arXiv:2602.04250v1 [math.PR] 4 Feb 2026

关于三角阵列的物理依赖与混合条件的札记

Florian Heinrichs

亚琛应用科学大学医药工程与技术数学系，海因里希-穆斯曼街*1*号，于利希，*52428*，德国

摘要

在边缘密度和条件密度满足温和的结构假设与正则性条件时，给出了以物理依赖度量表示的 β-混合系数显式上界。因此，弱物理依赖意味着三角阵列满足 β-混合和强混合；这一结果补充了Hill（2025）的研究，后者在矩假设下证明了逆向蕴含。

关键词： 弱依赖，强混合， β-混合，物理依赖，三角阵列，局部平稳性*2020*年数学主题分类： 62M10, 60G10, 60F05

1. 引言

当超越独立性假设时，控制依赖结构至关重要，这样才能推导出有意义的结果，例如中心极限定理。人们提出了多种依赖度量，包括混合条件 Rosenblatt（1956）；Ibragimov（1959）；Blum 等人（1963）；Philipp（1969）；Kolmogorov 和 Rozanov（1960）；Volkonskii 和 Rozanov（1959）、累积量条件 Brillinger（2001），以及物理依赖度量 Wu （ 2005 ）。这些依赖度量已从（平稳）时间序列推广到三角阵列。许多性质，例如协方差界，对三角阵列仍然成立。

电子邮箱地址： f.heinrichs@fh-aachen.de(Florian Heinrichs)

不同依赖度量之间的关系已得到广泛研究，尤其是混合条件之间的联系，可用以下图示概括 Doukhan（1995）；Bradley（2007）；Dedecker 等人（2007）。

$$\psi\text{-mixing} \overset{\Rightarrow}{\not\Leftarrow} \varphi\text{-mixing} \overset{\Rightarrow}{\not\Leftarrow} \begin{cases} \beta\text{-mixing} & \overset{\Rightarrow}{\not\Leftarrow} & \alpha\text{-mixing} \\ \not\Updownarrow & & \\ \rho\text{-mixing} & \overset{\Rightarrow}{\not\Leftarrow} & \alpha\text{-mixing} \end{cases}$$

众所周知，在满足某些矩界的条件下，α-混合蕴含累积量条件（参见 Bücher 等人（2020）中的引理 2）。另一方面，关于混合条件与物理依赖度量之间联系的研究较少。这可能是因为混合条件纯粹是分布性质，只取决于时间序列的分布。相比之下，弱物理依赖要求时间序列具有特定的表示，因此是一种结构性质。由于这种差异，难以直接比较这两种弱依赖概念。此外，已有弱依赖时间序列的例子表明，它们可能是 α-混合的，却不具备弱物理依赖性，反之亦然。这些例子为理解在何种条件下一个依赖度量可由另一个依赖度量控制提供了关键见解。

Hill（2025）最近证明，强混合条件可推出弱物理依赖。下文将给出反方向成立的条件。

2. 主要结果

首先，我们回顾 Zhou 和 Wu（2009）提出的局部平稳性框架。设 $\varepsilon = (\varepsilon_i)_{i\in\mathbb{Z}}$ 是一列独立同分布随机变量，且 $\varepsilon^* = (\varepsilon_i^*)_{i\in\mathbb{Z}}$ 是 ε 的一个独立副本。定义滤过 $\mathcal{F}_i = (\varepsilon_k)_{k\le i}$ 和 $\mathcal{F}_i^* = (\ldots, \varepsilon_{-2}, \varepsilon_{-1}, \varepsilon_0^*, \varepsilon_1, \ldots, \varepsilon_i)$。

若存在一个可能为非线性的滤波器 $G\colon [0,1]\times\mathbb{R}^{\mathbb{N}} \to \mathbb{R}$，它关于第一个自变量连续，并且满足 $X_{i,n} = G(i/n, \mathcal{F}_i)$，对所有 $i \in \{1, \ldots, n\}$ 和 $n \in \mathbb{N}$ 均成立，且 $(t,\omega) \mapsto G(t, \mathcal{F}_i(\omega))$ 可测，则三角阵 $\{(X_{i,n})_{i=1,\ldots,n}\}_{n\in\mathbb{N}}$ 是局部平稳。

