,X_{qr})\cdots\end{aligned}$$

$h_{l_4}(X_1,\ldots,X_r)\cdots h_{l_4}(X_{(q-1)r+1},\ldots,X_{qr})]$

(i) 的其余证明与Zhang 等人(2021 年中的引理 6.1 相同。这里我们给出一个自足的证明。我们证明一个略强的结果，与Zhang 等人(2021 年的结果一致，如下所示。

断言：在假设3 (b1)下，对于任意彼此不同的 $i_1^{(u)}, i_2^{(u)}, \ldots, i_{qr}^{(u)}$，且 u =1,2,3,4，我们有

$$\Big|\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}[h_{l_1}(X_{i_1^{(1)}},\ldots,X_{i_r^{(1)}})\cdots h_{l_1}(X_{i_{(q-1)r+1}^{(1)}},\ldots,X_{i_{qr}^{(1)}})\cdots h_{l_4}(X_{i_1^{(4)}},\ldots,X_{i_r^{(4)}})\cdots h_{l_4}(X_{i_{(q-1)r+1}^{(4)}},\ldots,X_{i_{qr}^{(4)}})]\Big|\lesssim\tilde{\Sigma}_r^2(q). \tag{15}$$

断言的证明：定义 $I_t^{(u)}=(i^{(u)}{}_{(r-1)t+1},\ldots,i^{(u)}{}_{(rt)})$。由Hölder不等式，我们有

$$\begin{aligned}
&\left|\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}[h_{l_1}(X_{I_1^{(1)}})\cdots h_{l_1}(X_{I_q^{(1)}}))\cdots h_{l_4}(X_{I_1^{(4)}})\cdots h_{l_4}(X_{I_q^{(4)}})]\right|\\
=&\left|\mathbb{E}\left[\prod_{u=1}^4\sum_{l_u\in\mathcal{L}}[h_{l_u}(X_{I_1^{(u)}})\cdots h_{l_u}(X_{I_q^{(u)}})\right]\right|\\
\le&\prod_{u=1}^4\left\{\mathbb{E}\left[\left(\sum_{l_u\in\mathcal{L}}[h_{l_u}(X_{I_1^{(u)}})\cdots h_{l_u}(X_{I_q^{(u)}})\right)^4\right]\right\}^{1/4}.
\end{aligned}$$

注意到

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{l\in\mathcal{L}}[h_l(X_1,\ldots,X_r)\cdots h_l(X_{(q-1)r+1},X_{qr})\right)^4\right]\\
=&\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}^q[h_{l_1}(X_1,\ldots,X_r)\cdots h_{l_4}(X_1,\ldots,X_r)]\\
\lesssim&\sum_{l_1,\ldots,l_4\in\mathcal{L}}\mathbb{E}^q[h_{l_1}(X_1,\ldots,X_r)h_{l_2}(X_1,\ldots,X_r)]\mathbb{E}^q[h_{l_3}(X_1,\ldots,X_r)h_{l_4}(X_1,\ldots,X_r)]\\
&+\sum_{l_1,\ldots,l_4\in\mathcal{L}}|\mathrm{cum}^q(h_{l_1}(X_1,\ldots,X_r),\ldots,h_{l_4}(X_1,\ldots,X_r))|\\
\lesssim&\Sigma_r^2(q)\asymp\tilde{\Sigma}_r^2(q).
\end{aligned}$$

上述不等式中倒数第二行是由CR 不等式和联合累积量的定义得到的。因此我们得出结论 (i)。

对于 (ii)，为简化记号，我们将左端写作

$$\sum_{l_1,l_2,l_3,l_4\in\mathcal{L}}\mathbb{E}\left[H_{q,l_1}^{(r)}(X_{i_1},\ldots,X_{i_{qr}})H_{q,l_2}^{(r)}(X_{i_{qr+1}},\ldots,X_{i_{2qr}})H_{q,l_3}^{(r)}(X_{i_{2qr+1}},\ldots,X_{i_{3qr}})H_{q,l_4}^{(r)}(X_{i_{3qr+1}},\ldots,X_{i_{4q}})\right],$$

