时空霍克斯过程中的后验集中

Xenia Miscouridou[1] 和 D´eborah Sulem[2]

[1]数学与统计系，塞浦路斯大学*[2]信息学学院，意大利语区瑞士大学´†

arXiv:2601.03719v1 [math.ST] 7 Jan 2026

摘要

我们为多元时空霍克斯过程中的推断建立了一个贝叶斯非参数框架，将现有理论结果从纯时间设定推广至更广泛的情形。该框架通过高斯过程先验，对霍克斯过程的背景强度和触发分量进行建模。在真实参数和非参数先验族满足适当的光滑性与正则性假设时，我们针对反复观测序列的渐近情形，推导出强度函数和参数的后验收缩速率。我们的分析将纯时间霍克斯过程的已知收缩结果推广至时空设定，从而能够联合建模时间和空间上的激发与聚类效应。据我们所知，这些结果首次为时空点数据中的贝叶斯非参数方法提供了理论保证。

关键词： 贝叶斯非参数方法、渐近理论、高斯过程

目录

1 引言 3

2 设置与方法论 4

2.1 设置 .4 2.2 方法论 .5

3 后验集中 7 3.1 一般结果 .7 3.2 高斯过程先验与 H¨older 光滑函数的应用 9

4 证明 10 4.1 命题1的证明 .10 4.2 命题 2 的证明 .11 4.3 命题 3 的证明 .13 4.4 命题4的证明 . 15 4.5 命题5的证明 .17

*miscouridou.xenia@ucy.ac.cy
†deborah.sulem@usi.ch

5 结论 **17**

A 技术引理 19
A.1 随机距离的界 .19A.2 伯恩斯坦不等式 .20

B 其他结果的证明 23
B.1 引理 4.1 的证明 .23B.2 引理 4.2 的证明 .24B.3 引理 4.3 的证明 .31B.4 引理 4.4 的证明 .36

1 引言

霍克斯过程是一类点过程模型，旨在刻画事件序列，其中事件发生的强度取决于过程的过去历史。其定义性特征是自激励：每个过去事件都会提高未来事件发生的可能性。霍克斯过程最初由 Alan G. Hawkes 于1971年提出，用于模拟地震的聚集现象 [Hawkes， 1971], 此后，霍克斯过程已广泛应用于多个领域，包括地震学 [Ogata、 1988], 社会与信息网络 [Crane and Sornette, 2008, Zhao et al.， 2015], 神经科学 [Reynaud-Bouret et al., 2014, Truccolo et al.， 2005], 动态网络分析 [Xu and Zha, 2016, Eichler et al., 2017],犯罪学 [Mohler et al., 2011, Miscouridou et al.， 2023], 以及流行病学 [Rizoiu et al.， 2018]。近年来，霍克斯过程也应用于机器学习，用于建模时间依赖关系、进行因果发现，并为大语言模型增添基于事件的动力学和记忆 [Mei and Eisner, 2017, Zuo et al., 2020, Huang et al., 2024, Hills et al.， 2024]。

霍克斯过程可视为非齐次簇泊松点过程，并具有自激强度函数。它也可以表示为具有潜在结构的分枝过程或簇过程。这种表示对于模拟和解释尤其有用，因为该过程可视为一系列事件级联，其中每个事件要么是外生生成的，要么由过去的某个事件（父事件）内生生成。霍克斯过程可以是单变量或多变量的。在后一种情况下，每个分量对应一种不同类型的事件，该过程等价于标记点过程。

最初，霍克斯过程被定义为单变量时间点过程 [霍克斯， 1971],其中，时刻 t_i 发生的每个事件都会提高时刻 $t > t_i$发生未来事件的概率。这种时间过程无法对时空事件数据中的空间效应进行建模；换言之，它假设事件的影响在整个空间中是均匀的，而这在许多应用中往往不切实际。在实践中，事件引起的激发效应可能同时取决于时间和空间邻近性，因此采用时空表述更为合适。事实上，近期研究强调了时空霍克斯过程在野火建模等应用中的作用 [Koh 等人，2023] 以及恐怖主义研究 [Jun 和 Cook， 2024]。Reinhart [2018]以及较新的 Bernabeu 等人 [2025]对相关研究进行了全面综述。尽管取得了这些进展，实际建模方法与理论理解之间仍存在显著差距，尤其是从贝叶斯视角来看。

从贝叶斯视角来看，确定后验收缩速率为模型随着数据量增加而学习真实自激机制的能力提供了基础性的理论验证。现有关于霍克斯过程后验收缩的研究几乎都集中于时间模型。Donnet 等人 [2020]在非参数设定下推导了多变量线性霍克斯过程的后验收缩速率，而 Sulem 等人 [2024] 则将分析扩展到考虑抑制效应的非线性霍克斯过程。关于点过程后验收缩的更一般数学框架见Donnet 等人 [2014]，；关于非齐次泊松过程以及使用高斯过程先验的类似结果，可参见Kirichenko 和 Van Zanten [2015] 以及Giordano 等人 [2025]。

在频率学派中，基于似然的霍克斯模型推断由来已久。Ogata[1978] 和 Ozaki [1979] 证明了平稳、单变量、指数型且纯时间霍克斯过程的最大似然估计（MLE）具有一致性和渐近正态性，而 Liniger [2009] 则将这些结果推广到了多变量情形。对于纯时间但非平稳的霍克斯模型，Chen 和 Hall [2013]以及 Kwan 等人 [2023] 研究了与本文密切相关的渐近设定下 MLE 的一致性。关于点过程 MLE 估计的更广泛结果，可参见 Daley 和 Vere-Jones [2003] 第 7 章。近期研究还针对多变量时间霍克斯过程的最小二乘估计，推导出了参数和非参数设定下的非渐近有限样本集中不等式 [Clinet 和

Yoshida，2017，Hansen 等人，2015年，Cai 等人， 2022]。

然而，据我们所知，现有的贝叶斯后验收缩结果和频率学派渐近结果均未涉及时空设定。尽管时空霍克斯过程在实证研究中的重要性日益提升，其贝叶斯推断的理论保证仍未得到探索。直到最近，相关研究才取得部分进展，例如 Siviero 等人 [2024], 提出了灵活的时空建模框架，但未给出渐近结果或收缩结果。这一空白促使我们开展本研究：对时空非平稳霍克斯过程进行严格的贝叶斯非参数分析，并建立其后验收缩性质。

为研究这类理论保证，可以采用不同的渐近设定，例如重复观测 [Dolmeta 和 Giordano, 2025b] ，或无限定义域 [Giordano 等人， 2025], ；针对每种设定，都需要不同的伯恩斯坦型不等式。

贡献。 在重复观测设定下，我们在贝叶斯非参数框架中，为非平稳时空霍克斯过程推导了后验收缩速率。该非参数框架通过对背景率和触发核分别采用定义在时空函数上的非参数先验族，使霍克斯过程的条件强度能够灵活设定。在真实参数满足适当正则性假设、先验族满足温和条件的情况下，我们推导出后验分布向真实值集中的显式速率。我们的结果适用于一般的非参数先验类别，尤其涵盖高斯过程先验；在现代点过程贝叶斯推断中，高斯过程先验是一种自然且广泛使用的选择 [Zhang 等人，2020年，Lloyd 等人，2015年，Malem-Shinitski 等人， 2022]。我们的证明结构与时间霍克斯过程相关论文的证明类似 [Donnet 等人，2020年，Sulem 等人， 2024] 但将其推广到时空情形并非易事，需要新的集中不等式。此外，与大多数以往的点过程研究不同，我们考虑的是重复观测设定，而非无限定义域。因此，我们的分析将现有时间霍克斯过程的后验收缩结果推广到了更一般的时空非平稳情形。

本文其余部分安排如下。第2 节介绍问题设定，并引入多元时空霍克斯过程模型，以及采用高斯过程先验描述背景强度和触发分量的贝叶斯非参数建模框架。第3节给出主要理论结果，在真实强度和先验满足适当正则条件的情况下，建立后验收缩速率。所有主要结果的证明见第4节。

2 设定与方法论

2.1 问题设定

我们假设在有界时空域 S 上，对点过程进行了重复观测。为简单起见，可以假设 $S = [0,1] \times [0,1]^d$ （注意，我们总可以将事件时间重新缩放至 $[0,1]$）[1]。在许多实际应用中，$d = 2$ （纬度和经度）。因此，数据由 n 个独立同分布的事件序列组成，每个序列包含带有时空坐标的事件，即，序列 i 是一组 m_i 个点 $N^i = (t^i_j,\ s^i_j)_{j \le m_i}$，其中 $s^i_j \in \mathbb{R}^d$ 且

$$t^i_1 < t^i_2 < \cdots < t^i_{m_i}$$

我们用 $N = (N^i)_{i \le n}$ 表示这些序列，并将后者建模为定义在 S 上的同一时空霍克斯过程 N (t, s) 的独立实现，如下所示。

[1]需要指出的是，只需改变参数的支撑集，我们的方法论便可轻松推广到一般有界定义域。

定义 1。 定义在定义域 S 上的时空点过程 $N(t,s)$ 是一个参数为 $f=(\mu,g)$ 的时空霍克斯过程，其中 $\mu \geq 0$ 和 $g \geq 0$ 分别为非负函数，称为背景率和触发核；若对于任意$(t,s) \in S$，其条件强度函数为

$$\begin{aligned}\lambda_{t,s}(f) = \lambda_{t,s}(\mu,g) &= \mu(t,s) + \int_{[0,t)\times[0,1]^d} g(t-t', s-s')dN(t',s') \\ &= \mu(t,s) + \sum_{(t_j,s_j)\in N, t_j<t} g(t-t_j, s-s_j),\end{aligned}$$

请注意，根据定义1 所定义的霍克斯过程，除非 μ随时间保持恒定，否则不是平稳的。

我们用$\mathbf{P}_f$表示参数为 f的霍克斯过程 $N(t,s)$的概率律，并用 $\mathbb{E}_f$表示相应的期望。对于子集 $A \subset S$，我们用 $N(A)$表示 A上的观测数量。我们还对触发核 g作出有限范围假设，即当 $t<0$ 或 $t>a$ 或$\|s\|_\infty > b$ 且 $0<a<1/2, 0<b<1/2$时， $g(t,s)=0$ 。这意味着我们可以将强度重写为

$$\lambda_{t,s}(f) = \mu(t,s) + \int_{t-a}^{t-} \int_{s'\in[0,1]^d:\|s-s'\|_\infty \leq b} g(t-t', s-s')dN(t',s'). \tag{1}$$

我们还作出另一项标准假设： $N(t,s)$的分枝比小于1，这意味着该过程是非爆炸的，即

$$\|g\|_1 := \int_0^a \int_{[0,1]^d \cap \{s:\|s\|_\infty \leq b\}} g(t,s)dtds < 1.$$

这里的统计目标是根据观测 N估计 f 。我们首先在关于背景率的温和假设下证明一个可辨识性结果，从而验证该估计问题的可行性。

假设1。 背景率 μ满足：

1. $\int_0^{1-a} \int_{[b,1-b]^d} \mu(t,s) > 0.$
2. $\mu(t,s) < +\infty, \quad \forall (t,s) \in S.$

假设1 确保背景率有限，且观测到至少一个事件的概率非零。

命题1（可识别性）。设 N 和 N' 是两个时空霍克斯过程，其对应参数分别为 $f=(\mu, g)$ 和 $f'=(\mu', g')$，且满足假设*1*。那么，

$$N \overset{d}{=} N' \iff f = f'.$$

命题1 的证明见第4节。

2.2 方法论

我们用 $f_0=(\mu_0, g_0)$ 表示真实参数，用$\mathbf{P}_0$ 和 $\mathbb{E}_0$ 分别表示霍克斯过程$N(t, s)$ 的概率分布及其期望。

现在我们来介绍针对时空霍克斯过程真实参数的贝叶斯非参数估计框架 f_0（定义1）。这里，我们重点讨论基于高斯过程（GP）变换的先验族，不过我们的理论结果（第3节）同样适用于更多

一般的先验族，例如贝塔密度混合或直方图先验（例如，参见Donnet 等人 [2020]和Sulem 等人 [2024]）。回顾一下，我们感兴趣的参数 $f = (\mu,\ g) \in \mathcal{F}$ 其中

$$\mathcal{F} = \{f = (\mu, g); \mu : S \to \mathbb{R}_+,\ g : [0, a] \times [0, b]^d \to \mathbb{R}_+\}.$$

我们在 $\mathcal{F}$ 上定义先验分布 Π，该分布在背景强度和触发核上进行因子分解，即：

$$\Pi(f) = \Pi_\mu(\mu)\Pi_g(g), \qquad f \in \mathcal{F}.$$

先验分布 Π_μ，Π_g 是定义在非负函数上的分布，通过对高斯过程进行变换而隐式构造。具体而言，

$$\begin{aligned} \mu &= \sigma(\nu), \qquad \nu \sim GP(0, k_\nu) \\ g &= \sigma(\phi), \qquad \phi \sim GP(0, k_\phi). \end{aligned}$$

