arXiv:2305.04290v1 [math.ST] 7 May 2023

非平稳与平稳条件下经验自协方差和交叉协方差正态近似的 Wasserstein 距离界

Andreas Anastasiou[1] 和 Tobias Kley[2]

摘要

自协方差函数和互协方差函数自然地出现在许多时间序列分析方法中（例如，自回归或预测）。在一定假设下，自协方差和交叉协方差的经验版本渐近服从正态分布，其协方差结构取决于二阶和四阶谱。在较宽松的假设下，我们推导出估计量的有限样本分布与高斯极限之间 Wasserstein距离的一个上界。估计量二阶矩的近似误差以及一个m相关的近似，是获得该上界的关键要素。作为一个具体示例，我们讨论如何计算创新项服从不同分布的 1 阶因果自回归过程的上界。为评估我们的结果，我们将该上界与通过模拟得到的 Wasserstein距离进行比较。

2010年数学学科分类：主要分类 62E17；次要分类 62F12。**关键词：**自协方差，时间序列，Wasserstein距离，Stein方法。

1 引言

近年来，评估各种渐近结果的质量引起了广泛关注。衡量分布近似误差的一种方法，是考察感兴趣的量的极限分布与实际分布之间 Wasserstein距离的显式上界；推导此类上界无疑是一项技术上颇为繁琐的工作。

[1] 塞浦路斯大学数学与统计系，邮政信箱：20537，尼科西亚1678，塞浦路斯。[2] 哥廷根乔治-奥古斯特大学数学随机过程研究所，Goldschmidtstraße 7，37077 哥廷根，德国。

我们考虑经验自协方差和交叉协方差

1. 不假设平稳性，并且

2. 针对弱平稳时间序列的情况。

我们的目标是给出一个上界，使得对于广泛的时间序列模型，不仅可以计算其收敛速率，还可以计算明确的常数。

我们考虑有一个 d 维时间序列 $\boldsymbol{X}(1)\dots,\boldsymbol{X}(n)$ 可用的情况，即，$\boldsymbol{X}(t)$ 取值于 $\mathbb{R}^d$，$t=1,\dots,n$。$\boldsymbol{X}(t)$ 的分量记为 $X_a(t)$，$a=1,\dots,d$。我们关注经验互协方差和自协方差，其定义如下：

$$\hat{\gamma}^*_{ab}(k) := \frac{1}{n}\sum_{t=1}^{n-k}(X_a(t+k)-\bar{X}_a)(X_b(t)-\bar{X}_b), \quad k=0,\dots,n-1, \tag{1.1}$$

其中 $\bar{X}_j := \frac{1}{n}\sum_{t=1}^{n} X_j(t)$，$j=a,b$ 表示经验均值。对于 $k=-n+1,\dots,-1$，我们定义 $\hat{\gamma}^*_{ab}(k) := \hat{\gamma}^*_{ba}(-k)$。文献中也存在其他定义，在正则条件下它们渐近等价。例如，关于自协方差的一些常见变体，可参见 Anderson（1971）第 8 章；特别是 Anderson（1971）中的推论 8.4.1，该推论在特定正则条件下证明了这些变体收敛到相同的高斯极限。

对于总体均值已知的平稳数据，我们可以用其总体均值替代 (1.1) 中的经验均值，如下所示。

$$\frac{1}{n}\sum_{t=1}^{n-k}(X_a(t+k)-\mathbb{E}X_a(t+k))(X_b(t)-\mathbb{E}X_b(t)), \quad k=0,\dots,n-1.$$

这相当于假设 $\{\boldsymbol{X}(t)\}$ 是中心化的（即 $\mathbb{E}\boldsymbol{X}(t)=0$），并采用以下经验互协方差的定义：

$$\hat{\gamma}_{ab}(k) := \frac{1}{n}\sum_{t=1}^{n-k} X_a(t+k)X_b(t), \quad k=0,\dots,n-1, \tag{1.2}$$

以及 $\hat{\gamma}_{ab}(k) := \hat{\gamma}_{ba}(-k)$，$k=-n+1$、.....、$-1$。自协方差和交叉协方差在许多时间序列方法中都具有重要作用，例如自回归（Jirak，2012，2014）和预测（Brockwell 和 Davis，2006；Kley 等，2019）。

在一定条件下，可以证明 $\hat{\gamma}^*_{ab}(k)$ 和 $\hat{\gamma}_{ab}(k)$ 是.....的相合估计。

$$\gamma_{ab}(k) := \mathbb{E}[X_a(t+k)X_b(t)]. \tag{1.3}$$

估计量分布的渐近正态性同样成立。我们有：

$$\sqrt{n}\big(\hat{\gamma}^*_{ab}(k) - \gamma_{ab}(k)\big) \xrightarrow[n\to\infty]{} N, \quad N \sim \mathcal{N}(0, \Sigma_{ab}(k)); \tag{1.4}$$

例如，参见 Brillinger (1975) 中的习题 7.10.36。渐近方差

$$\Sigma_{ab}(k) := \lim_{n\to\infty} \operatorname{var}\left(n^{-1/2} \sum_{t=1}^{n-k} X_a(t+k) X_b(t) \right) \tag{1.5}$$

取决于底层数据的二阶和四阶矩结构；参见 Brillinger (1975) 中的公式 (7.6.11)。通常很容易计算 $\Sigma_{ab}(k)$。我们在第 3 节中讨论的 AR(1) 时间序列情形的细节见在线补充材料的 D 节。为证明 (1.4)，通常会对依赖结构的强度以及所涉及随机变量的矩作出限制性假设（例如累积量可和性）；关于正态性失效的情形，参见 Horváth and Kokoszka (2008)。

下面给出全文采用的一般框架和符号。对于取值于 $\mathbb{R}^d$ 的随机向量 $\boldsymbol{U}$ 和 $\boldsymbol{V}$，我们使用如下定义的 1-瓦瑟斯坦度量：

$$d_{\mathrm{W}}(\mathcal{L}(\boldsymbol{U}), \mathcal{L}(\boldsymbol{V})) := \sup_{h\in\mathcal{H}} |\mathbb{E}[h(\boldsymbol{U})] - \mathbb{E}[h(\boldsymbol{V})]|, \quad \mathcal{H} = \{h : \mathbb{R}^d \to \mathbb{R} \,|\, \|h\|_{\mathrm{Lip}} \le 1\}, \tag{1.6}$$

其中，$\mathcal{L}(\boldsymbol{U})$ 表示 $\boldsymbol{U}$ 的分布律。此外，对于任意向量 $\boldsymbol{x} = (x_1, \ldots, x_d)$，我们用 $|\boldsymbol{x}|$ 表示其欧几里得范数：$= (\sum_{i=1}^d x_i^2)^{1/2}$，并且 $\|h\|_{\mathrm{Lip}} = \sup_{\boldsymbol{u}\neq\boldsymbol{v}} |h(\boldsymbol{u}) - h(\boldsymbol{v})|/|\boldsymbol{u} - \boldsymbol{v}|$。本文将 (1.6) 中的距离简称为 Wasserstein距离。本文的主要目的是通过给出 Wasserstein距离的上界，评估 (1.4) 中的分布近似质量；该距离衡量 (1.4) 左侧感兴趣的量的实际分布与其极限正态分布之间的差异。对于中心化数据，定理 2.1 和定理 2.3 分别针对非平稳序列和弱平稳序列实现了这一目标。结合定理 2.3 和引理 C.1，对平稳情形下 $\sqrt{n}\big(\hat{\gamma}^*_{ab}(k) - \gamma_{ab}(k)\big)$ 与 $\sqrt{n}\big(\hat{\gamma}_{ab}(k) - \gamma_{ab}(k)\big)$ 之间的 Wasserstein距离进行上界估计，可得出数据未中心化时的一个上界；详情见第 2.5 节和 C 节。

我们的方法依赖于一个 m 依赖序列的存在，这使我们能够在局部依赖结构下使用 Stein方法。Stein方法是一种强大的概率技术，最早由 Stein (1972) 提出。在随机变量之间存在此类依赖结构时，Stein方法尤其适用于评估给定随机变量的分布是否接近目标分布。通过 Stein方法得到的上界

在常数和样本量方面均有明确表达；例如，Anastasiou (2017) 在随机变量之间存在局部依赖结构的情况下，给出了最大似然估计量的正态近似上界。

近年来，人们对评估与求和 $\sum_{i=1}^{n} X_i$ 相关的正态近似质量颇感兴趣，其中 X_1、……、X_n 经过中心化，且遵循特定的依赖结构。乍看之下，经验自协方差和交叉协方差似乎符合这一框架（将 n 替换为 $n-k$，将 X_i 替换为 $n^{-1}X_a(t+k)X_b(t)$），但文献中关于 $\sum_{i=1}^{n} X_i$ 的结果并不能直接提供我们所需的结果；也就是评估式 (1.4) 中近似质量的显式有限样本上界。除其他原因外，这是因为经验自协方差和交叉协方差存在偏差。我们认为，经验自协方差和交叉协方差对实际应用至关重要，因此应当能够获得评估其有限样本分布近似的结果，并且这些结果应能根据底层过程/模型参数、区段大小 n 和滞后阶数 k 完全显式地表示。

我们现在继续讨论与评估相关的研究，即如何衡量相关数据求和的正态近似质量。在显式上界的框架下，但不再采用我们所使用的 m依赖结构，Röllin（2018）给出了 $\sum_{i=1}^{n} X_i$ 的分布与标准正态分布之间的 Wasserstein距离上界，其中 $X_1, \ldots, X_n$ 是一个离散时间鞅差序列。该上界的阶为$\mathcal{O}(n^{-1/2}\log n)$，其推导上界所采用的策略结合了Stein方法和Lindeberg论证。在与Polyak-Ruppert平均随机梯度下降相关的研究中，Anastasiou等人（2019）推导出了一个显式上界，用于衡量多元鞅差序列的求和分布与多元正态分布之间的分布距离。在最近的研究中，Fan和Ma（2020）通过放宽Röllin（2018）所采用的条件，扩展了其研究结果。除了离散时间鞅这一研究框架外，还有一些研究评估了满足特定混合条件的随机变量求和的正态近似；关于强混合条件和 φ-混合条件的情形，分别参见Sunklodas（2007）和Sunklodas（2011）。Dedecker和Rio（2008）给出了 $\sum_{i=1}^{n} X_i$ 的分布与正态分布之间的Wasserstein距离上界，其适用情形包括满足强混合假设，或者$(X_i)_{i\in\mathbb{Z}}$ 是一个遍历鞅差序列，或是一个满足特定投影判据的遍历平稳序列。

不再考虑上界中给出明确常数的情形，Dedecker 等人（2009）针对 $X_1, \ldots, X_n$ 为鞅差序列的情况，给出了 $\sum_{i=1}^{n} c_{n,i}X_n$ 形式的线性统计量之间最小距离的收敛速率，其中 $c_{n,i} \in \mathbb{R}$，且

它们的极限高斯分布。Fan（2019）在条件矩假设下，给出了鞅差序列中心极限定理的收敛速率。对于具有有限 $p \in (2,3]$ 矩的平稳序列 $X_1, \ldots, X_n$，Jirak（2016）在弱依赖条件下证明了一个 Berry-Esseen 定理，并给出了L^q-范数下的收敛速率，其中 $q \geq 1$。不过，所得上界并非显式上界，因为它们依赖于一个未明确给出的、会变化的绝对常数。

除了所采用的方法、遵循的证明思路以及对经验自协方差函数和交叉协方差函数这两类特定统计量的关注之外，本文所提出的结果还在三个主要方面具有创新性。首先，我们的结果适用于非平稳数据序列。其次，我们不仅关注收敛速率，而且本文推导出的 Wasserstein 距离界关于样本量 n、滞后阶数 k以及与底层数据相关的常数都是完全显式的，这使得该上界在示例中完全可计算。第三，我们所采用的假设不具有限制性，并且部分基于对原始时间序列的 m-依赖近似，这种近似在实际应用中便于处理，使得我们的结果能够适用于广泛的情形。在依赖范围有限的情况下，例如观测值相互独立或固定阶数的移动平均过程，我们的上界的阶为 $\mathcal{O}(n^{-1/2})$。在更一般的情况下，当序列依赖性在较大的滞后阶数下迅速消失且存在八阶矩时，我们的上界的阶为 $\mathcal{O}(n^{-1/2}\log n)$。关于该上界阶数的讨论可参见注2.5，并可在第2.6节中找到更详细的说明。

本文结构安排如下。第 2.1 节给出一般情形下的主要结果，即经验自协方差和交叉协方差函数 $\hat{}\gamma_{ab}(k)$（定义见式 (1.2)）的分布与其极限正态分布之间的 Wasserstein距离上界。需要指出的是，数据不一定来自平稳过程。第 2.2 节陈述并讨论弱平稳情形下的关键假设。第 2.3 节给出平稳性条件下的主要结果。第 2.4 节和第 2.5 节分别致力于计算基于 m 相关近似过程或原始中心化过程的矩所得到的上界；关于基于原始未中心化过程的矩计算上界的细节，将留待 C 节介绍。第2.6节详细解释了上界随样本量 n 变化的阶。在第 3 节，我们将一般结果应用于一个特定情形，即一阶因果自回归过程。第4节证明主要结果。第 5 节通过简要讨论结果为本文作结。关于计算第 2.4 节上界的技术细节、未纳入正文的逐步证明、计算方面的技术细节

以及模拟，还有第 3 节示例的附加表格，均收录于在线补充材料中。凡是以拉丁字母编号的章节、结果等，均可在补充材料中找到。

2 主要结果

2.1 一般情形的显式上界

在本节中，我们给出一个一般性结果，该结果不要求数据来自平稳过程。要应用该结果，需要存在一个长度相同的 m-相关序列 n 和维度为 d 且具有有限六阶矩的序列。数据与 m-相关近似之间的接近程度将决定上界的大小。

为便于表述，先介绍一些符号。对于任意向量 $\boldsymbol{x} = (x_1, \ldots, x_d)$，我们用 $|\boldsymbol{x}|$ 表示其欧几里得范数：$= (\sum_{i=1}^{d} x_i^2)^{1/2}$；对于随机向量 $\boldsymbol{X}$，其 L^q-范数记为 $\|\boldsymbol{X}\|_q$：$= (\mathbb{E}[|\boldsymbol{X}|^q])^{1/q}$，$q \geq 1$。我们记 $\mathbb{N}$：$= \{1, 2, \ldots\}$ 和 $\mathbb{N}_0$：$= \mathbb{N} \cup \{0\}$。回顾一下，随机向量 $(\zeta_1, \ldots, \zeta_r)$ 的 r^{th} 阶联合累积量定义为

$$\operatorname{cum}(\zeta_1, \ldots, \zeta_r) := \sum_{\nu} (-1)^{p-1} (p-1)! \Big(\mathbb{E} \prod_{j \in \nu_1} \zeta_j \Big) \cdots \Big(\mathbb{E} \prod_{j \in \nu_p} \zeta_j \Big), \tag{2.7}$$

其中求和遍及 $\{1, \ldots, r\}$ 的所有划分 $\nu := \{\nu_1, \ldots, \nu_p\}$；参见 Brillinger (1975)。对于不一定平稳的序列，其一般结果如下述定理 2.1 所示。证明推迟至第4节。

定理 *2.1.* 设 $\boldsymbol{X(1)}, \ldots, \boldsymbol{X(n)}$ 是一列 d 维随机向量，并假设 $\mathbb{E}\boldsymbol{(X(t))} = 0$ 对所有 $t \in \{1, \ldots, n\}$ 均成立。固定 a、$b \in \{1, \ldots, d\}$ 和 $k \in \{0, \ldots, n-1\}$，并令 $\hat{\gamma}_{ab}\boldsymbol{(k)}$ 如 ***(1.2)*** 中定义。固定 $m \in \mathbb{N}_0$，并令 $\boldsymbol{Y(1)}, \ldots, \boldsymbol{Y(n)}$ 是一列 m 相关的 d 维随机向量，使得 $\|\boldsymbol{Y(t)}\|_6 < \infty$ 对所有 $t = 1, \ldots, n$ 均成立，并假设

$$\tilde{\Sigma}_{ab}(k) := \operatorname{var}\left(n^{-1/2} \sum_{t=1}^{n-k} Y_a(t+k) Y_b(t) \right) > 0. \tag{2.8}$$

Let

$$\tilde{D}_{a,t} := \|X_a(t) - Y_a(t)\|_2 < \infty, \quad t \in \{1, \ldots, n\},$$

且若 $b \neq a$，则令 $\tilde{D}_{b,t}$ 以类似方式定义。此外，对于 $t = 1, \ldots, n-k$，令

$$\tilde{Z}(t) := Y_a(t+k) Y_b(t) - \mathbb{E}[Y_a(t+k) Y_b(t)]$$

并定义如下量

$$\tilde{Q}_t := \mathbb{E}\left|\left(\tilde{Z}(t)\tilde{Z}(A_t) - \mathbb{E}\left(\tilde{Z}(t)\tilde{Z}(A_t)\right)\right)\tilde{Z}(B_t)\right| + \frac{1}{2}\mathbb{E}\left|\tilde{Z}(t)\left(\tilde{Z}(A_t)\right)^2\right|, \tag{2.9}$$

其中 $\tilde{Z}(A) = \sum_{j\in A} Z(j)$ 对于 $A \subset \mathbb{N}$ 和

$$A_t := \{\ell = 1, \ldots, n-k : |\ell - t| \le m + k\},$$
$$B_t := \{\ell = 1, \ldots, n-k : |\ell - t| \le 2(m+k)\}.$$

最后，令

$$\tilde{K}_t := \tilde{D}_{a,t+k}\|X_b(t)\|_2 + \tilde{D}_{a,t+k}\tilde{D}_{b,t} + \|X_a(t+k)\|_2\tilde{D}_{b,t}.$$

那么，对于任意 $\gamma \in \mathbb{R}$ 和任意 $\sigma^2 > 0$,

$$d_{\mathrm{W}}\left(\mathcal{L}\left(n^{1/2}\left(\hat{\gamma}_{ab}(k) - \gamma\right)\right), \mathcal{N}(0, \sigma^2)\right) \le \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\tilde{K}_t + \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\left|\frac{n}{n-k}\gamma - \mathbb{E}[Y_a(t+k)Y_b(t)]\right|$$
$$+ \sqrt{\frac{2}{\pi\sigma^2}}\left|\sigma^2 - \tilde{\Sigma}_{ab}(k)\right| + \frac{2}{n^{3/2}\left(\tilde{\Sigma}_{ab}(k)\right)^{3/2}}\sum_{t=1}^{n-k}\tilde{Q}_t. \tag{2.10}$$

在定理及本文其余部分中，我们采用以下约定：使用波浪号符号区分与 m 相关的近似所对应的符号（例如 $\tilde{D}_{a,\,t}$，$\tilde{Z}(t)$ $\tilde{Q}_t$）。

由于假设条件宽松且常数表达明确，定理 2.1 例如可用于证明估计量序列的渐近正态性，其中底层模型或滞后阶数 k 取决于区段长度 n。该定理也可应用于具有时变系数的模型。鉴于这一结果可能适用的情形数不胜数，而本文篇幅有限，以下各节将重点讨论一个与实际应用最相关的情形：弱平稳数据。

2.2 平稳情形的关键假设

从本节开始，我们将 $\{\boldsymbol{X}(t) : t \in \mathbb{Z}\}$ 视为一个 d元、中心化且弱平稳的过程，记为 $\{\boldsymbol{X}(t)\}$, 并从中获得一个序列 $\boldsymbol{X}(1)\ldots, \boldsymbol{X}(n)$，其中 $\boldsymbol{X}(t)$ 取值于 $\mathbb{R}^d$，$t = 1, \ldots, n$。平稳情形下结果所用的主要假设如下。

假设 2.2。 对于给定的 $m \in \mathbb{N}_0$，存在一个 m-依赖的 d 维过程 $\{\boldsymbol{Y}(t)\}$，其中，对于 $a \in \{1, \ldots, d\}$，以及 $t \in \{1, \ldots, n\}$，

$$\tilde{D}_{a,t}^{(q)} := \|X_a(t) - Y_a(t)\|_q < \infty.$$

每当我们提及假设 *2.2* 时，均会指定 $q \geq 1$ 。

我们记

$$\tilde{D}_a^{(q)} := \sup_{1 \leq t \leq n} \tilde{D}_{a,t}^{(q)}. \tag{2.11}$$

我们强调，如果对依赖于 m 的序列存在选择，那么可以定义$\tilde{D}_a^{(q)} := \inf_{\boldsymbol{Y}} \sup_{1 \leq t \leq n} \|X_a(t) - Y_a(t)\|_q$，其中下确界是针对所有满足 $\|X_a(t) - Y_a(t)\|_q < \infty$ 的 $\{\boldsymbol{Y}(t)\}$ 的可能选择取得的；然而，为了得到一个易于计算的上界，我们针对一种特定选择陈述我们的结果。