对于 $p \ge 1$，滤波器 G 具有的阶数为 p 的物理依赖度量

$\sup_{t\in[0,1]}\mathbb{E}[|G(t,\ \mathcal{F}_i)|^p]<\infty$ 定义为

$$\delta_p(G,i)=\sup_{t\in[0,1]}\mathbb{E}\Big[\Big|G(t,\mathcal{F}_i)-G(t,\mathcal{F}_i^*)\Big|^p\Big]^{1/p},$$

对于 $i\in\mathbb{Z}$。若某个阶数 $p\geq 1$ 下，随着 $i\to\infty$， $\delta_p(G,i)$ 趋于零，则称三角阵 X 是弱物理依赖的。

在一个概率空间上，给定两个 σ-代数$\mathcal{F}$ 和$\mathcal{G}$ ，其强混合（或 α-混合）系数定义为

$$\alpha(\mathcal{F},\mathcal{G})=\sup\Big\{|\mathbb{P}(A\cap B)-\mathbb{P}(A)\mathbb{P}(B)|:A\in\mathcal{F},B\in\mathcal{G}\Big\}.$$

若强混合系数$_{n\in\mathbb{N}}$ 趋于零，则称三角阵 $X=\{(X_{i,n})_{i=1,\ldots,n}\}$是强混合的（或 α-混合），其中

$$\alpha(k)=\sup_{n\in\mathbb{N}}\sup_{j=1,\ldots,n-k}\alpha\Big(\sigma\{(X_{i,n})_{i=1}^{j}\},\sigma\{(X_{i,n})_{i=j+k}^{n}\}\Big)$$

当 k 趋于无穷时趋于零。 β-混合系数定义为

$$\beta(\mathcal{F},\mathcal{G})=d_{\mathrm{TV}}(\mathbb{P}_{\mathcal{F},\mathcal{G}},\mathbb{P}_{\mathcal{F}}\otimes\mathbb{P}_{\mathcal{G}}).$$

其中， d_{TV} 表示概率测度之间的全变差距离， $\mathbb{P}_{\mathcal{F}}$ 和 $\mathbb{P}_{\mathcal{G}}$ 分别表示 $\mathbb{P}$ 在相应 σ-代数上的限制，而$\mathbb{P}_{\mathcal{F},\mathcal{G}}$是定义在乘积 σ-代数上的概率律，其在矩形集上的定义为$\mathbb{P}_{\mathcal{F},\mathcal{G}}(A,B)\ =\mathbb{P}(A\cap B)$。注意，我们也可以等价地写成

$$\beta(\mathcal{F},\mathcal{G})=\mathbb{E}[d_{\mathrm{TV}}(\mathbb{P}_{\mathcal{F}|\mathcal{G}},\mathbb{P}_{\mathcal{F}})],\tag{1}$$

其中， $\mathbb{P}_{\mathcal{F}|\mathcal{G}}$ 表示 $\mathbb{P}$ 在$\mathcal{F}$ 上的限制，并以$\mathcal{G}$ 为条件（例如，参见Rio（2017）或Dedecker和Prieur（2005）对定理1.4的证明）。与前面一样，若 β-混合系数$_{n\in\mathbb{N}}$ 趋于零，则称三角阵 $X=\{(X_{i,n})_{i=1,\ldots,n}\}$是 β-混合的

$$\beta(k)=\sup_{n\in\mathbb{N}}\sup_{j=1,\ldots,n-k}\beta\Big(\sigma\{(X_{i,n})_{i=1}^{j}\},\sigma\{(X_{i,n})_{i=j+k}^{n}\}\Big)$$

当 k 趋于无穷时消失。众所周知， $\alpha(k)\leq\frac{1}{2}\beta(k)$，因此 β-混合意味着强混合（例如，参见 Doukhan（1995）第1.1节中的命题 1）。

如前所述，弱物理依赖与强混合这两个概念有着根本区别，因为前者涉及时间序列的结构，而后者仅涉及其分布。

强混合时间序列的一个简单例子是独立同分布随机变量序列，它们不存在矩，因此不具有弱物理依赖性，例如，当 X_i 服从帕累托分布时。在这种情况下，物理依赖度量没有定义。一个不那么简单的例子是随机环境中的随机游走。