上文中，ν 和 ϕ 是潜在函数，$\sigma:\ \mathbb{R} \to \mathbb{R}^+$ 是已知的连接函数，通常是在足够大的区间上严格递增且双射的函数，例如 Softplus 函数或 sigmoid 函数。此外，k_ν 和 k_ϕ 是定义在时空域 S 上的协方差函数（核）。为简单起见且不失一般性，我们在高斯过程先验中选择零均值函数。

高斯过程先验常用于点过程的贝叶斯非参数方法，例如用于非齐次泊松过程 [Adams 等人，2009年，Lloyd 等人，2015年，Kirichenko 等人，2015年，Palacios 和 Minin，2013年，Giordano 等人，2025年，Ng 和 Murphy，2019] 以及时间霍克斯过程 [Zhang 等人，2020年，Malem-Shinitski 等人，2022]。Miscouridou 等人 [2023], 使用带指数链接函数的高斯过程先验来估计霍克斯过程的时空背景率。相比之下，本文使用高斯过程先验，以非参数方式估计背景强度和触发核。关于高斯过程的介绍，可参见 Rasmussen 和 Williams[2005]。

在本节余下部分，我们将说明核函数和推断方法的可能选择。令 $u = (t,\ s) \in S$。常见的核包括平方指数核（RBF核）和 Mat´ern 核，其定义如下：

$$k_{RBF}(u, u') = \sigma \exp\left(-\frac{\|u - u'\|^2}{\ell^2}\right), \tag{2}$$

$$k_{Mat}(u, u') = \frac{\sigma^2}{\Gamma(\tau)2^{\tau-1}}\left(\frac{\|u - u'\|_2}{\ell}\right)^\tau B_\nu\left(\frac{\|u - u'\|_2}{\ell}\right), \tag{3}$$

其中包含超参数 $\sigma^2,\ \ell,\ \tau > 0$、$\Gamma$ 伽马函数，以及 B_τ第二类修正贝塞尔函数。我们注意到，当 $\tau \to \infty$ 时，Mat´ern 核等价于 RBF 核。在计算上，使用在时间和空间上可分离且平稳的核通常较为方便，即：

$$k_r(u, u') = k_{r,t}(|t - t'|)k_{r,s}(\|s - s'\|_2), \quad r \in \{\nu, \phi\}.$$

同样，RBF 核和 Mat´ern 核是时间核和空间核的常见选择$k_{r,\ t}$ 以及 $k_{r,\ s}$。我们注意到，可分离核并不意味着潜在函数 ν 或 ϕ（更不用说 μ 或 g）在空间和时间上是可分离函数。尽管如此，核及其超参数的选择决定了从高斯过程先验中采样得到的样本的光滑性（关于这一点的更多内容请参见第 3节）。

对 μ和 g 的推断随后通过后验分布进行。首先，我们定义观测集合 $N = (N^i)_{i=1,\dots,n}$ 的似然（例如参见 Daley 和 Vere-Jones [2003]）

$$L(N|f) = \prod_{i=1}^{n}\prod_{j=1}^{m_i}\lambda^i_{t^i_j,s^i_j}(f)\exp\left\{-\int_S \lambda^i_{t,s}(f)dtds\right\}, \qquad f \in \mathcal{F},$$

其中，对于每个 $i = 1, \ldots, n$,

$$\lambda^i_{t,s}(f) = \mu(t,s) + \int g(t-t', s-s')dN^i(t', s'). \tag{4}$$

后验分布定义为

$$\Pi(B|N) = \frac{\int_B L(N|f)d\Pi(f)}{\int_{\mathcal{F}} L(N|f)d\Pi(f)}, \qquad B \subset \mathcal{F}.$$

在实践中，后验分布的变分近似可能只能计算得到，例如定义为

$$\hat{Q}(f) = \arg\min_{Q \in \mathcal{Q}} KL(Q||\Pi(f|N)), \tag{5}$$

其中 KL是 Kullback-Leibler散度。在近似分布族 $\mathcal{Q}$上取最小值，例如，在 S上的高斯过程。该方法分别被 Lloyd 等人[2015] 、Zhang 等人 [2020], 和 [2020],Sulem 等人 [2022] 用于泊松过程和时间霍克斯过程。事实上，由于这里的后验分布不是共轭的，使用蒙特卡洛马尔可夫链技术从后验分布中采样的计算成本非常高。注意，在（5）中寻找最小化元等价于最大化定义为下式的证据下界（ELBO）：

$$ELBO(Q) = \mathbb{E}_Q[\log(L(N|f)\Pi(f))] - \mathbb{E}_Q[\log Q(f)].$$

3 后验集中

3.1 一般结果

本节分析观测序列数量为 $n \to \infty$ 时后验分布的渐近性质。具体而言，我们建立后验分布在强度函数 $\lambda_{t,s}(f)$ 以及参数 f 上的一般集中速率。为此，我们首先定义任意一对参数 f、$f' \in \mathcal{F}$之间的随机距离（强度函数上的 L_1-距离）如下

$$d_S(f, f') := \frac{1}{n}\sum_{i=1}^{n}\int_{[0,1]}\int_{[0,1]^d} |\lambda^i_{t,s}(f) - \lambda^i_{t,s}(f')|dtds = \frac{1}{n}\sum_{i=1}^{n}\|\lambda^i(f) - \lambda^i(f')\|_1,$$

w此处 $\lambda^i_{t,s}(f)$ i s defined as in (4) an^{d} $\lambda^i(f)$ d enotes t he correspon ding function from S to $\mathbb{R}^+$。然后，我们在参数上定义 L_1-距离如下

$$\|f - f'\|_1 := \|\mu - \mu'\|_1 + \|g - g'\|_1, \qquad f, f' \in \mathcal{F}.$$

我们的第一个结果是关于真实参数 f_0的 $\Pi(f|N)$ 在随机距离意义下的收缩，即,

$$\Pi(d_S(f, f_0) > M\epsilon_n|N_n) \xrightarrow[n\to\infty]{\mathbb{P}_0} 0. \tag{6}$$

其中 $\epsilon_n = o(1)$ 称为收缩率，而 $M > 0$ 是任意大的常数。在正式陈述我们的结果之前，我们先陈述关于 f_0， $\bar{\epsilon}_n$ 以及先验的假设。为将 RBF 核的情形（见第3.2节）纳入我们的理论，我们用两个序列ϵ_n， $\bar{\epsilon}_n = o(1)$ 来表述假设，使得 $\epsilon_n \geq \bar{\epsilon}_n$。

假设 2（有界参数）。 回顾一下， $f_0 = (\mu_0, g_0)$。我们假设 $\|g_0\|_1 < 1$ ， 且存在与 n 无关的常数 μ、 $\bar{\mu}$、 $\bar{g} > 0$ ， 使得对每个$(t, s) \in S$，

$$\underline{\mu} \leq \mu_0(t,s) \leq \bar{\mu}$$
$$0 \leq g_0(t,s) \leq \bar{g}.$$

假设3(先验质量**)**。设

$$B_\infty(\bar{\epsilon}_n) = \{f = (\mu, g); \|\mu - \mu_0\|_\infty + \|g - g_0\|_\infty \leq \bar{\epsilon}_n\}.$$

存在 $c_1 > 0$ 使得 $\Pi(B_\infty(\bar{\epsilon}_n)) \geq e^{-c_1 n\bar{\epsilon}_n^2}$。
假设 4 (筛集**)**。设 $\Lambda_{0,\ 2}: = \mathbb{E}_0 \left[\int (\lambda^1_{t,\ s}(f_0))^2 dtds\right]$ 且

$$\kappa := \frac{4\log 2}{\underline{\mu}} \left\{2 + 4\left(\frac{\bar{\mu}}{1 - \|g_0\|_1} + \Lambda_{0,2}\right)\right\}, \tag{7}$$

其中 $\underline{\mu}$, $\bar{\mu}$ 在假设 2中定义。存在 $\mathcal{F}_n \subset \mathcal{F}$ 和 $c_2 > c_1 + \kappa$， $\zeta_0 > 0$ 以及 $c_3 > 0$与 无关的常数，使得 $\Pi(\mathcal{F}_n) \geq 1 - e^{-c_2 n\bar{\epsilon}_n^2}$ 且

$$C(\zeta_0\epsilon_n, \mathcal{F}_n, \|.\|_1) \leq e^{c_3 n\epsilon_n^2}.$$

其中 $C(\epsilon_n, \mathcal{F}_n, \ \|.\|_1)$ 是 $\mathcal{F}_n$ 关于 L_1-范数、以半径 ϵ_n 的球构成的覆盖数。

假设3 和 4 与以下文献中的假设相似： Donnet 等人 [2020], Giordano 等人 [2025], Sulem 等人[2024]。在第 3.2 节中，我们将证明，在核函数满足温和条件时，这些假设对我们的基于高斯过程的先验均成立。假设 3 要求真实参数有界f_0 ，在实践中并不苛刻。点过程文献中通常也会假设类似的上界，例如参见 Giordano 等人 [2025]。下界确保任意点处事件发生的概率都非零 $(t, \ s) \in S$。

命题 2（随机距离上的集中）。 在假设 2、3 和 4 成立，且 $n\bar{\epsilon}_n^2 \to \infty$且 $\epsilon_n = o((\log n)^{-2})$ 时，(6) 成立。

虽然随机距离上的后验集中能够提供预测保证，但它并非参数空间上的显式距离。因此，为了获得关于参数解释的保证（例如，事件生成过程中内生效应/外生效应的大小），我们建立了第二个结果，即关于 $\mathcal{F}$ 上 L_1-距离的后验集中速率。

命题 3（在 L_1-距离下的集中性）。 在假设2、3、4下，且若 $n\bar{\epsilon}_n^2 \to \infty$且$\epsilon_n = o$（$(\log n)^{-2}$），则

$$\Pi(\|f - f_0\|_1 > M'\epsilon_n | N_n) \xrightarrow[n\to\infty]{\mathbb{P}_0} 0, \tag{8}$$

其中 $M' > 0$ 为任意大的常数。

命题的证明2 和3 见第4节。下一节中，我们将展示结果在我们的高斯过程先验（见第3.2节）以及 H¨较老的函数类中的应用。

3.2 高斯过程先验与 H¨older 光滑函数的应用

回顾第2.2 节中基于高斯过程的先验构造，其中潜在函数为 ν， ϕ，连接函数为 σ，协方差函数（核）为 k_ν， k_ϕ。我们将证明，在对 σ 和核作出温和假设的条件下，假设3 和4 均成立，且集中速率 ϵ_n 可由真实参数的光滑性明确表示 f_0。鉴于它们在实际应用中广受欢迎，本节将重点讨论 M´atern 核和平方指数核（定义见（2）、（3））。

在陈述假设之前，我们先介绍一些记号。对于 $\alpha > 0$，令 $C^\alpha(S)$ 表示 H¨older α-光滑函数空间，即函数可微 $\lfloor\alpha\rfloor$ 次，且其 $\lfloor\alpha\rfloor$ 阶导数是 $(\alpha - \lfloor\alpha\rfloor)$-连续的。也就是说，对于 $f \in C^\alpha(S)$，

$$|f^{\lfloor\alpha\rfloor}(x) - f^{\lfloor\alpha\rfloor}(y)| \le |x - y|^{\alpha - \lfloor\alpha\rfloor}.$$

对于 $\alpha \in \mathbb{N}$，我们用 $\mathcal{S}^\alpha(S)$ 表示阶数为 α 的 Sobole夫空间，即函数的（弱）导数$D^\gamma f$ 对于任意 $\|\gamma\|_1 \le \alpha$ 都是平方可积的。阶数为 $\alpha > 0$ 的 Sobole夫空间也可以通过傅里叶变换来定义，参见例如Ghosal和Van der Vaart [2017]中的定义 C.6。

我们的前两个假设是关于 σ 和 $(\mu_0,\ g_0)$ 的温和正则性与光滑性条件。

假设 5. 连接函数 σ： $\mathbb{R} \to \mathbb{R}^+$ 无限次可微、严格递增且为 *L-Lipschitz*，且 $L > 0$。此外，它是从 $\mathbb{R}$ 到 $(0, C)$ 的双射，满足 $C > \bar\mu \vee \bar g$ 且 $\bar\mu$， $\bar g$ 定义见假设2。

假设 6. 函数 μ_0 和 g_0 是 *H¨older* α-光滑的，满足 $\alpha > 0$，即 $\mu_0, g_0 \in C^\alpha(S)$。

注 3.1. *Softplus*函数和缩放*Sigmoid*函数 $\sigma(t) = \sigma^*(1 + e^{-\alpha^* t})^{-1}$ 满足 $\alpha^* > 0$ 和$\sigma^* > \bar\mu \vee \bar g$ 假设 5。事实上，在我们的框架中也可以采用常用的指数函数，因为很容易证明，若将 *Lipschitz*假设放宽为 局部-*Lipschitz*约束（如 *Dolmeta和Giordano [2025a]*, 在有界参数设定下（假设2）所述），我们的结果仍然成立。