尽管在本节中，为了能够对 $\gamma_{ab}(k)$ 进行有意义的定义，我们假设 $\{\boldsymbol{X}(t)\}$ 具有平稳性，但正如前文所述，我们的证明方法并不要求平稳性。我们详细陈述并讨论平稳情形下的结果，是因为它与实际应用密切相关，而且有助于理解更一般的结果。假设 2.2 意味着原始过程可以在 L^q 中通过一个 m-相依序列进行近似。对于定理 2.3，在 L^2 中进行近似就足够了；也就是说，我们要求满足 $q = 2$ 的假设 2.2。在第 2.5 节中，我们将解释如何根据原始过程 $\{\boldsymbol{X}(t)\}$ 的性质推导上界的一般步骤。引理 B.3 和 B.4 可用于推导此类上界；此时分别需要满足 $q = 4$ 和 $q = 6$ 的假设 2.2。我们不要求 $\{\boldsymbol{Y}(t)\}$ 具有平稳性或中心化性质，尽管在实际应用中通常如此。在第 3 节讨论的示例中，$\|X_a(t) - Y_a(t)\|_q$ 实际上独立于 t；有关该示例中如何计算式 (2.11) 所示的 $\tilde{D}_a^{(q)}$ 的详细信息，请参见 D 节。如果 $\{(\boldsymbol{X}(t)\ \boldsymbol{Y}(t))\}$ 在 q 阶矩意义下联合平稳，则 $\tilde{D}_a^{(q)} = \tilde{D}_{a\,0}^{(q)}$，并且可以省略式 (2.11) 中的上确界。我们的假设 2.2 在思想上类似于 Aue 等人（2009）中的假设 2.1；另请参见他们第 4 节提供的示例，这些示例说明了如何将此类框架应用于几种常见的时间序列模型。不过，请注意以下重要区别。量 $\tilde{D}_a^{(q)}$ 给出了 m-相依近似的优劣程度的上界，该程度通过 L^q 来衡量。尽管较大的 m 通常会带来更好的近似（即更小的 $\tilde{D}_a^{(q)}$），但我们并未对其衰减速率提出任何要求，而在推导渐近结果时通常会有此类要求。对于我们的主要结果，由于其本质上是有限样本结果，我们只要求 $X_a(t) - Y_a(t)$ 属于 L^q；也就是说，量 $\tilde{D}_{a\,t}^{(q)}$ 是有限的。

2.3 中心化平稳数据的显式上界

弱平稳序列情况下感兴趣的量的上界如下所示。

定理 *2.3.* 设 $\{\boldsymbol{X}(t)\}$ 是一个 d 维、中心化且弱平稳的过程。固定 $a, b \in \{1, \ldots, d\}$ 和 $k \in \{0, \ldots, n-1\}$，并令 $\hat{}\gamma_{ab}$ 和 $\gamma_{ab}(k)$ 分别如 *(1.2)* 和 *(1.3)* 中所定义。固定 $m \in \mathbb{N}_0$，并令 $\{\boldsymbol{Y}(t)\}$ 为假设 *2.2* 中的过程，假设该假设在 $q = 2$ 下成立，并且对所有 $t = 1, \ldots, n$，也有 $\|\boldsymbol{Y}(t)\|_6 < \infty$。对于给定的 $n \in \mathbb{N}$，假设$\tilde{}$ 分别在 *(1.5)* 和 *(2.8)* 中定义的 $\Sigma_{ab}(k)$ 和 $\Sigma_{ab}(k)$ 均为正。最后，令 $\tilde{D}_a^{(q)}$ 如 *(2.11)* 中所定义，则，

$$\tilde{K} := \tilde{D}_a^{(2)} \|X_b(0)\|_2 + \tilde{D}_a^{(2)} \tilde{D}_b^{(2)} + \|X_a(0)\|_2 \tilde{D}_b^{(2)}.$$

对于定理 *2.1* 中的 A_t 和 B_t，以及 (2.9) 中的 Q_t，我们有

$$d_{\mathrm{W}}\left(\mathcal{L}\left(\sqrt{n}\left(\hat{\gamma}_{ab}(k) - \gamma_{ab}(k)\right)\right), \mathcal{N}\left(0, \Sigma_{ab}(k)\right)\right) \leq \frac{k}{\sqrt{n}} |\gamma_{ab}(k)| + \frac{\sqrt{2}}{\sqrt{\pi \Sigma_{ab}(k)}} \left|\Sigma_{ab}(k) - \tilde{\Sigma}_{ab}(k)\right| + \frac{2(n-k)}{\sqrt{n}} \tilde{K} + \frac{2}{\left(n \tilde{\Sigma}_{ab}(k)\right)^{3/2}} \sum_{t=1}^{n-k} \tilde{Q}_t. \quad (2.12)$$

定理 2.3 的证明。选取 $\gamma := \gamma_{ab}(k)$ 和 $\sigma^2 := \Sigma_{ab}(k)$。注意，定理 2.3 的条件蕴含定理 2.1 的条件成立。因此，平稳情形下的上界由 $K_t \leq K$ 得出，对所有 $t = 1, \ldots, n-k$ 均成立。

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{t=1}^{n-k} \left| \frac{n}{n-k} \gamma - \mathbb{E}[Y_a(t+k) Y_b(t)] \right| \\
&\leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n-k} \left| \frac{n}{n-k} \gamma_{ab}(k) - \gamma_{ab}(k) \right| + \frac{1}{\sqrt{n}} \sum_{t=1}^{n-k} \mathbb{E} \left| X_a(t+k) X_b(t) - Y_a(t+k) Y_b(t) \right| \\
&\leq \frac{n-k}{\sqrt{n}} \left| \frac{k}{n-k} \gamma_{ab}(k) \right| + \frac{n-k}{\sqrt{n}} \tilde{K},
\end{aligned}$$

其中，我们使用了 (1.3) 和望远镜求和论证（引理 B.1，其中 $\alpha = 1$ 和 $p = 2$）。 □

注 2.4。构成(2.12) 右侧的四项大致可解释如下：*(i)* $kn^{-1/2}|\gamma_{ab}(k)|$，这是因为$\hat{}\gamma_{ab}(k)$ 是 $\gamma_{ab}(k)$的有偏估计，而极限正态分布的均值为零；*(ii)* $\sqrt{2/(\pi\Sigma_{ab}(k))}|\Sigma_{ab}(k) - \tilde{\Sigma}_{ab}(k)|$，与$\hat{}\gamma_{ab}(k)$ 的方差可能不同于 $n^{-1}\Sigma_{ab}(k)$ 这一事实有关；

(iii) $2(n-k)n^{-1/2}\tilde{K}$，源于我们的证明方法，其中使用了 m-依赖近似；*(iv)* $2(n\tilde{\Sigma}_{ab}(k))^{-3/2}\sum_{t=1}^{n-k}\tilde{Q}_t$，源于应用了*Stein*方法；参见引理*4.2*。

以下备注简要讨论了上界的计算方法及其阶数；详细说明见第2.6节。

注2.5。乍一看，该上界似乎有些复杂，尤其是由于 $\tilde{}$ $\tilde{}$表达式 Q_t。在第 *2.4* 节中，我们介绍两种方法，可以用其精确值能够在示例中计算出的表达式对 Q_t 给出上界。随后，在第 *3* 节中，对于因果自回归过程的情形，我们逐项计算这种上界项的精确值。为了获得一个收敛速率，我们将 m 选为 n的函数。使上界关于 n 的阶达到最优的选择取决于底层过程 $\{\boldsymbol{X}(t)\}$。在第*2.6*节中，我们讨论两种一般情形，在这两种情形下，上界分别为 $\mathcal{O}(n^{-1/2})$ 阶或 $\mathcal{O}(n^{-1/2}$ 对数 $n)$ 阶。

2.4 当已知 m 依赖近似时定理 2.3 中的上界

设 $\boldsymbol{X}(t)$ 满足：对于给定的 n、k 和 m，我们可以计算 $\gamma_{ab}(k)$、$\Sigma_{ab}(k)$、$\|X_a(0)\|_2$ 和 $\|X_b(0)\|_2$。进一步假设，我们可以选择 $\boldsymbol{Y}(t)$，使得 $\tilde{\Sigma}_{ab}(k)$、$\tilde{D}_a^{(2)}$ 和 $\tilde{D}_b^{(2)}$ $\tilde{}$ 均可计算。那么，要得到 (2.12) 中的上界，唯一缺少的部分就是在 (2.9) 中定义的 Q_t、$\tilde{}$。Q_t 定义中的绝对联合矩可能会带来不便。为了解决计算 Q_t 时可能遇到的问题，我们现在介绍两种方法，用可在示例中明确计算的量来给出 Q_t 的上界。首先，如果 $\{\boldsymbol{Y}(t)\}$ 是平稳的（否则请参见下文的方法1），我们用 $\|\boldsymbol{Y}(0)\|_6$ 来给出 Q_t 的上界；根据定理 2.3 的陈述，该值有限。其次，当 m 较大时，我们可以为 Q_t 得到一个更好的上界。采用第二种方法的代价是计算更加复杂，并且要求 $\|\boldsymbol{Y}(0)\|_8<\infty$。

$\tilde{}$**方法1用于给出** Q_t**的上界**。记为 $\mu_{jq}:=\|Y_j(0)\|_q$，我们有

$$\begin{aligned}\tilde{Q}_t &\le \sum_{j_1\in A_t}\sum_{j_2\in B_t}\Big(\mathbb{E}\big|\tilde{Z}(t)\tilde{Z}(j_1)\tilde{Z}(j_2)\big|+\mathbb{E}\big|\tilde{Z}(t)\tilde{Z}(j_1)\big|\mathbb{E}\big|\tilde{Z}(j_2)\big|\Big)+\frac{1}{2}\sum_{j_1\in A_t}\sum_{j_2\in A_t}\mathbb{E}\big|\tilde{Z}(t)\tilde{Z}(j_1)\tilde{Z}(j_2)\big|\\ &\le |A_t||B_t|\Big(\mu_{a6}^3\mu_{b6}^3+\mu_{a4}^2\mu_{b4}^2\mu_{a2}\mu_{b2}\Big)+\frac{1}{2}|A_t|^2\mu_{a6}^3\mu_{b6}^3\le\frac{5}{2}(4m+4k+1)^2\|\boldsymbol{Y}(0)\|_6^6.\end{aligned}\tag{2.13}$$

利用三角不等式、赫尔德不等式的推广形式以及 $\boldsymbol{Y}(t)$ 的平稳性，Q_t 定义中的联合矩被拆分为边际分布的矩

。如果 $\{\boldsymbol{Y}(t)\}$ 不具有平稳性，我们可以改用 $\sup_{t=1,\ldots,n}\|\boldsymbol{Y}(t)\|_6$。式（2.13）中的上界尤其简单且易于计算。实质上，我们看到边际矩 $\|Y_j(0)\|_q$（其中 $j=a,b$ 和 $q=2,4,6$）乘以 m^2 的某个倍数。m^2 阶的上界在 m 较小时最为有用。为了在 m 较大的情况下改进（2.13）~，接下来我们将推导一个关于 Q_t 的、以联合（非绝对）矩表示的上界。

~ ~**用于求上界的方法2** Q_t。我们应用柯西–施瓦茨不等式和 $\mathbb{E}(Z(t))=0$ 得到

$$\tilde{Q}_t \le \operatorname{var}\Big(\tilde{Z}(t)\tilde{Z}(A_t)\Big)^{1/2}\operatorname{var}\Big(\tilde{Z}(B_t)\Big)^{1/2}+\frac{1}{2}\Big[\mathbb{E}\Big(\tilde{Z}(t)\tilde{Z}(A_t)\Big)^2\Big]^{1/2}\operatorname{var}\Big(\tilde{Z}(A_t)\Big)^{1/2}. \qquad (2.14)$$

（2.14）右端与（2.13）中的第一个上界之间的一个关键区别是~，前者用 $Z(t)$ 的联合矩表示，而后者用~ $|Z(t)|$的联合矩表示。利用标准的组合论论证（参见 Brillinger (1975) 中的定理 2.3.2），可以根据 $\{\boldsymbol{Y}(t)\}$的累积量计算（2.14）的右端。这些论证直接明了但较为繁琐，因此推迟至 A 节。第二种方法的另一个重要优点是，在序列相关性通常会随较大滞后阶数而减弱、从而累积量可求和的情况下，第二种方法得到的上界阶数为 $\mathcal{O}(m)$；与方法1得到的 $\mathcal{O}(m^2)$ 上界相比，当 m 较大时，这一上界具有显著优势。直观地说，这是因为一个由 m 个元素组成的短程相关序列之和的方差阶数为 m。更多细节见命题 2.7 的证明，该证明可在 F.5 节中找到。

2.5 关于原始数据的上界

在第 2.4 节中，我们说明了当已知 $\{\boldsymbol{Y}(t)\}$（即 m-依赖近似）时，如何计算相应上界的细节。所述方法需要计算 $\{\boldsymbol{Y}(t)\}$ 的联合矩~，以便得到 $\Sigma_{ab}(k)$ 以及 (2.13) 或 (2.14) 右侧的值。如果能够进行此类计算，那么对于固定的 n 和 m，结合 (2.13) 或 (2.14) 对 (2.12) 得到的上界进行数值计算，是首选方法。本节旨在使我们的结果适用于以下情形：依赖于 $\{\boldsymbol{Y}(t)\}$ 的上界可能不便使用（例如，$\{\boldsymbol{Y}(t)\}$ 未知时）。

至少有两个充分理由，促使我们寻求一个只取决于以 $\{\boldsymbol{X}(t)\}$ 为定义依据的量的上界。第一个理由是哲学层面的。注意到感兴趣的统计量

$\gamma_{ab}(k)$ 是根据原始过程 $\{\boldsymbol{X}(t)\}$ 定义的，我们可以看到，(2.12) 左侧也只取决于 $\{\boldsymbol{X}(t)\}$。因此，(2.12)˜ ˜右侧所涉及的量，除其他项外，还包括以 $\Sigma_{ab}(k)$ 和 Q_t 定义的量，而二者都取决于 $\{\boldsymbol{Y}(t)\}$；这可以视为一种偏差。第二个理由是实用层面的。在第 2.6 节和 F.5 节中可以看到，当对 m 的依赖并非通过 $\boldsymbol{Y}(t)$ 的性质体现时，对该上界渐近性质的讨论可以得到简化。

为了得到一个关于 $\{\boldsymbol{X}(t)\}$ 的矩的上界，只需量化˜ ˜将 $\Sigma_{ab}(k)$ 替换为 $\Sigma_{ab}(k)$ 所产生的影响，以及将 Q_t 替换为

$$Q_t := \mathbb{E}\left|\left(Z(t)Z(A_t) - \mathbb{E}\Big(Z(t)Z(A_t)\Big)\right) Z(B_t)\right| + \frac{1}{2}\mathbb{E}\left|Z(t)\Big(Z(A_t)\Big)^2\right|, \tag{2.15}$$

其中，$Z(t) := X_a(t+k)X_b(t) - \mathbb{E}[X_a(t+k)X_b(t)]$、$t \in \mathbb{N}$，且 $Z(A) = \sum_{j\in A} Z(j)$ $A \subset \mathbb{N}$。

在 B 节中，我们给出了一些可用于推导关于$\{\boldsymbol{X}(t)\}$的矩的上界的结果。此外，在 C 节中，我们讨论非中心化数据的情形，并给出一个结果，用于在该情形下为 $d_{\mathrm{W}}(\mathcal{L}(\sqrt{n}(\hat{}\gamma^*_{ab}(k) - \gamma_{ab}(k))) \quad \mathcal{N}(0, \Sigma_{ab}(k)))$ 推导上界。

2.6 关于上界阶数的说明

在注2.5中，我们陈述了对上界进行渐近分析所得的结果。本节将给出详细说明。我们首先明确所采用的条件。为简便起见，我们仅考虑底层过程 $\{\boldsymbol{X}(t)\}$ 和滞后阶数 k 不允许随 n 改变的情形。我们考虑的两种情形为：

机制 1。设 $\{\boldsymbol{X}(t)\}$ 为 d 维中心化平稳过程，满足 $\|\boldsymbol{X}(0)\|_6 < \infty$，且为 M-相依（其中 $M \in \mathbb{N}_0$ 为固定值）。

机制 2。设 $\{\boldsymbol{X}(t)\}$ 为 d 维中心化平稳过程，满足 $\|\boldsymbol{X}(0)\|_8 < \infty$，且还满足假设 2.2（其中 $q = 8$），以及下文的 (2.16) 和 (2.17)。

我们分别假设情形 1 和情形 2 中的平稳性达到 6 阶或 8 阶矩。在情形 2 中，我们要求累积量的 8 阶可和性；即，对于 $p = 2$、.....、8，有

$$\sum_{k_1,\ldots,k_{p-1}=-\infty} (1 + |k_j|)|\mathrm{cum}(X_{a_1}(k_1), \ldots, X_{a_{p-1}}(k_{p-1}), X_{a_p}(0)| < \infty, \tag{2.16}$$

对于 $j = 1$，.....，$p-1$ 以及任意的 p 元组 a_1，.....，a_p。此外，我们要求假设 2.2 中依赖于 m 的近似足够精确，使得 L^q误差以指数速率趋于零；也就是说，存在常数 $K \geq 0$ 和 $\rho \in (0, 1)$，使得对于每个 $n \in \mathbb{N}$

我们可以选取 $m = m_n \in \mathbb{N}_0$ 以及一个依赖于 m、 d元的过程 $\{\boldsymbol{Y}^{(m)}(t)\}$，使其满足

$$\tilde{D}_a^{(q)} := \sup_{1\le t\le n} \|X_a(t) - Y_a^{(m)}(t)\|_q \le K\rho^m, \quad \text{for } a = 1,\dots,d;\ q = 8. \tag{2.17}$$

注 2.6。 *(i)* 机制 *1* 的示例包括有限阶移动平均过程和独立数据。在这一机制下，对于 $p = 2,\dots,6$，(2.16) 成立，因为只要某个变量独立于其他变量，累积量就会消失。此外，对于任意 M-依赖过程，如机制 *1* 中所述，假设 *2.2* 中 m-依赖近似的规范选择为$\boldsymbol{Y}_{(m)}(t) = \boldsymbol{X}(t)$，其中 $m \ge M$。在上界 (2.12) 中，将量 m 取为 $m = \min\{M, n\}$，可见(2.17) 的一个更强版本成立，其中对于 $n \ge M$，有 $\tilde{D}_a^{(q)} = 0$ 。

(ii) 机制 *2* 的一个示例是线性过程 $\boldsymbol{X}(t) = \sum_{j=0}^{\infty} \boldsymbol{\Psi}(j)\boldsymbol{\epsilon}(t-j)$，其中系数的谱范数满足 $\|\boldsymbol{\Psi}(j)\|_2 \le \rho^j$，对某个 $\rho \in (0,1)$ 成立，且创新项为独立同分布，满足 $\|\boldsymbol{\epsilon}(t)\|_8 < \infty$。由此可证明(2.16) 成立，并且(2.17)在 $C := \|\boldsymbol{\epsilon}(t)\|_8\rho/(1-\rho)$时成立。特别地，因果自回归过程也包括在内。

以下命题给出了定理 2.3 中上界 (2.12) 的阶。证明见 F.5 节。

命题 2.7。 *(i)* 在机制 *1* 中，取 $m := $ 最小值$\{M, n\}$，上界的阶为 $\mathcal{O}(n^{-1/2})$。*(ii)* 在机制 *2* 中，取 $m := C$对数 n, $C \ge 2^{3}$对数$(1/\rho)$，其中 ρ与(2.17)中相同，上界的阶为 $\mathcal{O}(n^{-1/2}$对数$n)$。

3 示例

3.1 一阶因果自回归过程

作为一个示例，为了讨论定理 2.3 的结果，我们现在考虑数据源自因果 AR(1) 过程 $\{X(t)\}$，该过程满足 $X(t) = \alpha X(t-1) + \varepsilon(t)$，其中$|\alpha| < 1$，$\{\varepsilon(t)\}$ 为独立同分布，且满足 $\mathbb{E}|\varepsilon(t)|^8 < \infty$。我们考虑 $\alpha \in \{0, 0.1, 0.3, 0.5, 0.7\}$以及关于创新项分布的三种情形：

- $\varepsilon(t) \sim \mathcal{N}(0,1)$，或
- $\varepsilon(t) \sim \nu^{-1/2}(\nu-2)^{1/2}t_\nu$，其中我们选择 $\nu \in \{9, 14\}$。

我们选取正态分布作为轻尾分布的例子，选取缩放后的 t_9-分布作为尾部更重、但仍满足存在 8 阶矩条件的分布，并选取缩放后的 t_{14}-分布作为介于两者之间的例子。请注意，对于这

三种情形，我们都将 1 阶和 2 阶累积量标准化；即 $\kappa_1 := \mathbb{E}(\varepsilon(t)) = 0$ and$\kappa_2 := \text{var}(\varepsilon(t)) = 1$。更高阶累积量取决于创新项的分布。若 $\varepsilon(t) \sim \mathcal{N}(0,1)$，则 3 阶及以上的累积量均为零；即$\kappa_p := \text{cum}_p(\varepsilon(t)) = 0$，其中 $p = 3, 4, \ldots$。当 $\varepsilon(t) \sim \nu^{-1/2}(\nu-2)^{1/2} t_\nu$, $\nu > 8$ 时，由于对称性，$p = 3, 5, 7$ 阶累积量为零，而 4、6 和 8 阶累积量则

$$\kappa_4 = \frac{6}{\nu - 4}, \quad \kappa_6 = \frac{240}{(\nu-4)(\nu-6)}, \quad \text{and} \quad \kappa_8 = \frac{5040(5\nu - 22)}{(\nu-4)^2(\nu-6)(\nu-8)}.$$

用于复现第 3 节结果的 R 代码和说明可在 https://github.com/tobiaskley/ccf_bounds_replication_package 获取。

3.2 计算上界

我们结合第二种方法的定理 2.3 和式 (2.14) 来计算上界 ˜，该方法在第 2.4 节中介绍；数据来自第 3.1 节所述的 AR(1) 过程。此例中上界的具体计算细节留待 D 节介绍。请注意，给定自回归参数 α、$\varepsilon(t)$ 的分布、区段长度 n 和滞后阶数 k 后，上界仍是 m 的函数。我们将上界记为 $B_n(m)$，以强调可以针对 m 的不同取值进行计算。此外，我们将 $m^* := \arg\min_{m=0,1,\ldots,m_{\max}} B_n(m)$ 记为使 m 达到最小值的取值。我们引入上界 $m_{\max}$ 作为计算的停止规则，并将其选得足够大，使得在此例的所有情况下都满足 $m^* < m_{\max}$，这意味着最小值不会在 $m = m_{\max}$ 处取得。我们选择了 $m_{\max} = 30$。图1展示了上界 $B_n(m)$ 随 m 变化的取值，并比较了不同的 n 以及$\varepsilon(t)$ 的不同分布。从图1中自左向右比较各图可以看出，随着 n 增大，m^* 增长得非常缓慢。这并不令人意外，因为在这个因果 AR(1) 过程的例子中，我们处于第 2.6 节所述的机制 2；回顾命题 2.7 中的渐近分析，其中 $m^* = \mathcal{O}(\log n)$，这意味着 $B_n(m^*) = \mathcal{O}(n^{-1/2} \log n)$。从图1中自上而下比较各图可以看出，尾部越重，上界的取值越大。这也符合预期，因为当创新项 $\varepsilon(t)$ 的分布具有更重的尾部时，其分布的累积量会变大。

