arXiv:2304.02421v1 [stat.ME] 5 Apr 2023

动态有向网络的时变 β-模型

Yuqing Du[1], Lianqiang Qu *[1]、Ting Yan[1], 和 Yuan Zhang[21]华中师范大学数学与统计学学院，中国湖北武汉，430079[2]俄亥俄州立大学统计学系，美国俄亥俄州哥伦布，43210

摘要

我们将著名的有向图 β-模型扩展到动态网络场景，其中观测的是不同时间点的邻接矩阵快照。我们提出一种核平滑似然方法，用于根据 N 个快照，估计包含 n 个节点的网络中的 $2n$ 个时变参数。我们证明了核平滑估计量的一致性和渐近正态性，这些性质在 n或 N 发散时均成立。我们的结果与单网络分析中的相应结论形成对比：在后者的渐近研究中，始终要求 $n \to \infty$ 。我们进行了全面的模拟研究，验证了理论预测，并从多个角度展示了我们的方法的性能。我们将我们的方法应用于一个电子邮件数据集，并得到了有意义的结果。

关键词： β-模型；有向网络；动态网络；核函数平滑

1 引言

在统计网络分析中，节点度数发挥着至关重要的作用，这既体现在其建模意义上，也体现在参数和非参数推断中的高计算效率上（Zhang andXia，2022 年）。在现有文献中，一类由度数驱动的模型，即 β-模型及其变体，受到了广泛关注和研究（Chatterjee 等人，2011；Hillarand Wibisono，2013；Olhede andWolfe，2012）。与随机块模型等其他常用网络模型相比，

*通信邮箱：qulianq@ccnu.edu.cn

β-模型为刻画具有潜在高度异质性的网络提供了一种简单而优雅的方式。 β-模型尤其具有吸引力的优势在于其参数估计便捷且效率极高（Chen 等，2021），另见 Hillar and Wibisono（2013）；Shao 等（2021）。值得注意的研究成果还包括：极大似然估计存在性（Rinaldo 等，2013）、中心极限定理（Yan 和 Xu，2013）、有向和二部 β-模型（Yan 等人，2016a；Fan等，2022）、加权边（Hillar and Wibisono，2013）、纳入协变量（Graham，2017）、正则化（Chen 等，2021；Stein 和 Leng，2021；Shao 等，2021）等。然而，关于β-模型的现有文献几乎都集中于静态网络；而时变数据在许多应用中十分常见，例如动态电子邮件网络和在线社交网络（Cui 和 Chen，2022）。因此，构建用于动态网络的时变 β-模型无疑具有重要意义。

本文填补了在建模、方法论和理论方面针对动态网络的 β-模型所存在的重要空白。我们的贡献有三方面。首先，我们通过扩展静态 β-模型，提出一种时变有向 β-模型，用于刻画有向动态网络中动态双度的演化。该模型包含 $2n$ 时变参数，每个节点分别关联一个出度变化参数和一个入度变化参数。其次，我们提出一种核平滑似然方法，利用邻近时间点的信息来估计 $2n$ 未知参数函数。这一方法受到静态网络经典 β-模型估计方法以及经典非参数统计中核方法的启发。第三，我们证明了所提出估计量的一致性和渐近正态性。与有向静态网络的结果（Yan 等人，2016a）相比，我们的分析表明，利用邻近时间点的信息能够显著提高估计精度，并放宽一致性和渐近正态性所需的假设。即使 n 保持固定，只要观测时间点的数量趋于无穷，这些结果依然成立。对于存在零入度或零出度的稀疏静态网络，有向β-模型的极大似然估计（MLE）不存在。只要这些节点在整个观测期内与其他节点相连，我们的核平滑估计量就能避免这一问题。数值研究和真实数据应用验证了我们的理论发现。

本文其余部分安排如下。首先，我们介绍一些记号；第 2节介绍我们提出的模型，设计估计方法和推断程序，并给出理论依据；第3节通过全面的模拟研究，探讨如何选择调优参数，从多个方面评估我们的方法的性能，并验证理论预测；第4节将我们的方法应用于一个真实数据集，并对结果进行解读；最后，我们在第 5节进行讨论。

1.1 记号

我们定义本文中使用的记号。对于向量 $\mathbf{x} = (x_1, x_2, \ldots x_n)^\top$，定义 $\|\mathbf{x}\|_\infty = \max_{1\le i\le n} |x_i|$。对于矩阵 $V := (v_{i,j})$，定义 $\|V\|_\infty$ 为 $\|V\|_\infty =$

$\sup_{\mathbf{x}\neq 0} \|V\mathbf{x}\|_\infty / \|\mathbf{x}\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |V_{i,\ j}|$. 定义 $\|V\|_{\max} := \max_{1\le\{i,j\}\le n} |V_{i,j}|$。我们需要近似求解某些具有特定结构的 V 矩阵的逆。具体而言，依照 Yan 等人（2016a）的做法，如果一个 $(2n-1)\times(2n-1)$ 矩阵 $V=(v_{i,j})$ 满足以下条件，则称其属于矩阵类 $\mathcal{L}_n(m,M)$: V satisfies: $m \le v_{i,j} - \sum_{j=n+1}^{2n-1} v_{i,j} \le M, \quad i=1,\ldots,n-1; v_{n,n} = \sum_{j=n+1}^{2n-1} v_{nj}; v_{i,j}=0, \quad i,j=1,\ldots,n, i\neq j; v_{i,j}=0, \quad i,j=n+1,\ldots,2n-1, i\neq j; m\le v_{i,j}=v_{ji}\le M, \quad i=1,\ldots,n, j=n+1,\ldots,2n-1, j\neq n+i; v_{i,n+i}=v_{n+i,i}=0, \quad i=1,\cdots,n-1;$ and $v_{i,i}=\sum_{k=1}^{n} v_{ki}=\sum_{k=1}^{n} v_{ik}, \quad i=n+1,\cdots,2n-1$。如果 $V\in\mathcal{L}_n(m,M)$，那么 V 是对称、逐元素非负且对角占优的，因此可逆。此外，根据 Yan 等人（2016a），V^{-1} 可以用 $S=\mathcal{S}(V)$ 很好地近似，如下所示

$$s_{i,j} = \begin{cases} \dfrac{\delta_{i,j}}{v_{i,i}} + \dfrac{1}{v_{2n,2n}}, & i,j=1,\ldots,n, \\ -\dfrac{1}{v_{2n,2n}}, & i=1,\ldots,n,\ j=n+1,\ldots,2n-1, \\ -\dfrac{1}{v_{2n,2n}}, & i=n+1,\ldots,2n-1,\ j=1,\ldots,n, \\ \dfrac{\delta_{i,j}}{v_{i,i}} + \dfrac{1}{v_{2n,2n}}, & i,j=n+1,\ldots,2n-1, \end{cases} \tag{1}$$

其中 $\delta_{i,\ j} := \mathbb{1}_{[i=j]}$ 以及 $v_{2n,\ i} = v_{i,\ 2n} := v_{ii} - \sum_{j=1,\ j\neq i}^{2n-1} v_{ij}$ 对于 $i=1,\ldots,2n-1$ 和 $v_{2n,\ 2n} = \sum_{i=1}^{2n-1} v_{2n,\ i}$。

2 我们的方法

2.1 数据结构及我们的模型

我们首先介绍数据结构。我们在 N 个随机选取的时间点 $T_1,\ldots,T_N$ 观察网络快照，这些时间点由概率密度函数 $f_T(t)$ 选取，其样本空间为 $[a,b]$，其中包含一些常数 $a<b$。所有网络快照在不同时间点共享同一个节点集 $[n]=\{1,\ldots,n\}$。在每个时间点 $t\in\mathcal{T}:=\{T_1,\ldots,T_N\}$，我们观察一个有向二元网络，并用其邻接矩阵 $A(t):=\{A_{i,j}(t)\}_{1\le\{i,j\}\le n}\in\{0,1\}^{n\times n}$ 表示，其中 $A_{i,\ j}(t)=1$ 表示在时间点 t 存在从 i 到 j 的有向边，而 $A_{i,\ j}(t)=0$ 表示不存在这种边。为简单起见，我们假设所有边彼此独立生成。我们假设不存在自环：$A_{i,\ i}(t)\equiv 0$。

为对这种数据结构建模，我们提出一种适用于有向网络的时变 β-模型。在任意时刻 $t\in\mathcal{T}$，每个节点 i 都与两个参数 $\alpha_i(t)$ 和 $\beta_i(t)$ 相关联，分别表示发送者效应和接收者效应的强度。从 i 到 j 的边概率为

$$P\big(A_{i,j}(t)=1\big) = \frac{e^{\alpha_i(t)+\beta_j(t)}}{1+e^{\alpha_i(t)+\beta_j(t)}} \quad \text{and} \quad P\big(A_{i,j}(t)=0\big) = 1 - P\big(A_{i,j}(t)=1\big). \tag{2}$$

与静态网络的有向网络和二部网络 β-模型类似（Yan 等人，2016a；Fan 等人，

2022），这里需要施加额外的正则条件以确保参数可识别性。为简单起见，我们设定 $\beta_n(t) = 0$(Yan 等人，2016a）。这也将保证我们的估计方程组（6）和（7）的解唯一，我们将在第 2.2 节介绍。令

$$\theta(t) := \Big(\alpha_1(t), \ldots, \alpha_n(t), \beta_1, \ldots, \beta_{n-1}(t)\Big)^\top,$$

表示 t 处的所有自由参数。

接下来，我们描述模型参数在不同时间点之间关系的假设。我们自然地预期，对于相近的时间点，$\theta(t)$ 和 $\theta(t')$ 应该相似 $t \approx t'$，除非它们之间存在变点；而相距较远的时间点上的参数则可能大不相同（也可能并非如此）。为简单起见，我们在方法开发和分析中采用连续的 $\theta(t)$设定。第5节末尾将讨论如何对我们的方法稍作调整以处理变点。

2.2 参数估计

我们采用半参数估计方法。我们方法的参数化特征体现在（2）中；而非参数化特征则在于，我们利用 $\theta(t)$ 在 t 上的平滑性来提高估计精度。给定观测到的时间点 $\mathcal{T}$，对数似然函数为

$$\begin{aligned}\mathcal{L}_0(\{(t,\theta(t))\}_{t\in\mathcal{T}}) &= \log\left[\prod_{t\in\mathcal{T}}\prod_{(i,j):1\le\{i\neq j\}\le n}\frac{e^{A_{i,j}(t)\{\alpha_i(t)+\beta_j(t)\}}}{1+e^{\alpha_i(t)+\beta_j(t)}}\right]\\ &= \sum_{t\in\mathcal{T}}\left\{\sum_{i=1}^{n}\Big(d_i(t)\alpha_i(t)+b_i(t)\beta_i(t)\Big) - \sum_{(i,j):1\le\{i\neq j\}\le n}\log\Big(1+e^{\alpha_i(t)+\beta_j(t)}\Big)\right\},\end{aligned} \tag{3}$$

其中，$d_i(t) = \sum_{j:1\le j\le n, j\neq i} A_{i,j}(t)$ 和 $b_j(t) = \sum_{i:1\le i\le n, i\neq j} A_{i,j}(t)$ 分别表示节点i 的出度和节点 j的入度。若在每个时间点分别对式（3）直接应用极大似然法，只会利用时间点 t处的 $A(t)$来估计 $\theta(t)$。也就是说，对于每个 $i \in [n]$ 和 $j \in [n-1]$，定义

$$F_{i;0}(t,\theta(t)) := d_i(t) - \sum_{\substack{j:1\le j\le n\\ j\neq i}}\frac{e^{\alpha_i(t)+\beta_j(t)}}{1+e^{\alpha_i(t)+\beta_j(t)}}, \quad t\in\{T_1,\ldots,T_N\} \tag{4}$$

$$F_{n+j;0}(t,\theta(t)) := b_j(t) - \sum_{\substack{i:1\le i\le n\\ i\neq j}}\frac{e^{\alpha_i(t)+\beta_j(t)}}{1+e^{\alpha_i(t)+\beta_j(t)}}, \quad t\in\{T_1,\ldots,T_N\} \tag{5}$$

$$F_0(t,\theta(t)) := (F_1(t,\theta(t)) \ldots, F_{2n-1}(t,\theta(t)))^\top \quad .$$

该方法通过求解 $\theta(t)$来估计 $F_0 = 0$。然后可以直接应用Yan 等人(2016a的算法和理论。本文将这种方法称为“逐点估计”，并在模拟研究中将其用作比较基准。

我们的目标是估计 $\theta(t)$，对所有 $t \in [a,b]$，而不只是对观测到的时间点 $t \in \mathcal{T}$。逐点估计量（4）和（5）无法实现这一目标。我们估计$\theta(t)$ 在所有 $t \in [a,b]$ 上的主要思路，是借助 T_ℓ 中接近 t 的部分所包含的信息——事实上，这也会提高时间点 $t \in \mathcal{T}$处的估计精度。基于这一思路，我们推广（4）和（5），将时间 $d_i(t)$ 处观测到的出度和入度 $b_i(t)$ 替换为其核平滑版本，以纳入邻近时间点的信息。更准确地说，定义

$$F_i(t,\theta(t)) := \sum_{\ell\in[N]} K_h(t-T_\ell)\left\{ d_i(T_\ell) - \sum_{\substack{j:1\le j\le n\\ j\neq i}} \frac{e^{\alpha_i(t)+\beta_j(t)}}{1+e^{\alpha_i(t)+\beta_j(t)}} \right\}, \quad t\in[a,b], 1\le i\le n \tag{6}$$

$$F_{n+j}(t,\theta(t)) := \sum_{\ell\in[N]} K_h(t-T_\ell)\left\{ b_j(T_\ell) - \sum_{\substack{i:1\le i\le n\\ i\neq j}} \frac{e^{\alpha_i(t)+\beta_j(t)}}{1+e^{\alpha_i(t)+\beta_j(t)}} \right\}, \quad t\in[a,b], 1\le j\le n-1 \tag{7}$$