对于高斯过程先验，收缩率取决于该过程样本路径的光滑性，以及与核函数相关联的再生核希尔伯特空间（RKHS）（详见，例如 Ghosal andVan derVaart [2017] 的第11章）。

对于定义在 S 上的 Mat´ern 核（维数为 $d+1$）及参数 $\tau > \frac{d+1}{2}$，样本路径具有 H¨older γ-光滑性，满足 $\gamma < \tau - \frac{d+1}{2}$，相应的再生核希尔伯特空间为 $\mathcal{S}^{\tau+\frac{d}{2}}$。我们将在下述命题中证明，在采用后一个核的情况下，后验分布以速率$n^{-\frac{\tau}{2\tau+d+1}}$ 集中，如果$\tau \le \alpha$；当 $\tau = \alpha$ 时，该速率对应于最优速率。

命题4 (Mat´ern 协方差核)。 在假设2、5、6以及采用参数为 ´ 的 *Mat* $\tau < \alpha$*ern* 核的基于高斯过程的先验下，假设3和4得到验证，其中 $\epsilon_n = \bar\epsilon_n \asymp n^{-\frac{\tau}{2\tau+d+1}}$，并且 (6) 和 (8)成立。

注3.2。 请注意，上述结果不是自适应的。对于采用 *Mat´ern* 核的高斯过程先验，获得自适应结果尤其困难。

相比之下，平方指数协方差核与长度尺度上的 ℓ^{d+1} 逆伽马超先验相结合，可以实现函数参数的自适应最优估计，误差仅相差对数因子。我们注意到，对于平方指数核，无论长度尺度取何值，其样本路径都是解析函数 ℓ，且再生核希尔伯特空间的定义更为复杂（参见 Ghosal和Van der Vaart [2017]中的引理11.35）。

命题5（平方指数协方差核）。 在假设 2、5、6 成立，且采用平方指数核的高斯过程先验，其中长度尺度参数的先验 服从逆伽马分布，即：

$$\ell^{d+1} \sim IG(a_0, b_0)$$

其中 a_0， $b_0 > 0$，则假设 3 和 4 得到验证，且

$$\bar{\epsilon}_n \asymp (\log n)^{\frac{d+1}{2+(d+1)/\alpha}} n^{-\frac{\alpha}{2\alpha+d+1}}, \qquad \epsilon_n \asymp (\log n)^{\frac{d+1}{2}} \bar{\epsilon}_n$$

且（6）且（8）成立。

备注3.3（可分离核）。补充一些在高斯过程中使用可分离核的论文。在时空数据中，通常会选择一个关于时间变量和空间变量可分解的核函数，因为这样可以实现计算加速。例如，可以使用带有超参数（τ_t、τ_s）的可分离 *Mat´ern* 核。相应的再生核希尔伯特空间是张量积 *Sobole* 夫空间 $\mathcal{S}^{\tau_t+\frac{1}{2}}([0,1]) \otimes \mathcal{S}^{\tau_s+\frac{d}{2}}([0,1]^d)$。回顾张量积 *Sobole* 夫空间的定义：

$$\mathcal{S}^\gamma([0,1]) \otimes \mathcal{S}^\gamma([0,1]^d) = \{f \in L_2([0,1]^{d+1}) : D^\alpha f \in L_2([0,1]^{d+1}), \quad \forall \alpha, \|\alpha\|_\infty \le \gamma\}.$$

与 *Sobole*夫空间 $\mathcal{S}^\gamma([0,1]^{d+1})$ 相比，后者要求对每个 $\|\zeta\|_1 \le \gamma$，阶数为 ζ的（偏）导数均平方可积，而张量积 *Sobole*夫空间 $\mathcal{S}^\gamma([0,1]) \otimes \mathcal{S}^\gamma([0,1]^d)$ 要求对每个 $\|\zeta\|_\infty \le \gamma$，阶数为 ζ 的偏导数均平方可积，这是一个严格更强的条件。因此，$\mathcal{S}^\gamma([0,1]) \otimes \mathcal{S}^\gamma([0,1]^d) \subset \mathcal{S}^\gamma([0,1]^{d+1})$ *[Zhang* 和 *Simon*，*2023]*。事实上，$\mathcal{S}^\gamma([0,1]) \otimes \mathcal{S}^\gamma([0,1]^d)$ 包含由 $\mathcal{S}^\gamma([0,1])$ 和 $\mathcal{S}^\gamma([0,1]^d)$ 中函数的乘积进行有限线性组合而成的函数，即：

$$f \in H^\gamma([0,1]) \otimes H^\gamma([0,1]^d) \iff f = \sum_j f_{tj}(t) f_{sj}(s).$$

换句话说，$f \in \mathcal{S}^\gamma([0,1]) \otimes \mathcal{S}^\gamma([0,1]^d)$ 具有加性形式，但通常不能在时间和空间上分离。

4 证明

4.1 命题1的证明

首先，回顾一下，N和N'分别是参数为 f和 f'的两个霍克斯过程，并且$N \overset{d}{=} N'$当且仅当对于几乎处处的(t,s)，$\lambda(t,s) \overset{d}{=} \lambda'(t,s)$。我们还注意到，$f = f'$直接意味着 $\lambda(t,s) \overset{d}{=} \lambda'(t,s)$ $\forall(t,s)$，因此只需证明反向蕴含。其次，注意对于任意 $t \ge 0$，满足 $A_t = [0,t] \times [0,1]^d$时，根据假设1，

$$\mathbb{P}\Big(N(A_t) = 0\Big) = e^{-\int_0^t \int_{[0,1]^d} \mu(u,s)\, du\, ds} > 0,$$

对于 N'，也有类似结论。因此，对于任意(t,s)，在给定 $N(A_t) = 0$的条件下，我们有 e

$$\lambda(t,s) = \mu(t,s).$$

对于 N'，类似地可得。因此，$N \overset{d}{=} N'$ 意味着

$$\lambda(t,s) \mid N(A_t) = 0 \overset{d}{=} \lambda'(s,t) \mid N'(A_t) = 0,$$

即，对于所有 (t,s)，$\mu(t,s) = \mu'(t,s)$，这等价于

$$\mu = \mu'.$$

现在，由假设 1，$\int_0^{1-a} \int_{[b,1-b]^d} \mu(t,\ s)dtds > 0$，事件 $\{N(\bar{S}) \ge 1\}$ 在 $\bar{S} = [0,1-a] \times [b,1-b]^d$下的概率非零（事件 $\{N'(\bar{S}) \ge 1\}$的情况类似）。给定

¯在 $\{N(S) \geq 1\}$的条件下，我们用 (T_1, S_1)（分别用 (T_1', S_1')）表示第一个¯事件 N（分别为 N'）的时空坐标。于是，在事件 $\{N(S) \geq 1\}$发生的条件下，对任意 $t \geq T_1$和 $s \in S$，

$$\lambda(t,s) = \mu(t,s) + g(t - T_1, s - S_1).$$

因此，

$$\lambda(t,s) \mid N(\bar{S}) = 1 \stackrel{d}{=} \lambda'(t,s) \mid N'(\bar{S}) = 1$$

这意味着

$$\mu(t,s) + g(t - T_1, s - S_1) \stackrel{d}{=} \mu'(t,s) + g'(t - T_1', s - S_1'),$$

并且还意味着

$$g(t - T_1, s - S_1) \stackrel{d}{=} g'(t - T_1', s - S_1'),$$

¯因为 $\mu(t,s) < +\infty$ 在假设 1下。由于 $(T_1, S_1) \in S$ 且上述等式对任意$1 \geq t \geq T_1$ 和 $s \in S$都成立，因此对任意 $u = t - T_1 \in [0, a]$ 和 $v = s - S_1$ 满足 $\|v\|_\infty \leq b$的也成立

$$g(u,v) = g'(u,v),$$

这等价于 $g = g'$。

4.2 命题2的证明

在证明命题 2之前，我们先陈述三个技术引理。第一个引理定义了一个高概率事件 Ω_n ，在该事件上， $S = [0,1]^{d+1}$ 上的点的平均数量和最大数量均有界。第二个引理给出了 f 与 f_0 之间的库尔贝克-莱布勒（KL）散度以及对数似然比偏差的上界，其中 f 与 f_0足够接近，具体而言为 $f \in B_\infty(\epsilon_n)$。最后一个引理证明了检验的存在性，使其第一类和第二类错误呈指数衰减。

L引理 4.1 （高概率事件）。 对于任意 $\alpha > 0$，存在 δ_0、 $c_\alpha > 0$ ，使得 t

$$\Omega_n := \left\{ \frac{\underline{\mu}}{1 - \|g_0\|_1} - \delta_n \leq \frac{1}{n} \sum_i N^i[0,1]^{d+1} \leq \frac{\bar{\mu}}{1 - \|g_0\|_1} + \delta_n \right\} \cap \left\{ \sup_{i=1,\ldots,n} N^i[0,1]^{d+1} \leq c_\alpha \log n \right\},$$

其中 $\delta_n = \delta_{0\log\sqrt{n}}{}^n$以及 $c > 0$。在假设 2下，

$$\mathbb{P}_0[\Omega_n] \geq 1 - 3n^{-\alpha}.$$

引理 4.2 (库尔贝克-莱布勒)。 在假设 2 下，且若 $\epsilon_n = o((\text{对数 } n)^{-2})$，则存在 b_1， $b_2 > 0$使得对于任意 $f \in B_\infty(\epsilon_n)$，

$$KL(f, f_0) := \mathbb{E}_0[\log L(N|f_0) - \log L(N|f)] \leq \kappa n \epsilon_n^2 (1 + o(1)) \tag{9}$$

$$\mathbb{P}_0(\log L(N|f_0) - \log L(N|f) > b_1 n \epsilon_n^2) \leq \frac{b_2}{n\epsilon_n^2}, \tag{10}$$

with

$$\kappa := \frac{4 \log 2}{\underline{\mu}} \left\{ 2 + 4 \left(\frac{\bar{\mu}}{1 - \|g_0\|_1} + \Lambda_{0,2} \right) \right\}. \tag{11}$$

引理 4.3 (检验)。 在假设 *4* 和 *2*下，存在一个检验函数 $\phi := \phi(N, \epsilon_n)$ 使得

$$\mathbb{E}_0[\phi \mathbb{1}_{\Omega_n}] = o(1)$$
$$\sup_{f \in \mathcal{F}_n} \mathbb{E}_f[(1-\phi)\mathbb{1}_{\Omega_n}\mathbb{1}_{f \in A_n}] \leq e^{-b_2 n \epsilon_n^2}$$

其中 $b_2 > c_1$， A_n 定义见（12），且Ω_n 的定义见引理*4.1*。

这些技术结果的证明见附录B。现在我们证明该命题。对于$M > 0$ 和 $\epsilon_n > 0$，定义关注子集

$$A_n = \{f \in \mathcal{F} : d_S(f, f_0) > M\epsilon_n\}. \tag{12}$$

注意，由于 $\Pi(A_n|N_n) \in [0,1]$， (6) 等价于 $\mathbb{E}_0[\Pi(A_n|N_n)] = o(1)$，即依期望收敛。给定一个检验函数 $\phi := \phi(N, \epsilon_n) \in \{0,1\}$ 和一个高概率事件Ω_n,

$$\begin{aligned}
\mathbb{E}_0[\Pi(A_n|N_n)] &= \mathbb{E}_0[\Pi(A_n|N_n)(\mathbb{1}_{\Omega_n} + \mathbb{1}_{\Omega_n^c})] \\
&\leq \mathbb{E}_0[\Pi(A_n|N_n)(\phi + 1 - \phi)\mathbb{1}_{\Omega_n}] + \mathbb{E}_0[\mathbb{1}_{\Omega_n^c}] \\
&\leq \mathbb{E}_0[\phi \mathbb{1}_{\Omega_n}] + \mathbb{E}_0[(1-\phi)\Pi(A_n|N_n)\mathbb{1}_{\Omega_n}] + \mathbb{P}_0[\Omega_n^c],
\end{aligned} \tag{13}$$

利用 $\Pi(A_n|N_n) \leq 1$。令Ω_n 如引理4.1所定义，则有 $\mathbb{P}_0[\Omega_n^c] = o(1)$。此外，我们可以将 $\Pi(A_n|N_n)$ 写为

$$\Pi(A_n|N_n) = \frac{\int_{A_n} L(N|f)\Pi(f)df}{\int_{\mathcal{F}} L(N|f)\Pi(f)df} = \frac{\int_{\mathcal{F}} \mathbb{1}_{f \in A_n} \frac{L(N|f)}{L(N|f_0)}\Pi(f)df}{\int_{\mathcal{F}} \frac{L(N|f)}{L(N|f_0)}\Pi(f)df}, \tag{14}$$