表1展示了在 $\varepsilon(t) \sim \sqrt{7/9}\, t_9$的情况下，不同 k、α和 n 取值对应的上界 $B_n(m^*)$。当 $\varepsilon(t) \sim \mathcal{N}(0, 1)$或$\varepsilon(t) \sim \sqrt{12/14}\, t_{14}$ 时的数值分别见G节的表5和表7。我们选择展示

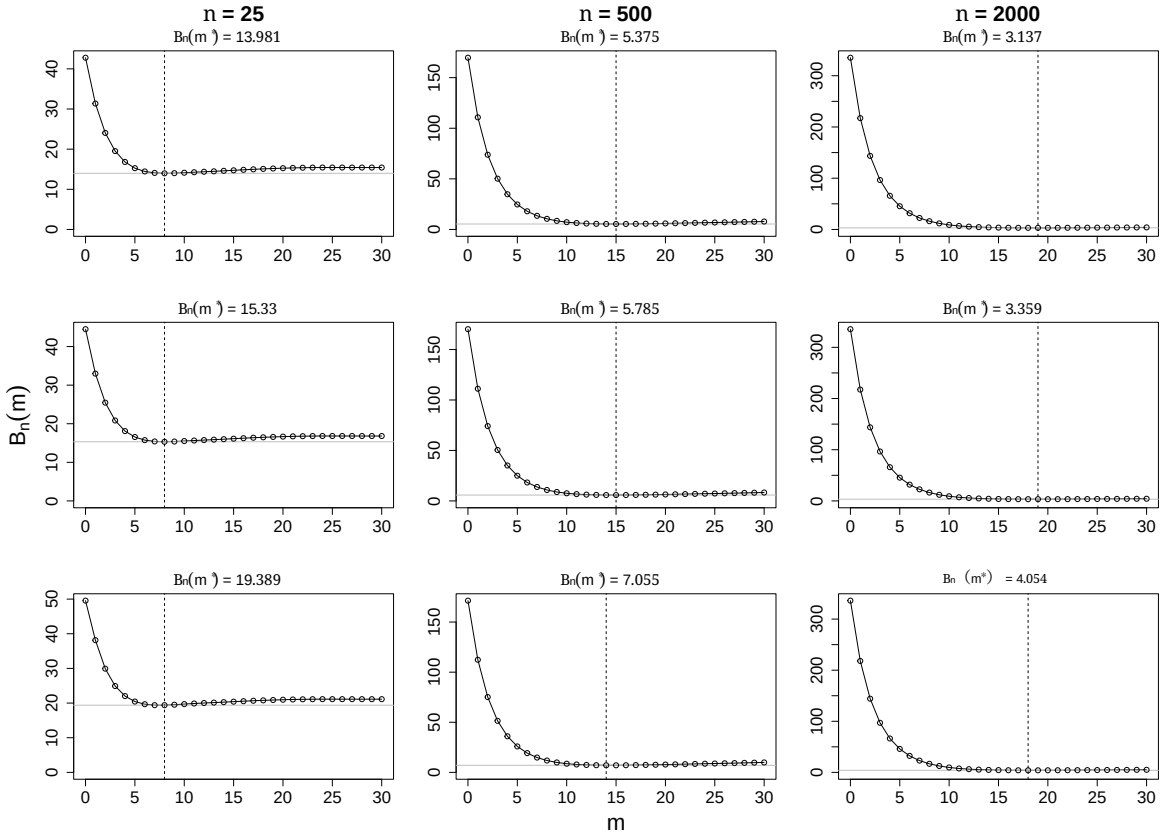

图1： 定理 *2.3* 中经验自协方差上界的取值，其中滞后阶数为 $k = 0$，*AR(1)* 过程的参数为 $\alpha = 0.7$，该上界是 m 的函数。竖直虚线表示达到最小值时的 m^*。灰色水平线表示最小值。顶行、中间行和底行分别显示 $\varepsilon(t) \sim \mathcal{N}(0,\ 1)$、$\varepsilon(t) \sim \sqrt{12/14}\ t_{14}$ 和 $\varepsilon(t) \sim \sqrt{7/9}\ t_9$。左列、中列和右列分别显示 $n = 25$、500 和 2000。

主论文中尾部最重的情况，因为在这种情况下，自协方差和交叉协方差函数的估计量收敛到高斯极限的速度最慢。我们没有考虑负的 α，因为在所考虑的情况下，结果与 $-\alpha$时相同。可以看出，上界的值随 $|\alpha|$ 增大而增大。比较各表中的上界可知，在大多数情况下，尾部越重，上界的值越大。可以看出，上界的值随 n 增大而减小。

表1： 结合式 (2.14) 使用定理 2.3 得到的上界值，其中 $m = m^*$ 按照第 3.2 节所述取值以使上界最小，适用于不同滞后阶数 k 和样本量 n下的经验自协方差。数据来自一个 AR(1) 过程，其 $\varepsilon(t) \sim \sqrt{7/9}\, t_9$，其中 α 取不同值。

k	$\alpha \mid n$	25	50	75	100	150	200	250	500	1000	2000
0	0	0.912	0.645	0.527	0.456	0.372	0.322	0.288	0.204	0.144	0.102
	0.1	11.003	9.294	8.822	8.707	7.509	6.658	6.091	4.779	3.770	2.773
	0.3	16.088	12.932	11.481	10.287	8.937	8.192	7.484	5.751	4.386	3.375
	0.5	16.952	13.518	11.760	10.689	9.245	8.365	7.706	5.955	4.579	3.514
	0.7	16.871	14.042	12.434	11.367	9.945	9.018	8.343	6.531	5.087	3.961
1	0	2.564	1.818	1.485	1.286	1.050	0.909	0.813	0.574	0.406	0.287
	0.1	7.711	5.701	4.808	4.285	3.686	3.350	3.118	2.283	1.708	1.326
	0.3	9.811	7.567	6.601	5.939	5.063	4.554	4.217	3.213	2.491	1.908
	0.5	12.512	9.980	8.716	7.861	6.825	6.133	5.671	4.394	3.402	2.641
	0.7	14.968	12.828	11.376	10.387	9.092	8.247	7.644	5.983	4.668	3.647
2	0	4.088	2.916	2.385	2.067	1.688	1.462	1.308	0.924	0.653	0.462
	0.1	10.398	7.659	6.409	5.667	4.804	4.309	3.925	2.833	2.074	1.561
	0.3	10.801	8.273	7.175	6.405	5.424	4.850	4.467	3.353	2.557	1.913
	0.5	12.211	9.739	8.459	7.632	6.592	5.921	5.471	4.210	3.233	2.486
	0.7	14.392	12.610	11.215	10.236	8.954	8.111	7.509	5.866	4.565	3.556

表 2： 定理 2.3 中针对经验自协方差所考虑的真实 1-Wasserstein 距离的值，涵盖一系列滞后阶数 k 和样本量 n。数据源自一个 AR(1) 过程，其中$\varepsilon(t) \sim \sqrt{7/9}\, t_9$，$\alpha$ 取一系列不同的值。

k	$\alpha \mid n$	25	50	75	100	150	200	250	500	1000	2000
0	0	0.288	0.218	0.184	0.163	0.136	0.120	0.109	0.080	0.058	0.041
	0.1	0.294	0.222	0.188	0.166	0.139	0.123	0.111	0.081	0.059	0.042
	0.3	0.354	0.266	0.224	0.198	0.165	0.145	0.131	0.095	0.069	0.049
	0.5	0.536	0.401	0.336	0.296	0.246	0.216	0.194	0.140	0.101	0.072
	0.7	1.185	0.891	0.746	0.655	0.544	0.475	0.428	0.307	0.219	0.156
1	0	0.072	0.040	0.028	0.021	0.015	0.011	0.009	0.005	0.002	0.001
	0.1	0.103	0.069	0.055	0.047	0.038	0.032	0.029	0.020	0.014	0.010
	0.3	0.256	0.187	0.155	0.135	0.111	0.097	0.087	0.062	0.044	0.031
	0.5	0.524	0.384	0.319	0.279	0.230	0.200	0.180	0.128	0.091	0.065
	0.7	1.282	0.951	0.791	0.693	0.572	0.499	0.448	0.320	0.227	0.161
2	0	0.083	0.045	0.031	0.024	0.016	0.013	0.010	0.005	0.003	0.001
	0.1	0.088	0.049	0.034	0.026	0.018	0.014	0.012	0.006	0.004	0.002
	0.3	0.167	0.113	0.091	0.078	0.063	0.055	0.049	0.034	0.024	0.017
	0.5	0.449	0.329	0.272	0.237	0.195	0.170	0.152	0.109	0.077	0.055
	0.7	1.307	0.966	0.802	0.701	0.578	0.504	0.452	0.322	0.229	0.162

为了与我们的上界进行比较，表1、表5和表7展示了该上界；我们还在表2、表6和表8中给出了通过模拟得到的真实Wasserstein距离数值，其中表5–表8见G节。关于真实Wasserstein距离模拟的更多细节见E节。观察这些数值可以发现，正如预期，我们的上界始终大于通过模拟得到的真实Wasserstein距离。

4 定理2.1的证明

在本节主要证明之前，我们先讨论一个有用的引理，它概述了本文所用的Stein方法结果，该结果适用于一般的局部依赖条件。考虑由随机变量 $\{\xi_i, i \in J\}$ 构成的集合，其指标集 J 为有限集。那么，局部依赖条件为

(LD) 对每个 $i \in J$，都存在 $A_i \subset B_i \subset J$，使得 ξ_i 与 $\{\xi_j$ 独立：$j \notin A_i\}$；并且 $\{\xi_j$ 与 $\{\xi_k$ 独立：$k \notin B_i\}$。

对于任意 $A \subset J$，我们现在记为

$$\xi(A) = \sum_{j \in A} \xi_j. \tag{4.18}$$

注4.1。考虑一个由随机变量 $X_1, \ldots, X_n$ 组成的 m-依赖序列。对于每个$i = 1, \ldots, n$，随机变量集合 $\{X_j : \ j \le i\}$ 和 $\{X_j : \ j > i + m\}$ 相互独立。因此，*(LD)* 满足条件，其中 $J := \{1, \ldots, n\}$, $A_i := \{\ell \in J : \ |\ell - i| \le m\}$，且 $B_i := \{\ell \in J : \ |\ell - i| \le 2m\}$。

下面的引理给出了满足上述条件 (LD) 的随机变量之和的分布与正态分布之间的Wasserstein距离上界。假设这些随机变量的均值为零，且方差不一定等于 1。证明见F.1节，其思路基于 Chen 等人（2011）第 134 页中定理4.13的证明步骤。

引理4.2。设 $\{\xi_i, i \in J\}$ 是一个取值于 $\mathbb{R}$ 的均值为零的随机场，并满足条件（*LD*）。记 $W := \sum_{i \in J} \xi_i$，并假设 $0 < \sigma^2 := \mathrm{var}(W) < \infty$。则对于(4.18)中所定义的 $\xi(A)$，其 *Wasserstein*距离 d_{W}（定义见(1.6)）满足 *Erotiske*

$$d_{\mathrm{W}}(\mathcal{L}(W), \mathcal{N}(0, \sigma^2)) \le \frac{2}{\sigma^3} \sum_{i \in J} \left\{ \mathbb{E}\left|\left(\xi_i \xi(A_i) - \mathbb{E}(\xi_i \xi(A_i))\right) \xi(B_i)\right| + \frac{1}{2} \mathbb{E}\left|\xi_i (\xi(A_i))^2\right| \right\}.$$

定理 2.1 的证明。首先，对于 $Z^*(t) = X_a(t+k)X_b(t) - \frac{n}{n-k}\gamma$，我们有人人操

$$\sqrt{n}\left(\hat{\gamma}_{ab}(k) - \gamma\right) = \sqrt{n}\left(\frac{1}{n}\sum_{t=1}^{n-k} X_a(t+k)X_b(t) - \gamma\right) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} Z^*(t). \tag{4.19}$$

For $\tilde{Z}(t) = Y_a(t+k)Y_b(t) - \mathbb{E}[Y_a(t+k)Y_b(t)]$, 三角不等式和 (4.19) 得出

$$d_{\mathrm{W}}\left(\mathcal{L}\left(\sqrt{n}\left(\hat{\gamma}_{ab}(k) - \gamma\right)\right), \mathcal{N}(0,\sigma^2)\right) = d_{\mathrm{W}}\left(\mathcal{L}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} Z^*(t)\right), \mathcal{N}(0,\sigma^2)\right) \le D_1 + D_2 + D_3,$$

其中

$$D_1 := d_{\mathrm{W}}\left(\mathcal{L}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} Z^*(t)\right), \mathcal{L}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} \tilde{Z}(t)\right)\right), \tag{4.20}$$

$$D_2 := d_{\mathrm{W}}\left(\mathcal{L}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} \tilde{Z}(t)\right), \mathcal{N}\left(0, \tilde{\Sigma}_{ab}(k)\right)\right), \tag{4.21}$$

$$D_3 := d_{\mathrm{W}}\left(\mathcal{N}\left(0, \tilde{\Sigma}_{ab}(k)\right), \mathcal{N}\left(0, \sigma^2\right)\right), \tag{4.22}$$

其中 $\tilde{\Sigma}_{ab}(k)$ 如下 如 (2.8) 所示。现在我们继续求 (4.20)、(4.21) 和 (4.22) 的上界).

（4.20）的上界： 对于 $h \in \mathcal{H}$，由于 $\|h\|_{\mathrm{Lip}} \le 1$，因此

$$\begin{aligned}
&\left|\mathbb{E}\left[h\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} Z^*(t)\right) - h\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} \tilde{Z}(t)\right)\right]\right| \le \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\mathbb{E}\left|Z^*(t) - \tilde{Z}(t)\right| \\
&\le \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\mathbb{E}\left|X_a(t+k)X_b(t) - Y_a(t+k)Y_b(t)\right| \qquad (4.23)\\
&\quad + \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\left|\frac{n}{n-k}\gamma - \mathbb{E}[Y_a(t+k)Y_b(t)]\right|.
\end{aligned}$$

接下来，我们通过望远镜求和论证对 (4.23) $\le \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} K_t$ 进行上界估计，具体由引理 F.1 给出，其中 $\alpha = 1$、$p = 2$、$X_1 := X_a(t+k)$、$X_2 := X_b(t)$、$Y_1 := Y_a(t+k)$ 以及 $Y_2 := Y_b(t)$。因此，我们有

$$(4.20) \le \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k} K_t + \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\left|\frac{n}{n-k}\gamma - \mathbb{E}[Y_a(t+k)Y_b(t)]\right|. \tag{4.24}$$

(4.21)的上界： 这里我们使用Stein方法。令 $W := \frac{1}{\sqrt{n}}\sum_{t=1}^{n-k}\tilde{Z}(t)$，并注意到

~ $\mathbb{E}(Z(t)) = 0$ 且 $\text{var}(W) = \Sigma_{ab}(t)$。由引理4.2可得

$$(4.21) \leq \frac{2}{\left(\tilde{\Sigma}_{ab}(k)\right)^{3/2}} \sum_{t=1}^{n-k} \tilde{Q}_t. \tag{4.25}$$

（4.22）的上界：利用 Nourdin 和 Peccati（2012）第 69 和 70 页中的结果，可以直接得出

$$(4.22) \leq \frac{\sqrt{2}}{\sqrt{\pi \Sigma_{ab}(k)}} \left| \Sigma_{ab}(k) - \tilde{\Sigma}_{ab}(k) \right|. \tag{4.26}$$

结合 （4.24）、（4.25）和（4.26），即可得到定理 2.1 中式（2.10）的结果。 □

5 讨论

本文得到了非平稳和平稳数据下，自协方差函数和交叉协方差函数估计量的真实分布与其极限高斯分布之间的 Wasserstein距离上界。与文献中针对一般线性统计量的现有结果相比，除所采用的工具（部分基于 Stein方法）和证明方法外，本文的结果主要在三个方面具有新意。首先，本文结果适用于非平稳数据序列，见定理 2.1。其次，我们关注的不仅是收敛速率，所推导的上界还以样本量、滞后阶数以及依赖于时间序列模型的常数表示，且完全显式。这使得我们能够在示例中计算上界。第三，我们采用的假设不具限制性，且部分基于对原始时间序列的 m-依赖近似，这种方法在应用中便于使用。相比之下，现有结果关注的结构限制性更强，通常具有概率性质且难以在实践中验证，例如强混合和 φ-混合条件，或离散时间鞅情形。

参考文献

Anastasiou, A.（2017）。相依随机变量的最大似然估计量正态近似上界。《统计与概率快报》，129：171–181。

Anastasiou, A.、Balasubramanian, K. 和 Erdogdu, M. A.（2019）。关于……的正态近似

通过鞅中心极限定理的非渐近速率实现随机梯度下降。在第三十二届学习理论会议论文集中，机器学习研究论文集第99卷，第115–137页。Anderson, T. W.（1971）。时间序列的统计分析。Wiley，纽约。Aue, A.、Hörmann, S.、Horváth, L. 和 Reimherr, M.（2009）。非平稳性与平稳性下多元时间序列模型协方差结构的突变检测。《统计年刊》，37（6B）：4046–4087。Brillinger, D. R.（1975）。时间序列：数据分析与理论。Holt, Rinehart and Winston, Inc.，纽约。Brockwell, P. J. 和 Davis, R. A.（2006）。时间序列：理论与方法。施普林格，纽约，第2版。Chen, L. H. Y.、Goldstein, L. 和 Shao, Q.-M.（2011）。通过*Stein*方法进行正态近似。概率及其应用。施普林格。Dedecker, J.、Merlevède, F. 和 Rio, E.（2009）。投影判据下中心极限定理中最小距离的收敛速率。《电子概率期刊》，14：978–1011。Dedecker, J. 和 Rio, E.（2008）。关于平稳序列的均值中心极限定理。《亨利·庞加莱研究所年刊：概率与统计》，44（4）：693–726。Fan, X.（2019）。某些鞅中心极限定理中的精确收敛速率。《数学分析与应用杂志》，469：1028–1044。Fan, X. 和 Ma, X.（2020）。关于鞅中心极限定理的Wasserstein距离。《统计与概率快报》，167：108892。Horváth, L. 和 Kokoszka, P.（2008）。长记忆时间序列的样本自协方差。《伯努利》，14（2）：405–418。Jirak, M.（2012）。Yule-Walker估计量和阶数选择的同时置信带。《统计年刊》，40（1）：494–528。Jirak, M.（2014）。序贯自回归拟合的同时置信带。《多元分析杂志》，124：130–149。Jirak, M.（2016）。弱依赖下的Berry-Esseen定理。《概率年刊》，44：2024–2063。

Kley, T., Preuß, P., and Fryzlewicz, P.（2019）。平稳性和非平稳性下时间序列的预测型有限样本模型选择。《电子统计学杂志》，13(2):3710 – 3774。Munk, A. and Czado, C.（1998）。相似分布的非参数验证与拟合优度评估。《皇家统计学会杂志》。*B*辑（统计方法学），60(1):223–241。Nourdin, I. and Peccati, G.（2012）。*Malliavin* 微积分下的正态近似：从 *Stein* 方法到普适性。剑桥数学论丛第192卷。剑桥大学出版社。Röllin, A.（2018）。鞅中心极限定理均值的定量上界。《统计与概率快报》，138:171–176。Stein, C.（1972）。相依随机变量求和的分布正态近似误差上界。载于第六届伯克利数学统计与概率研讨会论文集，第**2**卷，第586–602页。伯克利：加利福尼亚大学出版社。Sunklodas, J.（2007）。强混合随机变量的正态近似。《应用数学学报》，97:251–260。Sunklodas, J.（2011）。关于 φ-混合随机变量正态近似的一些估计。《立陶宛数学杂志》，51:260–273。

……的补充材料

“非平稳性和平稳性下经验自协方差与交叉协方差正态近似的Wasserstein距离上界”

第 2.4 节的技术细节

这里，我们说明如何根据 $\{\boldsymbol{Y}(t)\}$（其为中心化且弱平稳的过程）的 8 阶及以下累积量计算 (2.14) 的右端。首先，注意到

$$\mathbb{E}\Big(\tilde{Z}(t)\tilde{Z}(A_t)\Big)^2 = \mathrm{var}\Big(\tilde{Z}(t)\tilde{Z}(A_t)\Big) + \bigg[\mathbb{E}\Big(\tilde{Z}(t)\tilde{Z}(A_t)\Big)\bigg]^2,$$

and

$$\begin{aligned}
&\mathrm{var}\Big(\tilde{Z}(t)\tilde{Z}(A_t)\Big) = \sum_{j_1\in A_t}\sum_{j_2\in A_t}\mathrm{cov}\big(\tilde{Z}(t)\tilde{Z}(j_1),\tilde{Z}(t)\tilde{Z}(j_2)\big)\\
&= \sum_{j_1\in A_t}\sum_{j_2\in A_t}\Big[\mathrm{cum}(\tilde{Z}(t),\tilde{Z}(j_1),\tilde{Z}(t),\tilde{Z}(j_2)) + \mathrm{cum}(\tilde{Z}(t),\tilde{Z}(t))\mathrm{cum}(\tilde{Z}(j_1),\tilde{Z}(j_2))\\
&\qquad\qquad +\mathrm{cum}(\tilde{Z}(t),\tilde{Z}(j_2))\mathrm{cum}(\tilde{Z}(j_1),\tilde{Z}(t))\Big]\\
&= \sum_{j_1\in A_t}\sum_{j_2\in A_t}\mathrm{cum}(\tilde{Z}(t),\tilde{Z}(t),\tilde{Z}(j_1),\tilde{Z}(j_2)) + \tilde{C}(0)\mathrm{var}\Big(\sum_{j\in A_t}\tilde{Z}(j)\Big) + \Big[\mathbb{E}\Big(\tilde{Z}(t)\sum_{j\in A_t}\tilde{Z}(j)\Big)\Big]^2,
\end{aligned}$$