满足约束条件 $i_{qr}=i_{2qr}$ 和 $i_{3qr}=i_{4qr}$。由于 $H_{q\,l}^{(r)}$,是对不同参数下的 $\otimes^q h_l^{(r)}$ 取平均，因此我们可以将左侧的期望分解为若干项的线性组合，

这些项的形式为

$$[\otimes^q h_{l_1}^{(r)}(X_{I_1})][\otimes^q h_{l_2}^{(r)}(X_{I_2})][\otimes^q h_{l_3}^{(r)}(X_{I_3})][\otimes^q h_{l_4}^{(r)}(X_{I_4})],$$

其中，I_1、I_2、I_3、I_4分别是$\{i_t\}_{t=1}^{qr}$、$\{i_t\}_{t=qr+1}^{2qr}$、$\{i_t\}_{t=2qr+1}^{3qr}$、$\{i_t\}_{t=3qr+1}^{4qr}$的排列。我们进一步记 $I_u=$（I_{u1}，……，I_{uq}），其中 I_{uc}是一个包含 r个元素的索引集，对于 $u=1$，……，4；$c=1$，……，q。由于 $\{I_1,\ldots,I_4\}$中的每个元素恰好出现两次，我们考虑以下两种情况。

- 索引集 I_{uc}两两配对。在这种情况下，约束条件 $I_1\neq I_2$和 $i_{qr}=i_{2qr}, i_{3qr}=i_{4qr}$意味着存在 c_1,c_2,c_3,c_4，使得 $I_{1,c_1}=I_{3,c_3}, I_{1,c_1}=I_{4,c_4}$，或者 $I_{1,c_1}=I_{4,c_4}, I_{2,c_2}=I_{3,c_3}$。此外，关于索引的假设意味着 $I_{1q}=I_{2q}$和 $I_{3q}=I_{4q}$。根据假设3(a)和 Hölder不等式，该项为 $o_p(\tilde{\Sigma}_r^2(q))$。论证与 Zhang 等人（2021）补充材料中的引理6.3证明基本相同。因此，为便于说明，我们这里只概述证明思路；完整证明请参阅 Zhang 等人（2021）的补充材料。

- 假设索引集 $\{I_{uc}\}$ 并非两两配对。接下来，我们利用累积量分解期望，并将其写成累积量乘积的求和形式。我们使用累积量分解，将该项写成多个累积量与期望的乘积，形式如下

$$\sum_{l_1,l_2,l_3,l_4\in\mathcal{L}}\prod_{k=1}^{K}\mathrm{cum}(h_{u_1^{(k)}}(I_{u_1^{(k)},c_1^{(k)}}),\ldots,h_{u_{m_k}^{(k)}}(X_{I_{u_{m_k}^{(k)},c_{m_k}^{(k)}}})),\tag{16}$$

其中，$(B_1,B_2,\ldots,B_K)$是$(I_1,\ldots,I_4)$的一个划分，且满足 $B_k=(I_{u_1^{(k)},c_1^{(k)}},\ldots,I_{u_{m_k}^{(k)},c_{m_k}^{(k)}})$。我们假设这些块是极小的，即 $B_1,\ldots,B_K$彼此没有重叠。否则，由于独立性，我们可以进一步将某些块分解为更小块的乘积。由于 I_{uc}并非全部两两配对，因此至少存在一个块 B_{k_0}，其中包含至少三个不同 u 对应的 I_{uc}，这些 u 属于{1,2,3,4}。我们将其记为

$$\mathrm{cum}^{q/m_k}(h_{l_1}(X_{i_1},\ldots,X_{i_r}),\ldots,h_{l_u}(X_{i_{(c-1)r+1}},\ldots,X_{i_{cr}}),\tag{17}$$