定义

$$D_n := \int_{\mathcal{F}} \frac{L(N|f)}{L(N|f_0)}\Pi(f)df$$

注意，对于任意确定性序列 $\eta_n > 0$,

$$\begin{aligned}
\mathbb{E}_0[(1-\phi)\Pi(A_n|N_n)\mathbb{1}_{\Omega_n}] &= \mathbb{E}_0[(1-\phi)\Pi(A_n|N_n)\mathbb{1}_{\Omega_n}(\mathbb{1}_{D_n \geq \eta_n} + \mathbb{1}_{D_n < \eta_n})] \\
&\leq \mathbb{P}_0[D_n < \eta_n] + \mathbb{E}_0[(1-\phi)\Pi(A_n|N_n)\mathbb{1}_{\Omega_n}\mathbb{1}_{D_n > \eta_n}].
\end{aligned}$$

然后，利用（13）、（14）以及上述不等式,

$$\begin{aligned}
\mathbb{E}_0[\Pi(A_n|N_n)] &\leq \mathbb{E}_0[\phi \mathbb{1}_{\Omega_n}] + \mathbb{P}_0[D_n < \eta_n] \\
&\quad + \mathbb{E}_0\left[\frac{1}{D_n}(1-\phi)\int_{\mathcal{F}} \mathbb{1}_{f \in A_n}\frac{L(N|f)}{L(N|f_0)}\Pi(f)df\mathbb{1}_{D_n \geq \eta_n}\mathbb{1}_{\Omega_n}\right] + \mathbb{P}_0[\Omega_n^c] \\
&\leq \mathbb{E}_0[\phi \mathbb{1}_{\Omega_n}] + \mathbb{P}_0[D_n < \eta_n] + \frac{1}{\eta_n}\mathbb{E}_0\left[(1-\phi)\int_{\mathcal{F}_n} \frac{L(N|f)}{L(N|f_0)}\mathbb{1}_{f \in A_n}\Pi(f)df\mathbb{1}_{\Omega_n}\right] \\
&\quad + \mathbb{P}_0[\Omega_n^c] + \frac{1}{\eta_n}\mathbb{E}_0\left[\int_{\mathcal{F}_n^c} \frac{L(N|f)}{L(N|f_0)}\Pi(f)df\right].
\end{aligned}$$

在第二个不等式中使用 $\mathcal{F} = \mathcal{F}_n \cup \mathcal{F}_n^c$。根据富比尼定理,

$$\mathbb{E}_0\left[\int_{\mathcal{F}_n^c} \frac{L(N|f)}{L(N|f_0)}\Pi(f)df\right] = \int_{\mathcal{F}_n^c} \mathbb{E}_0\left[\frac{L(N|f)}{L(N|f_0)}\right]\Pi(f)df = \int_{\mathcal{F}_n^c} \Pi(f)df = \Pi(\mathcal{F}_n^c)$$

利用 $\mathbb{E}_0[\frac{L(N|f)}{L(N|f_0)}] = 1$。定义 $\eta_n = \Pi(B_\infty(\bar{\epsilon}_n))e^{-b_1 n\bar{\epsilon}_n^2} \geq e^{-(b_1+c_1)n\bar{\epsilon}_n^2}$在假设 3 下，存在某个 $b_1 > 0$使得 $b_1 < c_2 - c_1$，那么，利用在假设 4下， $\Pi(\mathcal{F}_n^c) \leq e^{-c_2 n\bar{\epsilon}_n^2}$，我们有

$$\frac{1}{\eta_n}\Pi(\mathcal{F}_n^c) \leq e^{(b_1+c_1-c_2)n\bar{\epsilon}_n^2} = o(1).$$

采用与Donnet 等人 [2020] （定理 1 的证明）中相同的计算，

$$\begin{aligned}&\mathbb{P}_0[D_n < \Pi(B_\infty(\bar{\epsilon}_n))e^{-b_1 n\bar{\epsilon}_n^2}] \\ &\quad\leq \frac{1}{\Pi(B_\infty(\bar{\epsilon}_n))(1-e^{-b_1 n\bar{\epsilon}_n^2})}\int_{B_\infty(\bar{\epsilon}_n)} \mathbb{P}_0(\log L(f_0|N) - \log L(f|N) > b_1 n\bar{\epsilon}_n^2)\Pi(f)df.\end{aligned}$$

根据引理 4.2，若 $b_1 > \kappa$，则对任意 $f \in B_\infty(\bar{\epsilon}_n)$，

$$\mathbb{P}_0(\log L(f_0|N) - \log L(f|N) > b_1 n\bar{\epsilon}_n^2) \leq \frac{b_2}{n\bar{\epsilon}_n^2},$$

其中 $b_2 = \frac{2\kappa}{(b_1-\kappa)^2}$ 由此可得

$$\mathbb{P}_0[D_n < \Pi(B_\infty(\epsilon_n))e^{-b_1 n\epsilon_n^2}] \leq \frac{b_2}{n\bar{\epsilon}_n^2(1-e^{-b_1 n\bar{\epsilon}_n^2})} = o(1),$$

在假设 $n\bar{\epsilon}_n^2 \to \infty$成立的条件下。注意，由于 $c_2 > c_1 + \kappa$ 根据假设 4，因此存在 $b_1 \in (\kappa, c_2 - c_1)$ （例如，$\kappa + \frac{c_2-c_1-\kappa}{2}$）。

此外，利用引理 4.3，我们可以找到 ϕ使得

$$\begin{aligned}&\mathbb{E}_0[\phi \mathbb{1}_{\Omega_n}] = o(1) \\ &\sup_{f\in\mathcal{F}_n} \mathbb{E}_f[(1-\phi)\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in A_n}] \leq e^{-b_2 n\epsilon_n^2} = o(\Pi(B_\infty(\epsilon_n))e^{-b_1 n\bar{\epsilon}_n^2}) = o(\eta_n).\end{aligned}$$

因此，再次利用富比尼定理，

$$\begin{aligned}\frac{1}{\eta_n}\mathbb{E}_0\left[(1-\phi)\int_{\mathcal{F}_n}\frac{L(N|f)}{L(N|f_0)}\mathbb{1}_{f\in A_n}\Pi(f)df\mathbb{1}_{\Omega_n}\right] &= \frac{1}{\eta_n}\int_{\mathcal{F}_n}\mathbb{E}_0\left[\frac{L(N|f)}{L(N|f_0)}(1-\phi)\mathbb{1}_{f\in A_n}\mathbb{1}_{\Omega_n}\right]\Pi(f)df \\ &= \frac{1}{\eta_n}\int_{\mathcal{F}_n}\mathbb{E}_f[(1-\phi)\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in A_n}]\Pi(f)df \\ &\leq \frac{\Pi(\mathcal{F}_n)}{\eta_n}e^{-b_2 n\bar{\epsilon}_n^2} \leq \frac{e^{-b_2 n\bar{\epsilon}_n^2}}{\eta_n} = o(1),\end{aligned}$$

至此，证明完毕。

4.3 命题3的证明

我们首先陈述一个引理，该引理给出了在任意 $f \in A_n^c$ 下随机变量 (Z_i) （如（15）中所定义）的期望上界。其证明见附录 B。

引理 4.4。对于任意 $f \in A_n^c$，存在一个常数 $p_0 > 0$，使得在 Ω_n 上，

$$\mathbb{E}_f[Z_1] \geq p_0\|f - f_0\|_1.$$

对于 $M' > 0$，令

$$A_{n,1} = \{f \in \mathcal{F} : \|f - f_0\|_1 > M'\epsilon_n\}.$$

利用 $A_{n,\ 1} = (A_n \cap A_{n,\ 1})\ \cup (A_n^c \cap A_{n,\ 1})$，我们有

$\mathbb{E}_0[\Pi(A_{n,\ 1}|N_n)] \leq \mathbb{E}_0[\Pi(A_n|N_n)] + \mathbb{E}_0[\Pi(A_n^c \cap A_{n,\ 1}|N_n)] = o(1)\ + \mathbb{E}_0[\Pi(A_n^c \cap A_{n,\ 1}|N_n)]$.利用命题2。应用命题 2证明中的相同分解，我们得到

$$\begin{aligned}\mathbb{E}_0[\Pi(A_n^c \cap A_{n,1}|N_n)] &\leq \mathbb{P}_0[D_n < \eta_n] + \frac{1}{\eta_n}\int_{\mathcal{F}_n} \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in A_{n,1}\cap A_n^c}]\Pi(f)df \\ &\quad + \mathbb{P}_0[\Omega_n^c] + \frac{1}{\eta_n}\Pi(\mathcal{F}_n^c) \\ &\leq o(1) + \frac{1}{\eta_n}\sup_{f\in\mathcal{F}_n\cap A_{n,1}} \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in A_n^c}].\end{aligned}$$

回忆一下，

$$\begin{aligned}f \in A_n^c &\iff d_S(f, f_0) \leq M\epsilon_n \\ &\iff \frac{1}{n}\sum_{i=1}^n \int |\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0)|dtds \leq M\epsilon_n\end{aligned}$$

其中，上述积分是在整个观测域 $[0,1]\times[0,1]^{d+1}$ 上进行的。回忆一下，对于任意序列 N^i，我们将 t_1^i，t_2^i、.....称为事件发生的时刻。定义

$$Z_i := \int_0^{t_2^i}\int_{[0,1]^d} |\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0)|dtds. \tag{15}$$

按照惯例，若 t_2^i 不存在（即该序列仅包含一个或不包含任何事件），则设 $t_2^i = 1$。注意，$(Z_i)_i$ 独立同分布，并且

$$\mathbb{P}_f\left[\{f \in A_n^c\} \cap \{\Omega_n\}\right] \leq \mathbb{P}_f\left[\frac{1}{n}\sum_{i=1}^n Z_i \leq M\epsilon_n \cap \{\Omega_n\}\right]$$

对于任意 $f \in A_n^c$，利用引理 4.4，存在一个常数 $p_0 > 0$使得在 Ω_n 上，

$$\mathbb{E}_f[Z_i] \geq p_0\|f - f_0\|_1.$$

因此，

$$\begin{aligned}\mathbb{P}_f\left[\frac{1}{n}\sum_{i=1}^n Z_i \leq M\epsilon_n \cap \{\Omega_n\}\right] &= \mathbb{P}_f\left[\frac{1}{n}\sum_{i=1}^n (Z_i - \mathbb{E}_f[Z_i]) \leq M\epsilon_n - \frac{1}{n}\sum_i \mathbb{E}_f[Z_i] \cap \{\Omega_n\}\right] \\ &\leq \mathbb{P}_f\left[\frac{1}{n}\sum_{i=1}^n (Z_i - \mathbb{E}_f[Z_i]) \leq M\epsilon_n - p_0\|f - f_0\|_1\right] \\ &\leq \mathbb{P}_f\left[\frac{1}{n}\sum_{i=1}^n (Z_i - \mathbb{E}_f[Z_i]) \leq -p_0\|f - f_0\|_1/2\right],\end{aligned}$$

利用 $f \in A_{n,\ 1}$ 并且对于 $M' > 2M/p_{0\circ}$ 此外，

$$\begin{aligned}
Z_i &= \int_0^{t_2^i} \int_{[0,1]^d} |\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0)| dtds \\
&= \int_0^{t_1^i} \int_{[0,1]^d} |\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0)| dtds + \int_{t_1^i}^{t_2^i} \int_{[0,1]^d} |\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0)| dtds \\
&= \int_0^{t_1^i} \int_{[0,1]^d} |\mu(t,s) - \mu_0(t,s)| dtds \\
&+ \int_{t_1^i}^{t_2^i} \int_{[0,1]^d} |\mu(t,s) + g(t - t_1^i, s - s_1^i) - \mu_0(t,s) - g_0(t - t_1^i, s - s_1^i)| dtds \\
&\leq \int_0^{t_2^i} \int_{[0,1]^d} |\mu(t,s) - \mu_0(t,s)| dtds \\
&+ \int_{t_1^i}^{t_2^i} \int_{[0,1]^d} |g(t - t_1^i, s - s_1^i) - g_0(t - t_1^i, s - s_1^i)| dtds \\
&\leq \|\mu - \mu_0\|_1 + \int_0^{t_2^i - t_1^i} \int_{[0,1]^d} |g(u, s - s_1^i) - g_0(u, s - s_1^i)| duds \\
&\leq \|\mu - \mu_0\|_1 + \|g - g_0\|_1 = \|f - f_0\|_1.
\end{aligned}$$

因此，

$$\mathbb{E}_f[Z_i^k] \leq \|f - f_0\|^k, \qquad k \geq 2.$$

应用伯恩斯坦不等式，我们得到

$$\begin{aligned}
\mathbb{P}_f \left[\frac{1}{n} \sum_{i=1}^n Z_i - \mathbb{E}_f[Z_i] \leq -p_0 \|f - f_0\|_1 / 2 \right] &\leq \exp\left(- \frac{p_0^2 n \|f - f_0\|_1^2}{8 \|f - f_0\|_1^2 (1 + p_0/3)} \right) \\
&= \exp\left(- \frac{p_0^2 n}{8(1 + p_0/3)} \right) \\
&= o(e^{-cn\epsilon_n^2}),
\end{aligned}$$