其中，核函数 $K_h(u) := (1/h)\cdot K\big(u/h\big)$ 为偶函数，支撑集为 $[-1,1]$，且满足$\int_{\mathbb{R}} K(u)du = 1$。

$\widehat{}$我们的估计量记为 $\theta(t)$，它是估计方程组 $F(t,\theta(t)) = 0$.的解。为了使叙述流畅，带宽 h 的选择留待第 3.1 节讨论；这里我们重点求解 (6) 和 (7)，并采用牛顿法。$F(t,\theta(t))/(Nn)$ 的雅可比矩阵记为 $V(t,\theta(t))$，其形式如下：

$$V(t,\theta(t)) := \frac{1}{N}\sum_{\ell\in[N]} K_h(t-T_\ell) V_0(t,\theta(t)),$$

其中，$V_0(t,\ \theta(t)) = V_0(t,\ (\alpha(t)\ \beta(t)))$ 进一步定义如下：

$$\{V_0(t,\theta(t))\}_{i,n+j} = \{V_0(t,\theta(t))\}_{n+i,j} := \frac{1}{n}\frac{e^{\alpha_i(t)+\beta_j(t)}}{\left\{1+e^{\alpha_i(t)+\beta_j(t)}\right\}^2}, \quad i\in[n],\ j\in[n-1],\ i\neq j,$$

$$\{V_0(t,\theta(t))\}_{i,i} := \sum_{\substack{i:1\le i\le n\\ j\neq i}} \{V_0(t,\theta(t))\}_{i,n+j}, \qquad i\in[n],$$

$$\{V_0(t,\theta(t))\}_{n+j,n+j} := \sum_{\substack{j:1\le j\le n-1\\ i\neq j}} \{V_0(t,\theta(t))\}_{i,n+j}, \qquad j\in[n-1],$$

并且 $\{V_0(t,\theta(t))\}_{i,\ j} = 0$ 否则。由于 $V(t,\theta(t))$逐元素非负且严格对角占优，因此它是正定的。因此，当估计方程（6）和（7）的解存在时，该解必定唯一，并可通过梯度下降或$\widehat{}$牛顿法（Bubeck，2015）求得。我们将在定理1中讨论 $\theta(t)$的存在性，并正式刻画其精度。

2.3 估计量的理论性质

$\widehat{}$回顾一下，我们的估计量 $\theta(t)$通过求解（6）和（7）得到。本节将从$\widehat{}$三个方面建立 $\theta(t)$ 的理论保证：存在性、误差率和渐近正态性。令 $\theta^*(t)$：$=\left(\alpha_1^*(t)\ldots,\alpha_n^*(t)\ \beta_1^*(t)\ldots,\beta_{n-1}^*(t)\right)^\top$ 为真实参数。定义

$$Q_{Nnh}=\sup_{t\in[a,b]}\max_{i,j:1\le\{i\neq j\}\le n}\frac{(1+e^{\alpha_i^*(t)+\beta_j^*(t)})^2}{e^{\alpha_i^*(t)+\beta_j^*(t)}}.$$

熟悉 β-模型文献的读者可能会立即注意到，我们的Q_{Nnh} 与其对应形式之间存在类比；后者通常记为 b_n (Fan 等，2022），对应于静态网络情形。现在，我们描述一些所需的温和正则条件。

条件 1. 参数 $\theta^*(t)$ 关于各个分量均为二阶连续可微。此外，存在一个常数或发散的确定性序列 C_{Nn}，使得

$$\max_{i\in[2n-1]}\left|\frac{d\theta_i^*(t)}{dt}\right|\le C_{Nn}\quad\text{and}\quad\max_{i\in[2n-1]}\left|\frac{d^2\theta_i^*(t)}{dt^2}\right|\le C_{Nn},$$

其中 C_{Nn} 满足

$$C_{Nn}h\to 0\ \text{ as }\ Nn\to\infty.$$

Co件 2. 假设 $f(t)>0$ 对所有 $t\in[a,b]$ 均成立，且为二阶连续可微 .

注 1. 条件 1 允许 $\theta^*(t)$ 随着 $Nn\to\infty$ 的增大而变化得越来越快。这并不奇怪，因为增长的 Nn 会带来更多信息。条件 2 是一个非常温和的条件：我们以 $[a,b]$ 均衡方式选择观测时间，使得整个时间区间 $[a,b]$ 的任何部分都不会被忽略。令 $k_{21}:=\int v^2K(v)dv<\infty$. 现在，我们可以陈述关于估计量存在性和一致相合性的主要定理。

定理1。 假设条件 *1*和 *2*成立。设 $h\to 0$，$Nh\to\infty$ 且以下任一条件成立：n 保持不变，或 $n\to\infty$。若

$$Q_{Nnh}=o\left\{\frac{1}{\left(\sqrt{C_1\log(Nnh)/(Nnh)}+C_2C_{Nn}h^2\right)^{1/6}}\right\},\tag{8}$$

$\widehat{}$则估计量 θ（t）存在且满足

$$\sup_{t\in[a,b]}\|\widehat{\theta}(t)-\theta^*(t)\|_\infty=O_p\left\{Q_{Nnh}^3\left(\sqrt{C_1\frac{\log(Nnh)}{Nnh}}+C_2C_{Nn}h^2\right)\right\},\tag{9}$$

其中，$C_1>6$ 为常数，且 $C_2>k_{21}$ 上确界$t\in[a,b]$（f'（t）$+f$（t）$/2$）。

定理（8）中的条件 1 似乎比使（9）右端项趋于零所需的条件更强。人们自然会想知道它是否可以得到实质性改进。答案是“并非易事”。具体原因较为技术性，简而言之，（8）不仅$\widehat{}$有助于一致相合性，也保证解 $\theta(t)$ 的存在。事实证明，在分析中，解的存在性要求对 Q_{Nnh} 作出更强的假设，但这一点

并未在（9）中明确体现。关于 Q_{Nnh} 的假设强于表面上所需条件（即误差界趋于零所需条件）的现象，在许多相关 β-模型的研究中都有报道，包括（Chatterjee et al.，2011；Yan 等人，2016a；Chenetal.，2021；Steinand Leng，2021；Shaoet al.，2021）。

定理 1 中的一致收敛速率具有熟知的偏差—方差权衡形式，因此应谨慎选择带宽，以适当平衡偏差与方差。从理论上讲，如果我们知道 C_{Nn}，就可以通过解析方法优化 h 的选择。例如，如果 C_{Nn} 是常数，则可以选择 $h \asymp (Nn)^{-1/5}$，从而保证（9）的右端项为$o_p(1)$。但在实践中，可能并不总是容易确定 C_{Nn}。因此，在第 3节中，我们提出了一种留一交叉验证方法来选择 h。数值研究表明，该方法表现良好。

注 2. 关于 β-模型的经典文献通常假设网络非常稠密；关于这一问题的综述，参见 Chen 等人(2021)；Shao 等人(2021) 及其中的参考文献。在我们的研究中，为保证一致性，需要满足 $Q_{Nnh} = o\Big\{1/(\sqrt{C_1 \log(Nnh)/(Nnh)} + C_2 C_{Nn} h^2)^{1/6}\Big\}$。由此得到网络稀疏性的如下界限

$$\min_{1\le i<j\le n} P\big(A_{i,j}(t)=1\big) \ge \frac{1}{Q_{Nnh}} = \omega\left(\left\{\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2 C_{Nn} h^2\right\}^{1/6}\right). \tag{10}$$

为了更好地理解（10），令 $N = O(n^{\psi})$，并注意到当 $\psi > 1/4$ 时，我们可以设 $h = O\{(NnC_{Nn})^{-1/5}\}$，并进一步将（10）简化为

$$\left\{\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2 C_{Nn} h^2\right\}^{1/6} = \frac{\log(Nn)^{1/12} C_{Nn}^{1/30}}{(Nn)^{1/15}}.$$

考虑一个简单的情形，其中 $C_{Nn} = O(1)$。与 Yan 和 Xu（2013）中针对某个 C 所要求的 $\rho_n \gtrsim \log^{-C} n$，以及 Yan 等人（2016a）中针对 $\rho_n \gtrsim n^{-1/12}$ 的要求相比，我们允许网络更加稀疏（回想一下 $N \gg n^{1/4}$）。请注意，在这个简单的例子中，我们对 h 的选择仍旨在最小化（9）的右端项。如果网络变得更加稀疏，我们就需要选择更大的 h 值——这很好理解，因为如果我们从某个时间点附近的网络快照中观测到的边太少，自然的做法是扩大 h，以纳入更宽时间窗口内的数据；但这当然可能会增大偏差。未来一个值得探讨的方向是处理非常稀疏的网络，为此可能需要进行某种正则化（Chen 等人，2021；Stein 和 Leng，2021；Shao 等人，2021）。

接下来，我们将在温和条件下证明我们的估计量具有渐近正态性。为此，先引入一些简写。定义 $k_{02} = \int K^2\ v\ dv\ \ u_{ij}\ t\ \ = \{e^{\alpha_i^*\ t\ +\beta_j^*\ t}\}/\{1+: e^{\alpha_i^*\ t\ +\beta_j^*\ t}\}$ $\mu_{ij}\ t\ \ = u'_{ij}\ t\ f'\ t\ +\frac{1}{2}u''_{ij}\ t\ f\ t$ ( ), ( ): ( ) ( ) ( ) ( )，以及 ( ): ( ) ( ) ( ) ( )。为明确写出渐近分布公式，定义 $S(t,\theta(t)) = \mathcal{S}(V(t,\theta(t)))$，其中回顾式（1）中的 $\mathcal{S}(\cdot)$，并定义 $V(t,\theta(t)) = f(t)V_0(t,\theta(t))$。现在可以陈述结果了。

定理2。 假设条件*1*和*2*成立。设 $Nnh^5 C_{Nn}^2 Q_{Nnh}^2 = o(1)$，$h \to 0$，$Nh \to$

∞, n 要么保持不变，要么发散，并且

$$Q_{Nnh} = o\left\{\frac{1}{(Nnh)^{1/18}\left(\sqrt{C_1\log(Nnh)/(Nnh)} + C_2C_{Nn}h^2\right)^2}\right\},$$

其中， C_1 和 C_2 在定理*1*中定义。于是，对于任意固定的 $1 \le p \le 2n-1$ 和 $t \in [a,b]$，$\sqrt{Nnh}\Big[\{\widehat{\theta}_1(t) - \theta_1^*(t)\ldots,\widehat{\theta}_p(t) - \theta_p^*(t)\}^\top - k_{21}h^2\{\mu_1(t)\ldots,\mu_p(t)\}^\top\Big]$ 依分布收敛于一个 p 维多元正态分布，其均值为零，协方差矩阵由 $k_{02}\bar{S}(t,\theta^*(t))\bar{V}(t,\theta^*(t))\,\bar{S}(t,\theta^*(t))^\top$的左上角 $p\times p$ 块给出，其中

$$\mu_q(t) = \left\{\frac{\sum_{j\in[n],j\neq q}\mu_{qj}(t)/n}{f(t)\{V_0(t,\theta^*(t))\}_{q,q}}\right\}\{1 - \mathbb{1}_{n+1\le q\le 2n-1}\} + \left\{\frac{\sum_{j\in[n],j\neq q-n}\mu_{j,q-n}(t)/n}{f(t)\{V_0(t,\theta^*(t))\}_{q,q}}\right\}\mathbb{1}_{n+1\le q\le 2n-1}$$
$$+ (-1)^{n+1\le q\le 2n-1}\frac{\sum_{j\in[n-1]}\mu_{jn}(t)/n}{f(t)\{V_0(t,\theta^*(t))\}_{2n,2n}},$$

对所有 $q = 1,\ldots,p$, $\{V_0(t,\theta^*(t))\}_{2n,2n} = 2\sum_{i\in[n]}\{V_0(t,\theta^*(t))\}_{i,i} - \sum_{i\in[n]}\sum_{j\in[2n-1]}\{V_0(t,\theta^*(t))\}_{i,j}$.