取 $u=3$或取 4，并取 $c=m_k$，以简化记号。利用 Hölder不等式，我们可以将（16）的（绝对值）界定为

$$\prod_{k=1}^{K}\left(\sum_{l_1,l_2,l_3,l_4\in\mathcal{L}}\mathrm{cum}^{q/m_k}(h_{u_1^{(k)}}(I_{u_1^{(k)},c_1^{(k)}}),\ldots,h_{u_{m_k}^{(k)}}(X_{I_{u_{m_k}^{(k)},c_{m_k}^{(k)}}}))\right)^{m_k/q}.$$

每一项都是 $O(\tilde{\Sigma}_r^{m_k/2q})$，因此乘积为 $O(\tilde{\Sigma}_r^2)$，因为 $\sum_{k=1}^K m_k=4q$。然而，请注意，右侧与（17）相关的项，即

$$\mathrm{cum}^{q/c}(h_{l_1}(X_{i_1},\ldots,X_{i_r}),\ldots,h_{l_u}(X_{i_{(c-1)r+1}},\ldots,X_{i_{cr}}),$$

根据假设 3(b2)，有 $o(\tilde{\Sigma}_r^{c/2q}(q))$。这意味着（16）为 $o_p(\tilde{\Sigma}_r^2(q))$，从而完成证明。

♢

注记。注意，由于阶数为 $r>2$ 的核函数具有特定结构，除了条件 (b1) 外，我们还必须作出更强的假设 (b2)。事实上，对于 $r=1$，引理 B.1 (ii) 可由 (b1) 推出，而这又蕴含鞅中心极限定理。然而，对于一般的 r，类似 (b1) 的直接扩展假设并不能显然地推出该引理，因此我们必须施加更强的假设 (b2)。

B.4. 定理3.7 的证明。为简化记号，我们定义

$$U_{n,q,l} = (P_{qr}^n)^{-1} \sum_{1\le i_1,i_2,\cdots,i_{qr}\le n} (\otimes^q h_l)(X_1,\ldots,X_{qr}),$$

使得 $U_{n,\ q} = \sum_{l\in\mathcal{L}} U_{n,\ q,\ l}$，以及

$$U_{n,\boldsymbol{\alpha},l} = (P_m^n)^{-1} \sum_{1\le i_1,i_2,\cdots,i_m\le n} h_l^{\boldsymbol{\alpha}}(X_1,\ldots,X_m),$$

其中，m 是 $h_l^{\boldsymbol{\alpha}}$的阶数，即 $m = \sum_t t\alpha_t$，因此 $U_{n,\ \boldsymbol{\alpha}} = \sum_{l\in\mathcal{L}} U_{n,\ \boldsymbol{\alpha},\ l}$。

~回顾一下，$U_{n,\ q}$ 是与中心化核 h_l 相关联的 U 统计量。我们还~将与 $U_{n,\ q,\ l}$ 以及 $U_{n,\ \boldsymbol{\alpha},\ l}$ 相关联的 h_l 对应量定义为 $U_{n,\ q,\ l}$ 以及 $U_{n,\ \boldsymbol{\alpha},\ l}$。

注意

$$U_{n,q} = \|\Theta\|_q^q + \tilde{U}_{n,q} + \sum_{c=1}^{q-1}\binom{q}{c}\sum_{l\in\mathcal{L}}\theta_l^{q-c}\tilde{U}_{n,c,l}.$$

~注意，$U_{n,\ c,\ l}$ 是与 $\otimes^c h_l$ 相关联的 U 统计量。我们可以应用引理 3.2 中的 Hoeffding 分解，将 $U_{n,\ c,\ l}$ 分解为 $U_{n,\ \boldsymbol{\alpha},\ l}$，其中所有满足 $\sum_{t=1}^r \alpha_t = c$的有效 $\boldsymbol{\alpha}$ 均适用。特别地，对于任意 $t<s$，我们有 h_l，因为 s具有退化阶数。