对于任意常数 $c > 0_\circ$ 因此我们可以得出，对于任意 $f \in \mathcal{F}_n \cap A_{n,\ 1}$，

$$\mathbb{E}_f[\mathbb{1}_{\Omega_n} \mathbb{1}_{f \in A_n^c}] = o(e^{-(b_1 + c_1) n \epsilon_n^2}) = o(\eta_n),$$

回忆起 $\eta_n = \Pi(B_\infty(\bar{\epsilon}_n)) e^{-b_1 n \bar{\epsilon}_n^2} \geq e^{-(b_1 + c_1) n \bar{\epsilon}_n^2}$，其中 $\bar{\epsilon}_n \leq \epsilon_{n\circ}$ 因此，

$$\frac{1}{\eta_n} \sup_{f \in \mathcal{F}_n \cap A_{n,1}} \mathbb{E}_f[\mathbb{1}_{\Omega_n} \mathbb{1}_{f \in A_n^c}] = o(1),$$

由此完成了该命题的证明。

4.4 命题4的证明

我们首先证明假设3 成立。注意到

$$B_\infty(\epsilon_n) \supset \{\mu : \|\mu - \mu_0\|_\infty \leq \epsilon_n / 2\} \times \{g : \|g - g_0\|_\infty \leq \epsilon_n / 2\},$$

因此

$$\Pi(B_\infty(\epsilon_n)) \geq \Pi_\mu(\|\mu-\mu_0\|_\infty \leq \epsilon_n/2)\Pi_g(\|g-g_0\|_\infty \leq \epsilon_n/2).$$

我们考虑第一项 $\Pi_\mu(\|\mu-\mu_0\|_\infty \leq \epsilon_n/2)$。由于 $\mu_0(x) \in [\mu, \bar{\mu}]$ 且 σ 是从$\mathbb{R}$到 $(0,\ C)$的双射，令 σ^{-1} 为 σ 在 $(0,\ C)$上的逆函数，并令 $\nu_0 := \sigma^{-1}(\mu_0)$。此外，$\mu_0 \in C^\alpha(S)$ 和 σ无限光滑意味着 $\nu_0 \in C^\alpha(S)$。由于 σ 是 L-Lipschitz 的，

$$\Pi_\mu(\|\mu-\mu_0\|_\infty \leq \epsilon_n/2) \geq \Pi_\nu(\|\nu-\nu_0\|_\infty \leq \epsilon_n/(2L)),$$

其中 Π_ν 是核为 k_ν 的高斯过程先验。

回顾一下，Mat´ern 核的再生核希尔伯特空间（参数为 $\tau \leq \alpha$），记作 $\mathcal{H}^\tau_{MAT}$，是光滑性为 $\tau + \frac{d+1}{2}$ 的 Sobole夫空间。由于 $\alpha \geq \tau$，$w_0 \in \mathcal{H}^\tau_{MAT}$。利用 Giordano 等人的引理 B.1，Giordanoetal. [2025]存在 $L_1 > 0$使得

$$\Pi_\nu(\|\nu-\nu_0\|_\infty \leq \epsilon_n/(2L)) \geq e^{-L_1 n\epsilon_n^2},$$

对于任意满足 $\epsilon_n \gtrsim n^{-\tau/(2\tau+d+1)}$ 的 $\epsilon_n \to 0$。现在考虑第二项 $\Pi_g(g:\ \|g-g_0\|_\infty \leq \epsilon_n/2)$。令 $g_{0,\ n}:= \max(g_0,\ \epsilon_n/4)$，使得 $\|g_0 - g_{0,\ n}\|_\infty \leq \epsilon_n/4$。因此，

$$\Pi_g(g : \|g-g_0\|_\infty \leq \epsilon_n/2) \geq \Pi_g(g : \|g-g_{0,n}\|_\infty \leq \epsilon_n/4).$$

定义 $\phi_{0,\ n} = \sigma^{-1}(g_{0,\ n})$，采用与 ν_0 相同的论证，并且由于 σ 是 L-Lipschitz 的，存在 $L_2 > 0$ 使得

$$\Pi_g(\|g-g_{0,n}\|_\infty \leq \epsilon_n/4) \geq \Pi_\phi(\|\phi-\phi_{0,n}\|_\infty \leq \epsilon_n/(4L)) \geq e^{-L_2 n\epsilon_n^2},$$

with Π_ϕ 是定义在 $[0,a]\times[0,b]^d$ 上、具有 Mat´ern 核的高斯过程先验，ϵ_n 同前。因此 e,

$$\begin{aligned}\Pi(B_\infty(\epsilon_n)) &\geq \Pi_\nu(\|\nu-\nu_0\|_\infty \leq \epsilon_n/(2L))\Pi_\phi(\|\phi-\phi_{0,n}\|_\infty \leq \epsilon_n/(4L)) \\ &\geq \Pi_\nu(\|\nu-\nu_0\|_\infty \leq \epsilon_n/(4L))\Pi_\phi(\|\phi-\phi_{0,n}\|_\infty \leq \epsilon_n/(4L)) \\ &\geq e^{-(L_1+L_2)n\epsilon_n^2},\end{aligned}$$

这证明了假设 3 满足 $c_1 :\ = L_1 + L_2$ 以及 $\bar{\epsilon}_n = \epsilon_n \asymp n^{-\tau/(2\tau+d+1}$).

现在证明假设 4 成立。对于 $M_1,\ M_2 > 0$，令

$$\mathcal{M}_n = M_1\epsilon_n\mathbb{B}_1 + M_2\sqrt{n}\epsilon_n\mathbb{H}_1,$$

其中 $\mathbb{B}_1,\ \mathbb{H}_1$ 分别是 $L_\infty(S)$ 中的 $\mathcal{H}^\tau_{MAT}$（相对于相应的索伯列夫范数）。利用 Giordano 等人 [2025], 中的引理 B.2，对于任意满足 $\epsilon_n \to 0$ 使得 $\epsilon_n \gtrsim n^{-\tau/(2\tau+d+1)}$ 且$R_1 > 0$的 $M_1,\ M_2 > 0$ 存在 $R_2 > 0$ 使得

$$\begin{aligned}&\Pi_\nu(\mathcal{M}_n^c) \leq e^{-R_1 n\epsilon_n^2} \\ &\log C(\epsilon_n, \mathcal{M}_n, \|\cdot\|_\infty) \leq R_2 n\epsilon_n^2. \end{aligned} \tag{16}$$

类似地，我们构造

$$\mathcal{N}_n = M_1\epsilon_n\bar{\mathbb{B}}_1 + M_2\sqrt{n}\epsilon_n\bar{\mathbb{H}}_1,$$

其中 $\bar{\mathbb{B}}_1,\ \bar{\mathbb{H}}_1$ 分别是在 $L_\infty([0,a]\times[0,b]^d)$ 中的单位球，以及在 $[0,a]\times[0,b]^d$ 上的 $\mathcal{H}^\tau_{MAT}$ 的类似定义（关于相应的索伯列夫范数），并且我们有

$$\Pi_\phi(\mathcal{N}_n^c) \leq e^{-R_1 n\epsilon_n^2} \tag{17}$$

$$\log C(\epsilon_n, \mathcal{N}_n, \|\cdot\|_\infty) \leq R_2 n\epsilon_n^2. \tag{18}$$

接下来，我们将筛集构造为

$$\mathcal{F}_n = \sigma(\mathcal{M}_n) \times \sigma(\mathcal{N}_n).$$

我们得到

$$\begin{aligned}\Pi(\mathcal{F}_n) = \Pi_\mu(\sigma(\mathcal{M}_n))\Pi_g(\sigma(\mathcal{N}_n)) &\geq \Pi_\nu(\mathcal{M}_n)\Pi_\phi(\mathcal{N}_n) \\ &\geq (1 - e^{-R_1 n\epsilon_n^2})^2 \\ &\geq 1 - 2e^{-R_1 n\epsilon_n^2} \geq 1 - e^{-\frac{R_1}{2}n\epsilon_n^2}.\end{aligned}$$

因此，选择 $R_1 > 2(c_1 + \kappa)$ 我们得到假设 4的第一部分。此外，由于 σ 是L-利普希茨的，

$$\begin{aligned}C(\epsilon, \mathcal{F}_n, \|\cdot\|_\infty) &= C(\epsilon, \sigma(\mathcal{M}_n), \|\cdot\|_\infty)C(\epsilon, \sigma(\mathcal{M}_n), \|\cdot\|_\infty) \\ &\leq C(L\epsilon, \mathcal{M}_n, \|\cdot\|_\infty)C(L\epsilon, \mathcal{M}_n, \|\cdot\|_\infty).\end{aligned}$$

因此，利用（16）和（18），我们得到

$$C(\epsilon_n, \mathcal{F}_n, \|\cdot\|_\infty) \leq e^{2R_2L^2n\epsilon_n^2},$$

这证明了假设 4的第二部分，其中 $c_3 := 2R_2L^2$ 且 $\epsilon_n \asymp n^{-\tau/(2\tau+d+1)}$。

4.5 命题5的证明

为了验证平方指数核在长度尺度上采用逆伽马先验时满足假设 3 和4 ，我们遵循命题 4 的证明中的相同步骤，并应用van derVaartandvan Zanten [2009]中的定理3.1。

事实上，如果 $\epsilon_n = 2LK(\log n)^{\frac{d+1}{2+d+1()/\alpha}} n^{-\frac{\alpha}{2\alpha+d+1}}$ 足够大，则

$$\Pi_\nu(\|\nu - \nu_0\|_\infty \leq \epsilon_n/(2L)) \geq e^{-\frac{n}{4L^2}\epsilon_n^2}.$$

采用与前面相同的论证，我们得到假设 3 成立，其中 $c_1 := L_1 + L_2$ 且 $\bar{\epsilon}_n \asymp (\log n)^{\frac{d+1}{2+d+1()}}$ $^{\alpha} n^{-\tau/(2\tau+d+1)}$。 此外，存在 $\mathcal{M}_n := \{\mu : S \to \mathbb{R}_+\}, \mathcal{G}_n : [0, a] \times [0, b]^d \to \mathbb{R}_+, M_1, M_2 > 0$ ，使得

$$\begin{aligned}&\Pi(\mathcal{M}_n^c) \leq e^{-\frac{n}{L}\bar{\epsilon}_n^2} \\ &\Pi(\mathcal{G}_n^c) \leq e^{-\frac{n}{L}\bar{\epsilon}_n^2} \\ &C(M_1\epsilon_n, \mathcal{M}_n, \|\cdot\|_\infty) \leq e^{M_1^2 n\epsilon_n^2} \\ &C(M_2\epsilon_n, \mathcal{G}_n, \|\cdot\|_\infty) \leq e^{M_2^2 n\epsilon_n^2},\end{aligned}$$

其中 $\epsilon_n \asymp (\log n)^{\frac{d+1}{2}} \bar{\epsilon}_n$，这使得可以验证假设 4。

5 结论

本研究通过确立多变量、非平稳时空霍克斯过程的后验收缩速率，推进了点过程贝叶斯非参数推断的理论认识。我们将现有针对时间霍克斯过程的分析扩展至一般时空设定，从而首次为自激机制的贝叶斯非参数学习提供理论保证；在这些机制中，背景强度和触发分量均随时间和空间演变

。该框架广泛适用于涉及复杂依赖结构的应用，例如地震活动、神经元脉冲序列、社会或金融互动，以及网络上的信息扩散；在这些应用中，事件同时呈现时间和空间上的激发效应。

在真实强度和先验满足温和正则性条件的情况下，我们证明，灵活的先验，尤其是采用平方指数核的层次高斯过程，能够达到渐近最优的集中速率。未来的工作可将这些结果扩展到多变量或非线性霍克斯过程；在这类过程中，多个潜在分量之间的相互作用或非线性激发效应会带来新的理论和计算挑战。此外，纳入协变量也是一个值得探索的方向，因为这将提升空间学习能力。