例 1. 设$(\varepsilon_i)_{i\in\mathbb{Z}}$ 和$(\eta_i)_{i\in\mathbb{Z}}$ 是相互独立的拉*ademacher*随机变量独立同分布序列，因此 $\mathbb{P}(\varepsilon_i \pm 1) = \mathbb{P}(\eta_i \pm 1) = \frac{1}{2}$，其中 $i \in \mathbb{Z}$。我们首先定义随机位置

$$S_0 = 0, \quad , S_n = \sum_{i=1}^{n} \varepsilon_i, \quad S_{-n} = \sum_{i=0}^{n-1} \varepsilon_{-i},$$

其中 $n \in \mathbb{N}$。由 $X_i = \eta_{S_i}$ 定义的时间序列 $X = (X_i)_{i\in\mathbb{Z}}$，根据 *den Hollander and Steif*（*2006*）的定理*2.5 (ii)*，是 β-混合的，因此具有强混合性。此外，根据 *den Hollander and Steif*（*2006*）的定理*2.2*， X 不存在因果表示，这正如物理依赖度量的定义所要求的。

前面的例子表明，对于具有有限矩的任意（弱依赖）时间序列，物理依赖度量并不一定有定义。对于具有所需表示的（局部平稳）时间序列，只要其矩在阶数 $r > p$ 上一致有界，根据 Hill（2025）的注2.4，强混合即可推出 p 阶的弱物理依赖。

例2. 设$(\varepsilon_i)_{i\in\mathbb{Z}}$ 为一列相互独立的伯努利分布随机变量，成功概率为 $p \in (0,1)$，且 $\mathbb{P}(\varepsilon_i = 1) = 1 - \mathbb{P}(\varepsilon_i = 0) = p$，对 $i \in \mathbb{Z}$，并且 $\rho \in (0, \frac{1}{2}]$。那么，由$(X_i)_{i\in\mathbb{Z}}$定义的时间序列为

$$X_i = \sum_{k=0}^{\infty} \rho^k \varepsilon_{i-k},$$

由 *Andrews*（*1984*）可知其并非强混合。相反，对于滤波器 $G(x) = \sum_{k=0}^{\infty} \rho^k x_{i-k}$ 与滤过 $F_i = (\varepsilon_k)_{k\le i}$ 的物理相依测度，成立

$$\delta_2(G,h) = \left(\mathbb{E}[|G(\mathcal{F}_h) - \mathcal{G}(\mathcal{F}_h^*)|^2]\right)^{1/2} = \left(\mathbb{E}[|\rho^h(\varepsilon_0 - \varepsilon_0^*)|^2]\right)^{1/2} \le C\rho^h,$$

当 $h \to \infty$ 时趋于零。

这个例子表明，即使在线性过程这一简单情形下，我们也不应期待弱物理依赖能推出强混合。对于线性过程，强混合的一个充分条件是创新项的边缘密度有有界变差（例如，参见 Doukhan（1995）第2.3节中的推论1）。为了从物理依赖推出β-混合，我们需要类似的正则性条件，如下所述。

假设1。 设 $X = (X_{i,n})_{i=1,\dots,n}\}_{n\in\mathbb{N}}$ 是一个中心化局部平稳时间序列，其表示为 $X_{i,n} = G(i/n, \mathcal{F}_i)$。 X 是*1*阶弱物理依赖的，即 $\Theta_1 < \infty$，其中 $\Theta_k = \sum_{h=k}^{\infty} \delta_1(G, h)$ 对 $k \in \mathbb{N}$成立。

假设 2。 设 $(\Omega, \mathcal{A}\ \mathbb{P})$ 表示一个概率空间， $X = (X_{i,n})_{i=1,\dots,n}\}_{n\in\mathbb{N}}$是一个中心化局部平稳时间序列，并且$(w_m)_{m\in\mathbb{N}}$ 是满足 $\sum_{m=1}^{\infty} w_m \le 1$ 的正权重。