注 3. 我们对定理 2 的结果作两点简要解读。首先，该定理允许 n 取固定值或趋于无穷大。若 $n\to\infty$，则 $p \le n$ 在某个点之后始终成立。其次，定理 $\{\sqrt{Nnh}k_{21}h^2\mu_p(t)\}$ 中的偏差项 2 一致有界于$O(\sqrt{Nnh}C_{Nn}Q_{Nnh}h^2)$，因此渐近趋于零。

为了将定理 2的结果付诸实践，我们估计渐近方差。注意，k_{02} 是已知的，且

$$\Sigma(t,\theta(t)) := \bar{S}(t,\theta(t))\bar{V}(t,\theta(t))\bar{S}(t,\theta(t))^\top$$

确实由 $V(t,\ \theta(t))$ 决定，其中回顾第 $\mathcal{S}(\cdot)$ 节中 1.1 的定义。令

$$\widehat{V}(t,\widehat{\theta}(t)) = \frac{1}{N}\sum_{\ell\in[N]}K_h(t-T_\ell)V_0(t,\widehat{\theta}(t)) \tag{11}$$

以及 $S(t,\widehat{\theta}_{\theta(t)})$，其 $\widehat{\ }(t)) = \mathcal{S}(V(t,\ \theta(t)))$，于是，估计的渐近方差为 $k_{02}\cdot\Sigma(t,$ e

$$\widehat{\Sigma}(t,\widehat{\theta}(t)) = S(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))S(t,\widehat{\theta}(t))^\top.$$

现在我们确认方差估计量确实是一致的。

命题 1. 假设条件*1* 和*2* 成立。若 $Nh\to\infty$ 满足 $N\to\infty$，则有 e

$$\|\widehat{\Sigma}(t,\widehat{\theta}(t)) - \Sigma(t,\theta^*(t))\|_{\max} = o_p(1).$$

在实践中，我们首先估计 $V(t,\theta(t))$，使用公式（11）；随后将其代入近似求逆公式 $\mathcal{S}(\cdot)$，在第1.1 节中得到 $S(t,\theta(t))$和 $\Sigma(t,\theta(t))$。

3 模拟

本节通过数值实验，从多个方面对我们的方法进行经验性考察，包括：不同 h 选择的影响、估计精度、

计算速度，以及数值结果与理论预测的吻合程度。代码可在在线补充材料中获取。在所有实验中，我们设定$[a,b]=[0.1,0.9]$和 f（t）$\sim$均匀分布（a，b）。为评估我们的方法对不同参数演化动态的估计效果，我们将α_i^*（t）的类型和β_j^*（t）的类型各设为四种，如表1所示。

表 1：模拟参数设置。

索引	$1\le i\le n/4$	$n/4+1\le i\le 2n/4$	$2n/4+1\le i\le 3n/4$	$3n/4+1\le i\le n-1$
$\alpha_i^*(t)$	$-0.8(3t-0.6)$	$-3.6(t-0.3)^2/(t-3)$	-0.4^t	$-3.2(t-0.5)^2/(1+t^2)$
$\beta_i^*(t)$	$2(t-0.5)^2/(t-2)$	$-1.6t(t-0.3)^3+(t-0.2)^2+0.2t$	$-1.8(t+0.6)$ 正弦$(0.2\pi t)$	$-1.6(t-0.2)^2\sin(\pi t)$

为确保模型的可识别性，设定 $\alpha_n^*(t)=\alpha_{n-1}^*(t)$ 和 $\beta_n^*(t)\equiv 0$ 。为便于叙述，在所有模拟中，我们均采用核函数 $K(x)=0.75(1-x^2)\mathbb{1}_{[|x|<1]}$。

3.1 模拟1：选择调优参数 h

在本次模拟中，我们考虑两种设置：$(n,\ N)=(40,\ 100)$ 和 $(160,\ 200)$。本次模拟旨在比较这两种设置下最优 h，并检验其比值是否符合定理 1 的理论建议。对于每种设置，我们通过留一交叉验证，从范围 h中选择最优 $[0.05,0.3]$，其损失函数为

$$L_{\mathrm{CV}}(h):=\sum_{\ell=1}^{N}\|A(T_\ell)-\widehat{W}_h(T_\ell)\|_F^2,$$

其中，$A(T_\ell)$ 表示时间 $t=T_\ell$ 的观测邻接矩阵，$\widehat{W}_h(T_\ell)$ 表示根据时间 $\{T_{\ell'}:\ \ell'=1、\ldots、N、\ \ell'\neq\ell\}$ 的数据估计得到的边概率矩阵。

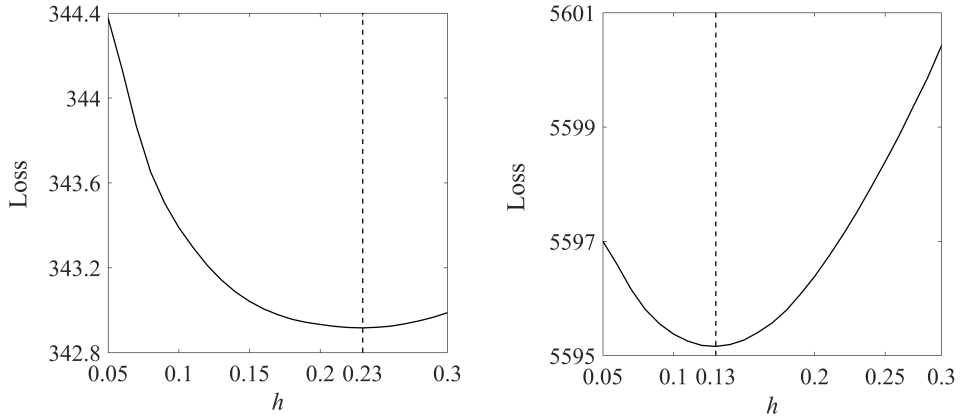

图1：左：$n=40$，$N=100$；右：$n=160$，$N=200$。两幅图均仅展示一次实验。

我们在两种设置下都重复进行了 50 次实验。交叉验证选出的平均最优 $h^{(n,N)}_{\mathrm{opt}}$

值分别为 $\bar{h}^{(40,100)}{}_{\text{opt}} = 0.225$ 和 $\bar{h}^{(160,200)}{}_{\text{opt}} = 0.140$。它们的比值为 0.622，与定理 1 的预测 $\{(40 \times 100)/(160 \times 200)\}^{(1/5)} = 0.659$。

3.2 模拟 2：均方根误差的递减速率与时间开销

在本次模拟中，我们重点评估网络规模 n 和时间点数量 N 变化时的估计误差。根据模拟1的结果和我们的理论，从本次模拟开始，我们设定 $h = 0.23\{(40 \times 100)/(Nn)\}^{1/5}$。我们改变 $n \in \{40, 80, 160, 320\}$ 和 $N \in \{25, 50, 100, 200\}$。估计误差通过以下均方根误差衡量：

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{\ell=1}^{N}(\widehat{\theta}(T_\ell) - \theta^*(T_\ell))^2}$$

我们将所有实验重复 100 次，并记录均方根误差的平均值。估计误差和运行时间报告于图 2中。从图中可以看出，log(均方根误差)

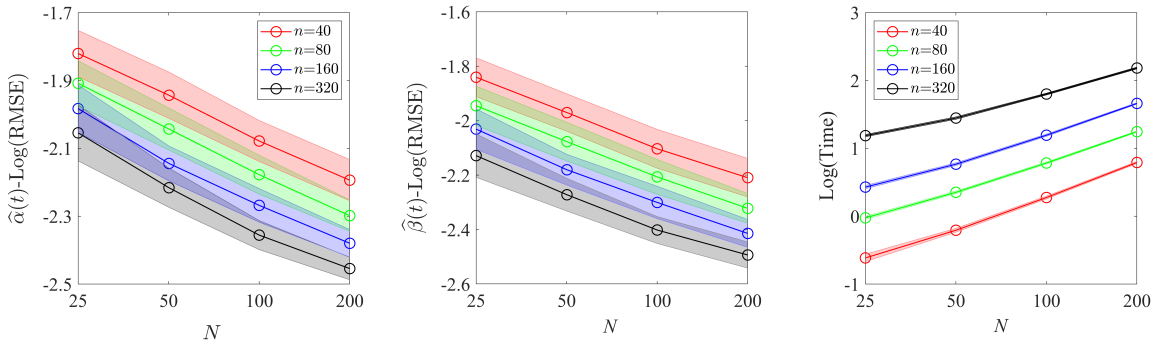

图 2：左图和中图分别为 $\alpha(t)$和 $\beta(t)$参数的对数估计误差。右图为对数时间开销。阴影表示重复实验中蒙特卡洛标准差。

以大约 $n^{-1/5}$ 的速率递减。我们的理论预测误差率为 $O(Q^3_{Nnh} \cdot n^{-2/5})$，与本次实验观察到的结果一致。具体而言，时间开销似乎与 N 呈线性关系，这也符合我们的理论预测；此外，当 n 加倍时，时间开销对数的纵向增量与 $\log 2 = 0.69$ 相当接近，这符合我们关于计算复杂度随 n 呈二次增长的理论认识。

3.3 模拟 3：拟合曲线与点态方法的比较

$\widehat{}$为评估我们的方法的性能，我们绘制真实曲线 $\theta(t)$，并将其与 $\theta(t)$进行比较。我们将我们的方法与两种替代方法进行比较：(1) 点态方法，该方法在每个 $\theta(t$处估计 $t \in \{T_1, \ldots, T_N\}$；以及 (2) 第2.2节所述的平滑点态方法，然后使用核平滑过程对所得估计进行平滑。

这里，我们根据模拟1的结果，将 $n = 160$、$N = 200$ 和 $h = 0.13$ 设为建议值。我们重复实验100次。各项设置见表1，其中给出了 $\alpha_i^*(t)$和 $\beta_j^*(t)$的设置。

图 3 和 4展示了结果；数值结果汇总于表2 和 3中。与平滑点态方法相比，我们的方法方差更小，偏差显著更低；相较于点态方法，我们的方法优势更为明显。如前所述，我们的方法可以利用相邻时间点的数据来改进估计；而由于数据有限，点态方法的曲线会在不同时间点之间波动 n ——这与我们的理论认识一致：将定理1 与单个时间点的经典误差率结果进行比较即可看出这一点，参见 Yan 等人 (2016a)。

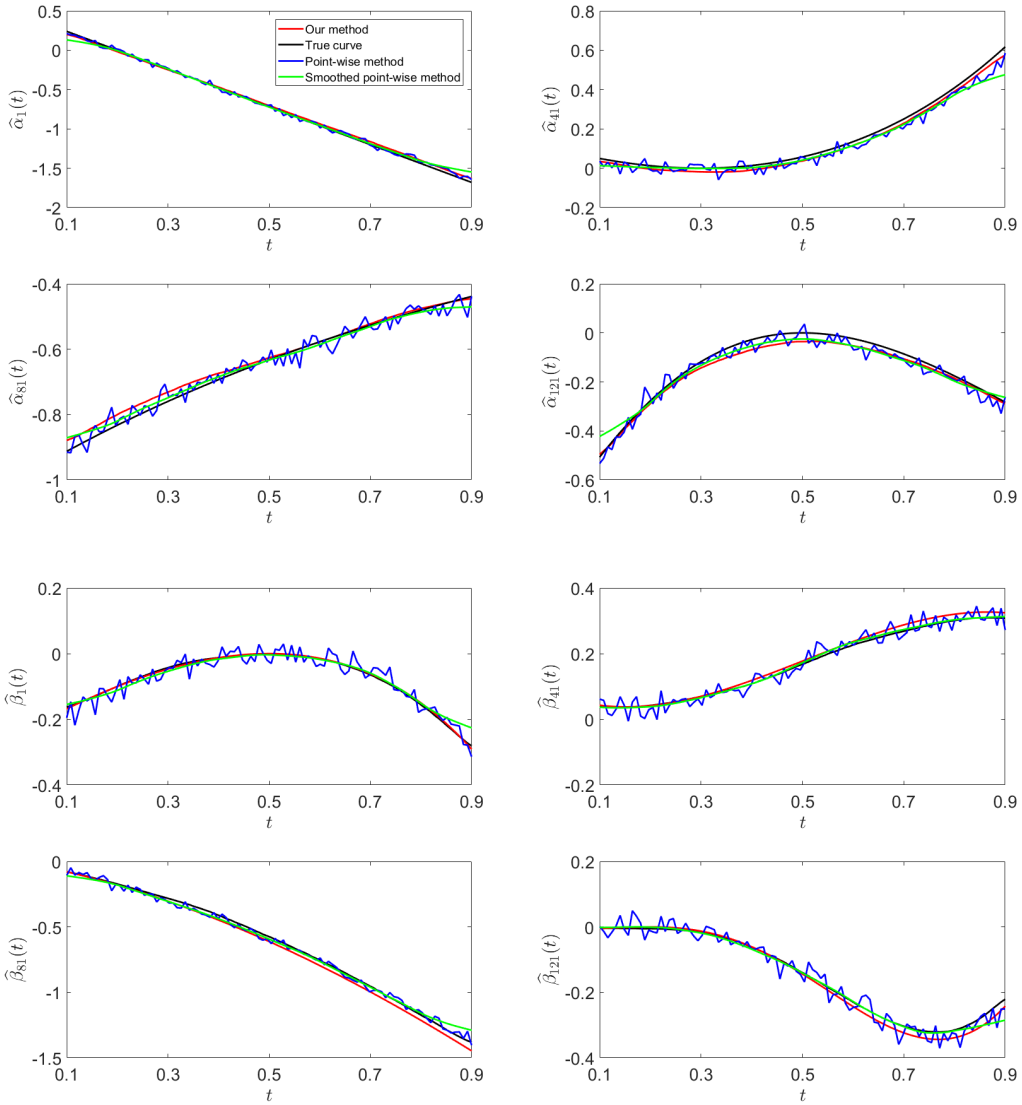

图 3： $\alpha_i^*(t)$ 和 $\beta_i^*(t)$（$i = 1$，41，81，121）的估计结果，其中 $n = 160$，N =200 和 $h = 0.13$.