其中

$$\begin{aligned}
\tilde{C}(u) := \mathrm{cum}(\tilde{Z}(0),\tilde{Z}(u)) &= \mathrm{cum}(Y_a(k),Y_b(0),Y_a(u+k),Y_b(u))\\
&\quad + \tilde{\gamma}_{aa}(u)\tilde{\gamma}_{bb}(u) + \tilde{\gamma}_{ab}(k-u)\tilde{\gamma}_{ab}(k+u),
\end{aligned} \tag{A.27}$$

其中 $\tilde{\gamma}_{ab}(k) := \mathrm{cum}(Y_a(k)\ Y_b(0))$。

注意到 (2.9) 中的 A_t 和 B_t 是形如 $\{s,\ldots,e\}$、˜的有限连续整数集，因此我们只需计算或给出 $C(0)$的上界，以及以下三个量

$$\sum_{j_1=s}^{e}\sum_{j_2=s}^{e}\mathrm{cum}(\tilde{Z}(t),\tilde{Z}(t),\tilde{Z}(j_1),\tilde{Z}(j_2)),\quad \mathrm{var}\Big(\sum_{j=s}^{e}\tilde{Z}(j)\Big),\quad \text{and}\quad \mathbb{E}\Big(\tilde{Z}(t)\sum_{j=s}^{e}\tilde{Z}(j)\Big), \tag{A.28}$$

其中 $s \le t \le e$。(A.28) 中的第二个和第三个量只取决于二阶 an d

(1,1)	(1,2)
(2,1)	(2,2)
(3,1)	(3,2)
(4,1)	(4,2)

(a) 待划分的指标表

$\tilde{Y}_{1,1} := Y_a(t+k)$	$\tilde{Y}_{1,2} := Y_b(t)$
$\tilde{Y}_{2,1} := Y_a(t+k)$	$\tilde{Y}_{2,2} := Y_b(t)$
$\tilde{Y}_{3,1} := Y_a(j_1+k)$	$\tilde{Y}_{3,2} := Y_b(j_1)$
$\tilde{Y}_{4,1} := Y_a(j_2+k)$	$\tilde{Y}_{4,2} := Y_b(j_2)$

(b) 考虑其累积量的变量

表3： 式 (A.29) 中求和各加项的待划分指标表和变量。

四阶 $\boldsymbol{Y}(t)$ 的 r 阶矩，在这个意义上更容易获得。我们有：

$$\begin{aligned}\operatorname{var}\Big(\sum_{j=s}^{e}\tilde{Z}(j)\Big) &= \sum_{j_1=s}^{e}\sum_{j_2=s}^{e}\operatorname{cov}(\tilde{Z}(0),\tilde{Z}(j_2-j_1)) = \sum_{|u|\le e-s}(e-s+1-|u|)\operatorname{cov}(\tilde{Z}(0),\tilde{Z}(u))\\ &= (e-s+1)\sum_{|u|\le e-s}\Big(1-\frac{|u|}{e-s+1}\Big)\tilde{C}(u),\end{aligned}$$

其中 $\tilde{C}(u)$ 如 (A.27) 中所定义，类似地

$$\mathbb{E}\Big(\tilde{Z}(t)\sum_{j=s}^{e}\tilde{Z}(j)\Big) = \sum_{j=s}^{e}\operatorname{cov}(\tilde{Z}(0),\tilde{Z}(j-t)) = \sum_{u=s-t}^{e-t}\operatorname{cov}(\tilde{Z}(0),\tilde{Z}(u)) = \sum_{u=s-t}^{e-t}\tilde{C}(u).$$

对于 (A.28) 中的第一项，由 Brillinger (1975) 的定理 2.3.2 可知，

$$\sum_{j_1=s}^{e}\sum_{j_2=s}^{e}\operatorname{cum}(\tilde{Z}(t),\tilde{Z}(t),\tilde{Z}(j_1),\tilde{Z}(j_2)) = \sum_{j_1=s}^{e}\sum_{j_2=s}^{e}\sum_{\nu}\prod_{r=1}^{R}\operatorname{cum}(\tilde{Y}_{i,j}:(i,j)\in\nu_r) \tag{A.29}$$

其中 $\tilde{Y}_{i,j}$ 在表 3(b) 中定义，且求和遍及表 3(a) 中所有不可分解划分 $\nu := \{\nu_1,\ldots,\nu_R\}$，其中 $|\nu_r|\ge 2$。

将一个含有 8 个条目的表格划分的方式数，即第 8 个贝尔数 $B_8 = 4140$。通过算法枚举所有划分非常简单，例如使用我们作为复现包一部分提供的 R 函数 all_不可分解_划分()，并验证其中只有 545 个是不可分解的，且每个划分中的各集合都至少包含两个元素，因此在考虑分布 $\mathbb{E}(\boldsymbol{Y}(t)) = 0$ 时，只需考虑这些划分。对于所有奇数阶累积量均为零的情形（例如第 3 节中的示例），我们只考虑各集合元素个数为偶数的划分，这样的划分共有 249 个。此外，对于 $\boldsymbol{Y}(t)$ 为中心化正态分布的情形，我们只考虑各集合恰好包含 2 个元素的不可分解划分，这样的划分共有 48 个。

B 第 2.5 节的技术细节

我们给出引理 B.3 和引理 B.4，它们可用于根据仅通过 $\{\boldsymbol{X}(t)\}$ 定义的量以及来自假设 2.2 的近似误差 $\tilde{D}_j^{(q)}$，得到 (2.12) 的上界。为陈述本节的结果，我们将在 (B.30) 中定义量 $\tilde{K}_p^{(\alpha)}$。接下来，我们陈述引理 B.1，其中指出 $\tilde{K}_p^{(\alpha)}$ 是 $\{\boldsymbol{X}(t)\}$ 的乘积与相应的 $\{\boldsymbol{Y}(t)\}$ 乘积之间近似误差（以 L^α 衡量）的上界。

引理 B.1。设 $\alpha \geq 1$, $p \in \mathbb{N}$, $a_1, \ldots, a_p \in \{1, \ldots, d\}$，以及 $t_1, \ldots, t_p \in \mathbb{Z}$。此外，设 $\{\boldsymbol{X}(t)\}$ 为 d 维、中心化且弱平稳，并假设 $\|\boldsymbol{X}(0)\|_{\alpha p} < \infty$。设假设 *2.2* 成立，其中 $q = \alpha p$。定义

$$\tilde{K}_p^{(\alpha)} = \tilde{K}_p^{(\alpha)}(\{a_1, \ldots, a_p\}) := \sum_{(\ell_1,\ldots,\ell_p)\in\Lambda_p} \prod_{i=1}^{p} \left(\tilde{D}_{a_i}^{(\alpha p)}\right)^{\ell_i} \|X_{a_i}(0)\|_{\alpha p}^{1-\ell_i}, \tag{B.30}$$

其中 $\tilde{D}_{a_i}^{(\alpha p)}$与(2.11)中的相同，$\Lambda_1 := \{1\}$, $\Lambda_2 := \{(1,0)、(1,1)、(0,1)\}$，并且

$$\Lambda_p := \{1\} \times \{0,1\}^{p-1} \cup \bigcup_{j=2}^{p-1} \{0\}^{j-1} \times \{1\} \times \{0,1\}^{p-j} \cup \{0\}^{p-1} \times \{1\}, \quad p = 3, 4, \ldots.$$

于是，我们有

$$\left\| \prod_{i=1}^{p} X_{a_i}(t_i) - \prod_{i=1}^{p} Y_{a_i}(t_i) \right\|_\alpha \leq \tilde{K}_p^{(\alpha)}. \tag{B.31}$$

引理 B.1 的证明见第 F.2 节。当 $p := 2$, $a_1 := a$，且$a_2 := b$ 时的情形

$$\tilde{K}_2^{(\alpha)} := \tilde{K}_2^{(\alpha)}(\{a,b\}) := \tilde{D}_a^{(2\alpha)} \|X_b(0)\|_{2\alpha} + \tilde{D}_a^{(2\alpha)} \tilde{D}_b^{(2\alpha)} + \tilde{D}_b^{(2\alpha)} \|X_a(0)\|_{2\alpha},$$

在后续证明、引理以及主要定理中反复出现。

注 B.2。*(i)* (B.30) 中的量 $\tilde{K}_p^{(\alpha)}$取决于指标集合，但与指标的顺序无关。这是因为集合 Λ_p 在指标置换下封闭；而 Λ_p 在换位（交换任意两个指标会得到 Λ_p中的另一个元素）下封闭，因此满足这一条件。因此，上界具有一个理想性质：无论(B.31)左侧指标 a_i 的顺序如何，上界都保持不变。

(ii) 在应用中，$\tilde{D}_a^{(\alpha p)}$的阶通常为 $o(1)$，且 $\|X_a(0)\|_{\alpha p} = \mathcal{O}(1)$，其中 $m \to \infty$。因此，$\tilde{K}_{p,\ m}^{(\alpha)} = \mathcal{O}(\max_{a\in\{a_1,\ldots,a_q\}} \tilde{D}_a^{(\alpha p)})$，这可由以下事实看出：$\Lambda_p$ 的每个元素至少有一个指标为 *1*。

(iii) 为了获得 $\{\boldsymbol{X}(t)\}$ 的联合矩与 m-依赖序列 $\{\boldsymbol{Y}(t)\}$ 的联合矩之间接近程度的上界，我们可以在 $\alpha = 1$ 下使用引理 ***B.1***

$$\left|\mathbb{E}\Big[\prod_{i=1}^{p} X_{a_i}(t_i)\Big] - \mathbb{E}\Big[\prod_{i=1}^{p} Y_{a_i}(t_i)\Big]\right| \le \Big\|\prod_{i=1}^{p} X_{a_i}(t_i) - \prod_{i=1}^{p} Y_{a_i}(t_i)\Big\|_1 \le \tilde{K}_p^{(1)}.$$

~(2.12) 的第二项通过 $\Sigma_{ab}(k)$ 依赖于 $\{\boldsymbol{Y}(t)\}$ 。下文的引理 B.3 量化了将根据 $\{\boldsymbol{Y}(t)\}$ 定义的 $\tilde{\Sigma}_{ab}(k)$ 替换为根据 $\{\boldsymbol{X}(t)\}$ 定义的 $n\mathrm{var}\big(\hat{}\gamma_{ab}(k)\big)$ 所产生的误差。引理 B.3 的证明见 F.3 节。

引理 B.3. 设假设 *2.2* 成立且 $q = 4$，则

$$|\tilde{\Sigma}_{ab}(k) - \Sigma_{ab}(k)| \le \Big|n\mathrm{var}\big(\hat{\gamma}_{ab}(k)\big) - \Sigma_{ab}(k)\Big| + 2\frac{(n-k)^2}{n}\Big(\max_{i=a,b} \tilde{D}_i^{(4)}\Big)\tilde{F},$$

其中

$$\begin{aligned}\tilde{F} := {} & |\gamma_{ab}(k)|\big(\|X_a(0)\|_2 + \|X_b(0)\|_2 + \min_{i=a,b} \tilde{D}_i^{(2)}\big) \\ & + \frac{1}{2}\Big(\max_{i=a,b} \tilde{D}_i^{(2)}\Big)\big(\|X_a(0)\|_2 + \|X_b(0)\|_2 + \min_{i=a,b} \tilde{D}_i^{(2)}\big)^2 \\ & + (\|X_a(0)\|_4 + \tilde{D}_a^{(4)})(\|X_b(0)\|_4 + \tilde{D}_b^{(4)}) \\ & \quad \times \Big(\|X_a(0)\|_4 + \|X_b(0)\|_4 + \tilde{D}_a^{(4)} + \tilde{D}_b^{(4)}\Big).\end{aligned} \tag{B.32}$$

~ ~为了给 (2.12) 中通过 Q_t 和 $\Sigma_{ab}(k)$ 依赖于 $\{\boldsymbol{Y}(t)\}$ 的第四项求上界，我们将使用上文的引理 B.3，将 $\Sigma_{ab}(k)$ 替换为 $\Sigma_{ab}(k)$；并使用下文的引理 B.4，将根据 $\{\boldsymbol{Y}(t)\}$ 定义的 Q_t 替换为根据 $\{\boldsymbol{X}(t)\}$ 在 (2.15) 中定义的 Q_t。我们先陈述引理 B.4，并将有关 (2.12) 中第四项上界的详细讨论留到本节末尾。引理 B.4 的证明见 F.4 节。

~ ~**引理 B.4**。 在 $q = 6$ 的条件下，假设 *2.2* 成立。设 Q_t 和 $Z(t)$ 分别如式 (2.9) 中及其下方所定义，且设 Q_t 和 $Z(t)$ 分别如式 (2.15) 中及其下方所定义。则，

$$|Q_t - \tilde{Q}_t| \le \tilde{K}_2^{(3)}\Big(|A_t||B_t| + \frac{1}{2}|A_t|^2\Big)\tilde{C}_1 + \tilde{K}_2^{(2)}|A_t||B_t|\tilde{C}_2 + \tilde{K}_2^{(1)}|B_t|\tilde{C}_3,$$

其中 $\tilde{K}_2^{(\alpha)} := \tilde{K}_2^{(\alpha)}(\{a,b\})$，$\alpha = 1, 2, 3$，如引理 *B.1* 中所定义，以及

$$\begin{aligned}
\tilde{C}_1 &:= 6\big(2\|X_a(0)\|_6\|X_b(0)\|_6 + \tilde{K}_2^{(3)}\big)^2 + 2(\tilde{K}_2^{(3)})^2,\\
\tilde{C}_2 &:= 8\Big(2\|X_a(0)\|_4\|X_b(0)\|_4 + \tilde{K}_2^{(2)}\Big)\Big(\|X_a(0)\|_2\|X_b(0)\|_2 + \tilde{K}_2^{(1)}\Big),\\
\tilde{C}_3 &:= 2\Big|\sum_{j\in A_t}\mathbb{E}(Z(t)Z(j))\Big|.
\end{aligned}$$

备注 B.5. $|Q_t - Q_t|$ 和引理 *B.4* 中的上界都依赖于 m。该上界包含三项加项，每一项都根据其相对于m的变化特征分解为三个因子：

(i) $\tilde{K}_2^{(\alpha)}$，在典型应用中会随着 m 增大而减小（参见备注 *B.2(ii)*）；

(ii) 关于变量 $|A_t|$ 和 $|B_t|$的多项式，它们会随着 m 增大而增大；

(iii) $\tilde{C}_1$、$\tilde{C}_2$和 $\tilde{C}_3$，它们均为非负且通常有界：若$|\tilde{K}_2^{(\alpha)}|$ 有界，则 $\tilde{C}_1$ 和 $\tilde{C}_2$ 有界；若 $\{\boldsymbol{X}(t)\}$ 的四阶及以下累积量可求和，则 $\tilde{C}_3$ 有界；参见(A.28)中对第三项的讨论。

在应用中，当 $\tilde{K}_2^{(\alpha)} = o(1/m^2)$时，该上界将趋于零，且对 t一致成立。

利用引理 B.3 和 B.4，我们现在将 (2.12) 中的 $2(n\tilde{\Sigma}_{ab}(k))^{-3/2}\sum_{t=1}^{n-k}\tilde{Q}_t$ 上界为 $2(n\Sigma_{ab}(k))^{-3/2}\sum_{t=1}^{n-k}Q_t$，的倍数，其中最后一个表达式仅依赖于 $\{\boldsymbol{X}(t)\}$。我们有

$$\begin{aligned}
\frac{2n^{-3/2}}{\big(\tilde{\Sigma}_{ab}(k)\big)^{3/2}}\sum_{t=1}^{n-k}\tilde{Q}_t &\le \frac{2n^{-3/2}}{\big(\tilde{\Sigma}_{ab}(k)\big)^{3/2}}\sum_{t=1}^{n-k}Q_t + \frac{2n^{-3/2}}{\big(\tilde{\Sigma}_{ab}(k)\big)^{3/2}}\sum_{t=1}^{n-k}|Q_t - \tilde{Q}_t| \qquad \text{(B.33)}\\
&= \left(\frac{2n^{-3/2}}{(\Sigma_{ab}(k))^{3/2}}\sum_{t=1}^{n-k}Q_t + \frac{2n^{-3/2}}{(\Sigma_{ab}(k))^{3/2}}\sum_{t=1}^{n-k}|Q_t - \tilde{Q}_t|\right)\left(\frac{\Sigma_{ab}(k)}{\tilde{\Sigma}_{ab}(k)}\right)^{3/2}.
\end{aligned}$$

如果对于某个具体示例，$\Sigma_{ab}(k)$是可处理的，但 Q_t 不可处理，那么只需使用引理 B.4，此时可以使用 (B.33) 中的上界。如果 $\Sigma_{ab}(k)$不可处理，但 $\Sigma_{ab}(k)$可以处理，则我们还需使用引理 B.3，以求得足够大的 n 和 m，使得 $|\Sigma_{ab}(k) - \Sigma_{ab}(k)| \le C\cdot\Sigma_{ab}(k)$，其中某个固定的 $C\in(0,1)$，因为这意味着 $\big(\Sigma_{ab}(k)/\tilde{\Sigma}_{ab}(k)\big)^{3/2} \le 1/(1-C)^{3/2}$。最后，注意，我们可以采用第 2.4 节中描述的两种策略之一来界定 Q_t，从而用一个可计算的量来界定 Q_t。采用这些方法所需的唯一调整，是在计算中将 $\{\boldsymbol{Y}(t)\}$ 替换为 $\{\boldsymbol{X}(t)\}$。

C 非中心化平稳数据的上界

定理 2.3 中的上界针对的是 $\hat{}\gamma_{ab}(k)$，其定义见 (1.2)，并假设数据已经过中心化。然而，在实际应用中，我们通常会遇到 $\hat{}\gamma^*_{ab}(k)$，其定义见 (1.1)，其中针对数据未经中心化的情况，额外进行了经验中心化。在本节中，我们将严格描述在实际应用中通常采用经验均值而非总体均值进行中心化时，对自协方差和交叉协方差的有限样本分布所产生的影响。下文的引理 C.1 给出了 $\hat{}\gamma^*_{ab}(k)$ 和 $\hat{}\gamma_{ab}(k)$ 的分布之间 Wasserstein距离的一个上界，其中二者分别表示经过经验中心化和未经经验中心化的自协方差函数与交叉协方差函数。该引理可用于推导 $d_{\mathrm{W}}(\mathcal{L}(\sqrt{n}(\hat{}\gamma^*_{ab}(k) - \gamma_{ab}(k))) \ \mathcal{N}(0, \Sigma_{ab}(k)))$ 的上界，其中数据未经中心化；具体说明见注 C.2(i)。

引理 C.1. 假设 $\{\boldsymbol{X}(t)\}$ 是弱平稳序列，且 $\mathbb{E}\boldsymbol{X}(t) = 0$ 和 $\sum_{u=-\infty}^{\infty} |u||\gamma_{jj}(u)| < \infty$ 对 $j = a, b$ 成立。那么，

$$
\begin{aligned}
&d_{\mathrm{W}}\big(\mathcal{L}\left(n^{1/2}(\hat{\gamma}^*_{ab}(k) - \gamma_{ab}(k))\right), \mathcal{L}\left(n^{1/2}(\hat{\gamma}_{ab}(k) - \gamma_{ab}(k))\right)\big) \\
&\le n^{-1/2}\Big(\sum_{u=-\infty}^{\infty} |\gamma_{aa}(u)|\Big)^{1/2}\Big(\sum_{u=-\infty}^{\infty} |\gamma_{bb}(u)|\Big)^{1/2} \\
&\quad + \frac{|k|}{n^{3/2}}\left(\sum_{u=-\infty}^{\infty} |\gamma_{aa}(u)| + \frac{1}{n}\sum_{u=-\infty}^{\infty} |u||\gamma_{aa}(u)|\right)^{1/2} \\
&\qquad \times \left(\sum_{u=-\infty}^{\infty} |\gamma_{bb}(u)| + \frac{1}{n}\sum_{u=-\infty}^{\infty} |u||\gamma_{bb}(u)|\right)^{1/2}
\end{aligned}
$$

证明. 令 $h \in \mathcal{H}$ 如 (1.6) 中所定义。由于 $\|h\|_{\mathrm{Lip}} \le 1$，我们有

$$
\begin{aligned}
&\left|h\left(n^{1/2}(\hat{\gamma}^*_{ab}(k) - \gamma_{ab}(k))\right) - h\left(n^{1/2}(\hat{\gamma}_{ab}(k) - \gamma_{ab}(k))\right)\right| \le n^{1/2}|\hat{\gamma}^*_{ab}(k) - \hat{\gamma}_{ab}(k)| \\
&= \frac{n-k}{n^{1/2}}\left|\bar{X}_{a,1:n}\bar{X}_{b,1:n} - \bar{X}_{a,1:n}\bar{X}_{b,1:n-k} - \bar{X}_{a,k+1:n}\bar{X}_{b,1:n}\right| \\
&= \frac{n-k}{n^{1/2}}\left|(\bar{X}_{a,1:n} - \bar{X}_{a,k+1:n})(\bar{X}_{b,1:n} - \bar{X}_{b,1:n-k}) - \bar{X}_{a,k+1:n}\bar{X}_{b,1:n-k}\right|.
\end{aligned}
$$