1. 关于勒贝格测度的$p_{j+k,\ n}$边际密度$X_{j+k,\ n}$，…，$X_{n,\ n})$ 具有关于 $L^1(\mathbb{R}^{n-j-k+1})$的偏导数，使得

$$\sum_{m=1}^{n-j-k+1} w_m \|\partial_m p_{j+k,n}\|_1 \le D_1,$$

对于某个常数 $D_1 \ge 0$ 以及任意 $j \in \{1, \dots, n\}$ $k \in \{1, \dots, n-j\}$和 $n \in \mathbb{N}$。

2. 在给定$(X_{1,\ n}, \dots, X_{j,\ n})$ 的条件下，向量$(X_{j+k,\ n}, \dots, X_{n,\ n})$的条件概率测度关于勒贝格测度具有密度$q_{j,\ k,\ n}$，使得偏导数属于 $L^1(\mathbb{R}^{n-j-k+1})$ ，并且

$$\sum_{m=1}^{n-j-k+1} w_m \|\partial_m q_{j,k,n}\|_1 \le D_2$$

$\mathbb{P}$-几乎处处，对于某个常数 $D_2 \ge 0$ 以及任意 $j \in \{1, \dots, n\}$ $k \in \{1, \dots, n-j\}$和 $n \in \mathbb{N}$。

定理1. 设 X 是满足假设*1*和假设*2*的三角阵。存在一个绝对常数 $C \ge 0$，使得对于每个 $k \in \mathbb{N}$

$$\beta(k) \le C\sqrt{D\Theta_k},$$

其中 $D =$ 最大值$\{D_1, D_2\}$。

该定理的一个直接推论是，弱物理依赖可以推出强混合。这使得基于强混合条件的统计方法（例如序列相关或离群值检验）（Bücher 等人，2023年；Heinrichs等人，2025年）能够应用于具有弱物理依赖的时间序列。

注1. 假设1规定，X 必须具备所需的结构，物理依赖度量才能存在。这一要求相当宽松，因为大多数实际感兴趣的时间序列都具有这种结构。此外，我们在定理1的证明中并未使用局部平稳性，因此，只要物理依赖度量存在且如假设中所述有界，该结果对任意三角阵列都成立。

假设 2 规定了边缘密度所需的正则性。正如例2所示，对于一般的创新项 ε，该定理并不成立。

3. 辅助结果

下文中，令 $\|f\|_1 = \int |f(x)| \, \mathrm{d}x$ 表示从 $\mathbb{R}^m \to \mathbb{R}$ 和 $m \in \mathbb{N}$上的函数所构成的 L^1-范数。此外，令 $W_{1,d}$ 表示相对于某个度量 d的 Wasserstein 1距离。

在证明定理1之前，我们首先需要一些辅助结果，以便用相应概率分布之间的瓦瑟斯坦距离来界定密度函数的 L^1-距离。

令 $\varphi \in C_c^\infty(\mathbb{R})$ 为 $\mathbb{R}$ 上具有紧支撑的对称光滑函数，$\int \varphi(x)\mathrm{d}x = 1$，$C_{\varphi,1} = \int |\varphi'(x)| \, \mathrm{d}x < \infty$，$C_{\varphi,2} = \int |x|\varphi(x) \, \mathrm{d}x < \infty$ 且 $\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\varphi(x/\varepsilon) = \delta(x)$，其中 δ 表示狄拉克δ函数。

对于 $\varepsilon > 0$，定义 $\varepsilon_m = \varepsilon w_m$，并将 $\mathbb{R}^\nu$ 上的乘积磨光函数定义为

$$\Phi_\varepsilon^{(w)}(x) = \prod_{m=1}^{\nu} \frac{1}{\varepsilon_m} \varphi\Big(\frac{x_m}{\varepsilon_m}\Big) = \prod_{m=1}^{\nu} \frac{1}{\varepsilon w_m} \varphi\Big(\frac{x_m}{\varepsilon w_m}\Big).$$

对于任意可积密度 f，定义磨光 $f_\varepsilon = f * \Phi_\varepsilon^{(w)}$，其中 $*$ 表示卷积。乘积结构使得可以逐坐标进行估计。$\lim_{\varepsilon \to 0} f_\varepsilon = f$ 这一事实是泛函分析中的一个标准结果，参见 Evans（1998年）第 5.3 章。不过，我们需要明确控制 f_ε 与 f 之间的距离，这由以下命题保证。