接下来，我们在稀疏网络设置下进行实验。参数设置见表4。这里我们按照模拟 $n = 40$，$N = 100$ 和 $h = 0.23$ 的结果建议进行设置

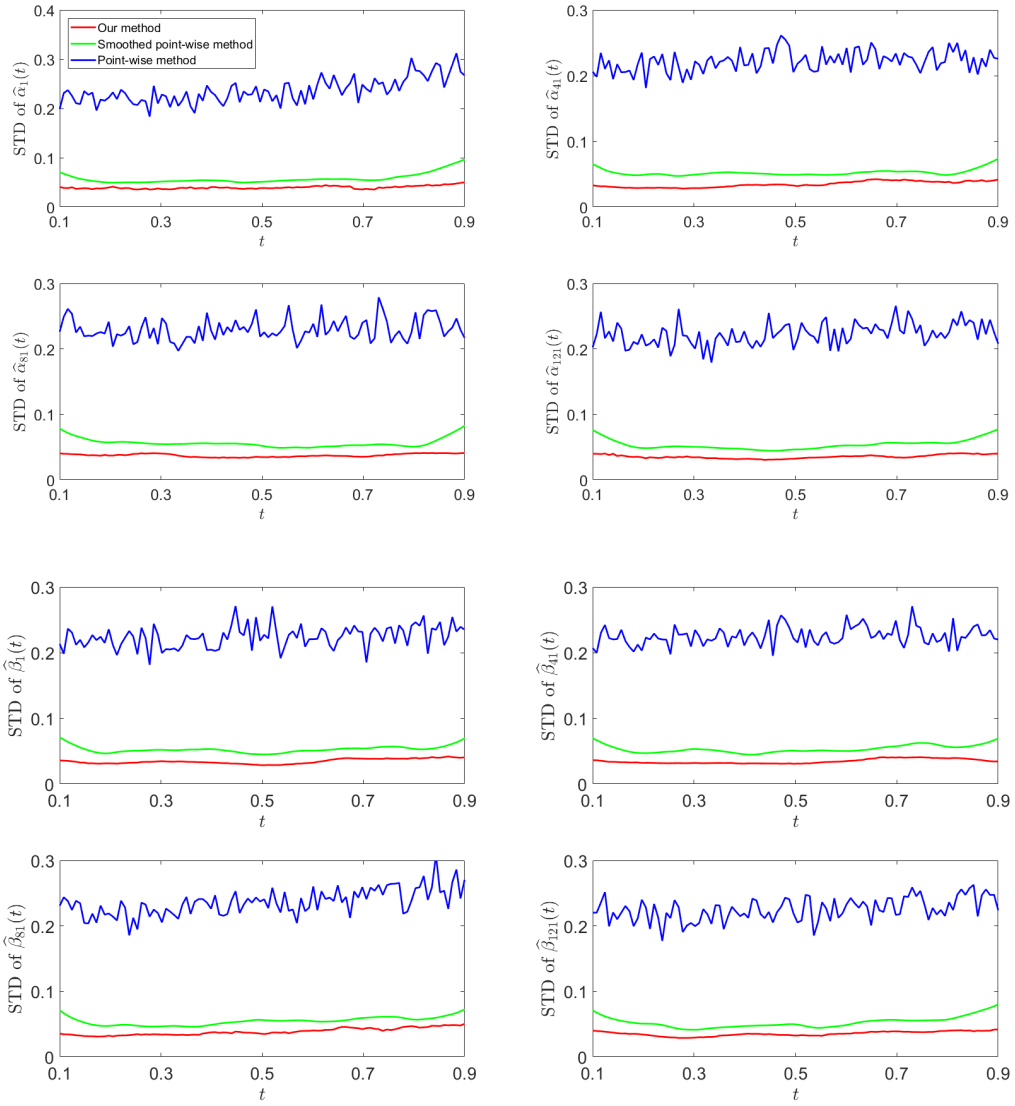

图4：估计量 $\widehat{\alpha}_i(t)$和 $\beta_i(t)$的样本标准差（标准差）（i =1, 41, 81, 121），其中 $n = 160, N = 200$ 和 $h = 0.13$.

1。结果如图5所示。我们看到，点态方法的估计量在某些时间点不存在，这也会影响平滑点态方法，因为它对逐点估计方法的结果进行平滑。与此形成鲜明对比的是，我们的方法在估计的最早阶段就纳入了相邻时间点的数据，因此表现稳定。

3.4 模拟4：渐近正态性

本次模拟旨在验证我们的渐近正态性理论，即定理2。为简便起见，我们考虑一种特定情形（n, N） =（160，200）。根据模拟1的结果，我们设定 $h = 0.13$。我们进行了1000次蒙特卡洛重复。图6和7 展示了结果。图示表明，模拟结果与以下内容高度吻合：

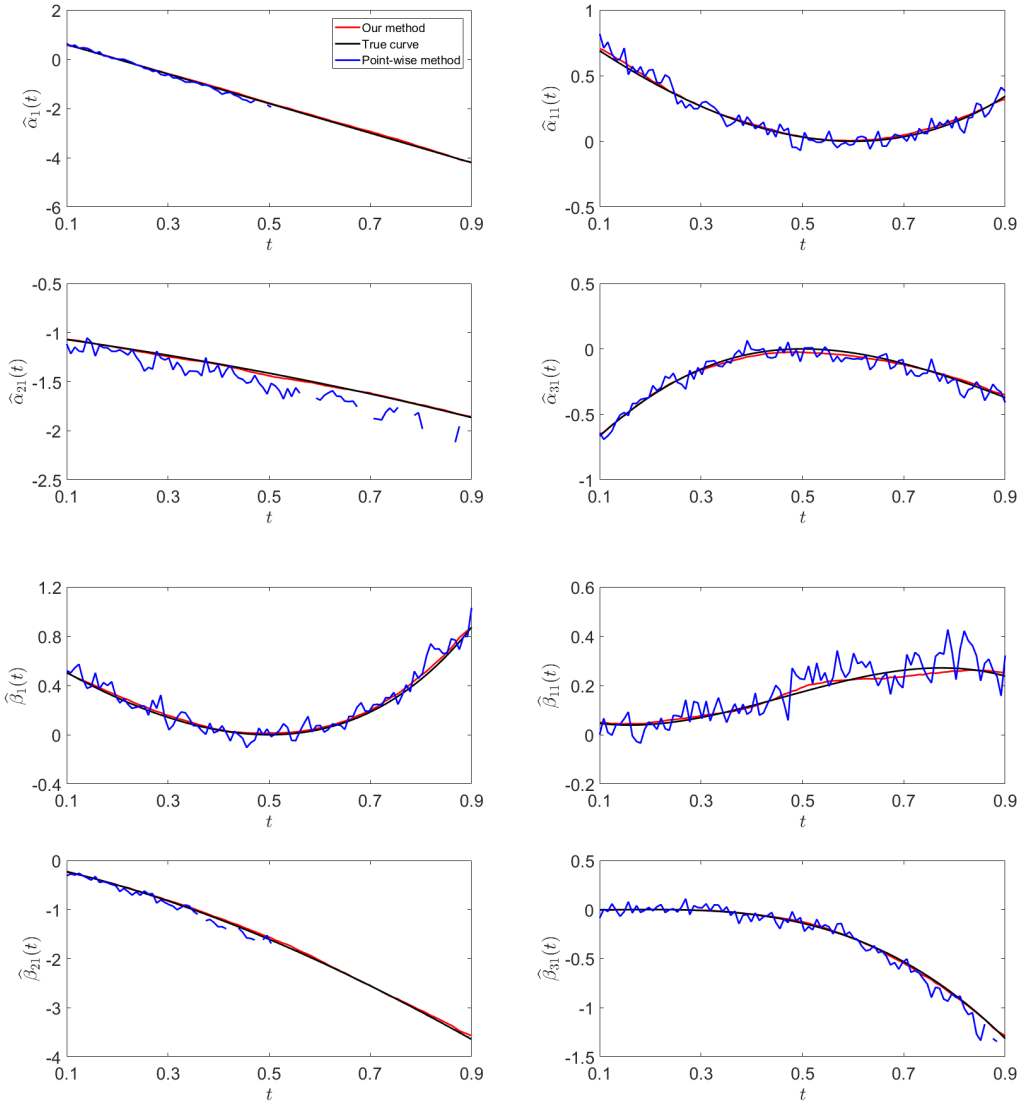

图5： $\alpha_i^*(t)$ 和 $\beta_i^*(t)$ 的估计结果（$i = 1$，11，21，31），其中 $n = 40$，$N = 100$和 $h = 0.23$.

定理2的预测。

4 数据示例

我们分析了 Leskovec 等人（2007 年）；Paranjape 等人（2017 年）的电子邮件数据集，该数据集可从 http: //snap.stanford.edu/data/email-Eu-core-temporal.html 获取。它记录了一家大型欧洲机构成员之间的电子邮件往来。每条记录（i, j, t）表示 i 在时间 t 给 j 发了电子邮件，其中 t 的精度为“天”。我们将各条记录汇总为按月划分的网络快照——时间单位过小会导致网络快照过于稀疏。这样得到 $N = 19$ 个快照，分别对应研究机构中四个选定部门在（n_1, n_2, n_3, n_4）=（298、154、89、140）个节点。为方便起见，我们将整个时间跨度重新缩放为

表 2: 当 $n = 160$, $N = 200$, $h = 0.1$ 时, $\widehat{\alpha}_i(t)$ 的偏差和标准差模拟结果 3.

		我们的方法		点态方法		平滑点态方法	
t	$\alpha^*(t)$	Bias	SD	Bias	SD	Bias	SD
0.2	$i = 1$	-0.004	0.041	0.019	0.233	-0.019	0.050
	$i = 41$	-0.008	0.029	-0.026	0.181	-0.017	0.049
	$i = 81$	0.013	0.037	0.016	0.204	0.033	0.057
	$i = 121$	-0.008	0.033	-0.022	0.205	-0.011	0.048
0.4	$i = 1$	-0.002	0.040	-0.037	0.227	0.012	0.054
	$i = 41$	-0.011	0.034	-0.047	0.216	-0.023	0.052
	$i = 81$	0.008	0.034	-0.016	0.214	0.018	0.055
	$i = 121$	-0.026	0.032	-0.056	0.209	-0.041	0.047
0.6	$i = 1$	0.004	0.042	-0.025	0.218	0.024	0.056
	$i = 41$	-0.019	0.038	-0.026	0.242	-0.019	0.050
	$i = 81$	-0.008	0.037	-0.018	0.238	0.000	0.049
	$i = 121$	-0.030	0.036	-0.045	0.208	-0.029	0.053

表 3: ^当 $n = 160$, $N = 200$, $h = 0.1$ 时, $\beta_i(t)$ 的偏差和标准差模拟结果 3.

		我们的方法		点态方法		平滑点态方法	
t	$\beta^*(t)$	Bias	SD	Bias	SD	Bias	SD
0.2	$i = 1$	0.001	0.031	0.018	0.212	-0.012	0.048
	$i = 41$	0.001	0.032	-0.007	0.223	-0.003	0.047
	$i = 81$	0.008	0.031	0.001	0.214	0.000	0.048
	$i = 121$	-0.005	0.034	-0.015	0.228	-0.001	0.051
0.4	$i = 1$	0.000	0.032	0.031	0.203	-0.004	0.053
	$i = 41$	-0.011	0.031	0.018	0.224	-0.011	0.045
	$i = 81$	0.037	0.038	0.056	0.247	0.010	0.050
	$i = 121$	-0.001	0.035	-0.003	0.212	-0.007	0.047
0.6	$i = 1$	-0.003	0.032	-0.018	0.220	-0.003	0.051
	$i = 41$	-0.010	0.035	-0.011	0.258	-0.002	0.050
	$i = 81$	0.040	0.040	0.059	0.260	0.031	0.054
	$i = 121$	0.014	0.035	-0.001	0.233	0.016	0.049

表4：模拟 3 中真实参数的第二组设定。

索引	$1 \le i \le n/4$	$n/4+1 \le i \le 2n/4$	$2n/4+1 \le i \le 3n/4$	$3n/4+1 \le i \le n-1$
$\alpha_i^*(t)$	$-2(3t-0.6)$	$-8(t-0.6)^2/(t-3)$	-2^t	$-4.2(t-0.5)^2/(1+t^2)$
$\beta_i^*(t)$	$-6(t-0.5)^2/(t-2)$	$-2(t-0.3)^3+(t-0.2)^2+0.2t$	$-4(t+0.8)\sin(0.2\pi t)$	$-5(t-0.2)^2$正弦$(0.2\pi t)$

$[0, 1]$。

我们使用第 3.1 节中的交叉验证方法调整带宽 h ，并分别为四个部门选择 $(h_1, h_2, h_3, h_4) = (0.21, 0.15, 0.46, 0.18)$。使用选定的$h$ 值，我们估计 $\{\widehat{\alpha}_i(t)\}_{i=1}^n$ 和 $\{\widehat{\beta}_j(t)\}_{j=1}^{n-1}$（回想一下，$\widehat{\beta}_n(t)$: $= \beta_n(t) \equiv 0$）。

然而，在实际数据分析中，我们需要非常谨慎地解释估计得到的 $\alpha_i(t)$ 和 $\beta_j(t)$。我们考虑一个简单示例，其中 $n = 2$；随着时间推移，节点 1 作为发送者和接收者都变得更加活跃，而节点 2 则变得越来越不活跃。人们可能会认为，这会导致 $\alpha_1(t)$ 和 $\beta_1(t)$ 增大，而 $\alpha_2(t)$ 和 $\beta_2(t)$ 减小。然而，出于可识别性方面的考虑，我们沿用有向 β-模型分析中的惯例，将$\beta_2(t) \equiv 0$。节点 2 作为接收者日益不活跃的行为会“传导”到所有其他参数 $\alpha_1(t)$ $\alpha_2(t)$ 和 $\beta_1(t)$ 上。因此，我们甚至无法确定 $\alpha_1(t)$ 和 $\alpha_2(t)$