通过计算方差并应用Hölder不等式，对于任意$1\le c\le q-1$，我们有：

$$\begin{aligned}
&\text{var}\left(\sum_{l\in\mathcal{L}}\theta_l^{q-c}\tilde{U}_{n,\boldsymbol{\alpha},l}\right)\\
\lesssim&\sum_{l_1,l_2\in\mathcal{L}}\theta_{l_1}^{q-c}\theta_{l_2}^{q-c}n^{-\sum_{t=s}^r t\alpha_t}\text{cov}(\tilde{h}_{l_1}^{\boldsymbol{\alpha}},\tilde{h}_{l_2}^{\boldsymbol{\alpha}})\\
\lesssim&n^{-\sum_{t=s}^r t\alpha_t}\text{var}\left(\sum_{l\in\mathcal{L}}\theta_l^{q-c}\tilde{h}_l^{\boldsymbol{\alpha}}\right)\\
\le&n^{-\sum_{t=s}^r t\alpha_t}\text{var}\left(\sum_{l\in\mathcal{L}}\theta_l^{q-c}h_l^{\boldsymbol{\alpha}}\right)\quad\text{(AM-GM inequality)}\\
=&n^{-\sum_{t=s}^r t\alpha_t}\sum_{l_1,l_2\in\mathcal{L}}\theta_{l_1}^{q-c}\theta_{l_2}^{q-c}\prod_{t=s}^r\sigma_t^{\alpha_t}(l_1,l_2)\\
\le&n^{-\sum_{t=s}^r t\alpha_t}\left(\sum_{l_1,l_2\in\mathcal{L}}\theta_{l_1}^q\theta_{l_2}^q\right)^{(q-c)/q}\prod_{t=s}^r\left(\sum_{l_1,l_2\in\mathcal{L}}\sigma_t^q(l_1,l_2)\right)^{\alpha_t/q}\quad\text{(Hölder)}\\
=&n^{-\sum_{t=s}^r t\alpha_t}\|\Theta\|_q^{2(q-c)}\prod_{t=s}^r\Sigma_t^{\alpha_t/q}(q).
\end{aligned}\tag{18}$$

其中，第一个Hölder不等式的应用方式与引理3.2的证明相同。如果 $\alpha_s < c$ 使得 $\alpha_t > 0$ 对于某个 $t > s$，假设 2 意味着右侧为

$$n^{-\sum_{t=s}^{r} t\alpha_t} \|\Theta\|_q^{2(q-c)} \prod_{t=s}^{r} o\big([n^{q(t-s)}\Sigma_s(q)]^{\alpha_t/q}\big) = o\big(n^{-cs}\Sigma_s^{c/q}(q)\|\Theta\|_q^{2(q-c)}\big).$$

另一方面，如果 $\alpha_s = c$ 使其成为 $\boldsymbol{\alpha}$ 中唯一的非零分量，则右侧等于 $n^{-cs}\|\Theta\|_q^{2(q-c)}\Sigma_s^{c/q}(q)$。

总之，我们始终有（回忆一下， $\Sigma_s(q)$ 和 $\tilde{\Sigma}_s(q)$ 具有相同的阶数）

$$\operatorname{var}\left(\sum_{l\in\mathcal{L}} \theta_l^{q-c}\tilde{U}_{n,\boldsymbol{\alpha},l}\right) \lesssim \big(\|\Theta\|_q^{2q}\big)^{(q-c)/q}\left(n^{-qs}\tilde{\Sigma}_s(q)\right)^{c/q}.$$

这是 $\|\Theta\|_q^{2q}$ 与 $n^{-qs}\tilde{\Sigma}_s(q)$ 的几何平均。因此，如果 $\gamma_{n,q} \to 0$， $\operatorname{var}^{1/2}(\tilde{U}_{n,q}) \propto n^{-qs/2}\tilde{\Sigma}_s^{1/2}(q)$ 占主导地位，超过 $\|\Theta\|_q^q$ 和所有 $\sum_{l\in\mathcal{L}} \theta_l^{q-c}\tilde{U}_{n,\boldsymbol{\alpha},l}$，这意味着 $U_{n,q}$ 与 $U_{n,q}$ 具有相同的渐近分布。