命题 1 (平滑误差**)**。 对于 $\mathbb{R}^\nu$ 上任意具有偏导数 $\partial_m f \in L^1$ 的密度 f，

$$\|f - f_\varepsilon\|_1 \leq C_{\varphi,2}\varepsilon \sum_{m=1}^{\nu} w_m \|\partial_m f\|_1.$$

证明。对每个 $x \in \mathbb{R}^\nu$，都有

$$f(x) - f_\varepsilon(x) = \int_{\mathbb{R}^\nu} \Big(f(x) - f(x-y)\Big)\Phi_\varepsilon^{(w)}(y)\,\mathrm{d}y.$$

根据梯度定理，也就是将微积分基本定理应用于从 $x-y$ 到 x 的线段，可得

$$f(x) - f(x-y) = \int_0^1 \nabla f(x-ty)\cdot y\,\mathrm{d}t = \sum_{m=1}^{\nu} \int_0^1 \partial_m f(x-ty) y_m\,\mathrm{d}t,$$

现在，特别地，根据富比尼定理以及勒贝格测度的平移不变性，

$$\begin{aligned}\|f - f_\varepsilon\|_1 &\leq \int_{\mathbb{R}^\nu}\int_0^1 \sum_{m=1}^{\nu} \|\partial f(\cdot - ty)\|_1 |y_m|\,\mathrm{d}t\Phi_\varepsilon^{(w)}(y)\,\mathrm{d}y \\ &= \sum_{m=1}^{\nu} \|\partial_m f\|_1 \int_0^1 \int_{\mathbb{R}^\nu} |y_m|\Phi_\varepsilon^{(w)}(y)\,\mathrm{d}y\,\mathrm{d}t.\end{aligned}$$

现在，根据 $\Phi_\varepsilon^{(w)}$的定义，我们有

$$\int_{\mathbb{R}^\nu} |y_m|\Phi_\varepsilon^{(w)}(y)\,\mathrm{d}y = \int_{\mathbb{R}} |x| \frac{1}{\varepsilon w_m}\varphi\Big(\frac{x}{\varepsilon w_m}\Big)\,\mathrm{d}x = \varepsilon w_m \int_{\mathbb{R}} |z|\varphi(z)\,\mathrm{d}z,$$

这里用到了如下事实：对每个 $\ell \neq m$，关于每个坐标 y_ℓ 的积分都等于 1。特别地，

$$\|f - f_\varepsilon\|_1 \leq \varepsilon\left(\int_{\mathbb{R}} |z|\varphi(z)\,\mathrm{d}z\right)\sum_{m=1}^{\nu} w_m\|\partial_m f\|_1 \leq C_{\varphi,2}\varepsilon\sum_{m=1}^{\nu} w_m\|\partial_m f\|_1.$$

□

接下来，我们根据底层测度之间的$W_{1,\,d}$-距离，估计两个磨光密度的 L^1-差。

命题 2 （$W_{1、d}$下的平滑差）。设 p 和 q 是定义在 $\mathbb{R}^\nu$ 上的两个概率密度，对应的概率分布分别为 P 和 Q。那么，对每个 $\varepsilon > 0$,

$$\|p_\varepsilon - q_\varepsilon\|_1 \le \frac{C_{\varphi,1}}{\varepsilon}\, W_{1,d}(P,Q),$$

其中 d定义为

$$d(y,z) = \sum_{m=1}^{\nu} w_m |y_m - z_m|$$

权重如假设 *2* 所述。

证明。 令 π 为 P 和 Q的任意耦合。于是,

$$\begin{aligned}\|p_\varepsilon - q_\varepsilon\|_1 &= \int_{\mathbb{R}^\nu} \iint \Phi_\varepsilon^{(w)}(x-y) - \Phi_\varepsilon^{(w)}(x-z)\pi(dy,dz)dx \\ &\le \iint \int_{\mathbb{R}^\nu} \Big|\Phi_\varepsilon^{(w)}(x-y) - \Phi_\varepsilon^{(w)}(x-z)\Big| dx\pi(dy,dz),\end{aligned}$$