其中 $\bar{X}_{a,u:v} := \frac{1}{v-u+1}\sum_{t=u}^{v} X_a(t)$，类似地，将 a 替换为 b。根据前面的

等式可得，对于 $k=1,\dots,n-1$，有

$$
\begin{aligned}
&d_{\mathrm{W}}\big(\mathcal{L}\left(n^{1/2}(\hat{\gamma}^*_{ab}(k)-\gamma_{ab}(k))\right),\mathcal{L}\left(n^{1/2}(\hat{\gamma}_{ab}(k)-\gamma_{ab}(k))\right)\big)\\
&\le \frac{n-k}{n^{1/2}}\Big(\|\bar{X}_{a,1:n}-\bar{X}_{a,k+1:n}\|_2\|\bar{X}_{b,1:n}-\bar{X}_{b,1:n-k}\|_2+\|\bar{X}_{a,k+1:n}\|_2\|\bar{X}_{b,1:n-k}\|_2\Big)\\
&\le \frac{n-k}{n^{1/2}}\Big(\frac{k}{n}\Big)^2\Bigg(\frac{n}{k(n-k)}\sum_{u=-\infty}^{\infty}|\gamma_{aa}(u)|+\frac{1}{k(n-k)}\sum_{u=-\infty}^{\infty}|u||\gamma_{aa}(h)|\Bigg)^{1/2}\\
&\qquad\qquad\times\Bigg(\frac{n}{k(n-k)}\sum_{u=-\infty}^{\infty}|\gamma_{bb}(u)|+\frac{1}{k(n-k)}\sum_{u=-\infty}^{\infty}|u||\gamma_{bb}(u)|\Bigg)^{1/2}\\
&\quad+\frac{n-k}{n^{1/2}}\frac{1}{n-k}\Big(\sum_{u=-\infty}^{\infty}|\gamma_{aa}(u)|\Big)^{1/2}\Big(\sum_{u=-\infty}^{\infty}|\gamma_{bb}(u)|\Big)^{1/2}\\
&=\frac{k}{n^{3/2}}\Bigg(\sum_{u=-\infty}^{\infty}|\gamma_{aa}(u)|+\frac{1}{n}\sum_{u=-\infty}^{\infty}|u||\gamma_{aa}(u)|\Bigg)^{1/2}\\
&\qquad\qquad\times\Bigg(\sum_{u=-\infty}^{\infty}|\gamma_{bb}(u)|+\frac{1}{n}\sum_{u=-\infty}^{\infty}|u||\gamma_{bb}(u)|\Bigg)^{1/2}\\
&\qquad+\frac{1}{n^{1/2}}\Big(\sum_{u=-\infty}^{\infty}|\gamma_{aa}(u)|\Big)^{1/2}\Big(\sum_{u=-\infty}^{\infty}|\gamma_{bb}(u)|\Big)^{1/2}
\end{aligned}
$$

其中我们用到了

$$
\begin{aligned}
\mathbb{E}\bar{X}^2_{j,u:v}&=\frac{1}{(v-u+1)^2}\sum_{s=u}^{v}\sum_{t=u}^{v}\mathbb{E}(X_j(s-t)X_j(0))\\
&=\frac{1}{(v-u+1)^2}\sum_{\ell=-(v-u+1)}^{v-u+1}(v-u+1-|\ell|)\gamma_{jj}(\ell)\\
&=\frac{1}{v-u+1}\sum_{\ell=-(v-u+1)}^{v-u+1}\Big(1-\frac{|\ell|}{v-u+1}\Big)\gamma_{jj}(\ell)\\
&\le\frac{1}{v-u+1}\sum_{\ell=-\infty}^{\infty}|\gamma_{jj}(\ell)|
\end{aligned}
$$

以及对于 $k=1$，……，$n-1$，这一事实，

$$\begin{aligned}
\|\bar{X}_{a,1:n}-\bar{X}_{a,k+1:n}\|_2^2 &= \mathbb{E}\Big(\frac{1}{n}\sum_{i=1}^{n}X_a(i)-\frac{1}{n-k}\sum_{i=k+1}^{n}X_a(i)\Big)^2\\
&= \mathbb{E}\Big(\frac{k}{n}\frac{1}{n-k}\sum_{i=k+1}^{n}X_a(i)-\frac{k}{n}\frac{1}{k}\sum_{i=1}^{k}X_a(i)\Big)^2\\
&= \Big(\frac{k}{n}\Big)^2\Big(\|\bar{X}_{a,k+1:n}\|_2^2+\|\bar{X}_{a,1:k}\|_2^2-2\mathrm{cov}(\bar{X}_{a,k+1:n},\bar{X}_{a,1:k})\Big)\\
&\le \Big(\frac{k}{n}\Big)^2\Bigg(\Big(\frac{1}{n-k}+\frac{1}{k}\Big)\Big(\sum_{u=-\infty}^{\infty}|\gamma_{aa}(u)|\Big)+2\frac{1}{k(n-k)}\sum_{u=1}^{\infty}u|\gamma_{aa}(u)|\Bigg)\\
&\le \Big(\frac{k}{n}\Big)^2\Bigg(\frac{n}{k(n-k)}\sum_{u=-\infty}^{\infty}|\gamma_{aa}(u)|+\frac{1}{k(n-k)}\sum_{u=-\infty}^{\infty}|u||\gamma_{aa}(u)|\Bigg).
\end{aligned}$$

同理，我们推导出

$$\begin{aligned}
\|\bar{X}_{b,1:n}-\bar{X}_{b,1:n-k}\|_2^2 &= \Big(\frac{k}{n}\Big)^2\|\bar{X}_{b,n-k+1:n}-\bar{X}_{b,1:n-k}\|_2^2\\
&\le \Big(\frac{k}{n}\Big)^2\Bigg(\frac{n}{k(n-k)}\sum_{u=-\infty}^{\infty}|\gamma_{bb}(u)|+\frac{1}{k(n-k)}\sum_{u=-\infty}^{\infty}|u||\gamma_{bb}(u)|\Bigg).
\end{aligned}$$

□

注 C.2。 *(i)* 引理 *C.1* 有助于在非中心平稳时间序列的情形下应用定理 *2.3*，例如 $\boldsymbol{N}(t)$。如下所示。定义 $\boldsymbol{X}(t):=\boldsymbol{N}(t)-\mathbb{E}(\boldsymbol{N}(t))^{-}$ $^{-}$并注意到 $X_j(t)-X_j=N_j(t)-N_j$。因此，

$$\hat{c}_{ab}^{*}(k):=\frac{1}{n}\sum_{t=1}^{n-k}(N_a(t+k)-\bar{N}_a)(N_b(t)-\bar{N}_b)=\hat{\gamma}_{ab}^{*}(k)=\frac{1}{n}\sum_{t=1}^{n-k}(X_a(t+k)-\bar{X}_a)(X_b(t)-\bar{X}_b),$$

其中 $\hat{}\gamma_{ab}^{*}(k)$ 是用中心化过程 $\boldsymbol{X}(t)$ 定义的，因此我们可以对其应用所得结果。由三角不等式即可得到 $d_{\mathrm{W}}\left(\mathcal{L}\left(\sqrt{n}\left(\hat{}c_{ab}^{*}(k)\right)-\gamma_{ab}(k)\right)\right)\ \mathcal{N}(0,\Sigma_{ab}(k))$的一个上界。

$$\begin{aligned}
&d_{\mathrm{W}}\left(\mathcal{L}\left(\sqrt{n}\left(\hat{c}_{ab}^{*}(k)-\gamma_{ab}(k)\right)\right),\mathcal{N}\left(0,\Sigma_{ab}(k)\right)\right)\\
&\le d_{\mathrm{W}}\left(\mathcal{L}\left(\sqrt{n}(\hat{\gamma}_{ab}^{*}(k)-\gamma_{ab}(k))\right),\mathcal{L}\left(\sqrt{n}(\hat{\gamma}_{ab}(k)-\gamma_{ab}(k))\right)\right) \qquad \text{(C.34)}\\
&\quad+d_{\mathrm{W}}\left(\mathcal{L}\left(\sqrt{n}\left(\hat{\gamma}_{ab}(k)-\gamma_{ab}(k)\right)\right),\mathcal{N}\left(0,\Sigma_{ab}(k)\right)\right). \qquad \text{(C.35)}
\end{aligned}$$

分别将引理 *C.1* 和定理 *2.3* 的结果应用于（C.34）和（C.35）中的各项，

从而得到所需的上界。
(ii) 引理 *C.1* 断言，存在一个阶为 $\mathcal{O}(n^{-1/2})$ 的上界，且该上界关于 k一致。*(iii)* 注意以下有趣的特殊情形：*(a)* 若 $k = 0$，则上界简化为

$$n^{-1/2}\Big(\sum_{u=-\infty}^{\infty} |\gamma_{aa}(u)| \sum_{u=-\infty}^{\infty} |\gamma_{bb}(u)|\Big)^{1/2}.$$

(b) 若数据在各分量之间不相关（即 $\gamma_{aa}(u) = \gamma_{bb}(u) = 0$ 对于 $u \neq 0$），则上界简化为

$$\Big(n^{-1/2} + k\, n^{-3/2}\Big)\Big(\gamma_{aa}(0)\gamma_{bb}(0)\Big)^{1/2}.$$

D 因果 AR(1) 上界的计算细节

这里，我们给出第 3.1 节所讨论的 AR(1) 示例中上界的推导细节。定理 2.3 中的上界为

$$\frac{k}{\sqrt{n}}|\gamma_{ab}(k)| + \frac{\sqrt{2}}{\sqrt{\pi\Sigma_{ab}(k)}}\left|\Sigma_{ab}(k) - \tilde{\Sigma}_{ab}(k)\right| + \frac{2(n-k)}{\sqrt{n}}\tilde{K} + \frac{2n^{-3/2}}{\Big(\tilde{\Sigma}_{ab}(k)\Big)^{3/2}}\sum_{t=1}^{n-k}\tilde{Q}_t.$$

由于 我们处于单变量时间序列的情形，因此省略下标 a 和 b .

在本 节中，我们讨论单变量（即 $a = b$）AR(1) 时间序列的情形

$$X(t) = \alpha X(t-1) + \varepsilon_t = \sum_{j=0}^{\infty} \alpha^j \varepsilon_{t-j},$$

其参数为 $\alpha \in (-1,\ 1)$，且 $\{\varepsilon_t\}$ 独立同分布，满足 $\kappa_p := \mathrm{cum}_p(\varepsilon_t)$ 对于 $p = 8$ 以及 $\kappa_p = 0$ 在 p 为奇数时存在。我们假设 $\kappa_1 = \mathbb{E}\varepsilon_t = 0$ ，并使用以下移动平均过程作为 m-依赖近似

$$Y(t) = \sum_{j=0}^{m} \alpha^j \varepsilon_{t-j}.$$

我们首先处理上界中的第一项和第三项，然后处理第二项和第四项。对于第一项和第三项，只需计算

- $\tilde{D}^{(q)} := \left[\mathbb{E}\left(X(0) - Y(0)\right)^q\right]^{1/q}$, for $q = 2, 4, 6, 8$, and

- $\mathbb{E}(X(t+k)X(t))^q$，对于 $q=1$，2，3，4。

注意， $\|X(t)\|_{2q}$ 由 $\mathbb{E}(X(t+k)X(t))^q$ 可得，其中 $k=0$

对于 $\tilde{D}^{(q)}$，利用 Brillinger (1975) 中关于乘积的累积量的定理 2.3.2，我们得到

$$\begin{aligned}
&\mathbb{E}\left(X(0)-Y(0)\right)^q = \mathbb{E}\left(\sum_{j=m+1}^{\infty}\alpha^j\varepsilon(-j_\ell)\right)^q \\
&= \sum_{j_1=m+1}^{\infty}\cdots\sum_{j_q=m+1}^{\infty}\alpha^{j_1+\cdots+j_q}\sum_{\{\nu_1,\ldots,\nu_R\}}\prod_{r=1}^{R}\mathrm{cum}(\varepsilon(-j_\ell):\ell\in\nu_r) \\
&= \sum_{\{\nu_1,\ldots,\nu_R\}}\prod_{r=1}^{R}\sum_{j=m+1}^{\infty}\alpha^{|\nu_r|j}\kappa_{|\nu_r|} = \sum_{\{\nu_1,\ldots,\nu_R\}}\prod_{r=1}^{R}\kappa_{|\nu_r|}\alpha^{|\nu_r|(m+1)}\sum_{j=0}^{\infty}\alpha^{|\nu_r|j} \\
&= \alpha^{q(m+1)}\sum_{\{\nu_1,\ldots,\nu_R\}}\prod_{r=1}^{R}\frac{\kappa_{|\nu_r|}}{1-\alpha^{|\nu_r|}}
\end{aligned}$$

其中，求和是针对 $\{1,\ldots,q\}$的划分 $\{\nu_1,\ldots,\nu_R\}$ 进行的。特别地，对于$q=2$，4，8，并且由于所考虑分布的奇数阶累积量均为零，我们有

$$\begin{aligned}
\tilde{D}^{(2)} &= \alpha^{m+1}\Big(\frac{\kappa_2}{1-\alpha^2}\Big)^{1/2}, \\
\tilde{D}^{(4)} &= \alpha^{m+1}\Big(\frac{\kappa_4}{1-\alpha^4}+3\Big(\frac{\kappa_2}{1-\alpha^2}\Big)^2\Big)^{1/4}, \\
\tilde{D}^{(8)} &= \alpha^{m+1}\Big(\frac{\kappa_8}{1-\alpha^8}+28\frac{\kappa_6}{1-\alpha^6}\frac{\kappa_2}{1-\alpha^2}+35\Big(\frac{\kappa_4}{1-\alpha^4}\Big)^2 \\
&\qquad +210\frac{\kappa_4}{1-\alpha^4}\Big(\frac{\kappa_2}{1-\alpha^2}\Big)^2+105\Big(\frac{\kappa_2}{1-\alpha^2}\Big)^4\Big)^{1/8}
\end{aligned}$$

类似地，我们得到，对于 $\alpha\neq 0$，有

$$\mathbb{E}\left(X(t+k)X(t)\right)^q = \alpha^{q|k|}\sum_{\{\nu_1,\ldots,\nu_R\}}\prod_{r=1}^{R}\frac{\kappa_{|\nu_r|}}{1-\alpha^{|\nu_r|}},$$

其中求和是关于 $\{1,\ldots,2q\}$ 的划分 $\{\nu_1,\ldots,\nu_R\}$ 进行的。

特别地，对于 $q = 1,2,3,4$，我们有，

$$\begin{aligned}
\mathbb{E}\left(X(t+k)X(t)\right) &= \alpha^{|k|}\frac{\kappa_2}{1-\alpha^2} =: \gamma(k) \\
\mathbb{E}\left(X(t+k)X(t)\right)^2 &= \alpha^{2|k|}\Big(\frac{\kappa_4}{1-\alpha^4} + 3\Big(\frac{\kappa_2}{1-\alpha^2}\Big)^2\Big) \\
\mathbb{E}\left(X(t+k)X(t)\right)^3 &= \alpha^{3|k|}\Big(\frac{\kappa_6}{1-\alpha^6} + 15\frac{\kappa_4}{1-\alpha^4}\frac{\kappa_2}{1-\alpha^2} + 30\Big(\frac{\kappa_2}{1-a^2}\Big)^3\Big) \\
\mathbb{E}\left(X(t+k)X(t)\right)^4 &= \alpha^{4|k|}\Big(\frac{\kappa_8}{1-\alpha^8} + 28\frac{\kappa_6}{1-\alpha^6}\frac{\kappa_2}{1-\alpha^2} + 35\Big(\frac{\kappa_4}{1-\alpha^4}\Big)^2 \\
&\qquad + 210\frac{\kappa_4}{1-\alpha^4}\Big(\frac{\kappa_2}{1-\alpha^2}\Big)^2 + 105\Big(\frac{\kappa_2}{1-\alpha^2}\Big)^4\Big)
\end{aligned} \tag{D.36}$$

此外，对于 $\alpha = 0$，我们有

$$\mathbb{E}\left(X(t+k)X(t)\right)^q = \begin{cases} \mathbb{E}\varepsilon_t^{2q} & k = 0 \\ \left(\mathbb{E}\varepsilon_t^q\right)^2 & k \neq 0. \end{cases}$$

因此，对于 $\alpha = 0$ 和 $k = 0$ 的情形，若采用 $0^0 = 1$ 这一约定，则 (D.36) 中的表达式是正确的。对于 $\alpha = 0$ 和 $k \neq 0$，我们有

$$\begin{aligned}
\mathbb{E}\left(X(t+k)X(t)\right) &= 0 \\
\mathbb{E}\left(X(t+k)X(t)\right)^2 &= (\kappa_2)^2 \\
\mathbb{E}\left(X(t+k)X(t)\right)^3 &= 0 \\
\mathbb{E}\left(X(t+k)X(t)\right)^4 &= \Big(\kappa_4 + 3(\kappa_2)^2\Big)^2
\end{aligned}$$

这涵盖了上界中第一项和第三项的相关部分，接下来我们将关注第二项和第四项。为此，我们首先讨论如何计算本例中 m 依赖近似的联合累积量，然后再转向上界的第二项。

的联合累积量 $\{Y(t)\}$。为了计算上界，我们将经常需要 $\{Y(t)\}$的联合累积量。对于 u_1，**...** **...**，$u_p \in \mathbb{Z}$，记 $M := \min\{u_1, \ldots, u_p\}$, $R := \max\{u_1, \ldots, u_p\} - M$,

以及 $S:=\sum_{i=1}^{p}(u_i-M)$。于是，我们有

$$\begin{aligned}
&\operatorname{cum}(Y(u_1),\ldots,Y(u_p))\\
&=\sum_{j_1=0}^{m}\cdots\sum_{j_p=0}^{m}\alpha^{j_1+\ldots+j_p}\operatorname{cum}(\varepsilon(u_1-M-j_1),\ldots,\varepsilon(u_p-M-j_p))\\
&=\begin{cases}\sum\limits_{j_s=0}^{m-R}\alpha^{pj_s+S}\operatorname{cum}(\varepsilon(-j_s),\ldots,\varepsilon(-j_s))=\kappa_p\alpha^S\dfrac{1-\alpha^{p(m-R+1)}}{1-\alpha^p} & R\le m\\ 0 & R>m.\end{cases}
\end{aligned}\tag{D.37}$$

对于第二个等式，注意到 $u_\ell-M\ge 0$ 对所有 $\ell=1$、.....、p 均成立，并且 $u_s-M=0$ 对至少一个 $s=1$、.....、p 成立。现在我们来说明，对于满足 $(j_1$、.....、$j_p)\in\{1,\ldots,m\}^p$ 且 j_s 的 $\varepsilon_{u_1-M-j_1}$、.....、$\varepsilon_{u_p-M-j_p}$ 的累积量 $\operatorname{cum}(\varepsilon_{u_1-M-j_1}$、.....、$\varepsilon_{u_p-M-j_p})$，至多只有一个非零。为此，注意到对所有 $\ell\in 1$、.....、p，包括 $\ell=s$，都必须有 $u_s-M-j_s=-j_s=u_\ell-M-j_\ell$，这意味着对于不为零的累积量，$j_\ell=j_s+u_\ell-M\in\{1,\ldots,m\}$。因此，为确保对所有 ℓ 都有 $j_\ell\le m$，必须排除 $j_s>m-R$。此外，我们有 $u_{\ell_1}-j_{\ell_1}=u_{\ell_2}-j_{\ell_2}$，这意味着必须要求

$$\max_{\ell_1,\ell_2\in\{1,\ldots,p\}}|j_{\ell_1}-j_{\ell_2}|=\max_{\ell_1,\ell_2\in\{1,\ldots,p\}}|u_{\ell_1}-u_{\ell_2}|=R\le m.$$

可以采用一种非常相似但更简单的论证来获得

$$\operatorname{cum}(X(u_1),\ldots,X(u_p))=\kappa_p\frac{\alpha^S}{1-\alpha^p}.$$

$\tilde{\Sigma}_{ab}(k)$ 和 $\Sigma_{ab}(k)$的计算。为此，我们注意到

$$\begin{aligned}
\Sigma_{ab}(k)&=\sum_{u=-\infty}^{\infty}\gamma^2(u)+\sum_{u=-\infty}^{\infty}\gamma(u-k)\gamma(u+k)+\frac{\kappa_4}{(\kappa_2)^2}\gamma^2(k)\\
&=(\kappa_2)^2\frac{1+\alpha^2+\alpha^{2|k|}\big(1+\alpha^2+2k(1-\alpha^2)\big)}{(1-\alpha^2)^3}+\kappa_4\frac{\alpha^{2|k|}}{(1-\alpha^2)^2}.
\end{aligned}$$

此外，我们有

$$\tilde{\Sigma}_{ab}(k) = n^{-1}\mathrm{var}\Big(\sum_{t=1}^{n-k} Y(t+k)Y(t)\Big)$$
$$= \frac{1}{n}\sum_{t_1=1}^{n-k}\sum_{t_2=1}^{n-k}\tilde{C}(t_1-t_2) = \frac{n-k}{n}\sum_{|u|\le\min\{n-k,m\}}\Big(1-\frac{|u|}{n-k}\Big)\tilde{C}(u),$$

~ ~，其中 $C(u)$ 定义见 (2.7)。在 $C(u)$ 的定义中出现的 $\{Y(t)\}$ 的累积量，可以按照上一节中的 (D.37) 所述进行计算。