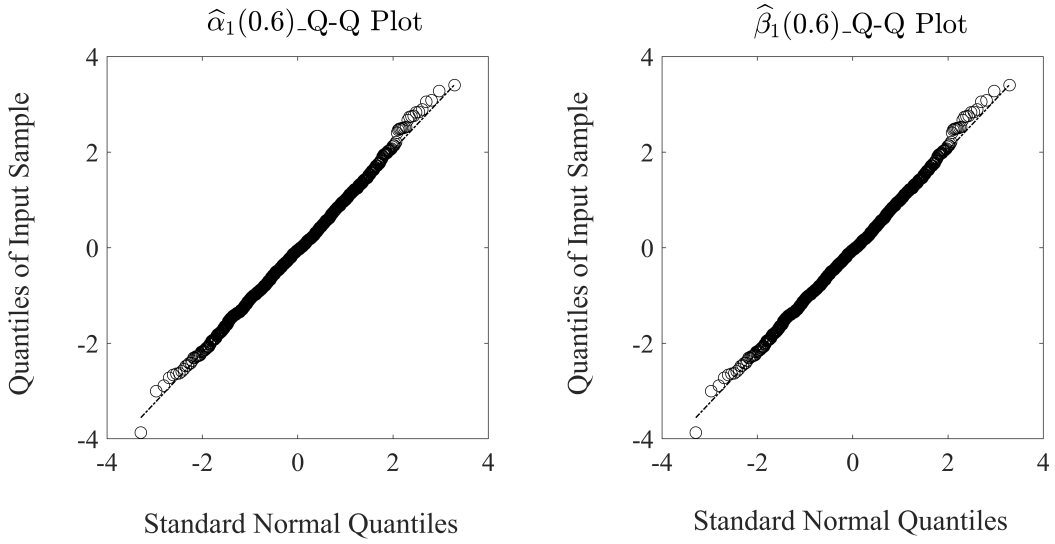

图6：一维边缘正态性。

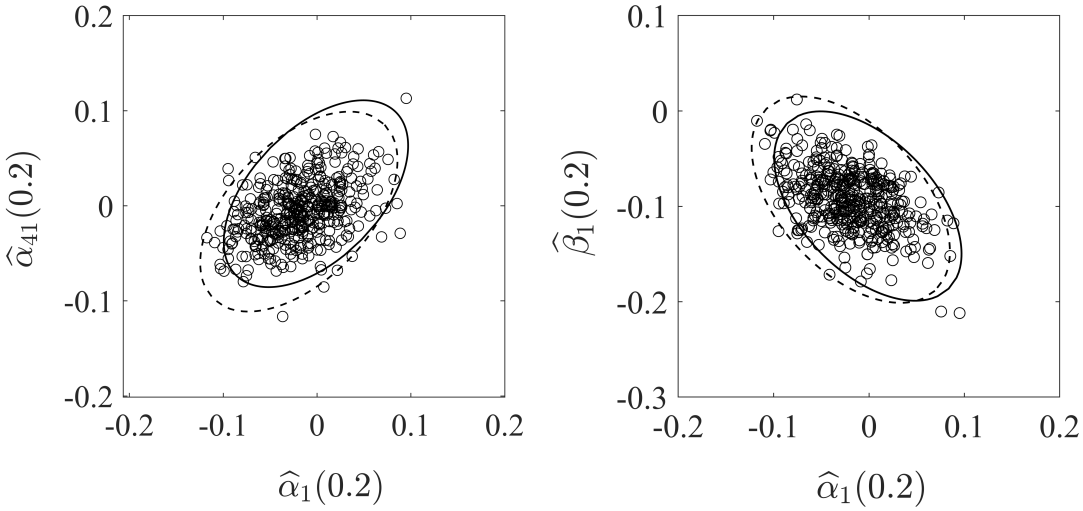

图7：二维边缘正态性。实线椭圆是 $(x-\mu)^{\top}\Sigma^{-1}(x-\mu)=9$，其中 (μ, Σ) 分别是定理 2预测的渐近中心和协方差矩阵。虚线椭圆是类似的等高线，但 (μ, Σ) 分别为样本均值和协方差矩阵。

是否会增大。

因此，我们不应仅仅分别考察每个单独的 $\alpha_i(t)$ 和 $\beta_j(t)$ 的趋势。幸运的是，上述可识别性问题不会影响任何 $i \neq j$ 下的 $\alpha_i(t)-\alpha_j(t)$ 或$\beta_i(t)-\beta_j(t)$。鉴于此，在获得 $\widehat{\alpha}_i(t)$ 的估计后，我们按如下方式计算成对距离矩阵 $\mathcal{D}_\alpha$

$$\left(\mathcal{D}_\alpha\right)_{i,j}^2 := \int_a^b \left|\widehat{\alpha}_i(t)-\widehat{\alpha}_j(t)\right|^2 dt \tag{12}$$

然后，我们使用 $\mathcal{D}_\alpha$ 进行 K 均值聚类，并为每个节点获取一个 α簇标签。需要说明的是，此处的聚类是根据节点的 $\alpha_i(t)$ 曲线随时间的变化对节点进行分类，其含义与动态随机块模型中的“变化的社区成员身份”概念大不相同。此外，对 $\mathcal{D}_\alpha$进行多维尺度分析（MDS），并投影到 $\mathbb{R}$，可为每个节点生成一维潜在空间位置。$\widehat{}$使用 $\beta_j(t)$ 的结果进行相同操作，可获得每个节点的 β 簇标签和 β-MDS 位置坐标。一个自然的问题是，如何为每个部门的两种聚类结构（α-聚类和 β-聚类）选择 K，即簇的数量。为此，我们针对不同的 K，计算平均簇间距离与平均簇内距离之比。这表明，当 $N = 19.$ 时，部门 1 的 α-聚类和 β-聚类应选择 $K = 4$；部门 2–4 的聚类应选择 $K = 5$。当 $N = 38.$ 时，部门 1 和部门 4 的 α-聚类和 β-聚类应选择 $K = 4$；部门 2–3 的聚类应选择 $K = 5$。当 $N = 57.$ 时，部门 2 的 α-聚类和 β-聚类应选择$K = 5$；其他部门的聚类应选择 $K = 4$。这一步的所有细节均见图8。接下来，我们在图9中展示

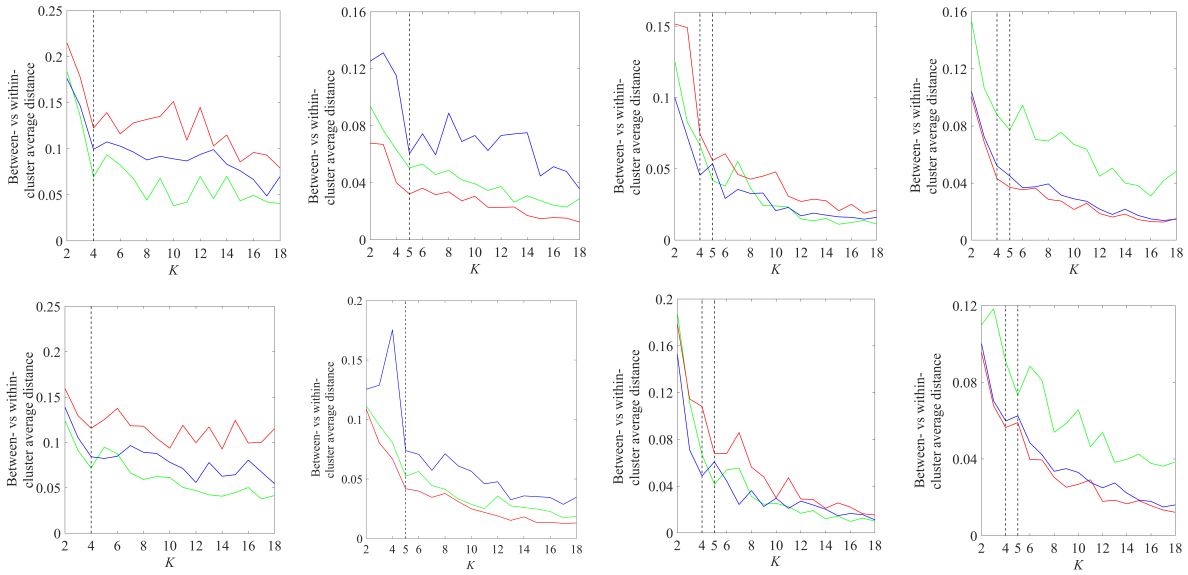

图8：绿色线条： $N = 19$；红色线条： $N = 38$；蓝色线条： $N = 57.$ 选择电子邮件数据集中各部门进行 α聚类和 β聚类的簇数。我们的选择在两方面之间进行了平衡：(i) 选择较小的 K；(ii) 最小化簇间平均距离与簇内平均距离之比。

所有节点的聚类结果以不同颜色和标记显示在散点图中，图中展示的是当 $N = 19$ 时的结果，以及当 $N = 38$ 和 $N = 57$ 时的聚类结果，详见补充材料。每个节点的 X-和 Y-坐标分别是其 α-和 β-多维尺度分析（MDS）坐标。每个节点的颜色表示其 α-聚类所属类别，标记表示其 β-聚类所属类别。下面我们对图9中的结果作一些解读。首先，我们注意到， α-和 β-聚类结果主要反映了节点的 α-和 β-多维尺度分析（MDS）坐标，但两者并不完全一致——回想一下， α-和 β-聚类使用各自对应的距离矩阵，而其MDS坐标可视为将距离矩阵非线性投影到一维子空间后的结果。图9 中的四个散点图均表明，节点的发送者和接收者效应随时间变化的趋势总体上彼此一致

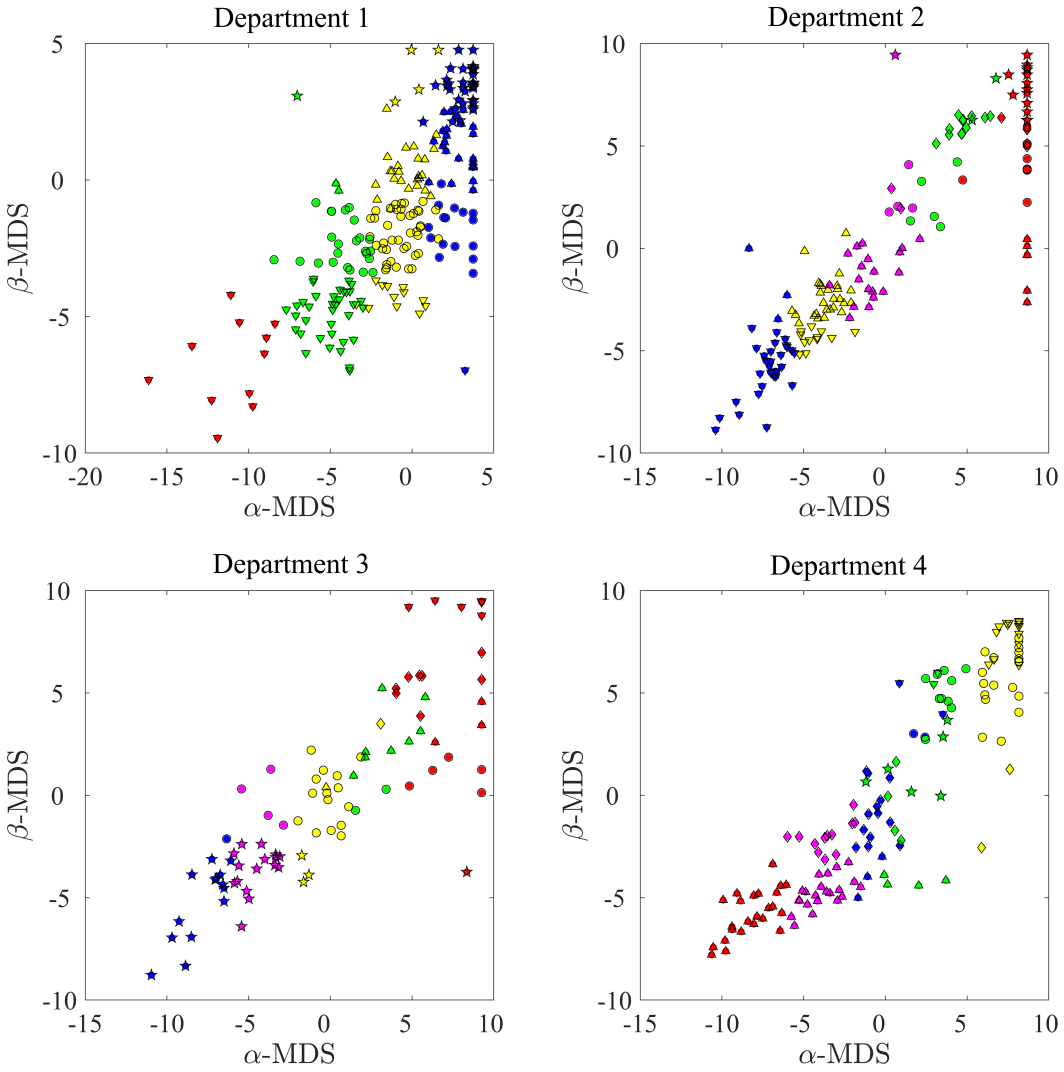

图9： $\widehat{\alpha}_i(t)$和 $\beta_j(t)$的示意图。每个点的 X坐标由 $\widehat{\alpha}_i(t)$之间 ℓ_2 距离矩阵的多维尺度分析得到；同理可得其Y坐标。颜色表示 α聚类结果，标记表示 β聚类结果。

因为大多数点沿对角直线排列；但这些趋势并非完全一致，这反映在 α-和 β-聚类结果之间的差异上。

5 讨论

本文通过令参数随时间演化，将有向二元网络的 β-模型扩展到了动态设置。可以将这项工作视为探索这一有趣方向的第一步。仍有许多引人入胜的问题有待研究。这里，我们选取几个未来研究方向进行讨论。

首先，我们的研究聚焦于二元边。回顾我们的方法可以发现，关键在于利用相邻时间点的对应项，对每个时间点的似然函数偏导公式进行核平滑。例如，我们可以将边生成方案从伯努利分布扩展到泊松分布，其中 $P(a_{ij}=x)=e^{x(\alpha_i+\beta_j)}/x!\cdot e^{-e^{\alpha_i+\beta_j}}$，$x=0,1,2,\ldots$. 这一思路与 Yan 等人 (2016b) 将普通的 β-模型从伯努利分布扩展到泊松分布及许多其他一般边分布的做法非常相似。详细的表述和理论研究留待未来工作。

第二个讨论主题是放宽对所有 $\alpha_i(t)$'s 和 $\beta_j(t)$'s 的全局光滑性假设，以适应变点（Wang 等人，2021；Bhamidi 等人，2018；Yu 等人，2021）。为此调整我们的方法，我们考虑单变点设定：$\alpha_i(t)$'s 和 $\beta_j(t)$'s 在 $t=t_0$ 处有一个共同的跳跃点，其余部分则满足条件 1 所要求的光滑性。假设 $t_0\in[a_1,b_1]\subset[a+C_a,b-C_b]$，其中 C_a 和 $C_b>0$ 是全局常数（我们假设 a_1 和 b_1 已知，但可以将其保守地设为接近 a 和 b 的值）。在这种情形下，将我们的方法简单调整用于变点检测，可得

$$\widehat{t}_0:=\arg\max_{t:t\in[a_1,b_1]}\left\|\left(\widehat{\alpha}_+(t)-\widehat{\alpha}_-(t),\widehat{\beta}_+(t)-\widehat{\beta}_-(t)\right)\right\|_2^2,$$

$$\text{where } \alpha_i(t)=\left(\widehat{\alpha}_+(t)\right)_i \text{ solves: } \sum_{\ell\in[N]:T_\ell\geq t}K_h(t-T_\ell)\left\{d_i(T_\ell)-e^{\alpha_i(t)}\left(\sum_{j:j\neq i}e^{\beta_j(t)}\right)\right\}=0,$$

$$\text{and } \alpha_i(t)=\left(\widehat{\alpha}_-(t)\right)_i \text{ solves: } \sum_{\ell\in[N]:T_\ell<t}K_h(t-T_\ell)\left\{d_i(T_\ell)-e^{\alpha_i(t)}\left(\sum_{j:j\neq i}e^{\beta_j(t)}\right)\right\}=0, \quad (13)$$