根据富比尼定理。与命题 **1** 的证明类似，由梯度定理以及勒贝格测度的平移不变性可得,

$$\int_{\mathbb{R}^\nu} \Big|\Phi_\varepsilon^{(w)}(x-y) - \Phi_\varepsilon^{(w)}(x-z)\Big| dx \le \sum_{m=1}^{\nu} \|\partial_m \Phi_\varepsilon^{(w)}\|_1 |y_m - z_m|.$$

根据 $\Phi_\varepsilon^{(w)}$的定义,

$$\begin{aligned}\|\partial_m \Phi_\varepsilon^{(w)}\|_1 &= \int_{\mathbb{R}^\nu} \frac{1}{(\varepsilon w_m)^2} \varphi'\Big(\frac{y_m}{\varepsilon w_m}\Big) \prod_{\ell \ne m} \frac{1}{\varepsilon w_\ell} \varphi\Big(\frac{y_\ell}{\varepsilon w_\ell}\Big)\, \mathrm{d}y \\ &= \int_{\mathbb{R}} \frac{1}{(\varepsilon w_m)^2} \varphi'\Big(\frac{x}{\varepsilon w_m}\Big)\, \mathrm{d}x = \frac{C_{\varphi,1}}{\varepsilon w_m},\end{aligned}$$

因为关于 y_ℓ的积分（对 $\ell \ne m$而言）等于 **1**。特别地,

$$\int_{\mathbb{R}^\nu} \Big|\Phi_\varepsilon^{(w)}(x-y) - \Phi_\varepsilon^{(w)}(x-z)\Big| dx \le \frac{C_{\varphi,1}}{\varepsilon} d(y,z).$$

对 π 积分并对所有耦合 π 取最小值，即可得到该结论。

□

结合命题 1 和命题 2，可得

$$\|p-q\|_1 \le 2C_{\varphi,2}\varepsilon D + \frac{C_{\varphi,1}}{\varepsilon} W_{1,d}(P,Q),$$

对于任意 $\varepsilon > 0$，其中我们使用假设 2 将加权导数和控制为 D。右端在 $\varepsilon = \sqrt{C_{\varphi,1} W_{1,d}(P,Q) / 2DC_{\varphi,2}}$ 时取得最小值，并得到界

$$\|p-q\|_1^2 \le CDW_{1,d}(P,Q), \tag{2}$$

对于某个常数 $C \ge 0$，该常数仅取决于 φ。

4. 定理1的证明

任取但固定的 $j \in \{1, \ldots, n\}$ $k \in \{1, \ldots, n-j\}$ 和$n \in \mathbb{N}$。令 $\nu = n - j - k + 1$，并在 $\mathbb{R}^\nu$ 上定义加权 ℓ^1 度量：

$$d(x,y) = \sum_{m=1}^{\nu} w_m |x_m - y_m|,$$

对于 x, $y \in \mathbb{R}^\nu$。此外，令

$$\mathcal{F}_i^{(k,\ell)} = (\ldots, \varepsilon_{i-\ell-k-2}, \varepsilon_{i-\ell-k-1}, \varepsilon^*_{i-\ell-k}, \ldots, \varepsilon^*_{i-\ell}, \varepsilon_{i-\ell+1}, \ldots, \varepsilon_i),$$

对于 ℓ, $k \in \mathbb{N}_0$和 $i \in \mathbb{Z}$，以及 $\mathcal{F}_i^{(-1,\ell)} = \mathcal{F}_i$。注意

$$\mathcal{F}_{j+i}^{(\infty,i)} = (\ldots, \varepsilon^*_{j-1}, \varepsilon^*_j, \varepsilon_{j+1}, \ldots, \varepsilon_{j+i}),$$

对任意 $i \in \mathbb{N}$。定义 $\tilde{X}_{i,n} = G(i/n, \mathcal{F}_{j+i}^{(\infty,i)})$，其中 $i \ge k$。令 $\mathbf{X} = (X_{j+k,n}, \ldots, X_{n,n})$~ ~ ~以及 $\mathbf{X} = (X_{j+k,n}, \ldots, X_{n,n})$。