另一方面，如果 $\gamma_{n,q} \to \infty$， $\|\Theta\|_q^q$ 是 $U_{n,q}$ 的主导项，并且其发散至 ∞。

对于满足 $\gamma_{n,q} \to \gamma \in (0,\infty)$ 的局部备择，只需证明交叉乘积项仍然被 $\|\Theta\|_q^q$ 和 $\tilde{U}_{n,q}$ 所支配；在这种情况下，它们具有相同的阶数。根据假设 4，注意到我们已经证明，若 $\alpha_t > 0$ 对某个 $t > s$ 成立，则

$$\operatorname{var}\left(\sum_{l\in\mathcal{L}} \theta_l^{q-c}\tilde{U}_{n,\boldsymbol{\alpha},l}\right) = o\big(n^{-cs}\tilde{\Sigma}_s^{c/q}(q)\|\Theta\|_q^{2(q-c)}\big) = o\big(\operatorname{var}(\tilde{U}_{n,q})\big).$$

因此，剩下只需证明在假设 4 下，结果对于 $\boldsymbol{\alpha} = ce_s(r)$ 也成立。在这种情况下，我们已在式（18）的第三行证明了

$$\operatorname{var}\left(\sum_{l\in\mathcal{L}} \theta_l^{q-c}\tilde{U}_{n,\boldsymbol{\alpha},l}\right) \lesssim n^{-cs} \sum_{l_1,l_2\in\mathcal{L}} \theta_{l_1}^{q-c}\theta_{l_2}^{q-c}\tilde{\Sigma}_s^c(l_1,l_2),$$

根据假设 $o\big(n^{-qs}\tilde{\Sigma}_s(q)\big)$，这就是 4。因此，我们在局部备择假设下得到了所需的渐近结果。

B.5. 命题 **3.9** 的证明。定义 $f(q) = [(qs)!]^{1/2q}(\sqrt{N}R/d)^{1/q}$，它去除了与 q 在 $\delta(q)$ 中无关的所有项。

假设 $d \geq \sqrt{N}R$。那么 $\sqrt{N}R/d$ 是 q 中的递增函数。由于 $[(qs)!]^{1/2q}$ 也是关于 q 的递增函数，因此对于所有 $q \in 2\mathbb{Z}_+$， $f(q)$ 在 $q = 2$ 处取得最小值。

当 $d < \sqrt{N}R$ 时，我们沿用 He 等人 (2021 年) 中命题 2.3 的证明思路。定义 $D = \sqrt{N}R/d > 1$ 以简化记号。注意，我们有

$$\frac{f(q+1)}{f(q)} = \frac{\{[(q+1)s]!\}^{1/(2q+1)}D^{1/(q+1)}}{[(qs)!]^{1/2q}D^{1/q}} = \left(\frac{\left(P_s^{(q+1)s}\right)^q}{(qs)!D^2}\right)^{\frac{1}{2q(q+1)}}.$$

因此，当且仅当 $f(q+1) > f(q)$ 时， $g(q) \triangleq \left(P_s^{(q+1)s}\right)^q/(qs)! > D^2$。不难看出， $g(q)$ 关于 q 递增，因为

$$\frac{g(q+1)}{g(q)} = \left(\frac{P_s^{(q+2)s}}{P_s^{(q+1)s}}\right)^{q+1} > 1.$$

这意味着，如果 $f(q+1) \geq f(q)$，那么必有 $f(q+2) > f(q+1)$，而最优的 q 是满足 $g(q_0) \geq D^2$ 的最小整数 q_0。注意到，对于任意 $q < q_0$，$g(q) < D^2$ 意味着 $f(q)$ 关于 $q < q_0$递减，且 $f(q)$ 关于 $q \geq q_0$严格递增。由于我们已证明 $g(q)$ 严格递增，因此得出 q_0 关于 D递增。