$\widehat{\ }$ $\widehat{\ }$并类似地定义 $\beta_+(t)$ 和 $\beta_-(t)$。除（13）之外，扩展该模型的另一种方式是推广权重函数，不仅考虑时间上的接近程度（由$|t-T_\ell|$ 对每个 t 和 T_ℓ所反映），还考虑其估计参数之间的接近程度（Li 等人，2011）。这两种有趣的扩展都需要进一步研究。

致谢

我们感谢主编 Jaakko Peltonen 教授、副主编以及两位审稿人提出富有见地且具有建设性的意见，这些意见显著提升了本文的质量。我们还感谢 Chenlei Leng 教授和 Subhabrata Sen 教授给予的宝贵讨论。Qu 获国家自然科学基金委员会资助（项目编号：11771171 和 12001219）。如需获取补充材料，请发送电子邮件至 qulianq@ccnu.edu.cn。

参考文献

Bhamidi, S., Jin, J. 和 Nobel, A.（2018）。网络模型中的变点检测：优先连接与长程依赖。《应用概率年刊》，28(1):35–78。Bubeck, S.（2015）。凸优化：算法与复杂性。《机器学习基础与趋势》，8(3-4):231–357。Chatterjee, S., Diaconis, P. 和 Sly, A.（2011）。具有给定度序列的随机图。《应用概率年刊》，21(4):1400–1435。

Chen, M., Kato, K. 和 Leng, C.（2021）。通过稀疏 β-模型分析网络。《英国皇家统计学会会刊：*B* 辑（统计方法学）》，83(5):887–910。Cui, X. 和 Chen, Y.（2022+）。推断动态网络中的社会影响。《统计学中国》（即将刊出），（即将刊出）。Fan, Y., Jiang, B., Yan, T. 和 Zhang, Y.（2022）。参数数量递增的二部图模型渐近理论。《加拿大统计学杂志》（即将刊出）。Graham, B. S.（2017 年）。具有度异质性的网络形成计量经济学模型。《计量经济学》，85(4):1033–1063。Hillar, C. 和 Wibisono, A.（2013）。图上的最大熵分布。*arXiv* 预印本 *arXiv:1301.3321*。Leskovec, J., Kleinberg, J. 和 Faloutsos, C.（2007 年）。图的演化：稠密化与直径缩小。《*ACM* 数据知识发现汇刊（*TKDD*）》，1(1):2–es。Li, Y., Zhu, H., Shen, D., Lin, W., Gilmore, J. H. 和 Ibrahim, J. G.（2011）。神经影像数据的多尺度自适应回归模型。《英国皇家统计学会会刊：*B* 辑（统计方法学）》，73(4):559–578。Masry, E.（1996）。时间序列的多元局部多项式回归：一致强一致性与收敛速率。《时间序列分析杂志》，17(6):571–599。Olhede, S. C. 和 Wolfe, P. J.（2012）。基于度的网络模型。*arXiv* 预印本 *arXiv:1211.6537*。Paranjape, A., Benson, A. R. 和 Leskovec, J.（2017）。时序网络中的模体。载于第十届 *ACM* 国际网络搜索与数据挖掘会议论文集，第 601–610 页。Rinaldo, A., Petrovic, S. 和 Fienberg, S. E.（2013）。β-模型中的极大似然估计，补充材料。《统计年鉴》，41:1085–1110。Shao, M., Zhang, Y., Wang, Q., Zhang, Y., Luo, J. 和 Yan, T.（2021）。大型稀疏网络中 β-模型的 L-2 正则化极大似然估计。*arXiv* 预印本 *arXiv:2110.11856*。Stein, S. 和 Leng, C.（2021）。用于稀疏有向网络的稀疏随机图模型。*arXiv* 预印本 *arXiv:2108.09504*。Wang, D., Yu, Y. 和 Rinaldo, A.（2021）。稀疏动态网络中的最优变点检测与定位。《统计年鉴》，49(1):203–232。Yan, T., Leng, C. 和 Zhu, J.（2016a）。有向指数随机图模型在双度序列递增时的渐近性质。《统计年鉴》，44(1):31–57。Yan, T., Qin, H. 和 Wang, H.（2016b）。以顶点强度为参数的无向随机图模型渐近性质。《统计学中国》，第 273–293 页。Yan, T. 和 Xu, J.（2013）。顶点数发散的无向随机图 β-模型中心极限定理。《生物计量学》，100(2):519–524。Yu, Y., Padilla, O. H. M., Wang, D. 和 Rinaldo, A.（2021）。网络在线变点定位的最优方法。*arXiv* 预印本 *arXiv:2101.05477*。Zhang, Y. 和 Xia, D.（2022）。网络矩的 Edgeworth 展开。《统计年鉴》，50(2):726–753。

附录：定理1，定理2 和命题1

定理1的证明。我们通过检验在线补充材料中引理1的条件，证明这些估计量的一致性。首先，我们计算

$$r = \|(F'(t,\theta^{(0)}(t))/(Nn))^{-1}F(t,\theta^{(0)}(t)/(Nn))\|_\infty.$$

注意，$-F'(t,\theta(t))/(Nn) \in \mathcal{L}_n(m,M)$. 如引理6所示，我们有

$$\frac{1}{Nn}F_i'(t,\theta(t)) \le \epsilon + \frac{1}{Nn}\mathbb{E}F_i'(t,\theta(t)),\ i \in [2n-1],$$

其中，$\epsilon > 0$ 是某个常数。此外，通过与引理6的证明类似的一些论证，我们有

$$\begin{aligned}
m &= \frac{1}{NnQ_{Nnh}}\sum_{\ell\in[N]} K_h(t-T_\ell) \ge (f(t) + \frac{1}{2}k_{21}f''(t)h^2 + o(h^2))/(2nQ_{Nnh}),\\
M &= \frac{1}{4Nn}\sum_{\ell\in[N]} K_h(t-T_\ell) \le (f(t) + \frac{1}{2}k_{21}f''(t)h^2 + o(h^2))/(2n).
\end{aligned}$$

由引理2和引理3，我们有

$$\begin{aligned}
r =& \|(F'(t,\theta^{(0)}(t))/(Nn))^{-1}F(t,\theta^{(0)}(t)/(Nn))\|_\infty\\
\le& \left(\frac{2c_4(2n-1)M^2\|F(t,\theta^{(0)}(t)/(Nn))\|_\infty}{m^3(n-1)^2} + \max_i \frac{|F_i(t,\theta^{(0)}(t))/(Nn)|}{\{V(t,\theta^{(0)}(t))\}_{i,i}} + \frac{|F_{2n}(t,\theta^{(0)}(t))/(Nn)|}{\{V(t,\theta^{(0)}(t))\}_{2n,2n}}\right)\\
\le& \left(\frac{2c_4(2n-1)M^2}{m^3(n-1)^2} + \frac{c_5}{(2n-1)mN}\right)\|F(t,\theta^{(0)}(t))/(Nn)\|_\infty\\
\le& O\left(\frac{Q_{Nnh}^3 + Q_{Nnh}}{f(t) + k_{21}f''(t)h^2/2 + o(h^2)}\right)\left(\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2C_{Nn}h^2\right)\\
\le& O\left(Q_{Nnh}^3\left(\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2C_{Nn}h^2\right)\right),
\end{aligned}$$

其中，c_4 和 c_5 是某些正常数，并且

$$\{V(t,\theta(t))\}_{2n,2n} = 2\sum_{i\in[n]}\{V(t,\theta(t))\}_{i,i} - \sum_{i\in[n]}\sum_{j\in[2n-1]}\{V(t,\theta(t))\}_{i,j}.$$

最后一个等式在条件2下成立。由引理1可知

$$\rho = \frac{c_1(2n-1)M^2H_1}{2m^3n^2} + \frac{H_2}{(n-1)m} \le O\left(Q_{Nnh}^3 + Q_{Nnh}\right)$$

其概率趋于1，其中 c_1 是某个正常数。因此，

$$\begin{aligned}\rho r &= O\Big(Q_{Nnh}^3 + Q_{Nnh}\Big) \times O\Big(Q_{Nnh}^3\big(\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2C_{Nn}h^2\big)\Big)\\ &= O\Big(Q_{Nnh}^6\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + Q_{Nnh}^4\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2C_{Nn}Q_{Nnh}^6h^2 + C_2C_{Nn}Q_{Nnh}^4h^2\Big)\\ &= o(1).\end{aligned}$$

若 $Q_{Nnh} = o\Big(1/(\sqrt{C_1\log(Nnh)/(Nnh)} + C_2C_{Nn}h^2)^{1/6}\Big)$，其小于 $1/2$，因为 $Nnh \to \infty$. 因此，

$$\sup_{t\in[a,b]}\|\widehat{\theta}(t) - \theta^*(t)\|_\infty = O_p\left(Q_{Nnh}^3\Big(\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2C_{Nn}h^2\Big)\right) = o_p(1).$$

证明至此完成。 □

定理 2 的证明。 根据定理 1，有

$$\widehat{b}_{Nnh} := \sup_{t\in[a,b]}\max_{i\in[2n-1]}\Big|\{\widehat{\theta}(t) - \theta^*(t)\}_i\Big| = O_p\Big\{Q_{Nnh}^3(\sqrt{C_1\frac{\log(Nnh)}{Nnh}} + C_2C_{Nn}h^2)\Big\}.$$

Let $\widehat{\omega}_{ij}(t) = \widehat{\alpha}_i(t) + \beta_j(t) - \alpha_i^*(t) - \beta_j^*(t)$。根据 $F(t,\ \theta(t))$，由泰勒展开可得

$$\begin{aligned}F(t,\widehat{\theta}(t)) =& F(t,\theta^*(t)) + F'(t,\theta^*(t))(\widehat{\theta}(t) - \theta^*(t)) + \widehat{\gamma}(t,\theta^*(t))\\ =& F(t,\theta^*(t)) + \frac{1}{Nn}F'(t,\theta^*(t))(Nn)(\widehat{\theta}(t) - \theta^*(t)) + \widehat{\gamma}(t,\theta^*(t))\\ =& F(t,\theta^*(t)) - V(t,\theta^*(t))(Nn)(\widehat{\theta}(t) - \theta^*(t)) + \widehat{\gamma}(t,\theta^*(t)),\end{aligned}$$

这意味着

$$\begin{aligned}&\widehat{\theta}(t) - \theta^*(t) - \frac{1}{Nn}V^{-1}(t,\theta^*(t))\mathbb{E}F(t,\theta^*(t))\\ =&\frac{1}{Nn}V^{-1}(t,\theta^*(t))\big\{F(t,\theta^*(t)) - \mathbb{E}F(t,\theta^*(t))\big\} + \frac{1}{Nn}V^{-1}(t,\theta^*(t))\widehat{\gamma}(t,\theta^*(t)).\end{aligned}$$

这里 $\widehat{\gamma}(t,\theta^*(t)) = \Big(\widehat{\gamma}_1(t,\theta^*(t))\ \widehat{\gamma}_2(t,\theta^*(t))\ldots,\widehat{\gamma}_{2n-1}(t,\theta^*(t))\Big)^\top$，其中

$$\begin{aligned}\widehat{\gamma}_i(t,\theta^*(t)) &= \sum_{j:1\le j\le n\atop j\ne i}\sum_{\ell\in[N]}K_h(t-T_\ell)\widehat{\gamma}_{ij}(t,\theta^*(t)), \quad i\in[n],\\ \widehat{\gamma}_{n+j}(t,\theta^*(t)) &= \sum_{i:1\le i\le n\atop i\ne j}\sum_{\ell\in[N]}K_h(t-T_\ell)\widehat{\gamma}_{ij}(t,\theta^*(t)), \quad j\in[n-1],\end{aligned}$$

with

$$\widehat{\gamma}_{ij}(t,\theta^*(t)) = -\frac{e^{\alpha_i^*(t)+\beta_j^*(t)+\widehat{\omega}_{ij}(t)\phi_{ij}}(1-e^{\alpha_i^*(t)+\beta_j^*(t)+\phi_{ij}\widehat{\omega}_{ij}(t)})}{2(1+e^{\alpha_i^*(t)+\beta_j^*(t)+\phi_{ij}\widehat{\omega}_{ij}(t)})^3}\widehat{\omega}_{ij}^2(t),$$

以及 $0 \le \phi_{ij} \le 1$. 由于 $\left|e^x(1-e^x)/(1+e^x)^3\right| \le 1$, 因此我们有

$$
\begin{aligned}
|\widehat{\gamma}_{ij}(t,\theta^*(t))| &\le |\widehat{\omega}_{ij}^2(t)/2| \le 2\widehat{b}_{Nnh}^2, \\
|\widehat{\gamma}_i(t,\theta^*(t))| &= \sum_{\substack{j:1\le j\le n \\ j\ne i}} \sum_{\ell\in[N]} K_h(t-T_\ell)|\widehat{\gamma}_{ij}(t,\theta^*(t))| \le 2N(n-1)\widehat{b}_{Nnh}^2/h, \quad i\in[n], \\
|\widehat{\gamma}_{n+j}(t,\theta^*(t))| &= \sum_{\substack{i:1\le i\le n \\ i\ne j}} \sum_{\ell\in[N]} K_h(t-T_\ell)|\widehat{\gamma}_{ij}(t,\theta^*(t))| \le 2N(n-1)\widehat{b}_{Nnh}^2/h, \quad j\in[n-1].
\end{aligned}
$$