利用望远镜求和论证，我们可以将其改写为

$$G(t, \mathcal{F}_{j+i}) - G(t, \mathcal{F}_{j+i}^{(\infty,i)}) = \sum_{\ell=0}^{\infty} G(t, \mathcal{F}_{j+i}^{(\ell-1,i)}) - G(t, \mathcal{F}_{j+i}^{(\ell,i)}),$$

因为$\mathcal{F}_{j+i} = \mathcal{F}_{j+i}^{(-1,i)}$。根据三角不等式，成立

$$\begin{aligned}\mathbb{E}[|G(t, \mathcal{F}_{j+i}) - G(t, \mathcal{F}_{j+i}^{(\infty,i)})|] &\le \sum_{\ell=0}^{\infty} \mathbb{E}[|G(t, \mathcal{F}_{j+i}^{(\ell-1,i)}) - G(t, \mathcal{F}_{j+i}^{(\ell,i)})|] \\ &\le \sum_{\ell=0}^{\infty} \delta_1(G, i+\ell) = \Theta_i,\end{aligned}$$

˜根据假设1。特别地，对任意 $i \geq k$， $\mathbb{E}[|X_{j+i,n} - X_{j+i,n}|] \leq \Theta_i$。根据度量 d的定义，可得

$$\mathbb{E}[d(\mathbf{X}, \tilde{\mathbf{X}})] = \sum_{m=1}^{\nu} w_m \mathbb{E}[|X_{j+m,n} - \tilde{X}_{j+m,n}|] \leq \Theta_k.$$

˜ ˜此外，（$\mathbf{X}$,$\mathbf{X}$）是一个耦合，其中 $\mathbf{X}$ 与$\mathbf{X}$ 具有相同的边缘˜分布，且$\mathbf{X}$ 独立于$\mathcal{G}_j$: = （$X_{1,n}, \ldots, X_{j,n}$）。令 $P_{j+k,n}$ 表示$\mathbf{X}$的分布，$Q_{j,k,n}$ 表示在给定$\mathcal{G}_j$时$\mathbf{X}$ 的条件分布。根据康托罗维奇-鲁宾斯坦对偶，对$Q_{j,k,n}$的每个实现，我们可以将瓦瑟斯坦距离表示为

$$W_{1,d}(P_{j+k,n}, Q_{j,k,n}) = \sup_{\|f\|_L \leq 1} \Big| \mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{X})|\mathcal{G}_j] \Big|, \quad \mathbb{P} - \text{a.e.}$$

其中上确界取遍所有具有˜利普希茨常数小于或等于1的利普希茨连续函数。由构造可知，$\mathbf{X}$独立˜ ˜ ˜于$\mathcal{G}_j$，且 $\mathbf{X}$的分布为$P_{j+k,n}$，从而 $\mathbb{E}[f$（$\mathbf{X}$）$|\mathcal{G}_j] = \mathbb{E}[f$（$\mathbf{X}$）$] = \mathbb{E}[f$（$\mathbf{X}$）$]$。特别地，

$$|\mathbb{E}[f(\mathbf{X})] - \mathbb{E}[f(\mathbf{X})|\mathcal{G}_j]| = |\mathbb{E}[f(\tilde{\mathbf{X}})|\mathcal{G}_j] - \mathbb{E}[f(\mathbf{X})|\mathcal{G}_j]| \leq \mathbb{E}\Big[|f(\tilde{\mathbf{X}}) - f(\mathbf{X})|\Big|\mathcal{G}_j\Big].$$

对所有利普希茨˜常数以1为界的利普希茨连续函数取上确界，得到$W_{1,d}$ （$P_{j+k,n}, Q_{j,k,n}$） $\leq \mathbb{E}[d$（$\mathbf{X}$,$\mathbf{X}$）$|\mathcal{G}_j]$。通过取期望，我们最终得到

$$\mathbb{E}[W_{1,d}(P_{j+k,n}, Q_{j,k,n})] \leq \mathbb{E}[d(\mathbf{X}, \tilde{\mathbf{X}})] \leq \Theta_k.$$