接下来讨论上界中的第四项。回顾一下，在第 2.3 节中，我们讨论了~计算 Q_t 上界的两种方法。在本节中，我们采用第二种方法，并讨论如何计算 (2.14) 右端的精确值。

~ Q_t**的上界计算**。我们有

$$\tilde{Q}_t \le \mathrm{var}\Big(\tilde{Z}(t)\sum_{j\in A_t}\tilde{Z}(j)\Big)^{1/2}\mathrm{var}\Big(\sum_{j\in B_t}\tilde{Z}(j)\Big)^{1/2} + \frac{1}{2}\Big[\mathbb{E}\Big(\tilde{Z}(t)\sum_{j\in A_t}\tilde{Z}(j)\Big)^2\Big]^{1/2}\mathrm{var}\Big(\sum_{j\in A_t}\tilde{Z}(j)\Big)^{1/2}$$
$$= \Big(M_{1,t}+\tilde{C}(0)\times M_{2a,t}+M_{3,t}^2\Big)^{1/2}M_{2b,t}^{1/2} + \frac{1}{2}\Big(M_{1,t}+\tilde{C}(0)\times M_{2a,t}+2M_{3,t}^2\Big)^{1/2}M_{2a,t}^{1/2},$$

其中

$$M_{1,t} := \sum_{j_1\in A_t}\sum_{j_2\in A_t}\sum_{\nu}\prod_{r=1}^{R}\mathrm{cum}(\tilde{Y}_{ij}:(ij)\in\nu_r) = \sum_{\ell_1\in A_t-t}\sum_{\ell_2\in A_t-t}\tilde{D}(\ell_1,\ell_2),$$
$$M_{2a,t} := |A_t|\sum_{|u|\le|A_t|-1}\Big(1-\frac{|u|}{|A_t|}\Big)\tilde{C}(u), \quad M_{2b,t} := |B_t|\sum_{|u|\le|B_t|-1}\Big(1-\frac{|u|}{|B_t|}\Big)\tilde{C}(u),$$
$$M_{3,t} := \sum_{u\in A_t}\tilde{C}(u-t), \qquad \tilde{D}(u_1,u_2) := \sum_{\nu}\prod_{r=1}^{R}\mathrm{cum}(\tilde{Y}_{ij}:(ij)\in\nu_r)$$

~其中，$M_{1,t}$ 和 $D(u_1,u_2)$ 的定义中所涉及的求和，是针对表 3(a) 中的不可分解~划分 ν 进行的，而项 Y_{ij} 的定义见表 3(b)。对于 $M_{1,t}$ 的定义，我们采用原始~索引；对于 $D(u_1,u_2)$ 的定义，则采用偏移了 t 的索引。~更确切地说，在以下定义中，我们在 $M_{1,t}$ 中采用第一个，在 $D(u_1,u_2)$ 中采用第二个：

$$\tilde{Y}_{1,1} := Y(t+k) \stackrel{d}{=} Y(k) \qquad \tilde{Y}_{1,2} := Y(t) \stackrel{d}{=} Y(0)$$
$$\tilde{Y}_{2,1} := Y(t+k) \stackrel{d}{=} Y(k) \qquad \tilde{Y}_{2,2} := Y(t) \stackrel{d}{=} Y(0)$$
$$\tilde{Y}_{3,1} := Y(j_1+k) \stackrel{d}{=} Y(u_1+k) \quad \tilde{Y}_{3,2} := Y(j_1) \stackrel{d}{=} Y(u_1)$$
$$\tilde{Y}_{4,1} := Y(j_2+k) \stackrel{d}{=} Y(u_2+k) \quad \tilde{Y}_{4,2} := Y(j_2) \stackrel{d}{=} Y(u_2)$$

直接计算 ˜算 Q_t 的成本非常高。因此，为了提高计算效率，我们注意到 t

- $\tilde{C}(u) = \tilde{C}(-u)$，
- $\tilde{D}(u_1, u_2) = D(u_2, u_1)$,
- $M_{1,t}$, $M_{2a,t}$和 $M_{3,t}$在 $t = m+k+1, \ldots\ldots, n-m-2k$上为常数，且
- $M_{2b,t}$在 $t = 2(m+k)+1, \ldots\ldots, n-2m-3k$上为常数。

˜ ˜此外，加项 $C(u)$ 和 $D(\ell_1, \ell_2)$ 不依赖于 t。因此，对于给定的 a ˜ ˜、 m 和 k，我们首先计算 $C(u)$，其中 $u = 0, \ldots\ldots, 2(m+k)$，以及 $D(u_1, u_2)$，其中 $u_1, u_2 = -(m+k), \ldots\ldots, m+k$ 满足 $u_1 \le u_2$。然后，可以相当高效地计算上界。

E 模拟 Wasserstein距离的详细信息

我们通过模拟得到 Wasserstein距离，以便与我们的上界进行比较。将 $\sqrt{n}(\hat{\gamma}_{ab}(k) - \gamma_{ab}(k))$ 的累积分布函数记为 F，将 $N \sim \mathcal{N}(0, \Sigma_{ab}(k))$ 的累积分布函数记为 G，则有

$$W := d_{\mathrm{W}}\left(\mathcal{L}\left(\sqrt{n}\left(\hat{\gamma}_{ab}(k) - \gamma_{ab}(k)\right)\right), \mathcal{N}\left(0, \Sigma_{ab}(k)\right)\right) = \int_0^1 |F^{-1}(u) - G^{-1}(u)| \mathrm{d}u, \qquad \text{(E.38)}$$

其中， $F^{-1}(u) := \inf\{x \in \mathbb{R} : F(x) \ge u\}$ 是 F的分位数函数；类似地， G^{-1} 是 G的分位数函数。我们通过对 i. i.d. $Z_1, \ldots, Z_R \sim F$进行抽样，然后将 (E.38) 中的 F替换为经验分布 $\mathbb{F}(x) := \frac{1}{R}\sum_{r=1}^{R} I\{Z_r \le x\}$. 更确切地说， −

$$\begin{aligned}(\mathrm{E.38}) &\approx \int_0^1 |\mathbb{F}^{-1}(u) - G^{-1}(u)| \mathrm{d}u = \sum_{r=1}^{R} \int_{(i-1)/R}^{i/R} |Z_{(r)} - G^{-1}(u)| \mathrm{d}u \\ &\approx \frac{1}{R}\sum_{r=1}^{R} \left| Z_{(r)} - \sqrt{\Sigma_{ab}(k)}\Phi^{-1}\Big(\frac{2i-1}{2R}\Big)\right| =: \hat{W}_i.\end{aligned}$$

如果 $F \neq G$ （并且满足额外的正则性条件），那么第一次近似的误差阶为 $O_P(R^{-1/2})$；例如参见 Munk 和 Czado（1998）。第二次近似的误差

阶为 $O(R^{-1})$。由于我们考虑的 n 数量级为数千，因此选择 $R = 4\times10^6$。请注意，我们选择了 R 作为 2000^2，而 $n = 2000$ 是最大样本长度。此外，为了降低（伪）随机 W_i 的方差，我们对其模拟独立副本（i =1，.....， $B = 50$），然后报告平均值。

F 更多技术性证明

F.1 引理4.2的证明

对于 $N \sim \mathcal{N}(0,\ \sigma^2)$ 和 $h \in \mathcal{H}_{\mathrm{W}}$，首先令 $f = f_h$ 为 Stein方程的解

$$\sigma^2 f'(w) - wf(w) = h(w) - \mathbb{E}[h(N)] \tag{F.39}$$

对应于 $\mathcal{N}(0,\ \sigma^2)$ 分布。由于 $\forall i \in J$、 ξ_i 和 $W - \xi(A_i)$（其中 $\xi(A)$ 如 (4.18) 所示）相互独立，且 $\mathbb{E}(\xi_i) = 0$，因此我们有

$$\mathbb{E}\left(Wf(W)\right) = \sum_{i\in J}\mathbb{E}\left(\xi_i f(W)\right) = \sum_{i\in J}\mathbb{E}\left(\xi_i\left(f(W) - f(W-\xi(A_i))\right)\right).$$

现在加上并减去量 $\sum_{i\in J}\mathbb{E}\left(\xi_i\xi(A_i)f'(W)\right)$，得到

$$\mathbb{E}\left(Wf(W)\right) = \sum_{i\in J}\mathbb{E}\left(\xi_i\left(f(W) - f(W-\xi(A_i)) - \xi(A_i)f'(W)\right)\right) + \mathbb{E}\left(\sum_{i\in J}\xi_i\xi(A_i)f'(W)\right).$$

对于 $\mathbb{E}(W^2) = \mathrm{var}(W) = \sigma^2$，我们现在有了

$$\sigma^2 = \mathbb{E}\left(W^2\right) = \mathbb{E}\left(\sum_{i\in J}\xi_i\sum_{j\in J}\xi_j\right) = \sum_{i\in J}\mathbb{E}\left(\xi_i\xi(A_i)\right)$$

因为 ξ_i 与不属于 A_i 的所有元素独立。现在使用（F.39）中的 Stein方程，我们有

$$
\begin{aligned}
&\mathbb{E}\left(\sigma^2 f'(W) - W f(W)\right) \\
&= \mathbb{E}\left(\sum_{i\in J} \mathbb{E}\left(\xi_i \xi(A_i)\right) f'(W)\right) \\
&\quad - \sum_{i\in J} \mathbb{E}\left(\xi_i \left(f(W) - f(W-\xi(A_i)) - \xi(A_i) f'(W)\right)\right) - \mathbb{E}\left(\sum_{i\in J} \xi_i \xi(A_i) f'(W)\right) \\
&= -\sum_{i\in J} \mathbb{E}\left(\left(\xi_i \xi(A_i) - \mathbb{E}(\xi_i \xi(A_i))\right) f'(W)\right) - \sum_{i\in J} \mathbb{E}\left(\xi_i \left(f(W) - f(W-\xi(A_i)) - \xi(A_i) f'(W)\right)\right) \\
&= -\sum_{i\in J} \mathbb{E}\left(\left(\xi_i \xi(A_i) - \mathbb{E}(\xi_i \xi(A_i))\right) \left(f'(W) - f'(W-\xi(B_i))\right)\right) \\
&\quad - \sum_{i\in J} \mathbb{E}\left(\xi_i \left(f(W) - f(W-\xi(A_i)) - \xi(A_i) f'(W)\right)\right) \qquad \text{(F.40)}
\end{aligned}
$$

最后一个等式成立，是因为 ξ_i 和 $\xi(A_i)$ 均与 $W - \xi(B_i)$ 独立，其中 $\xi(A)$ 的符号对任意 $A \subset J$ 均在 (4.18) 中给出。因此，

$$
\sum_{i\in J} \mathbb{E}\left(\left(\xi_i \xi(A_i) - \mathbb{E}(\xi_i \xi(A_i))\right) f'(W - \xi(B_i))\right) = \sum_{i\in J} \mathbb{E}\left(\xi_i \xi(A_i) - \mathbb{E}(\xi_i \xi(A_i))\right) \mathbb{E}\left(f'(W - \xi(B_i))\right) = 0.
$$

利用一阶泰勒展开，

$$
f'(W) - f'(W - \xi(B_i)) = \xi(B_i) f''(W^*), \qquad \text{(F.41)}
$$

当 W^* 介于 W 与 $W - \xi$ 之间时（B_i）。此外，现在对 $f(W - \xi(A_i))$ 在 W 处作二阶泰勒展开，可得

$$
f(W) - f(W - \xi(A_i)) - \xi(A_i) f'(W) = -\frac{(\xi(A_i))^2}{2} f''(\tilde{W}), \qquad \text{(F.42)}
$$

˜其中 W 介于 W 与 $W-\xi$ 之间（A_i）。我们用 $\|\cdot\|$ 表示函数的上确界范数，将（F.41）和（F.42）的结果应用于（F.40），得到

$$
\begin{aligned}
&|\mathbb{E}\left[h(W)\right]-\mathbb{E}\left[h(N)\right]| = \left|\mathbb{E}\left(\sigma^2 f'(W)-Wf(W)\right)\right| \\
&= \left|-\sum_{i\in J}\mathbb{E}\left(\left(\xi_i\xi(A_i)-\mathbb{E}(\xi_i\xi(A_i))\right)\xi(B_i)f''(W^*)\right)+\sum_{i\in J}\mathbb{E}\left(\xi_i\frac{(\xi(A_i))^2}{2}f''(\tilde{W})\right)\right| \\
&\le \sum_{i\in J}\mathbb{E}\left|\left(\xi_i\xi(A_i)-\mathbb{E}(\xi_i\xi(A_i))\right)\xi(B_i)f''(W^*)\right|+\sum_{i\in J}\mathbb{E}\left|\xi_i\frac{(\xi(A_i))^2}{2}f''(\tilde{W})\right| \\
&\le \|f''\|\sum_{i\in J}\mathbb{E}\left|\left(\xi_i\xi(A_i)-\mathbb{E}(\xi_i\xi(A_i))\right)\xi(B_i)\right|+\frac{1}{2}\|f''\|\sum_{i\in J}\mathbb{E}\left|\xi_i(\xi(A_i))^2\right| \\
&\le \|f''\|\sum_{i\in J}\left\{\mathbb{E}\left|\left(\xi_i\xi(A_i)-\mathbb{E}(\xi_i\xi(A_i))\right)\xi(B_i)\right|+\frac{1}{2}\mathbb{E}\left|\xi_i(\xi(A_i))^2\right|\right\}. \qquad \text{(F.43)}
\end{aligned}
$$

根据 Chen 等人（2011）的第 2.2 节，我们知道 (F.39) 中 Stein方程的解满足 $\|f''\|\le 2\frac{\|h'\|}{\sigma^3}$。将这一结果代入 (F.43) 可得 ——

$$
|\mathbb{E}\left[h(W)\right]-\mathbb{E}\left[h(N)\right]| \le 2\frac{\|h'\|}{\sigma^3}\sum_{i\in J}\left\{\mathbb{E}\left|\left(\xi_i\xi(A_i)-\mathbb{E}(\xi_i\xi(A_i))\right)\xi(B_i)\right|+\frac{1}{2}\mathbb{E}\left|\xi_i(\xi(A_i))^2\right|\right\}.
$$

由于我们按照 (1.6) 处理 $h\in H_{\mathrm{W}}$，因此有 $\|h'\|\le 1$，从而得出该引理的结论。□

F.2 引理 B.1 及 i 的一个广义版本的证明 t

引理 F.1 在下文中陈述并证明，它涵盖了引理 B.1 这一特殊情形。这两个引理的重要区别在于，引理 F.1 也可用于非平稳数据的情形。

引理 F.1. 对于 $p\in\mathbb{N}$ 和 $\alpha\ge 1$，设 $X_1,\ldots,X_p,Y_1,\ldots,Y_p$ 为取值于 $\mathbb{R}$ 的随机变量，满足 $\|X_i\|_{\alpha p}<\infty$ 和 $\|Y_i\|_{\alpha p}<\infty, i=1,\ldots,p$。那么，按照引理 *B.1* 中的定义，令 Λ_p，我们有

$$
\Big\|\prod_{i=1}^{p}X_i-\prod_{i=1}^{p}Y_i\Big\|_{\alpha}\le\sum_{(\ell_1,\ldots,\ell_p)\in\Lambda_p}\prod_{i=1}^{p}\|X_i-Y_i\|_{\alpha p}^{\ell_i}\|X_i\|_{\alpha p}^{1-\ell_i}, \qquad \text{(F.44)}
$$

引理 B.1 的证明。如果我们令 $X_i:=X_{a_i}(t_i)$ 和 $Y_i:=Y_{a_i}(t_i)$ 满足 $\|X_i-Y_i\|_{\alpha p}\le\tilde{D}_{a_i}^{(\alpha p)}$，则引理 B.1 可由更一般的引理 F.1 推出。

引理 F.1 的证明。对于 $p=1$ 该断言显然。对于 $p\geq 2$ 该断言可由以下不等式链推出，其中我们记 $D_i := \|X_i - Y_i\|_{\alpha p}$

$$\begin{aligned}
\Big\|\prod_{i=1}^{p} X_i - \prod_{i=1}^{p} Y_i\Big\|_\alpha &= \Big\|\sum_{j=1}^{p}\Big(\prod_{\substack{i=1\\ i<j}}^{p} X_i\Big)(X_j - Y_j)\Big(\prod_{\substack{i=1\\ i>j}}^{p} Y_i\Big)\Big]\Big\|_\alpha \\
&\leq \sum_{j=1}^{p}\Big\|\Big(\prod_{\substack{i=1\\ i<j}}^{p} X_i\Big)(X_j - Y_j)\Big(\prod_{\substack{i=1\\ i>j}}^{p} Y_i\Big)\Big\|_\alpha \\
&\leq \sum_{j=1}^{p}\Big(\big\|X_j - Y_j\big\|_{\alpha p}\prod_{\substack{i=1\\ i<j}}^{p}\|X_i\|_{\alpha p}\prod_{\substack{i=1\\ i>j}}^{p}\big(\|X_i\|_{\alpha p} + \big\|X_i - Y_i\big\|_{\alpha p}\big)\Big)
\end{aligned}$$

$$\begin{aligned}
&= D_1\big(\|X_2\|_{\alpha p} + D_2\big)\cdots\big(\|X_p\|_{\alpha p} + D_p\big) \\
&\quad + I\{p>2\}\sum_{j=2}^{p-1}\|X_1\|_{\alpha p}\cdots\|X_{j-1}\|_{\alpha p}D_j\big(\|X_{j+1}\|_{\alpha p} + D_{j+1}\big)\cdots\big(\|X_p\|_{\alpha p} + D_p\big) \\
&\quad + D_p\|X_1\|_{\alpha p}\cdots\|X_{p-1}\|_{\alpha p} \\
&= D_1\sum_{(\ell_2,\ldots,\ell_p)\in\{0,1\}^{p-1}}\|X_2\|_{\alpha p}^{1-\ell_2}D_2^{\ell_2}\cdots\|X_p\|_{\alpha p}^{1-\ell_p}D_p^{\ell_p} \\
&\quad + I\{p>2\}\sum_{j=2}^{p-1}\|X_1\|_{\alpha p}\cdots\|X_{j-1}\|_{\alpha p}D_j \\
&\qquad\quad \times\sum_{(\ell_{j+1},\ldots,\ell_p)\in\{0,1\}^{p-j}}\|X_{j+1}\|_{\alpha p}^{1-\ell_{j+1}}D_{j+1}^{\ell_{j+1}}\cdots\|X_p\|_{p}^{1-\ell_p}D_p^{\ell_p} \\
&\quad + \|X_1\|_{\alpha p}\cdots\|X_{p-1}\|_{\alpha p}D_p,
\end{aligned}$$

其中，第一个不等式由三角不等式得出，这是因为 $\alpha\geq 1$；第二个不等式则由（推广的）赫尔德不等式得出。□

F.3 引理 B.3 的证明

˜注意，根据三角不等式以及主论文中 (2.8) 对 $\Sigma_{ab}(k)$ 的定义，我们有

$$\begin{aligned}
|\tilde{\Sigma}_{ab}(k) - \Sigma_{ab}(k)| \leq &\Big|\mathrm{var}\left(n^{-1/2}\hat{\gamma}_{ab}^{(m)}(k)\right) - \mathrm{var}\left(n^{-1/2}\hat{\gamma}_{ab}(k)\right)\Big| \\
&+ \Big|n\mathrm{var}\big(\hat{\gamma}_{ab}(k)\big) - \Sigma_{ab}(k)\Big|,
\end{aligned}$$

其中

$$\hat{\gamma}_{ab}^{(m)}(k) := \frac{1}{n}\sum_{t=1}^{n-k} Y_a(t+k)Y_b(t) \text{ for } k = 0, \ldots, n-1.$$

接下来，我们得到如下更一般的上界，随后将用它来推导该引理的断言：

$$\begin{aligned}
&\Big|\mathrm{cov}\big(\sqrt{n}\hat{\gamma}_{a_1b_1}(k_1), \sqrt{n}\hat{\gamma}_{a_2b_2}(k_2)\big) - \mathrm{cov}\big(\sqrt{n}\hat{\gamma}^{(m)}_{a_1b_1}(k_1), \sqrt{n}\hat{\gamma}^{(m)}_{a_2b_2}(k_2)\big)\Big| \\
&\le \frac{1}{n}\sum_{t_1=1}^{n-k_1}\sum_{t_2=1}^{n-k_2}\Big|\mathbb{E}\big[X_{a_1}(t_1+k_1)X_{b_1}(t_1)X_{a_2}(t_2+k_2)X_{b_2}(t_2)\big] \\
&\qquad\qquad - \mathbb{E}\big[Y_{a_1}(t_1+k_1)Y_{b_1}(t_1)Y_{a_2}(t_2+k_2)Y_{b_2}(t_2)\big]\Big| \\
&\quad + \frac{1}{n}\sum_{t_1=1}^{n-k_1}\sum_{t_2=1}^{n-k_2}\Big|\Big(\mathbb{E}\big[X_{a_1}(t_1+k_1)X_{b_1}(t_1)\big] - \mathbb{E}\big[Y_{a_1}(t_1+k_1)Y_{b_1}(t_1)\big]\Big)\mathbb{E}\big[X_{a_2}(t_2+k_2)X_{b_2}(t_2)\big] \\
&\qquad + \Big(\mathbb{E}\big[X_{a_2}(t_2+k_2)X_{b_2}(t_2)\big] - \mathbb{E}\big[Y_{a_2}(t_2+k_2)Y_{b_2}(t_2)\big]\Big)\mathbb{E}\big[Y_{a_1}(t_1+k_1)Y_{b_1}(t_1)\big]\Big| \\
&\le \frac{(n-k_1)(n-k_2)}{n}\Big(K_{4,m}(\{a_1,b_1,a_2,b_2\}) + K_{2,m}(\{a_1,b_1\})|\gamma_{a_2b_2}(k_2)| \\
&\qquad\qquad + \big(|\gamma_{a_1b_1}(k_1)| + K_{2,m}(\{a_1,b_1\})\big)K_{2,m}(\{a_2,b_2\})\Big) \\
&\le \frac{(n-k_1)(n-k_2)}{n}\Big(\max_{i=a_1,b_1,a_2,b_2}\tilde{D}_i^{(4)}\Big) \\
&\quad\times\Bigg(\sum_{j\in\{a_1,b_1,a_2,b_2\}}\prod_{i\in\{a_1,b_1,a_2,b_2\}\setminus\{j\}}(\|X_i(0)\|_4 + \tilde{D}_i^{(4)}) \\
&\qquad + (\|X_{a_1}(0)\|_2 + \|X_{b_1}(t)\|_2 + \min_{i=a_1,b_1}\tilde{D}_i^{(2)})|\gamma_{a_2b_2}(k_2)| \\
&\qquad + (\|X_{a_2}(0)\|_2 + \|X_{b_2}(t)\|_2 + \min_{i=a_2,b_2}\tilde{D}_i^{(2)})|\gamma_{a_1b_1}(k_1)| \\
&\qquad + \min\Big\{\max_{i=a_1,b_1}\tilde{D}_i^{(2)}, \max_{i=a_2,b_2}\tilde{D}_i^{(2)}\Big\} \\
&\qquad\quad\times\Big(\|X_{a_1}(0)\|_2 + \|X_{b_1}(t)\|_2 + \min_{i=a_1,b_1}\tilde{D}_i^{(2)}\Big)\Big(\|X_{a_2}(0)\|_2 + \|X_{b_2}(t)\|_2 + \min_{i=a_2,b_2}\tilde{D}_i^{(2)}\Big)\Bigg)
\end{aligned}$$