根据下面给出的引理8-10, 我们有

$$
\begin{aligned}
\Big\|\frac{1}{Nn}V^{-1}(t,\theta^*(t))\widehat{\gamma}(t,\theta^*(t)\Big\|_{\max} &= o_p\big((Nnh)^{-1/2}\big), \\
\Big\|\frac{1}{Nn}V^{-1}(t,\theta^*(t))\mathbb{E}F(t,\theta^*(t)) - k_{21}h^2\mu(t)\Big\|_{\max} &= o_p\big((Nnh)^{-1/2}\big)
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{Nn}V^{-1}(t,\theta^*(t))\big\{F(t,\theta^*(t)) - \mathbb{E}F(t,\theta^*(t))\big\} \\
=&\frac{1}{Nn}\bar{S}(t,\theta^*(t))\big\{F(t,\theta^*(t)) - \mathbb{E}F(t,\theta^*(t))\big\} + o_p\big((Nnh)^{-1/2}\big),
\end{aligned}
$$

其中 $\mu(t)$ 定义见引理9。也就是说,

$$
\begin{aligned}
&\sqrt{Nnh}\Big[\{\widehat{\theta}(t) - \theta^*(t)\}_i - k_{21}h^2\mu_i(t)\Big] \\
=&\sqrt{Nnh}\Big[\Big\{\frac{1}{Nn}\bar{S}(t,\theta^*(t))(F(t,\theta^*(t)) - \mathbb{E}F(t,\theta^*(t)))\Big\}_i\Big] + o_p(1), \quad i\in[2n-1].
\end{aligned}
$$

对于任意 $i\in[n]$, 可以证明

$$
\begin{aligned}
&\sum_{\substack{j:1\le j\le n \\ j\ne i}} \sum_{\ell\in[N]} \mathbb{E}\Big|K_h(t-T_\ell)\big(a_{ij}(T_\ell) - u_{ij}(t)\big) - \mathbb{E}\big\{K_h(t-T_\ell)\big(a_{ij}(T_\ell) - u_{ij}(t)\big)\big\}\Big|^3 \\
\le &\sum_{\substack{j:1\le j\le n \\ j\ne i}} \sum_{\ell\in[N]} \mathbb{E}\Big|K_h(t-T_\ell)\big(a_{ij}(T_\ell) - u_{ij}(t)\big) - \mathbb{E}\big\{K_h(t-T_\ell)\big(a_{ij}(T_\ell) - u_{ij}(t)\big)\big\}\Big|^2/h \\
=&O\big(Nn/h^2\big),
\end{aligned}
$$

并且根据引理5, 我们有

$$
\mathrm{Var}\Big\{\sum_{\substack{j:1\le j\le n \\ j\ne i}} \sum_{\ell\in[N]} K_h(t-T_\ell)\big(a_{ij}(T_\ell) - u_{ij}(t)\big)\Big\} = O(Nn/h).
$$

因此，

$$\frac{\sum_{j\in[n],j\neq i}\sum_{\ell\in[N]}\mathbb{E}\Big|K_h(t-T_\ell)\big(a_{ij}(T_\ell)-u_{ij}(t)\big)-\mathbb{E}\big\{K_h(t-T_\ell)\big(a_{ij}(T_\ell)-u_{ij}(t)\big)\big\}\Big|^3}{\Big(\mathrm{Var}\Big\{\sum_{j\in[n],j\neq i}\sum_{\ell\in[N]}K_h(t-T_\ell)\big(a_{ij}(T_\ell)-u_{ij}(t)\big)\Big\}\Big)^{3/2}}$$
$$\leq O\Big(\frac{1}{\sqrt{Nnh}}\Big).$$

类似地，对于任意 $j\in[n-1]$,

$$\frac{\sum_{i\in[n],i\neq j}\sum_{\ell\in[N]}\mathbb{E}\Big|K_h(t-T_\ell)\big(a_{ij}(T_\ell)-u_{ij}(t)\big)-\mathbb{E}\big\{K_h(t-T_\ell)\big(a_{ij}(T_\ell)-u_{ij}(t)\big)\big\}\Big|^3}{\Big(\mathrm{Var}\Big\{\sum_{i\in[n],i\neq j}\sum_{\ell\in[N]}K_h(t-T_\ell)\big(a_{ij}(T_\ell)-u_{ij}(t)\big)\Big\}\Big)^{3/2}}$$
$$\leq O\Big(\frac{1}{\sqrt{Nnh}}\Big).$$

这意味着，在条件1-2、 $C_{Nn}^2Nnh^5Q_{Nnh}^2=o(1)$ $Nh\to\infty$ 和下，Lyapunov中心极限定理的条件成立

$$Q_{Nnh}=o\Big[1/\big\{(Nnh)^{1/18}(\sqrt{C_1\log(Nnh)/(Nnh)}+C_2C_{Nn}h^2)^2\big\}\Big].$$

因此，对于任意固定的 $1\leq p\leq 2n-1$ 和 $t\in[a,b]$, $\sqrt{Nnh}\Big[\{\widehat{\theta}_1(t)-\theta_1^*(t).....,\ \widehat{\theta}_p(t)-\theta_p^*(t)\}^\top-k_{21}h^2\{\mu_1(t).....,\ \mu_p(t)\}^\top\Big]$ 依分布收敛于一个均值为零、协方差为 $k_{02}\Sigma(t,\ \theta^*(t))$的 p维多元正态随机向量

直接计算可得

$$\mathrm{Cov}\left(\frac{\sqrt{Nnh}}{Nn}\bar{S}(t,\theta^*(t))\big\{F(t,\theta^*(t))-\mathbb{E}F(t,\theta^*(t))\big\}\right)=\bar{S}(t,\theta^*(t))\Omega(t,\theta^*(t))\bar{S}(t,\theta^*(t))^\top,$$

其中 $\Omega(t,\ \theta^*(t))=\mathrm{Cov}\big\{\sqrt{h/(Nn)}F(t,\ \theta^*(t))\big\}$.
接下来，我们计算 $\Omega(t,\ \theta^*(t))$。对于 $i\in[n]$，根据引理5，$\Omega(t,\ \theta^*(t))$ 的第 i 个对角元素为

$$\begin{aligned}
\big\{\Omega(t,\theta^*(t))\big\}_{i,i}=&\frac{h}{Nn}\mathrm{Var}\big\{F_i(t,\theta^*(t))\big\}\\
=&\frac{1}{n}\sum_{\substack{j:1\leq j\leq n\\ j\neq i}}\Big[k_{02}f(t)u_{ij}(t)(1-u_{ij}(t))+k_{12}h\tilde{\psi}_{ij}(t)+o(C_{Nn}h)\Big]\\
=&\frac{1}{n}\sum_{\substack{j:1\leq j\leq n\\ j\neq i}}k_{02}f(t)u_{ij}(t)(1-u_{ij}(t))+\frac{1}{n}\sum_{\substack{j:1\leq j\leq n\\ j\neq i}}k_{12}h\tilde{\psi}_{ij}(t)+o(C_{Nn}h)\\
=&k_{02}f(t)\big\{V_0(t,\theta^*(t))\big\}_{i,i}+k_{12}h\big\{\tilde{V}(t,\theta^*(t))\big\}_{i,i}+o(C_{Nn}h)\\
=&k_{02}\big\{\bar{V}(t,\theta^*(t))\big\}_{i,i}+k_{12}h\big\{\tilde{V}(t,\theta^*(t))\big\}_{i,i}+o(C_{Nn}h)+O(Nnh^5C_{Nn}^2),
\end{aligned}$$

其中 $k_{12} := \int vK^2(v)dv$, $\tilde{\psi}_{ij}(t) := u_{ij}(t)(1-u_{ij}(t))f'(t) + f(t)u'_{ij}(t)(1-2u_{ij}(t))$, 一个 d

$$\big\{\tilde{V}(t,\theta^*(t))\big\}_{i,i} = \frac{1}{n}\sum_{\substack{j:1\le j\le n\\ j\neq i}} \tilde{\psi}_{ij}(t).$$

对于任意 $j \in [n-1]$,

$$\begin{aligned}\big\{\Omega(t,\theta^*(t))\big\}_{n+j,n+j} =& \frac{h}{Nn}\mathrm{Var}\big\{F_{n+j}(t,\theta^*(t))\big\}\\ =& k_{02}\big\{\bar{V}(t,\theta^*(t))\big\}_{n+j,n+j} + k_{12}h\big\{\tilde{V}(t,\theta^*(t))\big\}_{n+j,n+j}\\ &+ o\big(C_{Nn}h\big) + O\big(Nnh^5C_{Nn}^2\big),\end{aligned}$$

其中 $\big\{\tilde{V}(t,\theta^*(t))\big\}_{n+j,n+j} = \frac{1}{n}\sum_{\substack{i:1\le i\le n\\ i\neq j}} \tilde{\psi}_{ij}(t)$. 此外,

$$\begin{aligned}\big\{\Omega(t,\theta^*(t))\big\}_{i,j} =& \frac{h}{Nn}\mathrm{Cov}\Big\{F_i(t,\theta^*(t)), F_j(t,\theta^*(t))\Big\} = 0,\ i\in[n],\ j\in[n],\ i\neq j,\\ \big\{\Omega(t,\theta^*(t))\big\}_{n+i,n+j} =& \frac{h}{Nn}\mathrm{Cov}\Big\{F_{n+i}(t,\theta^*(t)), F_{n+j}(t,\theta^*(t))\Big\} = 0,\ i,j\in[n-1],\ i\neq j,\\ \big\{\Omega(t,\theta^*(t))\big\}_{i,n+i} =& \frac{h}{Nn}\mathrm{Cov}\Big\{F_i(t,\theta^*(t)), F_{n+i}(t,\theta^*(t))\Big\} = 0,\ i\in[n-1],\\ \big\{\Omega(t,\theta^*(t))\big\}_{n+i,i} =& \frac{h}{Nn}\mathrm{Cov}\Big\{F_{n+i}(t,\theta^*(t)), F_i(t,\theta^*(t))\Big\} = 0,\ i\in[n-1].\end{aligned}$$

对于任意 $i \in [n], j \in [n-1], i \neq j$,

$$\begin{aligned}\big\{\Omega(t,\theta^*(t))\big\}_{i,n+j} =& \frac{h}{Nn}\mathrm{Cov}\Big\{F_i(t,\theta^*(t)), F_{n+j}(t,\theta^*(t))\Big\}\\ =& k_{02}\big\{\bar{V}(t,\theta^*(t))\big\}_{i,n+j} + k_{12}h\big\{\tilde{V}(t,\theta^*(t))\big\}_{i,n+j} + o\big(C_{Nn}h/n\big) + O\big(Nh^5C_{Nn}^2\big),\\ \big\{\Omega(t,\theta^*(t))\big\}_{n+j,i} =& \frac{h}{Nn}\mathrm{Cov}\Big\{F_{n+j}(t,\theta^*(t)), F_i(t,\theta^*(t))\Big\}\\ =& k_{02}\big\{\bar{V}(t,\theta^*(t))\big\}_{n+j,i} + k_{12}h\big\{\tilde{V}(t,\theta^*(t))\big\}_{n+j,i} + o\big(C_{Nn}h/n\big) + O\big(Nh^5C_{Nn}^2\big),\end{aligned}$$

其中 $\big\{\tilde{V}(t,\theta^*(t))\big\}_{n+j,i} = \big\{\tilde{V}(t,\theta^*(t))\big\}_{i,n+j} = \tilde{\psi}_{ij}(t)$. 现在我们证明

$$\Big\|\bar{S}(t,\theta^*(t))\Omega(t,\theta^*(t))\bar{S}(t,\theta^*(t)) - k_{02}\bar{S}(t,\theta^*(t))\bar{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))\Big\|_{\max} = o(1).$$

为此, 定义 $\tilde{V}(t,\theta^*(t)) \in \mathbb{R}^{(2n-1)\times(2n-1)}$, 其第 (i,j) 个元素为 $\big\{\tilde{V}(t,\theta^*(t))\big\}_{i,j}$. 于是, 我们有

$$\Omega(t,\theta^*(t)) = k_{02}\bar{V}(t,\theta^*(t)) + k_{12}h\tilde{V}(t,\theta^*(t)) + \Big[o\big(C_{Nn}h\big) + O\big(Nnh^5C_{Nn}^2\big)\Big]\mathbb{D}.$$

这里 $\mathbb{D} = (d_{ij}) \in \mathbb{R}^{(2n-1)\times(2n-1)}$, 且 $d_{ii} = 1\ (i\in[n])$ $d_{n+j,n+j} = 1\ (j\in[n-1])$ $d_{i,n+j} = d_{n+i,j} = 1/n\ (i\in[n],\ j\in[n-1],\ i\neq j)$, 且 $d_{ij} = 0$ 否则. 我们首先证明

$$\Big\|\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))\Big\|_{\max} = O(C_{Nn}Q_{Nnh}^2).$$

写出

$$\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t)) = \Big(\sum_{k=1}^{2n-1} \big\{\bar{S}(t,\theta^*(t))\big\}_{ik} \tilde{\psi}_{kj}(t,\theta^*(t))\Big).$$

具体而言，对于 $i=j\in[n]$，我们有

$$\begin{aligned}\Big| \sum_{k=1}^{2n-1} \big\{\bar{S}(t,\theta^*(t))\big\}_{i,k} \tilde{\psi}_{kj}(t,\theta^*(t))\Big| \leq & \frac{|\tilde{\psi}_{ii}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{i,i}} + \frac{|\tilde{\psi}_{ii}(t,\theta^*(t)) - \sum_{k=1,k\neq j}^{2n-1} \tilde{\psi}_{kj}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{2n,2n}} \\ = & \frac{|\tilde{\psi}_{ii}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{i,i}} + \frac{|\tilde{\psi}_{2n,j}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{2n,2n}} \\ \leq & \frac{C_2 C_{Nn} Q_{Nnh}}{f(t)}\Big(1+\frac{1}{n}\Big),\end{aligned}$$