注意到$p_{j+k,\ n}$ 表示P_{j+k}的$_n$ 密度，且 $q_{j,\ k,\ n}$ 表示条件分布 $Q_{j,\ k,\ n}$的密度。根据谢费定理（例如参见 Tsybakov，2009 年的引理2.1），有 $2d_{\mathrm{TV}}(P_{j+k,\ n},\ Q_{j,\ k,\ n})$ $= \|p_{j+k,\ n} - q_{j,\ k,\ n}\|_1$ （$\mathbb{P}$-几乎处处成立）。特别地，根据詹森不等式和（2），

$$\begin{aligned} 2\mathbb{E}[d_{\mathrm{TV}}(P_{j+k,n}, Q_{j,k,n})] &\leq \mathbb{E}[\|p_{j+k,n} - q_{j,k,n}\|_1^2]^{1/2} \\ &\leq \sqrt{CD\mathbb{E}[W_{1,d}(P_{j+k,n}, Q_{j,k,n})]} \leq \sqrt{CD\Theta_k}. \end{aligned}$$

最后，由 (1) 可得

$$\beta(k) = \sup_{n \in \mathbb{N}} \sup_{j=1,\ldots,n-k} \mathbb{E}[d_{\mathrm{TV}}(P_{j+k,n}, Q_{j,k,n})] \leq \frac{\sqrt{C}}{2} \sqrt{D\Theta_k}.$$

□

参考文献

Andrews, D.W., 1984年。非强混合自回归过程。《应用概率杂志》21，930–934。Blum, J.R., Hanson, D.L., Koopmans, L.H., 1963年。关于一类随机过程的强大数定律。第2卷。Bradley, R.C., 2007年。《强混合条件导论》。第1卷。Kendrick出版社。Brillinger, D.R., 2001年。《时间序列》。工业与应用数学学会。网址：https://epubs.siam.org/doi/abs/10.1137/1.9780898719246，doi:10.1137/1.9780898719246，arXiv:https://epubs.siam.org/doi/pdf/10.1137/1.9780898719246。Bücher, A., Dette, H., Heinrichs, F., 2020年。检测局部平稳函数型时间序列中偏离二阶平稳性的情况。《统计数学研究所年刊》72，1055–1094。Bücher, A., Dette, H., Heinrichs, F., 2023年。一种用于检测局部平稳函数型时间序列中序列相关的港口曼托型检验。《随机过程统计推断》26，255–278。Dedecker, J., Doukhan, P., Lang, G., José Rafael, L.R., Louhichi, S., Prieur, C., 2007年。《弱依赖：示例与应用》。第1卷。纽约施普林格出版社。Dedecker, J., Prieur, C., 2005。新的依赖系数：示例及其在统计学中的应用。《概率论及相关领域》132，203–236。Doukhan, P., 1995。混合，载于：《混合：性质与示例》。施普林格，第15–23页。Evans, L.C., 1998年。偏微分方程。《数学研究生教材》第19卷。美国数学学会，普罗维登斯，罗得岛州。Heinrichs, F., Bastian, P., Dette, H., 2025年。非平稳时间序列中的序贯异常值检测。《时间序列分析杂志》。

Hill, J.B., 2025年。混合鞅与物理依赖的等价性及其应用。《统计与概率快报》221，110380。den Hollander, F., Steif, J.E., 2006。随机环境中的随机游走：近期若干成果综述。讲义-专著丛书，53–65。Ibragimov, I.A., 1959。关于严格平稳随机过程的一些极限定理，载于：苏联科学院院刊，第711–714页。Kolmogorov, A.N., Rozanov, Y.A., 1960。关于平稳高斯过程的强混合条件。《概率论及其应用》5，204–208。Philipp, W., 1969。混合随机变量序列的中心极限定问题。《概率论及相关领域杂志》12，155–171。Rio, E., 2017。弱依赖随机过程的渐近理论。第80卷。施普林格。Rosenblatt, M., 1956。中心极限定理与强混合条件。《美国国家科学院院刊》42，43–47。Tsybakov, A.B., 2009。《非参数估计导论》。施普林格纽约，纽约州纽约市。Volkonskii, V.A., Rozanov, Y.A., 1959。关于随机函数的一些极限定理。I。《概率论及其应用》4，178–197。Wu, W.B., 2005。非线性系统理论：重新审视依赖。《美国国家科学院院刊》102，14150–14154。Zhou, Z., Wu, W.B., 2009。非平稳时间序列的局部线性分位数估计。《统计年刊》，2696–2729。