其中，对于第三个不等式，根据引理 B.1，取 $\alpha = 1$ 和 $p = 2$，可得

$$\begin{aligned}
\tilde{K}_2^{(1)}(\{a,b\}) &= \tilde{D}_a^{(2)}\|X_b(t)\|_2 + \tilde{D}_a^{(2)}\tilde{D}_b^{(2)} + \|X_a(0)\|_2\tilde{D}_b^{(2)} \\
&\le \Big(\max_{i=a,b}\tilde{D}_i^{(2)}\Big)(\|X_a(0)\|_2 + \|X_b(t)\|_2 + \min_{i=a,b}\tilde{D}_i^{(2)})
\end{aligned}$$

并且，根据引理 B.1，取 $\alpha = 1$ 和 $p = 4$，可得

$$
\begin{aligned}
&\tilde{K}_4^{(1)}(\{a_1, b_1, a_2, b_2\}) \\
&= \tilde{D}_{a_1}^{(4)}(\|X_{b_1}(0)\|_4 + \tilde{D}_{b_1}^{(4)})(\|X_{a_2}(0)\|_4 + \tilde{D}_{a_2}^{(4)})(\|X_{b_2}(0)\|_4 + \tilde{D}_{b_2}^{(4)}) \\
&\quad + \tilde{D}_{b_1}^{(4)}\|X_{a_1}(0)\|_4(\|X_{a_2}(0)\|_4 + \tilde{D}_{a_2}^{(4)})(\|X_{b_2}(0)\|_4 + \tilde{D}_{b_2}^{(4)}) \\
&\quad + \tilde{D}_{a_2}^{(4)}\|X_{a_1}(0)\|_4\|X_{b_1}(0)\|_4(\|X_{b_2}(0)\|_4 + \tilde{D}_{b_2}^{(4)}) \\
&\quad + \tilde{D}_{b_2}^{(4)}\|X_{a_1}(0)\|_4\|X_{b_1}(0)\|_4\|X_{a_2}(0)\|_4 \\
&\leq \Big(\max_{i=a_1,b_1,a_2,b_2} \tilde{D}_i^{(4)}\Big) \sum_{j\in\{a_1,b_1,a_2,b_2\}} \prod_{i\in\{a_1,b_1,a_2,b_2\}\setminus\{j\}} (\|X_i(0)\|_4 + \tilde{D}_i^{(4)}).
\end{aligned}
$$

对于第三个不等式，我们还利用了以下事实：

$$
\max_{i=j_1,j_2} \tilde{D}_i^{(2)} \leq \max_{i=a_1,b_1,a_2,b_2} \tilde{D}_i^{(4)}, \quad \text{for any } j_1, j_2 \in \{a_1, b_1, a_2, b_2\}.
$$

该引理的断言由对我们有 $k_1 = k_2$、$a_1 = a_2$ 和 $b_1 = b_2$ 的特殊情形作一些简化后得出。□

F.4 引理 B.4 的证明

我们通过以下不等式链得出该断言。

$$
\begin{aligned}
&|Q_t - \tilde{Q}_t| \\
&\leq \left| \|(Z_t Z(A_t) - \mathbb{E}(Z_t Z(A_t)))Z(B_t)\|_1 - \|(\tilde{Z}_t \tilde{Z}(A_t) - \mathbb{E}(\tilde{Z}_t \tilde{Z}(A_t)))\tilde{Z}(B_t)\|_1 \right| \\
&\quad + \left| \frac{1}{2}\|Z_t Z(A_t)^2\|_1 - \frac{1}{2}\|\tilde{Z}_t \tilde{Z}(A_t)^2\|_1 \right| \\
&\leq \mathbb{E}\Big| (Z_t Z(A_t) - \mathbb{E}(Z_t Z(A_t)))Z(B_t) - (\tilde{Z}_t \tilde{Z}(A_t) - \mathbb{E}(\tilde{Z}_t \tilde{Z}(A_t)))\tilde{Z}(B_t) \Big| \\
&\quad + \frac{1}{2}\mathbb{E}\left| Z_t Z(A_t)^2 - \tilde{Z}_t \tilde{Z}(A_t)^2 \right| \\
&\leq \mathbb{E}\left| Z_t Z(A_t) Z(B_t) - \tilde{Z}_t \tilde{Z}(A_t)\tilde{Z}(B_t) \right| + \mathbb{E}\left| \mathbb{E}(Z_t Z(A_t))Z(B_t) - \mathbb{E}(\tilde{Z}_t \tilde{Z}(A_t))\tilde{Z}(B_t) \right| \\
&\quad + \frac{1}{2}\mathbb{E}\left| Z_t Z(A_t)^2 - \tilde{Z}_t \tilde{Z}(A_t)^2 \right| \\
&\leq \sum_{j_1 \in A_t} \sum_{j_2 \in B_t} \mathbb{E}\left| Z_t Z_{j_1} Z_{j_2} - \tilde{Z}_t \tilde{Z}_{j_1} \tilde{Z}_{j_2} \right| + \frac{1}{2} \sum_{j_1 \in A_t} \sum_{j_2 \in A_t} \mathbb{E}\left| Z_t Z_{j_1} Z_{j_2} - \tilde{Z}_t \tilde{Z}_{j_1} \tilde{Z}_{j_2} \right| \qquad \text{(F.45)} \\
&\quad + \mathbb{E}\left| \mathbb{E}\left(Z_t \sum_{j \in A_t} Z_j \right) \sum_{j \in B_t} (Z_j - \tilde{Z}_j) \right| \qquad \text{(F.46)} \\
&\quad + \mathbb{E}\left| \left[\mathbb{E}\left(Z_t \sum_{j \in A_t} Z_j \right) - \mathbb{E}\left(\tilde{Z}_t \sum_{j \in A_t} \tilde{Z}_j \right) \right] \sum_{j \in B_t} (Z_j + \tilde{Z}_j - Z_j) \right|, \qquad \text{(F.47)}
\end{aligned}
$$

其中我们使用了三角不等式和反三角不等式。对于 $|Q_t - Q_t|$，所述上界˜将是我们现在推导出的式 (F.45)、(F.46) 和 (F.47) 的上界之和。首先，˜我们推导两个初步界，以控制 $\|Z_0\|_\alpha$ 和 $\|Z_0 - Z_0\|_\alpha$，$\alpha = 1$，2，3。利用三角不等式和赫尔德不等式，我们有

$$
\|Z_0\|_\alpha \leq \|X_a(t+k)X_b(t)\|_\alpha + |\mathbb{E}[X_a(t+k)X_b(t)]| \leq 2\|X_a(0)\|_{2\alpha}\|X_b(0)\|_{2\alpha} \qquad \text{(F.48)}
$$

并且，利用三角不等式和引理 B.1，我们得到

$$
\|Z_0 - \tilde{Z}_0\|_\alpha \leq 2\|X_a(t+k)X_b(t) - Y_a(t+k)Y_b(t)\|_\alpha \leq 2\tilde{K}_2^{(\alpha)}(\{a,b\}) \qquad \text{(F.49)}
$$

接下来，为了对 (F.45) 取上界，我们使用引理 F.1，其中 $\alpha = 1$, $p = 3$, X_1, X_2, X_3 分别等于˜ ˜ ˜Z_t, Z_{j_1}, Z_{j_2}，而 Y_1, Y_2, Y_3 分别等于 Z_t, Z_{j_1}, Z_{j_2}。

˜然后， $\|X_i\|_\alpha = \|Z_0\|_\alpha$ 和 $\|X_i - Y_i\|_\alpha = \|Z_0 - Z_0\|_\alpha$，使得

$$\begin{aligned}
&\mathbb{E}\left|Z_t Z_{j_1} Z_{j_2} - \tilde{Z}_t \tilde{Z}_{j_1} \tilde{Z}_{j_2}\right| \\
&\leq \|Z_0 - \tilde{Z}_0\|_3 \Big(3\|Z_0\|_3^2 + 3\|Z_0\|_3 \|Z_0 - \tilde{Z}_0\|_3 + \|Z_0 - \tilde{Z}_0\|_3^2\Big) \\
&\leq 2\tilde{K}_2^{(3)} \Big(3\big(2\|X_a(0)\|_6 \|X_b(0)\|_6 + \tilde{K}_2^{(3)}\big)^2 + (\tilde{K}_2^{(3)})^2\Big)
\end{aligned} \tag{F.50}$$

显然，这意味着

$$(\text{F.45}) \leq 2\tilde{K}_2^{(3)} \Big(3\big(2\|X_a(0)\|_6 \|X_b(0)\|_6 + \tilde{K}_2^{(3)}\big)^2 + (\tilde{K}_2^{(3)})^2\Big)\Big(|A_t||B_t| + \frac{1}{2}|A_t|^2\Big) \tag{F.51}$$

对于 (F.46)，我们有

$$(\text{F.46}) = \Big|\sum_{j\in A_t} \mathbb{E}(Z_t Z_j)\Big| \times \mathbb{E}\left|\sum_{j\in B_t} (Z_j - \tilde{Z}_j)\right| \leq \Big|\sum_{j\in A_t} \mathbb{E}(Z_t Z_j)\Big| \times |B_t| \times 2\tilde{K}_2^{(1)}, \tag{F.52}$$

其中我们使用了带有 $\alpha = 1$ 的 (F.49)。最后，为了给 (F.47) 建立上界，我们像在 (F.50) 中那样使用引理 F.1，但采用 $p = 2$，得到

$$\begin{aligned}
\mathbb{E}\left|Z_t Z_j - \tilde{Z}_t \tilde{Z}_j\right| &\leq \|Z_0 - \tilde{Z}_0\|_2 \Big(2\|Z_0\|_2 + \|Z_0 - \tilde{Z}_0\|_2\Big) \\
&\leq 2\tilde{K}_2^{(2)} \Big(4\|X_a(0)\|_4 \|X_b(0)\|_4 + 2\tilde{K}_2^{(2)}\Big),
\end{aligned}$$

其中第二个不等式由带有 $\alpha = 2$ 的初步界 (F.48) 和 (F.49) 得出。因此，我们有

$$\begin{aligned}
(\text{F.47}) &= \left|\mathbb{E}\left(Z_t \sum_{j\in A_t} Z_j\right) - \mathbb{E}\left(\tilde{Z}_t \sum_{j\in A_t} \tilde{Z}_j\right)\right| \times \mathbb{E}\left[\Big|\sum_{j\in B_t} (Z_j + \tilde{Z}_j - Z_j)\Big|\right] \\
&\leq \sum_{j\in A_t} \mathbb{E}\Big|Z_t Z_j - \tilde{Z}_t \tilde{Z}_j\Big| \times \sum_{j\in B_t} \Big(\|Z_0\|_1 + \|\tilde{Z}_0 - Z_0\|_1\Big) \\
&\leq 2\tilde{K}_2^{(2)} \Big(4\|X_a(0)\|_4 \|X_b(0)\|_4 + 2\tilde{K}_2^{(2)}\Big)|A_t| \\
&\quad \times 2\Big(\|X_a(0)\|_2 \|X_b(0)\|_2 + \tilde{K}_2^{(1)}\Big)|B_t|.
\end{aligned} \tag{F.53}$$

对 t 求和 (F.51)、(F.52) 和 (F.53) 中得到的三个上界给出了该断言 . □

F.5 命题 2.7 的证明

(2.12) 中给出的上界由四项构成，下面我们针对情形 1 和情形 2 逐一讨论。对于机制 1，我们假设 $\boldsymbol{Y}^{(m)}(t)$ 按照注 2.6 中的描述选取。在此证明中，我们将依次把该上界中的第一项、第二项、第三项和第四项称为相应的项，即 (2.12) 中相应的加项。对于上界的第一项，在两种情形下显然都有$kn^{-1/2}|\gamma_{ab}(k)| = \mathcal{O}(n^{-1/2})$。关于上界的第二项，下面我们证明，在两种情形下，若适当选取 m，则该项的阶为 $\mathcal{O}(n^{-1/2})$。利用引理 B.3，可以控制˜表达式 $|\Sigma_{ab}(k) - \Sigma_{ab}(k)|$。因此，

$$\frac{\sqrt{2}}{\sqrt{\pi\Sigma_{ab}(k)}}\left|\Sigma_{ab}(k) - \tilde{\Sigma}_{ab}(k)\right| \le \frac{\sqrt{2}}{\sqrt{\pi\Sigma_{ab}(k)}}\left|n\mathrm{var}\big(\hat{\gamma}_{ab}(k)\big) - \Sigma_{ab}(k)\right| + 2\frac{\sqrt{2}}{\sqrt{\pi\Sigma_{ab}(k)}}\frac{(n-k)^2}{n}\Big(\max_{i=a,b}\tilde{D}^{(4)}_{i,m}\Big)\tilde{F}_m, \tag{F.54}$$

其中 $\tilde{F}_m$ 由 (B.32) 给出。在机制 1 中，$\tilde{D}^{(q)}_{j\ ,\ m} = 0$ 对于 $n \ge M$ 成立，因此当 n 足够大时，(F.54) 为零。在机制 2 中，由于四阶矩有限，根据 (2.17) 和詹森不等式，我们有

$$\min_{i=a,b}\tilde{D}^{(2)}_{i,m} \le \max_{i=a,b}\tilde{D}^{(2)}_{i,m} \le \max_{i=a,b}\tilde{D}^{(4)}_{i,m} = o(1),$$

˜当 $m \to \infty$时，这意味着 $F_m = \mathcal{O}(1)$。现在，由于 $m = C\log\ n$，且 $C \ge -3/(2\log(\rho))$，我们有 (F.54) 在机制 2 中的阶为 $\mathcal{O}(n\rho^m) = \mathcal{O}(n^{-1/2})$；我们此前已看到，机制 1 中也如此。接下来需要估计 $|n\mathrm{var}\big(\hat{}\gamma_{ab}(k)\big) - \Sigma_{ab}(k)|$ 的阶，即经验交叉协方差的有限样本方差与渐近方差之差。在条件 (2.16) 下，该量的阶为 $\mathcal{O}(n^{-1})$；我们已论证该条件在两种机制下均成立。例如，可以根据 Brillinger (1975) 第 7.6 节中的论证得出这一结论。总之，在情形 1 和情形 2 中，上界的第二项的阶为 $\mathcal{O}(n^{-1/2})$。

关于上界中的第三项 $2(n-k)n^{-1/2}\tilde{K}_m$，注意到在机制 1 中，当 $\tilde{D}^{(q)}_{j\ ,m} = 0$对于 $n \ge M$ 时，$\tilde{K}_m$ 最终会消失；而在机制 2 中，根据 (2.17)，我们有˜$K_m = \mathcal{O}(\rho^m)$。因此，在机制 1 中，当 n 足够大时，上界中的第三项会消失；在机制 2 中，如果我们选择 $m = C\log n$，则该项的阶为 $\mathcal{O}(\sqrt{n}\rho^m) = \mathcal{O}(n^{-1/2})$，并且只需满足 $C \ge -1/\log(\rho)$。总之，在两种机制下，上界中的第三项均为 $\mathcal{O}(n^{-1/2})$ 阶。

现在我们继续讨论上界的第四项 $2n^{-3/2}(\tilde{\Sigma}_{ab}(k))^{-3/2}\sum_{t=1}^{n-k}\tilde{Q}_t$ 的收敛速率。

表 4： 指标表以及需对公式 (F.57) 中求和的加项进行划分的变量

(1,a)	(1,b)
(2,a)	(2,b)
(3,a)	(3,b)
(4,a)	(4,b)

(a) 待划分的指标表

$Y_{1,a} := X_a(t+k)$	$Y_{1,b} := X_b(t)$
$Y_{2,\ a} := X_a(t+k)$	$Y_{2,b} := X_b(t)$
$Y_{3,a} := X_a(j_1+k)$	$Y_{3,b} := X_b(j_1)$
$Y_{4,\ a} := X_a(j_2+k)$	$Y_{4,b} := X_b(j_2)$

(b) 所考虑累积量的变量

对于这一项，机制的不同会导致分析结果不同。根据对上界第二项的讨论，我们有 $\Sigma_{ab}(k) \to \Sigma_{ab}(k) > 0$，当 $n \to \infty$ 时。由 (B.33)、引理 B.4 以及条件 (2.17)（根据注 2.6，该条件在两种机制下均成立），只需证明在机制 1 下 $\sup_{t=1,\dots,n-k} Q_t$ 的阶为 $\mathcal{O}(1)$，而在机制 2 下其阶为 $\mathcal{O}(\log n)$。

机制 1：按照 (2.13) 对 Q_t 进行上界估计，但将 $\{\boldsymbol{Y}^{(m)}(t)\}$ 替换为 $\{\boldsymbol{X}(t)\}$，则有 $Q_t \le 3(4m+4k+1)^2\|\boldsymbol{X}(0)\|_6^6 = \mathcal{O}(m^2) = \mathcal{O}(1)$，因为 $m = \min\{n, M\} \le M$。

Regi机制 2：这里我们使用 (2.14) 的一个版本，将 $\{\boldsymbol{Y}^{(m)}(t)\}$ 替换为 $\{\boldsymbol{X}(t)\}$：

$$\begin{aligned} Q_t \le{}& \operatorname{var}\Big(Z_t \sum_{j\in A_t} Z_j\Big)^{1/2} \operatorname{var}\Big(\sum_{j\in B_t} Z_j\Big)^{1/2} \\ &+ \frac{1}{2}\Big[\mathbb{E}\Big(Z_t \sum_{j\in A_t} Z_j\Big)^2\Big]^{1/2} \operatorname{var}\Big(\sum_{j\in A_t} Z_j\Big)^{1/2} \end{aligned} \tag{F.55}$$

只需证明 (A.28) 中定义的三个量在将 $\boldsymbol{Y}^{(m)}(t)$ 替换为 $\boldsymbol{X}(t)$ 后为 $\mathcal{O}(m)$。现在我们证明这可由条件 (2.16) 和直至 8 阶的累积量可和性推出。按照 A 节中的论证，我们有

$$\operatorname{var}\Big(\sum_{j\in A_t} Z_j\Big) \le |A_t| \sum_{u=-\infty}^{\infty} |C(u)|, \quad \text{and} \quad \mathbb{E}\Big(Z_t \sum_{j\in A_t} Z_j\Big) \le \sum_{u=-\infty}^{\infty} |C(u)|, \tag{F.56}$$

其中

$$C(u) = \operatorname{cum}(X_a(k), X_b(0), X_a(u+k), X_b(u)) + \gamma_{aa}(u)\gamma_{bb}(u) + \gamma_{ab}(k-u)\gamma_{ab}(k+u),$$

根据条件 (2.16)，该项可求和。最后，采用表 4(b) 中的符号，我们

have

$$\sum_{j_1\in A_t}\sum_{j_2\in A_t}\mathrm{cum}(Z_t,Z_t,Z_{j_1},Z_{j_2})\le\sum_{\nu}S(\nu),$$
$$S(\nu):=\sum_{j_1=-\infty}^{\infty}\sum_{j_2=-\infty}^{\infty}\prod_{r=1}^{R}|\mathrm{cum}(Y_{ij}:(ij)\in\nu_r)|, \tag{F.57}$$

其中，求和遍及表 4(a) 中指标的所有不可分解划分 $\nu:=\{\nu_1,\dots,\nu_R\}$。为完成论证，只需证明对于表 4(a) 的每个不可分解划分 ν，在 $|\nu_r|\ge 2$ 的条件下，

$$\exists C_\nu:|S(\nu)|\le C_\nu, \tag{F.58}$$

其中 C_ν 与 t 独立。于是，由于 ν 只有有限多个，我们已经证明 (F.57) 左侧关于 t 一致地为 $\mathcal{O}(1)$。(F.58) 的严格证明直接但繁琐，下面将在命题 2.7 证明的其余步骤之后给出。

总之，结合 (F.55) 与关于如何计算 (2.14) 的讨论，再结合 (F.56)、(F.57)、(F.58)、$|A_t|\le 4(m+k)+1$ 以及针对上界第一、第二和第三项所选的 $m=C\log n$，我们证明了在机制 2 下，$\sup_{t=1,\dots,n-k}Q_t=\mathcal{O}(\log n)$，这意味着机制 2 下上界的第四项为 $\mathcal{O}(n^{-1/2}\log n)$ 阶。