致谢 XM 的这项研究获得了欧盟“地平线欧洲”研究与创新计划的资助，资助依据为玛丽·斯克沃多夫斯卡-居里资助协议 101151781。

A 技术引理

A.1 随机距离的界

接下来的两个引理利用 L_1-范数，在高概率事件 Ω_n （定义见引理4.1）上给出随机距离的下界和上界。

引理 A.1。 在 Ω_n 上，对于任意 f, f',

$$d_S(f, f') \le N_0 \|f - f'\|_1.$$

其中 $N_0 = \bar{\mu} + \|g_0\|_1 + 1$。

证明。

$$\begin{aligned}
d_S(f, f') &= \frac{1}{n}\sum_i \int |\lambda_f^i(t,s) - \lambda_{f'}^i(t,s)| dtds \\
&\le \frac{1}{n}\sum_i \int |\mu(t,s) - \mu'(t,s)| dtds + \frac{1}{n}\sum_i \int \left| \sum_{t_j<t} g(t-t_j, s-s_j) - g'(t-t_j, s-s_j) \right| dtds \\
&\le \|\mu(t,s) - \mu'(t,s)\|_1 + \frac{1}{n}\sum_i \sum_{t_j<t} \int |g(t-t_j, s-s_j) - g'(t-t_j, s-s_j)| \, dtds \\
&\le \|\mu(t,s) - \mu'(t,s)\|_1 + \frac{1}{n}\sum_i \sum_{t_j<t} \|g - g'\|_1 \\
&\le \|\mu - \mu'\|_1 + \|g - g'\|_1 \frac{1}{n}\sum_i N^i[0,1]^{d+1} \\
&\le N_0(\|\mu - \mu'\|_1 + \|g - g'\|_1) = N_0 \|f - f'\|_1,
\end{aligned}$$

其中 $N_0 = \bar{\mu} + \|g_0\|_1 + 1$ 在 Ω_n 上。 □

引理 A.2。 对于任意 $f \in \mathcal{F}$，在 Ω_n 上，

$$-d_S(f, f_0) + \|\mu_0\|_1 + \|g_0\|_1 \frac{e_0}{2} \le \|\mu\|_1 + \|g\|_1 \frac{\sum_{i=1}^n N^i(S)}{n} \le \|\mu_0\|_1 + \frac{3e_0}{2}\|g_0\|_1 + d_S(f, f_0) \tag{19}$$

证明。 一方面,

$$\begin{aligned}
d_S(f, f_0) &\le \int_{[0,1]} \int_{[0,1]^d} |\mu(t,x) - \mu_0(t,x)| dtdx + \frac{\sum_{i=1}^n N^i(S)}{n} \int_{[0,a]} \int_{[0,b]} |g(t,x) - g_0(t,x)| dtdx \\
&\le \|\mu - \mu_0\|_1 + \frac{\sum_{i=1}^n N^i(S)}{n} \|g - g_0\|_1
\end{aligned} \tag{20}$$

with

$$\begin{aligned}
\|\mu - \mu_0\|_1 &= \int_{[0,1]} \int_{[0,1]^d} |\mu(t,x) - \mu_0(t,x)| dtdx \\
\|g - g_0\|_1 &= \int_{[0,a]} \int_{[0,b]} |g(t,x) - g_0(t,x)| dtdx.
\end{aligned}$$

另一方面，

$$d_S(f,f_0) \geq \frac{1}{n}\left|\sum_i \int_{[0,1]}\int_{[0,1]^d}(\lambda^i_{t,x}(f)-\lambda^i_{t,x}(f_0))dtdx\right|$$
$$\geq \frac{1}{n}\left|\sum_i(\|\mu\|_1+N^i(S)\|g\|_1-\|\mu_0\|_1-N^i(S)\|g_0\|_1\right|$$

由此我们可以推得

$$-d_S(f,f_0)+\|\mu_0\|_1+\|g_0\|_1\frac{\sum_{i=1}^n N^i(S)}{n} \leq \|\mu\|_1+\|g\|_1\frac{\sum_{i=1}^n N^i(S)}{n} \leq \|\mu_0\|_1+\|g_0\|_1\frac{\sum_{i=1}^n N^i(S)}{n}+d_S(f,f$$

前一个不等式基本上意味着，如果 $d_S(f,\ f_0)$ 很小，那么只要 $\frac{\sum_{i=1}^n N^i(S)}{n}$ 集中在其期望附近，$(\mu,\ g)$ 的 L_1-范数也会被真实参数的 L_1-范数所界定。~~在 Ω_n 上~~，

$$\frac{\underline{\mu}}{1-\|g_0\|_1}-\delta_n \leq \frac{1}{n}\sum_i N^i[0,1]^{d+1} \leq \frac{\bar{\mu}}{1-\|g_0\|_1}+\delta_n,$$

因此，当 $e_0=\frac{\bar{\mu}}{1-\|g_0\|_1}$ 和 n 足够大时，我们得到（19）。

□

A.2 伯恩斯坦不等式

接下来的两个引理给出了点过程的两个有用的伯恩斯坦不等式形式。

引理 A.3。 设 $N=(N^i)_{i=1,\ldots,n}$ 为定义在 $[0,1]\times[0,1]^d$ 上、参数为 f 且满足假设 2 的独立同分布时空霍克斯点过程。设 $(S_i)_{i=1,\ldots,n}\subset[0,1]^{d+1}$ 为可能随机的独立子集，且 $v>0$ 为确定性常数，使得

$$\sum_{i=1}^n \Lambda^i(S_i)=\sum_i\int_{S_{1,i}}\lambda^i_{t,s}(f)dtds \leq v,$$

其中 $\Lambda^i(S_i)=\int_{S_i}\lambda^i_{t,s}(f)dtds$。那么，对于任意 $x>0$，

$$\mathbb{P}\left(\sum_{i=1}^n N^i(S_i)-\Lambda^i(S_i)\geq\sqrt{2vx}+\frac{x}{3}\right)\leq e^{-x}, \tag{21}$$
$$\mathbb{P}\left(\sum_{i=1}^n N^i(S_i)-\Lambda^i(S_i)\leq-\sqrt{2vx}-\frac{x}{3}\right)\leq e^{-x}, \tag{22}$$

其中 $\Lambda^i(S_i)=\int_{S_i}\lambda^i_{t,s}(f)dtds$。

证明。 我们证明右尾概率的界（21）。左尾概率（22）可采用相同策略证明。证明分为三个主要步骤：1）推导 $\sum_i N^i(S_i)$ 重新中心化版本的指数矩；2）使用切尔诺夫不等式对尾概率进行界定；3）对任意足够小的自由参数取一个下界，从而得到所需结果。

步骤1 令

$$E := e^{\theta \sum_i (N^i(S_i) - \Lambda^i(S_i)) - \phi(\theta) \sum_i \Lambda^i(S_i)}$$

其中 $\phi(x) = e^u - u - 1$，且对任意 $\theta > 0$。我们将证明 $\mathbb{E}[E] \leq 1$，其中使用简写记号 $\mathbb{E} := \mathbb{E}_f$。首先，由于这些变量

$$\theta(N^i(S_i) - \Lambda^i(S_i)) - \phi(\theta)\Lambda^i(S_i)$$

相互独立，则

$$\mathbb{E}[E] = \prod_{i=1}^{n} \mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i)) - \phi(\theta)\Lambda^i(S_i)}].$$

此外，由于 $\phi(\theta) \geq \frac{\theta^2}{2} + \frac{\theta^3}{6} = \frac{\theta^2}{2}(1 + \frac{\theta}{3})$，因此

$$\begin{aligned}
\mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i)) - \phi(\theta)\Lambda^i(S_i)}] &\leq \mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i)) - \frac{\theta^2}{2}(1+\frac{\theta}{3})\Lambda^i(S_i)}] \\
&= \mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i)(1 + \frac{\theta}{2}(1+\frac{\theta}{3})))}] \\
&\leq \mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i))}]\mathbb{E}[e^{-\Lambda^i(S_i)\frac{\theta^2}{2}(1+\frac{\theta}{3}))}] \qquad (23)
\end{aligned}$$

在最后一个不等式中使用柯西-施瓦茨不等式。根据假设2，N^i 是非爆炸的，并且我们可以轻易证明它具有指数矩。为此，首先注意到

$$N^i(S_i) - \Lambda^i(S_i) \leq N^i([0,1]^{d+1}) - \Lambda^i([0,1]^{d+1})$$

并且 $N^i([0,1]^{d+1})$ 在随机意义下被 $\bar{N}([0,1])$ 控制，其中 $\bar{N}$ 是一个背景强度为常数 $\bar{\mu}$、时间核为 $\bar{g}(t) = \int_s g(t, s)ds$ 且补偿子为 $\bar{\Lambda}$ 的时间点过程。因此，

$$\mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i))}] \leq \mathbb{E}[e^{\theta(N^i([0,1]^{d+1}) - \Lambda^i([0,1]^{d+1}))}] \leq \mathbb{E}[e^{\theta(\bar{N}([0,1]) - \bar{\Lambda}([0,1]))}].$$

根据Br´emaud 的定理2 [1981]，

$$\mathbb{E}[e^{\theta(\bar{N}([0,1]) - \bar{\Lambda}([0,1]))}] \leq 1,$$

这 因此意味着 $\mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i))}] \leq 1$。此外，若 $\Lambda_i(S_i) \geq 0$且 $\theta > 0$，则

$$\mathbb{E}[e^{-\theta\Lambda^i(S_i)(1+\frac{\theta}{2}(1+\frac{\theta}{3}))}] \leq 1.$$

鉴于 (23)，我们得到

$$\mathbb{E}[e^{\theta(N^i(S_i) - \Lambda^i(S_i)) - \phi(\theta)\Lambda^i(S_i)}] \leq 1,$$

从而得到 $\mathbb{E}[E] \leq 1$，这正是我们想要证明的。

步骤 2 根据标准的切尔诺夫不等式，对于任意 $x > 0$，我们有

$$\mathbb{P}\left(\theta \sum_i (N^i(S_i) - \Lambda^i(S_i)) - \phi(\theta) \sum_i \Lambda^i(S_i) \geq x\right) = \mathbb{P}\left(E \geq e^x\right) \leq \mathbb{E}[E]e^{-x} \leq e^{-x}.$$

此外，由于 $\phi(\theta) \leq 2 \frac{\theta^2}{(1-\theta/3)}$，且根据假设，$\sum_i \Lambda^i(S_i) \leq v$，我们得到

$$\begin{aligned}
&\mathbb{P}\left(\theta \sum_i (N^i(S_i) - \Lambda^i(S_i)) - \phi(\theta) \sum_i \Lambda^i(S_i) \geq x\right) \\
&= \mathbb{P}\left(\sum_i N^i(S_i) - \Lambda^i(S_i) \geq \theta^{-1}(x + \phi(\theta) \sum_i \Lambda^i(S_i))\right) \\
&\geq \mathbb{P}\left(\sum_i N^i(S_i) - \Lambda^i(S_i) \geq \theta^{-1}x + \frac{\theta}{2(1-\theta/3)}v\right),
\end{aligned}$$

因此，

$$\mathbb{P}\left(\sum_i N^i(S_i) - \Lambda^i(S_i) \geq \theta^{-1}x + \frac{\theta}{2(1-\theta/3)}v\right) \leq e^{-x}. \tag{24}$$

步骤 3。 注意，式（24）中的界对任意 $\theta > 0$均成立。但对任意 $\theta \in (0, 3)$，我们有 e

$$\theta^{-1}(x + \phi(\theta)v) \geq \sqrt{2vx} + \frac{x}{3}.$$

因此，结合式（24）可得

$$\mathbb{P}\left(\sum_i N^i(S_i) - \Lambda^i(S_i) \geq \sqrt{2vx} + \frac{x}{3}\right) \leq e^{-x}.$$

□

引理 A.4。 设 $N = (N^i)_{i=1,\dots,n}$ n 是定义在 $[0,1] \times [0,1]^d$ 上、具有参数 f 且满足假设 2 的独立同分布时空霍克斯点过程。设 $(S_i)_{i=1,\dots,n} \subset [0,1]^{d+1}$ 为可能是随机的独立子集。那么，对任意 $x > 0$，存在与 n 和 $(S_i)_i$ 无关的常数 σ^2、$b > 0$，使得

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n N^i(S_i) - \Lambda^i(S_i) \geq x\right) \leq e^{-\frac{nx^2}{2(\sigma^2 + bx)}}, \tag{25}$$

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n N^i(S_i) - \Lambda^i(S_i) \leq -x\right) \leq e^{-\frac{nx^2}{2(\sigma^2 + bx)}}. \tag{26}$$

证明。 此证明基于标准伯恩斯坦不等式的矩版本，以及 N^i 具有指数矩这一事实，从而推广了 Hansen 等人 [2015]的结果。

我们首先回顾满足矩条件的标准伯恩斯坦不等式：设 $(Z_i)_i$ 独立同分布且中心化的随机变量，并且 $b, \sigma^2 > 0$ 使得对于任意 $i \in [n]$,

$$\begin{aligned}
\mathbb{E}[Z_i^2] &\leq \sigma^2 \\
\mathbb{E}[Z_i^k] &\leq \frac{1}{2}k!b^{k-2}\sigma^2, \quad k \geq 2.
\end{aligned}$$

于是可得

$$\mathbb{P}\left(\frac{1}{n}\sum_i Z_i \geq x\right) \leq \exp\left(-\frac{nx^2}{2(\sigma^2 + bx)}\right) \tag{27}$$

$$\mathbb{P}\left(\frac{1}{n}\sum_i Z_i \leq -x\right) \leq \exp\left(-\frac{nx^2}{2(\sigma^2 + bx)}\right). \tag{28}$$