其中，$\{\bar{V}(t,\ \theta^*(t))\}_{2n,\ 2n} = 2\sum_{i=1}^{n}\{\bar{V}(t,\ \theta^*(t))\}_{i,\ i} - \sum_{i=1}^{n}\sum_{j=1}^{2n-1}\{\bar{V}(t,\ \theta^*(t))\}_{i,\ j}$和

$$\tilde{\psi}_{2n,i}(t) = \tilde{\psi}_{i,2n}(t) := \tilde{\psi}_{ii}(t) - \sum_{j=1,j\neq i}^{2n-1} \tilde{\psi}_{ij}(t),\quad i\in[2n-1].$$

对于 $i=j=n+1,\ \ldots\ldots,\ 2n-1$，我们有

$$\begin{aligned}\Big| \sum_{k=1}^{2n-1} \big\{\bar{S}(t,\theta^*(t))\big\}_{i,k} \tilde{\psi}_{kj}(t,\theta^*(t))\Big| \leq & \frac{|\tilde{\psi}_{ii}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{i,i}} + \frac{|\tilde{\psi}_{ii}(t,\theta^*(t)) - \sum_{k=1,k\neq j}^{2n-1} \tilde{\psi}_{kj}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{2n,2n}} \\ = & \frac{|\tilde{\psi}_{ii}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{i,i}} \leq C_2 C_{Nn} Q_{Nnh}/f(t).\end{aligned}$$

对于 $i\in[n],\quad j\in[n-1],\quad i\neq j$，我们有

$$\begin{aligned}\Big| \sum_{k=1}^{2n-1} \big\{\bar{S}(t,\theta^*(t))\big\}_{i,k} \tilde{\psi}_{kj}(t,\theta^*(t))\Big| = & \frac{|\tilde{\psi}_{jj}(t,\theta^*(t)) - \sum_{k=1,k\neq j}^{2n-1} \tilde{\psi}_{kj}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{2n,2n}} \\ \leq & C_2 C_{Nn} Q_{Nnh}/(nf(t)).\end{aligned}$$

对于 $i\in[n]$、$j=n+1$、$\ldots\ldots$、$2n-1$、$j\neq n+i$，我们有

$$\Big| \sum_{k=1}^{2n-1} \big\{\bar{S}(t,\theta^*(t))\big\}_{i,k} \tilde{\psi}_{kj}(t,\theta^*(t))\Big| = \frac{|\tilde{\psi}_{ij}(t,\theta^*(t))|}{\{\bar{V}(t,\theta^*(t))\}_{i,i}} \leq C_2 C_{Nn} Q_{Nnh}/(nf(t)).$$

对于 $i=n+1,\ldots,2n-1, j\in[n],$，我们有

$$\Big| \sum_{k=1}^{2n-1} \big\{\bar{S}(t,\theta^*(t))\big\}_{i,k} \tilde{\psi}_{kj}(t,\theta^*(t))\Big| \leq C_2 C_{Nn} Q_{Nnh}/(nf(t)).$$

否则，我们有

$$\Big|\sum_{k=1}^{2n-1}\big\{\bar{S}(t,\theta^*(t))\big\}_{i,k}\tilde{\psi}_{kj}(t,\theta^*(t))\Big|=0.$$

这意味着

$$\Big\|\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))-\frac{C_2C_{Nn}Q_{Nnh}}{f(t)}\mathbb{1}\Big\|_{\max}\leq 2C_2C_{Nn}Q_{Nnh}/(nf(t)),$$

其中 $\nVdash\in\mathbb{R}^{(2n-1)\times(2n-1)}$ 是单位矩阵。因此，我们有

$$\begin{aligned}
&\Big\|\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))\Big\|_{\max}\\
=&\Big\|\Big(\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))-\frac{C_2C_{Nn}Q_{Nnh}}{f(t)}\mathbb{1}\Big)\bar{S}(t,\theta^*(t))+\frac{C_2C_{Nn}Q_{Nnh}}{f(t)}\bar{S}(t,\theta^*(t))\Big\|_{\max}\\
\leq&\Big\|\Big(\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))-\frac{C_2C_{Nn}Q_{Nnh}}{f(t)}\mathbb{1}\Big)\bar{S}(t,\theta^*(t))\Big\|_{\max}+\frac{C_2C_{Nn}Q_{Nnh}}{f(t)}\|\bar{S}(t,\theta^*(t))\Big\|_{\max}\\
\leq&\Big\|\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))-\frac{C_2C_{Nn}Q_{Nnh}}{f(t)}\mathbb{1}\Big\|_{\max}\times\max_{1\leq j\leq 2n-1}\sum_{k=1}^{2n-1}\Big|\big\{\bar{S}(t,\theta^*(t))\big\}_{k,j}\Big|\\
&+\frac{C_2C_{Nn}Q_{Nnh}}{f(t)}\Big\|\bar{S}(t,\theta^*(t))\Big\|_{\max}\\
=&O(C_{Nn}Q_{Nnh}^2).
\end{aligned}$$

也就是说，

$$\Big\|\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))\Big\|_{\max}=O(C_{Nn}Q_{Nnh}^2).$$

同理，我们可以证明

$$\Big[o\big(C_{Nn}h\big)+O\big(Nnh^5C_{Nn}^2\big)\Big]\Big\|\bar{S}(t,\theta^*(t))\mathbb{D}\bar{S}(t,\theta^*(t))\Big\|_{\max}=O(hC_{Nn}Q_{Nnh}^2).$$

这些事实意味着

$$\begin{aligned}
&\Big\|\bar{S}(t,\theta^*(t))\Omega(t,\theta^*(t))\bar{S}(t,\theta^*(t))-k_{02}\bar{S}(t,\theta^*(t))\bar{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))\Big\|_{\max}\\
\leq&\Big\|\bar{S}(t,\theta^*(t))\tilde{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))\Big\|_{\max}+\Big[o\big(C_{Nn}h\big)+O\big(Nnh^5C_{Nn}^2\big)\Big]\Big\|\bar{S}(t,\theta^*(t))\mathbb{D}\bar{S}(t,\theta^*(t))\Big\|_{\max}\\
=&O(hC_{Nn}Q_{Nnh}^2).
\end{aligned}$$

因此，如果 $hC_{Nn}Q_{Nnh}^2=o(1)$, 那么我们有

$$\Big\|\bar{S}(t,\theta^*(t))\Omega(t,\theta^*(t))\bar{S}(t,\theta^*(t))-k_{02}\bar{S}(t,\theta^*(t))\bar{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))\Big\|_{\max}=o(1).$$

证明至此完成。 □

命题的证明*1*。与引理 10 的证明一样，我们有

$$S(t,\theta(t))=\frac{1}{\widehat{f}(t)}S_0(t,\theta(t)) \quad \text{and} \quad \bar{S}(t,\theta(t))=\frac{1}{f(t)}S_0(t,\theta(t)).$$

接着，直接计算可得

$$\begin{aligned}
&\Big\|\widehat{\Sigma}(t,\widehat{\theta}(t))-\Sigma(t,\theta^*(t))\Big\|_{\max}\\
=&\Big\|S(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))S(t,\widehat{\theta}(t))^\top-k_{02}\bar{S}(t,\theta^*(t))\bar{V}(t,\theta^*(t))\bar{S}(t,\theta^*(t))^\top\Big\|_{\max}\\
=&\Big\|\frac{1}{\widehat{f}^2(t)}S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))S_0(t,\widehat{\theta}(t))^\top-\frac{1}{f^2(t)}k_{02}S_0(t,\widehat{\theta}(t))\bar{V}(t,\theta^*(t))S_0(t,\widehat{\theta}(t))^\top\Big\|_{\max}\\
\le&\Big\|\Big(\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big)S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))S_0(t,\widehat{\theta}(t))^\top\Big\|_{\max}\\
&+\Big\|\frac{1}{f^2(t)}S_0(t,\widehat{\theta}(t))\big(\widehat{V}(t,\widehat{\theta}(t))-k_{02}\bar{V}(t,\theta^*(t))\big)S_0(t,\widehat{\theta}(t))^\top\Big\|_{\max}.
\end{aligned}$$

接下来我们证明

$$\begin{aligned}
\Big\|\Big(\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big)S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))S_0(t,\widehat{\theta}(t))^\top\Big\|_{\max}=&O_p\bigg(Q_{Nnh}\Big\{\sqrt{\frac{\log(1/h)}{Nh}}+h^2\Big\}\bigg),\\
f^{-2}(t)\Big\|S_0(t,\widehat{\theta}(t))\big(\widehat{V}(t,\widehat{\theta}(t))-k_{02}\bar{V}(t,\theta^*(t))\big)S_0(t,\widehat{\theta}(t))^\top\Big\|_{\max}=&O_p\bigg(Q_{Nnh}\Big\{\sqrt{\frac{\log(1/h)}{Nh}}+h^2\Big\}\bigg).
\end{aligned}$$

记

$$S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))=\Big(\sum_{k\in[2n-1]}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}\Big).$$

那么，如果 $i=j\in[n]$，我们有

$$\begin{aligned}
&\Big|\sum_{k\in[2n-1]}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}\Big|\\
\le&\frac{|\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{i,i}|}{\big\{V_0(t,\widehat{\theta}(t))\big\}_{i,i}}+\frac{|\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{i,i}-\sum_{k=1,k\neq j}^{2n-1}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}|}{\big\{V_0(t,\widehat{\theta}(t))\big\}_{2n,2n}}\\
=&\frac{|\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{i,i}|}{\big\{V_0(t,\widehat{\theta}(t))\big\}_{i,i}}+\frac{|\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{2n,j}|}{\big\{V_0(t,\widehat{\theta}(t))\big\}_{2n,2n}}.
\end{aligned}$$

可以证明，以趋于 1 的概率，

$$|\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{i,i}|\le 2|\mathbb{E}\big\{\widehat{V}(t,\theta^*(t))\big\}_{i,i}|.$$

那么，我们有

$$\Big|\sum_{k=1}^{2n-1}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}\Big|=O_p(1+1/n).$$

类似地，我们可以证明

$$
\begin{aligned}
&|\sum_{k=1}^{2n-1}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}| = O_p(1), \quad i=j=n+1,\ldots,2n-1,\\
&|\sum_{k=1}^{2n-1}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}| = O_p(1/n), \quad i\in[n], j\in[n-1], i\neq j,\\
&|\sum_{k=1}^{2n-1}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}| = O_p(1/n), \quad i\in[n], j=n+1,\ldots,2n-1, j\neq n+i,\\
&|\sum_{k=1}^{2n-1}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}| = O(1/n), \quad i=n+1,\ldots,2n-1, j\in[n],\\
&|\sum_{k=1}^{2n-1}\big\{S_0(t,\widehat{\theta}(t))\big\}_{i,k}\big\{\widehat{V}(t,\widehat{\theta}(t))\big\}_{k,j}| = 0, \quad \text{otherwise.}
\end{aligned}
$$

这些事实意味着

$$
\big\|S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t)) - C_4\mathbb{1}\big\|_{\max} = O(1/n).
$$

以概率趋于1，其中某个常数为 C_4. 根据Masry (1996) 中的定理6，我们有

$$
\sup_{t\in[a,b]}|\widehat{f}(t)-f(t)| = O\Big(\Big\{\frac{\log(1/h)}{Nh}\Big\}^{1/2}+h^2\Big).
$$

这意味着

$$
\begin{aligned}
&\Big\|\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big|\times\Big\|S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))S_0(t,\widehat{\theta}(t))^\top\Big\|_{\max}\\
\le&\Big|\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big|\times\Big\|\big(S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))-C_4\mathbb{1}\big)S_0(t,\widehat{\theta}(t))+C_4\mathbb{1}S_0(t,\widehat{\theta}(t))\Big\|_{\max}\\
\le&\Big|\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big|\times\Big\|\big(S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))-C_4\mathbb{1}\big)S_0(t,\widehat{\theta}(t))\Big\|_{\max}\\
&+\Big|\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big|\times C_4\Big\|S_0(t,\widehat{\theta}(t))\Big\|_{\max}\\
\le&\Big|\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big|\times\Big\|S_0(t,\widehat{\theta}(t))\widehat{V}(t,\widehat{\theta}(t))-C_4\mathbb{1}\Big\|_{\max}\times\max_{j\in[2n-1]}\sum_{k=1}^{2n-1}\Big|\big\{S_0(t,\widehat{\theta}(t))\big\}_{k,j}\Big|\\
&+\Big|\frac{1}{\widehat{f}^2(t)}-\frac{1}{f^2(t)}\Big|\times C_4\|S_0(t,\widehat{\theta}(t))\Big\|_{\max}\\
=&O_p\Big(Q_{Nnh}\big\{\sqrt{\frac{\log(1/h)}{Nh}}+h^2\big\}\Big).
\end{aligned}
$$

通过一些与上述类似的论证，我们可以证明

$$
\Big\|\frac{1}{f^2(t)}S_0(t,\widehat{\theta}(t))\big(\widehat{V}(t,\widehat{\theta}(t))-k_{02}\bar{V}(t,\theta^*(t))\big)S_0(t,\widehat{\theta}(t))^\top\Big\|_{\max} = O_p\Big(Q_{Nnh}\Big\{\sqrt{\frac{\log(1/h)}{Nh}}+h^2\Big\}\Big).
$$

因此，如果 $Nh \to \infty$，那么

$$\|\widehat{\Sigma}(t, \widehat{\theta}(t)) - \Sigma(t, \theta^*(t))\|_{\max} = o_p(1)$$

在定理 1和2的条件下。至此，证明完成。 □