现在我们通过证明：对于每个如（F.58）中的 ν，以下两个断言之一成立，来完成证明：

（a）存在一个集合 ν_ℓ 属于 ν，使得

$$\sum_{j_1=-\infty}^{\infty}\sum_{j_2=-\infty}^{\infty}|\mathrm{cum}(Y_{ij}:(ij)\in\nu_\ell)|\le c_\nu,$$

（b）或者存在两个不同的集合 ν_{ℓ_1} 和 ν_{ℓ_2} 属于 ν，使得

$$\sum_{j_1=-\infty}^{\infty}\sum_{j_2=-\infty}^{\infty}|\mathrm{cum}(Y_{ij}:(ij)\in\nu_{\ell_1})||\mathrm{cum}(Y_{ij}:(ij)\in\nu_{\ell_2})|\le c_\nu,$$

其中常数为 c_ν 不依赖于 t。所需的 $S(\nu)\le C_\nu$ 于是可通过应用

$$|\mathrm{cum}(Y_{ij}:(ij)\in\nu_r)|\le|\nu_r|^{|\nu_r|}(|\nu_r|-1)!\max\{1,\|\boldsymbol{X}(0)\|_{|\nu_r|}\}^{|\nu_r|}=:B(\nu_r), \tag{F.59}$$

于定义 $S(\nu)$ 中剩余的因子而得出。注意， $B(\nu_r)$ 不依赖于 t。(F.59) 中的不等式由定义（2.7）、三角不等式、 ν_r 的划分数量可由 $|\nu_r|^{|\nu_r|}$ 上界控制这一事实、 ν_r 的划分中集合的最大数量为 $|\nu_r|$ 这一事实、广义 Hölder 不等式、平稳性以及詹森不等式得出。下文中，我们将表 4(a) 中的 (3,1) 和 (3,2) 称为 j_1-指标，将 (4,1) 和 (4,2) 称为 j_2-指标。 j-指标指 (3,1)、(3,2)、(4,1) 和 (4,2)，非-j-指标则指其余的 (1,1)、(1,2)、(2,1) 和 (2,2)。现在，假设 $\nu := \{\nu_1, \ldots, \nu_R\}$ 是表 4(a) 的一个不可分解划分，其中每个 ν_r 至少包含两个元素。要证明 (a) 或 (b) 成立，显然只需考察 j-指标可能的排列。为此，注意到根据 (F.59)，我们可以忽略不包含至少一个 j-指标的集合。

在讨论 j-指标在具有 $|\nu_r| \geq 2$ 的不可分解划分ν 中的所有可能排列之前，我们现在解释如何证明 (a) 或 (b) 成立：为了证明 (a) 成立，对于给定的划分 $\nu := \{\nu_1, \ldots, \nu_R\}$，只需确定一个集合 ν_ℓ 属于 ν，且其中至少包含一个 j_1-指标、至少一个 j_2-指标以及至少一个非 j-指标。 于是我们有 $|\nu_\ell| =: p \geq 3$。设 $\nu_\ell := \{(u_1, v_1) \ldots, (u_p, v_p)\}$ 中的元素按如下方式排列：第一个是一个 j_1-指标，第二个是一个 j_2-指标，而第 p 个是一个非 j-指标。那么，由 (F.59) 以及 $\{\boldsymbol{X}(t)\}$ 的平稳性，我们有

$$
\begin{aligned}
S(\nu) &\leq \left(\prod_{\substack{r=1 \\ r \neq \ell}}^{R} B(\nu_r) \right) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} |\mathrm{cum}(X_{v_1}(j_1 + h_1 - s), X_{v_2}(j_2 + h_2 - s), \ldots, X_{v_p}(0))| \\
&\leq \left(\prod_{\substack{r=1 \\ r \neq \ell}}^{R} B(\nu_r) \right) \sum_{i_1=-\infty}^{\infty} \cdots \sum_{i_{p-1}=-\infty}^{\infty} |\mathrm{cum}(X_{v_1}(i_1), \ldots, X_{v_{p-1}}(i_{p-1}), X_{v_p}(0))| =: c_\nu,
\end{aligned}
$$

其中 h_1， $h_2 \in \{0, k\}$（取决于 v_1 和 v_2）、 $s \in \{t, t + k\}$（取决于 v_p），第一行中的 ... ...表示表 4(a) 中时间指标平移了s的 $|\nu_\ell| - 3$ 变量。

现在，为了证明 (b) 成立，对于给定的划分 $\nu := \{\nu_1, \ldots, \nu_R\}$，只需确定集合 ν_{ℓ_1} 和 ν_{ℓ_2}，使其满足以下条件：

(I) $\nu_{\ell_1} := \{(u_{1,1}, v_{1,1}) \ldots, (u_{1,p_1}, v_{1,p_1})\}$ 至少包含一个 j_1-指标；
(II) $\nu_{\ell_2} := \{(u_{2,1}, v_{2,1}) \ldots, (u_{2,p_2}, v_{2,p_2})\}$ 至少包含一个 j_2-指标；(III) ν_{ℓ_1} 或 ν_{ℓ_2} 至少包含一个非-j-指标；

(IV) 如果 ν_{ℓ_1} 同时包含两个 j_1-指标，那么它必须至少包含三个元素；同样，如果ν_{ℓ_2} 同时包含两个 j_2-指标，那么它必须至少包含三个元素。

注意，对于任意不可分解划分，(IV) 始终成立，因为包含恰好两个 j_1-指标或恰好两个 j_2-指标的集合的划分是可分解的。

首先，考虑根据 (III) 可知存在的非-j-指标属于 ν_{ℓ_1} 的情形。它属于 ν_{ℓ_2} 的情形可以类似处理。我们可以对集合和指标进行编号，使得 $(u_{1,1},v_{1,1})$ 是一个 j_1-指标，$(u_{2,1}, v_{2,1})$ 是一个 j_2-指标，而 (u_{1,p_1}, v_{1,p_1}) 是一个非-j-指标。此外，由于 (IV) 和 $|\nu_{\ell_2}| \geq 2$，(u_{2,p_2},v_{2,p_2}) 不是一个 j_2-指标。因此，对于满足 (I)–(IV) 的 ν_{ℓ_1} 和 ν_{ℓ_2}，我们有

$$
\begin{aligned}
S(\nu) &\leq \left(\prod_{\substack{r=1 \\ r\notin\{\ell_1,\ell_2\}}}^{R} B(\nu_r) \right) \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} |\mathrm{cum}(X_{v_{1,1}}(j_1+h_1-s_1), \ldots, X_{v_{1,p_1}}(0))| \\
&\qquad\qquad \times |\mathrm{cum}(X_{v_{2,1}}(j_2+h_2-s_2), \ldots, X_{v_{2,p_2}}(0))| \\
&\leq \left(\prod_{\substack{r=1 \\ r\notin\{\ell_1,\ell_2\}}}^{R} B(\nu_r) \right) \sum_{i_{1,1}=-\infty}^{\infty} \cdots \sum_{i_{1,p_1-1}=-\infty}^{\infty} |\mathrm{cum}(X_{v_{1,1}}(i_{1,1}), \ldots, X_{v_{1,p_1-1}}(i_{1,p_1-1}), X_{v_{1,p_1}}(0))| \\
&\quad \times \sum_{i_{2,1}=-\infty}^{\infty} \cdots \sum_{i_{2,p_2-1}=-\infty}^{\infty} |\mathrm{cum}(X_{v_{2,1}}(i_{2,1}), \ldots, X_{v_{2,p_2-1}}(i_{2,p_2-1}), X_{v_{2,p_2}}(0))| =: c_\nu,
\end{aligned}
$$

其中 h_1，$h_2 \in \{0, k\}$（取决于 $v_{1,\ 1}$和 $v_{2,\ 1}$ 的取值）、$s_1 \notin \{j_1, j_1+k, j_2, j_2+k\}$ 以及 $s_2 \notin \{j_2, j_2+k\}$。对于第二个不等式，我们将 $j_1+h_1-s_1$ 替换为 $i_{1,\ 1}$，这对应于 j_1-求和中指标的平移，因为 $(u_{1,\ p_1},\ v_{1,\ p_1})$ 是一个非 j-指标，随后分两步取上界

$$
\begin{aligned}
&|\mathrm{cum}(X_{v_{1,1}}(i_{1,1}), \ldots, X_{v_{1,p_1}}(0))| \\
&\qquad \leq \sum_{i_{1,2}=-\infty}^{\infty} \cdots \sum_{i_{1,p_1-1}=-\infty}^{\infty} |\mathrm{cum}(X_{v_{1,1}}(i_{1,1}), \ldots, X_{v_{1,p_1-1}}(i_{1,p_1-1}), X_{v_{1,p_1}}(0))|, \quad \text{(F.60)}
\end{aligned}
$$

其中，(F.60) 右侧不再依赖于 j_2。于是，$S(\nu) \leq c_\nu$ 成立

from

$$\sum_{j_2=-\infty}^{\infty} |\mathrm{cum}(X_{v_{2,1}}(j_2+h_2-s_2),\ldots,X_{v_{2,p_2}}(0))|$$
$$\leq \sum_{i_{2,1}=-\infty}^{\infty} \cdots \sum_{i_{2,p_2-1}=-\infty}^{\infty} |\mathrm{cum}(X_{v_{2,1}}(i_{2,1}),\ldots,X_{v_{2,p_2-1}}(i_{2,p_2-1}),X_{v_{2,p_2}}(0))|,$$

这是因为 $(u_{2,\ p_2},\ v_{2,\ p_2})$ 不是一个 j_2-指标。为完成证明，尚需讨论 j-指标的所有可能排列。为涵盖所有情况，我们根据划分中至少包含一个 j-指标的集合数量，组织如下讨论。

1. 划分中存在一个 ν_ℓ，其中包含所有 j-指标。在这种情况下，ν_ℓ 还必须至少包含一个额外的非 j-指标，否则该划分将是可分解的。因此，由上述内容可知，(a) 成立。

2. 存在两个集合 ν_{ℓ_1} 和 ν_{ℓ_2}，每个集合都至少包含一个 j-指标，且它们的并集包含所有 j-指标。若 $|\nu_{\ell_1}| = |\nu_{\ell_2}| = 2$，则会导致划分可分解，因此这种情况不可能发生。因此，我们要么有 $|\nu_{\ell_1}| \geq 3$ 和 $|\nu_{\ell_2}| \geq 2$，要么有 $|\nu_{\ell_1}| \geq 2$ 和 $|\nu_{\ell_2}| \geq 3$。有两种可能的子情况：(i) 一个集合包含三个 j-指标，另一个集合仅包含一个 j-指标；或者 (ii) ν_{ℓ_1} 和 ν_{ℓ_2} 均恰好包含两个 j-指标。

   若满足 (i)，那么只含一个 j指标的集合也必须至少含有一个非j指标（即满足 (III)）。此时 (b) 成立，因为如果只含一个 j指标的集合还含有一个 j_1指标，那么另一个集合就含有一个 j_2指标（反之亦然；即满足 (I) 和 (II)）。

   若满足 (ii)，那么至少含有三个元素的集合含有一个非j指标（即满足 (III)）。此外，由于每个集合恰好含有两个 j指标，且恰好有两个 j_1指标和两个 j_2指标，因此，若一个集合含有一个 j_1指标，另一个集合就必须含有一个 j_2指标（即满足 (I) 和 (II)）。因此，(b) 成立。

3. 存在一个恰好含有两个 j指标的集合，以及两个各自恰好含有一个 j指标的集合。于是，由于 $|\nu_r| \geq 2$，这两个各含一个 j指标的集合也至少含有一个非j指标。可能有两种子情形：恰好含有一个 j指标的集合中的 j指标要么都是 j_1指标，要么都是 j_2指标；或者 (ii) 一个集合中有一个 j_1指标，另一个集合中有一个j_2指标。

   若 (i) 成立，则恰含两个 j指标的集合中的 j指标，要么都是 j_2指标，要么都是 j_1指标。恰含两个 j指标的集合（即两个 j_1指标或

两个 j_2指标）因此还必须包含另一个非j指标，否则该划分就会是可分解的。取 ν_{ℓ_1} 为含两个 j指标的集合，并取 ν_{ℓ_2} 为恰含一个 j指标的两个集合之一，我们可以看到 (b) 成立：这两个集合中的指标分别为 j_1和 j_2 指标（即满足 (I) 和 (II)），且每个集合都至少含有一个非j指标（即满足 (III)）。

若满足 (ii)，则我们将 ν_{ℓ_1} 作为恰有一个 j_1-指标的集合之一，并将 ν_{ℓ_2} 作为恰有一个 j_2-指标的集合（即满足 (I) 和 (II)）。由于这两个集合各自也有一个非 j-指标（即满足 (III)），因此 (b) 得以满足。

4. 存在四个集合，每个集合恰有一个 j-指标。我们将 ν_{ℓ_1} 作为恰有一个 j_1-指标的集合之一，并将 ν_{ℓ_2} 作为恰有一个 j_2-指标的集合（即满足 (I) 和 (II)）。这两个集合各自也有一个非 j-指标（即满足 (III)），因此 (b) 得以满足。

至此，主论文中 (F.58) 的证明完成，同时也阐明了如何根据 (2.16) 中的求和以及 (F.59) 中定义的量 $B(\nu_r)$ 求出 C_ν；这些量均独立于 t。

□

G 论文中示例的补充表格

本节给出定理 2.3 中针对第 3.1 节示例在以下两种情形下所考虑的上界和真实 1-Wasserstein 距离的取值：$\varepsilon(t) \sim \mathcal{N}(0,\ 1)$，以及 $\varepsilon(t) \sim \sqrt{12/14}t_{14}$。

表5： 定理 2.3 中的上界结合式 (2.14) 所得的上界值，其中按照第 3.2 节所述取 $m = m^*$ 以使上界最小；针对经验自协方差，考察不同滞后阶数 k 和样本量 n。数据来自一个 AR(1) 过程，其中 $\varepsilon_t \sim \mathcal{N} \in (0, 1)$，且 α 取一系列不同的值。

k	$\alpha \mid n$	25	50	75	100	150	200	250	500	1000	2000
0	0	0.912	0.645	0.527	0.456	0.372	0.322	0.288	0.204	0.144	0.102
	0.1	4.438	3.784	3.219	2.894	2.533	2.340	2.222	1.661	1.239	0.967
	0.3	7.005	5.701	5.033	4.673	3.984	3.585	3.324	2.564	1.986	1.488
	0.5	9.842	8.080	7.164	6.484	5.676	5.110	4.715	3.650	2.798	2.128
	0.7	13.981	11.712	10.379	9.485	8.289	7.511	6.931	5.375	4.124	3.137
1	0	2.564	1.818	1.485	1.286	1.050	0.909	0.813	0.574	0.406	0.287
	0.1	6.778	5.045	4.259	3.801	3.279	2.988	2.752	1.996	1.477	1.135
	0.3	9.065	6.998	6.105	5.498	4.658	4.170	3.849	2.907	2.226	1.648
	0.5	11.388	9.075	7.926	7.126	6.179	5.522	5.084	3.886	2.947	2.221
	0.7	14.561	12.279	10.808	9.820	8.531	7.699	7.091	5.452	4.152	3.138
2	0	4.088	2.916	2.385	2.067	1.688	1.462	1.308	0.924	0.653	0.462
	0.1	8.272	6.135	5.159	4.582	3.917	3.540	3.249	2.348	1.726	1.312
	0.3	10.061	7.882	6.848	6.167	5.216	4.658	4.288	3.197	2.429	1.792
	0.5	12.843	10.204	8.804	7.902	6.803	6.068	5.575	4.234	3.196	2.400
	0.7	15.174	13.119	11.527	10.442	9.040	8.137	7.485	5.726	4.342	3.270

表6： 定理 2.3 中所考察的真实 1-Wasserstein 距离值，针对经验自协方差，考察不同滞后阶数 k 和样本量 n。数据来自一个 AR(1) 过程，其中 $\varepsilon_t \sim \mathcal{N} \in (0, 1)$，且 α 取一系列不同的值。

k	$\alpha \mid n$	25	50	75	100	150	200	250	500	1000	2000
0	0	0.129	0.091	0.074	0.065	0.053	0.046	0.041	0.029	0.020	0.014
	0.1	0.135	0.096	0.078	0.068	0.055	0.048	0.043	0.030	0.021	0.015
	0.3	0.191	0.136	0.112	0.097	0.080	0.069	0.062	0.044	0.031	0.022
	0.5	0.360	0.261	0.215	0.187	0.153	0.133	0.119	0.084	0.060	0.042
	0.7	0.970	0.717	0.595	0.519	0.428	0.372	0.333	0.237	0.168	0.119
1	0	0.051	0.026	0.018	0.014	0.009	0.007	0.005	0.003	0.001	0.001
	0.1	0.079	0.052	0.041	0.035	0.028	0.024	0.022	0.015	0.011	0.008
	0.3	0.209	0.149	0.122	0.106	0.087	0.075	0.067	0.048	0.034	0.024
	0.5	0.440	0.317	0.260	0.226	0.185	0.161	0.144	0.102	0.072	0.051
	0.7	1.132	0.828	0.684	0.596	0.490	0.425	0.381	0.271	0.192	0.136
2	0	0.062	0.032	0.022	0.016	0.011	0.008	0.007	0.003	0.002	0.001
	0.1	0.067	0.035	0.024	0.019	0.013	0.010	0.008	0.005	0.003	0.002
	0.3	0.144	0.098	0.078	0.067	0.054	0.047	0.042	0.029	0.021	0.015
	0.5	0.406	0.293	0.241	0.210	0.172	0.149	0.133	0.095	0.067	0.047
	0.7	1.206	0.881	0.727	0.633	0.520	0.451	0.404	0.287	0.203	0.144

表7： 根据第 3.2 节所述，取使上界最小化的 $m = m^*$，将定理 2.3 与 (2.14) 结合，得到不同滞后阶数 k 和样本量 n下经验自协方差的上界值。数据来自一个 AR(1) 过程，其中 $\varepsilon_t \sim \sqrt{12/14}\ t_{14}$，而 α 取一系列不同的值。

k	$\alpha \mid n$	25	50	75	100	150	200	250	500	1000	2000
0	0	0.912	0.645	0.527	0.456	0.372	0.322	0.288	0.204	0.144	0.102
	0.1	6.056	5.796	4.892	4.349	3.727	3.378	3.153	2.500	1.842	1.402
	0.3	9.007	7.383	6.423	5.888	5.109	4.573	4.217	3.270	2.512	1.905
	0.5	11.020	9.007	7.946	7.205	6.294	5.672	5.235	4.064	3.136	2.411
	0.7	14.064	11.794	10.500	9.621	8.441	7.673	7.099	5.554	4.312	3.335
1	0	2.564	1.818	1.485	1.286	1.050	0.909	0.813	0.574	0.406	0.287
	0.1	7.224	5.339	4.505	4.017	3.460	3.147	2.928	2.135	1.589	1.229
	0.3	8.995	6.966	6.092	5.490	4.675	4.202	3.890	2.950	2.274	1.719
	0.5	11.204	8.934	7.829	7.068	6.139	5.515	5.099	3.934	3.023	2.320
	0.7	14.097	11.893	10.569	9.656	8.449	7.658	7.089	5.518	4.267	3.290
2	0	4.088	2.916	2.385	2.067	1.688	1.462	1.308	0.924	0.653	0.462
	0.1	9.329	6.892	5.780	5.121	4.358	3.922	3.583	2.587	1.898	1.435
	0.3	10.481	8.065	6.983	6.240	5.282	4.721	4.347	3.262	2.485	1.847
	0.5	12.334	9.736	8.425	7.589	6.562	5.879	5.421	4.152	3.166	2.408
	0.7	14.490	12.447	10.980	10.019	8.750	7.917	7.322	5.676	4.369	3.354

表8： 定理 2.3 中所考察的经验自协方差的真实 1-Wasserstein 距离，适用于不同滞后阶数 k 和样本量 n。数据来自一个 AR(1) 过程，其中$\varepsilon_t \sim \sqrt{12/14}\ t_{14}$，而 α 取一系列不同的值。

k	$\alpha \mid n$	25	50	75	100	150	200	250	500	1000	2000
0	0	0.207	0.151	0.125	0.109	0.090	0.078	0.070	0.050	0.036	0.025
	0.1	0.213	0.155	0.128	0.112	0.092	0.080	0.072	0.051	0.037	0.026
	0.3	0.270	0.197	0.163	0.142	0.117	0.102	0.091	0.065	0.046	0.033
	0.5	0.444	0.325	0.270	0.235	0.194	0.168	0.151	0.108	0.076	0.054
	0.7	1.072	0.798	0.664	0.581	0.479	0.417	0.375	0.267	0.189	0.134
1	0	0.062	0.034	0.023	0.018	0.012	0.009	0.007	0.004	0.002	0.001
	0.1	0.092	0.061	0.048	0.041	0.033	0.028	0.025	0.018	0.012	0.009
	0.3	0.234	0.169	0.139	0.121	0.099	0.086	0.077	0.055	0.039	0.027
	0.5	0.483	0.351	0.289	0.252	0.207	0.179	0.161	0.114	0.081	0.057
	0.7	1.207	0.888	0.736	0.643	0.529	0.460	0.412	0.293	0.208	0.147
2	0	0.074	0.039	0.027	0.020	0.014	0.010	0.008	0.004	0.002	0.001
	0.1	0.078	0.042	0.029	0.022	0.016	0.012	0.010	0.005	0.003	0.002
	0.3	0.156	0.106	0.085	0.073	0.059	0.051	0.045	0.032	0.022	0.016
	0.5	0.428	0.311	0.257	0.224	0.184	0.159	0.143	0.101	0.072	0.051
	0.7	1.257	0.924	0.765	0.667	0.549	0.477	0.427	0.304	0.215	0.152