我们的目标是将前述不等式应用于 $Z_i := N^i(S_i) - \Lambda^i(S_i)$。根据 Λ^i的定义，$\mathbb{E}[N^i(S_i) - \Lambda^i(S_i)] = 0$。此外，由于 $S_i \subset [0,1] \times [0,1]^d$，并利用Hansen 等人 [2015], 中的命题 2，可知存在常数 θ，$C > 0$ 该常数只依赖于 f，使得

$$\mathbb{E}[e^{\theta_i N^i(S_i)}] \leq \mathbb{E}[e^{\theta N^i([0,1]^{d+1})}] \leq C,$$

由于 $\Lambda(S_i) \geq 0$，这意味着

$$\mathbb{E}[e^{\theta_i (N^i(S_i) - \Lambda(S_i))}] \leq C,$$

这意味着 $N^i(S_i) - \Lambda(S_i)$ 存在指数矩。由此我们可以推导出

$$\begin{aligned}
\mathbb{E}[Z_i^2] &\leq \frac{2}{\theta}\mathbb{E}[e^{\theta N^i(S_i)}] \leq \frac{2C}{\theta} =: \sigma^2 \\
\mathbb{E}[Z_i^k] &\leq \frac{k!}{\theta^k}\mathbb{E}[e^{\theta N^i(S_i)}] \leq \frac{k!C}{\theta^k} \leq \frac{1}{2}k!\sigma^2\frac{1}{\theta^{k-2}} =: \frac{1}{2}k!\sigma^2 b^{k-2},
\end{aligned}$$

其中 $b := \frac{1}{\theta}$。因此，应用伯恩斯坦不等式（27），我们得到

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} N^i(S_i) - \Lambda^i(S_i) \geq x\right) \leq e^{-\frac{nx}{2(\sigma^2 + bx)}}.$$

同理，应用（27）可得到（26）中的左尾概率。 □

B 其他结果的证明

B.1 引理4.1的证明

L引理 B.1 (高概率事件)。对于任意 $\alpha > 0$，存在 δ_0、$c_\alpha > 0$，使得 t

$$\Omega_n := \left\{\frac{\underline{\mu}}{1 - \|g_0\|_1} - \delta_n \leq \frac{1}{n}\sum_i N^i[0,1]^{d+1} \leq \frac{\bar{\mu}}{1 - \|g_0\|_1} + \delta_n\right\} \cap \left\{\sup_{i=1,\ldots,n} N^i[0,1]^{d+1} \leq c_\alpha \log n\right\},$$

满足 $\delta_n = \delta_0 \frac{\log n}{\sqrt{n}}$且 $c > 0$。在假设 2 下，

$$\mathbb{P}_0[\Omega_n] \geq 1 - 3n^{-\alpha}.$$

证明。定义：

$$\Omega_1 := \left\{\sup_{i=1,\ldots,n} N^i[0,1]^{d+1} \leq c_\alpha \log n\right\}.$$

我们首先证明，对于任意 $\alpha > 0$，存在 c_α，使得 $\mathbb{P}_0(\Omega_1^c) \leq n^{-\alpha}$。首先注意，在我们的假设 2 下，每个时空过程 N_t^i、$_s$都被强度为的时间平稳点过程 $\bar{N}_t^i$随机占优

$$\bar{\lambda}_t^i(f_0) = \bar{\mu} + \int_{t-a}^{t-} \bar{g}_0(t-u)d\bar{N}_u^i,$$

其中 $\bar{g}_0(u) = \int g_0(u, \ s)ds$。因此，根据 Reynaud-Bouret 和 Roy [2007], 存在 $c > 0$使得

$$\mathbb{P}_0(\Omega_1^c) = \mathbb{P}_0(\sup_{i=1,\ldots,n} N^i[0,1]^{d+1} > c_\alpha \log n) \leq \mathbb{P}_0(\sup_{i=1,\ldots,n} \bar{N}^i[0,1] > c_\alpha \log n) \leq e^{-cc_\alpha \log n} = n^{-cc_\alpha} \leq n^{-\alpha},$$

如果 $c_\alpha > \alpha/c_\circ$

现在我们定义

$$\Omega_2 := \left\{ \frac{\underline{\mu}}{1-\|g_0\|_1} - \delta_n \le \frac{1}{n}\sum_i N^i[0,1]^{d+1} \le \frac{\bar{\mu}}{1-\|g_0\|_1} + \delta_n \right\},$$

并证明存在 δ_0，使得 $\mathbb{P}_0[\Omega_2^c] \le 2n^{-\alpha}$。我们再次利用以下事实：每个 $N^i_{t,\ s}$ 都被 $\bar{N}^i_t$随机占优，且以强度随机占优 N^i_t —

$$\underline{\lambda}^i_t(f_0) = \underline{\mu} + \int_{t-a}^{t-} g_0(t-s)d\underline{N}^i_s.$$

因此，

$$\frac{\underline{\mu}}{1-\|g_0\|_1} = \mathbb{E}_0[\underline{N}^i[0,1]] \le \mathbb{E}_0[N^i[0,1]^{d+1}] \le \mathbb{E}_0[\bar{N}^i[0,1]] = \frac{\bar{\mu}}{1-\|g_0\|_1},$$

并利用引理 A.4 以及 $S=[0,1]$,

$$\begin{aligned}
\mathbb{P}_0\left(\frac{1}{n}\sum_i N^i[0,1]^{d+1} \ge \frac{\bar{\mu}}{1-\|g_0\|_1} + \delta_n\right) &\le \mathbb{P}_0\left(\frac{1}{n}\sum_i \bar{N}^i[0,1] \ge \frac{\bar{\mu}}{1-\|g_0\|_1} + \delta_n\right) \\
&= \mathbb{P}_0\left(\frac{1}{n}\sum_i \bar{N}^i[0,1] - \mathbb{E}_0[\bar{N}^i[0,1]] \ge \delta_n\right) \\
&\le e^{-\frac{n\delta_n^2}{2(\sigma^2+b\delta_n)}} \le e^{-\delta_0 \log n/(4\sigma^2)} = n^{-\delta_0/(4\sigma^2)} \le n^{-\alpha}
\end{aligned}$$

当 n 足够大且 $\delta_0 \ge 4\sigma^2\alpha$ 时。类似地,

$$\begin{aligned}
\mathbb{P}_0\left(\frac{1}{n}\sum_i N^i[0,1]^{d+1} \le \frac{\underline{\mu}}{1-\|g_0\|_1} - \delta_n\right) &\le \mathbb{P}_0\left(\frac{1}{n}\sum_i \underline{N}^i[0,1] \le \frac{\underline{\mu}}{1-\|g_0\|_1} - \delta_n\right) \\
&= \mathbb{P}_0\left(\frac{1}{n}\sum_i \underline{N}^i[0,1] - \mathbb{E}_0[\underline{N}^i[0,1]] \le -\delta_n\right) \\
&\le e^{-\frac{n\delta_n^2}{2(\sigma^2+b\delta_n)}} \le e^{-\delta_0 \log n/(4\sigma^2)} = n^{-\delta_0/(4\sigma^2)} \le n^{-\alpha}.
\end{aligned}$$

因此我们可以得出结论，$\mathbb{P}_0(\Omega_2^c) \le 2n^{-\alpha}$，从而得到

$$\mathbb{P}_0(\Omega_n^c) \le \mathbb{P}_0(\Omega_1^c) + \mathbb{P}_0(\Omega_2^c) \le 3n^{-\alpha}.$$

□

B.2 引理证明 4.2

引理 B.2（库尔贝克-莱布勒）。在假设 *2* 成立且 $\bar{\epsilon}_n = o((\log n)^{-2})$ 时，存在 b_1，$b_2 > 0$使得对于任意 $f \in B_\infty(\bar{\epsilon}_n)$，

$$KL(f, f_0) := \mathbb{E}_0[\log L(N|f_0) - \log L(N|f)] \le \kappa n\bar{\epsilon}_n^2(1+o(1)) \tag{29}$$

$$\mathbb{P}_0(\log L(N|f_0) - \log L(N|f) > b_1 n\bar{\epsilon}_n^2) \le \frac{b_2}{n\bar{\epsilon}_n^2}, \tag{30}$$

with

$$\kappa := \frac{4\log 2}{\underline{\mu}}\left\{2 + 4\left(\frac{\bar{\mu}}{1-\|g_0\|_1} + \Lambda_{0,2}\right)\right\}. \tag{31}$$

证明。我们首先证明第一个结论（29）。证明分为三个主要步骤。第一步，我们将 KL 散度改写为关于真实条件强度测度的单个积分，并将其分解为两项，分别考虑高概率事件 Ω_n 及其补集 Ω_n^c。第二步和第三步，我们分别控制这两项。第一项可以通过条件强度 $\lambda_{t,s}^1(f)$ 与 $\lambda_{t,s}^1(f_0)$ 之间的平方 ℓ_2-距离来控制，而该距离本身又可以通过参数上的平方 ℓ_2-距离来控制，即 $\|f-f_0\|_2^2$。接着，我们利用柯西-施瓦茨不等式以及对 N 的四阶矩的界来控制第二项。下文在计算时空域 $[0,1]\times[0,1]^d$ 上的积分时，为了简化记号，我们省略积分上下限。另需注意，由于定义域是 $(d+1)$ 维超立方体，我们会多次用到 $\int 1dtds=1$。

令 $f\in B_\infty(f_0,\epsilon_T)$。我们有

$$
\begin{aligned}
KL(f,f_0) &= \mathbb{E}_0[\log L(N|f_0)-\log L(N|f)] \\
&= \mathbb{E}_0\left[\sum_i\int -\log\frac{\lambda_{t,s}^i(f)}{\lambda_{t,s}^i(f_0)}dN_{t,s}^i+\int(\lambda_{t,s}^i(f)-\lambda_{t,s}^i(f_0))dtds\right] \\
&= n\mathbb{E}_0\left[\int -\log\frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)}\lambda_{t,s}^1(f_0)dtds+\int(\lambda_{t,s}^1(f)-\lambda_{t,s}^1(f_0))dtds\right] \\
&= n\mathbb{E}_0\left[\int \psi(\frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)})\lambda_{t,s}^1(f_0)dtds\right] \\
&= \underbrace{n\,\mathbb{E}_0\left[\mathbb{1}_{\Omega_n}\int\psi(\frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)})\lambda_{t,s}^1(f_0)dtds\right]}_{=:I_1}+\underbrace{n\,\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\psi(\frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)})\lambda_{t,s}^1(f_0)dtds\right]}_{=:I_2},
\end{aligned}
$$

其中，在第三个等式中，我们定义了 $\psi(x)=-\log x+x-1\geq 0, x>0$。

关于 I_1 我们利用 $\psi(x)\leq -4\log(r)(x-1)^2$ 对任意 $x\geq r$ 和 $r>0$ 均成立。因此，我们得到比值 $\lambda_{t,s}^1(f)/\lambda_{t,s}^1(f_0)$ 在事件 Ω_n 上的下界 r。首先注意到

$$
\frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)}=1+\frac{\lambda_{t,s}^1(f)-\lambda_{t,s}^1(f_0)}{\lambda_{t,s}^1(f_0)}\geq 1-\frac{|\lambda_{t,s}^1(f)-\lambda_{t,s}^1(f_0)|}{\lambda_{t,s}^1(f_0)}.
$$

我们利用假设 $\frac{|\lambda_{t,s}^1(f)-\lambda_{t,s}^1(f_0)|}{\lambda_{t,s}^1(f_0)}\leq 2$ 并且在 Ω_n 上对其进行上界估计。由于 $f\in B_\infty(f_0,\bar{\epsilon}_n)$ 且 f_0 满足假设 2，

$$
\mu(t,s)\geq\mu_0(t,s)-\|\mu-\mu_0\|_\infty\geq\underline{\mu}-\bar{\epsilon}_n\geq\underline{\mu}/2,
$$

对任意 t,s，以及任意足够大的 n，这意味着

$$
\lambda_{t,s}^1(f)\geq\mu(t,s)\geq\underline{\mu}/2. \tag{32}
$$

类似地，我们有

$$
\mu(t,s)\leq\mu_0(t,s)+\|\mu-\mu_0\|_\infty\leq\bar{\mu}+\bar{\epsilon}_n\leq\bar{\mu}+1. \tag{33}
$$

此外，在 Ω_n 上，$\sup_i\sup_{t\in[0,1]}N^i[t,t-A]\leq c_\alpha\log n$，因此，

$$
\begin{aligned}
\left|\lambda_{t,s}^1(f_0)-\lambda_{t,s}^1(f)\right| &= \left|\mu_0(t,s)-\mu(t,s)+\sum_{t_i^1\leq t}(g_0-g)(t-t_i,s-s_i)\right| \\
&\leq \|\mu_0-\mu\|_\infty+\|g_0-g\|_\infty N^1(t,t-a) \\
&\leq \|\mu_0-\mu\|_\infty+c_\alpha\|g_0-g\|_\infty\log n.
\end{aligned} \tag{34}
$$

由于 $f \in B_\infty(f_0, \bar{\epsilon}_n)$，这意味着

$$\frac{|\lambda^1_{t,s}(f) - \lambda^1_{t,s}(f_0)|}{\lambda^1_{t,s}(f_0)} \leq \frac{\|\mu_0 - \mu\|_\infty + c_\alpha \|g_0 - g\|_\infty \log n}{\underline{\mu}} \leq \frac{\epsilon_n(1 + c_\alpha \log n)}{\underline{\mu}} \leq \frac{1}{2},$$

对于足够大的 n，利用假设 $\bar{\epsilon}_n = o((\log n)^{-1})$，因此，

$$\frac{\lambda^1_{t,s}(f)}{\lambda^1_{t,s}(f_0)} \geq \frac{1}{2}. \tag{35}$$

因此，结合 $r = \frac{1}{2}$，我们得到 $\psi(\frac{\lambda^1_{t,s}(f)}{\lambda^1_{t,s}(f_0)}) \leq 4\log(2)(\frac{\lambda^1_{t,s}(f)}{\lambda^1_{t,s}(f_0)} - 1)^2$ 由此可得

$$\begin{aligned} I_1 = \mathbb{E}_0\left[\mathbb{1}_{\Omega_n} \int \psi(\frac{\lambda^1_{t,s}(f)}{\lambda^1_{t,s}(f_0)}) \lambda^1_{t,s}(f_0) dtds\right] &\leq 4\log(2)\mathbb{E}_0\left[\int \left(\frac{\lambda^1_{t,s}(f)}{\lambda^1_{t,s}(f_0)} - 1\right)^2 \lambda^1_{t,s}(f_0) dtds\right] \\ &\leq \frac{4\log 2}{\underline{\mu}} \mathbb{E}_0\left[\int (\lambda^1_{t,s}(f) - \lambda^1_{t,s}(f_0))^2 dtds\right], \end{aligned}$$

在假设 2 下，利用 $(x+y)^2 \leq 2x^2 + 2y^2$，我们有

$$\begin{aligned} &\mathbb{E}_0\left[\int (\lambda^1_{t,s}(f) - \lambda^1_{t,s}(f_0))^2 dtds\right] \\ &\leq 2\|\mu - \mu_0\|_2^2 + 2\mathbb{E}_0\left[\int \left(\sum_{t^1_i < t} g(t - t^1_i, s - s^1_i) - g_0(t - t^1_i, s - s^1_i)\right)^2 dtds\right] \\ &= 2\|\mu - \mu_0\|_2^2 + 2\mathbb{E}_0\left[\int \left(\int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} (g(t-u, s-v) - g_0(t-u, s-v)) dN^1_{u,v}\right)^2 dtds\right] \\ &\leq 2\|\mu - \mu_0\|_2^2 \qquad (36) \\ &\quad + 4\mathbb{E}_0\left[\int \left(\int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} (g(t-u, s-v) - g_0(t-u, s-v))(dN^1_{u,v} - \lambda^1_{u,v}(f_0) dudv)\right)^2 dtds\right] \\ &\quad + 4\mathbb{E}_0\left[\int \left(\int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} (g(t-u, s-v) - g_0(t-u, s-v)) \lambda^1_{u,v}(f_0) dudv\right)^2 dtds\right] \qquad (37) \end{aligned}$$

为了对右端的第二项和第三项进行界定（37），我们将使用以下恒等式（例如，参见 Karr [2017] 中的定理 B12）：对于任意确定的平方可积函数 h,

$$\begin{aligned} &\mathbb{E}_0\left[\int \left(\int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} h(t-u, s-v)(dN^1_{u,v} - \lambda^1_{u,v}(f_0) dudv)\right)^2 dtds\right] \\ &= \mathbb{E}_0\left[\int \int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} h^2(t-u, s-v) \lambda^1_{u,v}(f_0) dudvdtds\right] \\ &= \int \int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} h^2(t-u, s-v) \mathbb{E}_0\left[\lambda^1_{u,v}(f_0)\right] dudvdtds. \end{aligned} \tag{38}$$

我们还将使用 $\max_{u,v} \mathbb{E}_0\left[\lambda^1_{u,v}(f_0)\right]$ 的上界。注意，

$$\begin{aligned} \mathbb{E}_0[\lambda^1_{t,s}(f_0)] &= \mu_0(t,s) + \mathbb{E}_0\left[\int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} g_0(t-u, s-v) dN^1_{u,v}\right] \\ &= \mu_0(t,s) + \mathbb{E}_0\left[\int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} g_0(t-u, s-v) \lambda^1_{u,v}(f_0) dudv\right] \\ &\leq \bar{\mu} + \|g_0\|_1 \max_{u,v} \mathbb{E}_0\left[\lambda^1_{u,v}(f_0)\right], \end{aligned}$$

这意味着

$$\max_{t,s} \mathbb{E}_0 \left[\lambda^1_{t,s}(f_0)\right] \leq \frac{\bar{\mu}}{1-\|g_0\|_1} < \infty, \tag{39}$$

在假设 2下。利用 (38) 以及 $h = g - g_0$ 和 (39)，我们得到

$$\mathbb{E}_0 \left[\int \left(\int_{u:u\in[t-a,t)} \int_{v:\|v-s\|\leq b} (g(t-u,s-v) - g_0(t-u,s-v))(dN^1_{u,v} - \lambda^1_{u,v}(f_0)dudv) \right)^2 dtds \right]$$
$$\leq \|g-g_0\|_2^2 \frac{\bar{\mu}}{1-\|g_0\|_1}.$$

此外，由柯西-施瓦茨不等式可得，

$$\mathbb{E}_0 \left[\int \left(\int (g-g_0)(t-u,s-v)\lambda^1_{u,v}(f_0)dudv \right)^2 dtds \right]$$
$$\leq \mathbb{E}_0 \left[\int \left(\int (g-g_0)^2(t-u,s-v)dvdu \int (\lambda^1_{u,v}(f_0))^2 dvdu \right) dtds \right]$$
$$\leq \|g-g_0\|_2^2 \mathbb{E}_0 \left[\int \int (\lambda^1_{u,v}(f_0))^2 dvdudtds \right] = \|g-g_0\|_2^2 \Lambda_{0,2},$$

with

$$\Lambda_{0,2} := \mathbb{E}_0 \left[\int (\lambda^1_{u,v}(f_0))^2 dvdu \right].$$

我们声称$\Lambda_{0,\ 2} < \infty$，因为

$$\Lambda_{0,2} \leq 2\|\mu_0\|_2^2 + 2\|g_0\|_2^2 \mathbb{E}_0 \left[(N[0,1]^{d+1})^2\right],$$

$\|\mu_0\|_2^2 < \infty$在假设 2下，且 $\mathbb{E}_0 \left[(N[0,1]^{d+1})^2\right] < \infty$。 $N[0,1]^{d+1}$的二阶矩存在，这是因为时空点过程可以视为带标记的时间点过程（TPP），而任何非爆炸性时间点过程在有限域上都具有指数矩，这意味着

$$\mathbb{E}_0 \left[(N^1[0,1]^{d+1})^k\right] < \infty, \quad k > 0. \tag{40}$$

因此，由 (37)，我们得到

$$\mathbb{E}_0 \left[\int (\lambda^1_{t,x}(f) - \lambda^1_{t,x}(f_0))^2 dtdx \right] \leq 2\|\mu-\mu_0\|_2^2 + 4\left(\frac{\bar{\mu}}{1-\|g_0\|_1} + \Lambda_{0,2}\right)\|g-g_0\|_2^2$$
$$\leq \left\{2 + 4\left(\frac{\bar{\mu}}{1-\|g_0\|_1} + \Lambda_{0,2}\right)\right\}\|f-f_0\|_2^2, \tag{41}$$

因此，

$$I_1 \leq \frac{4\log 2}{\underline{\mu}} \left\{2 + 4\left(\frac{\bar{\mu}}{1-\|g_0\|_1} + \Lambda_{0,2}\right)\right\}\|f-f_0\|_2^2 \leq \frac{4\log 2}{\underline{\mu}} \left\{2 + 4\left(\frac{\bar{\mu}}{1-\|g_0\|_1} + \Lambda_{0,2}\right)\right\}\bar{\epsilon}_n^2 = \kappa\bar{\epsilon}_n^2,$$

利用 $f \in B_\infty(f_0,\ \bar{\epsilon}_n)$，$\|f-f_0\|_2^2 \leq \|f-f_0\|_\infty^2 \leq \bar{\epsilon}_n^2$ 并且

$$\kappa = \frac{4\log 2}{\underline{\mu}} \left\{2 + 4\left(\frac{\bar{\mu}}{1-\|g_0\|_1} + \Lambda_{0,2}\right)\right\}.$$

现在证明 $I_2 = o(\bar{\epsilon}_n^2)$。一方面，根据假设2 并利用（33），

$$\frac{\lambda_{t,s}^1(f_0)}{\lambda_{t,s}^1(f)} = 1 + \frac{\lambda_{t,s}^1(f_0) - \lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f)} \leq 1 + \frac{|\lambda_{t,s}^1(f) - \lambda_{t,s}^1(f_0)|}{\lambda_{t,s}^1(f)} \leq 1 + 2\frac{\|\mu - \mu_0\|_\infty + \|g - g_0\|_\infty N[0,1]^{d+1}}{\lambda_{t,s}^1(f)}$$
$$\leq 1 + \frac{2\bar{\epsilon}_n(1 + N[0,1]^{d+1})}{\lambda_{t,s}^1(f)} \leq 1 + \frac{2\bar{\epsilon}_n(1 + N[0,1]^{d+1})}{\underline{\mu}}, \tag{42}$$

另一方面，

$$\frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)} \leq 1 + \frac{\bar{\epsilon}_n(1 + N[0,1]^{d+1})}{\lambda_{t,s}^1(f_0)} \leq 1 + \frac{\bar{\epsilon}_n(1 + N[0,1]^{d+1})}{\underline{\mu}}. \tag{43}$$

因此，利用 $\log x \leq x - 1$,

$$\begin{aligned}
I_2 &= \mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\left\{-\log(\frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)}) + \frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)} - 1\right\}\lambda_{t,s}^1(f_0)dtds\right] \\
&= \mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\left\{\log(\frac{\lambda_{t,s}^1(f_0)}{\lambda_{t,s}^1(f)}) + \frac{\lambda_{t,s}^1(f)}{\lambda_{t,s}^1(f_0)} - 1\right\}\lambda_{t,s}^1(f_0)dtds\right] \\
&\leq \mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\left\{\log(1 + \frac{2\epsilon_n(1 + N^1[0,1]^{d+1})}{\underline{\mu}}) + \frac{\epsilon_n(1 + N^1[0,1]^{d+1})}{\underline{\mu}}\right\}\lambda_{t,s}^1(f_0)dtds\right] \\
&\leq \mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\left\{\frac{2\epsilon_n(1 + N^1[0,1]^{d+1})}{\underline{\mu}} + \frac{\epsilon_n(1 + N^1[0,1]^{d+1})}{\underline{\mu}}\right\}\lambda_{t,s}^1(f_0)dtds\right] \\
&= \frac{3}{\bar{\mu}}\bar{\epsilon}_n\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\lambda_{t,s}^1(f_0)dtds\right] + \frac{3}{\bar{\mu}}\bar{\epsilon}_n\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}N^1[0,1]^{d+1}\int\lambda_{t,s}^1(f_0)dtds\right] \\
&\leq C\bar{\epsilon}_n\sqrt{\mathbb{P}_0(\Omega_n^c)\mathbb{E}_0\left[\int(\lambda_{t,s}^1(f_0))^2dtds\right]} + C\bar{\epsilon}_n\sqrt{\mathbb{P}_0(\Omega_n^c)\mathbb{E}_0\left[(N^1[0,1]^{d+1})^2\int(\lambda_{t,s}^1(f_0)^2dtds\right]} \\
&\leq C\bar{\epsilon}_n\sqrt{\mathbb{P}_0(\Omega_n^c)\mathbb{E}_0\left[\int(\lambda_{t,s}^1(f_0))^2dtds\right]} + C\bar{\epsilon}_n\sqrt{\mathbb{P}_0(\Omega_n^c)}\sqrt{\mathbb{E}_0\left[(N^1[0,1]^{d+1})^4\right]\mathbb{E}_0\left[\int(\lambda_{t,s}^1(f_0))^4dtds\right]}
\end{aligned}$$

其中 $C = \frac{3}{\bar{\mu}}$ 并在最后两个不等式中使用柯西-施瓦茨不等式。此外，利用引理 4.1，$\mathbb{P}_0(\Omega_n^c) \leq n^{-\alpha}$ 对于任意 $\alpha > 0$。由（40），我们有 $\mathbb{E}_0\left[(N^1[0,1]^{d+1})^4\right] < \infty$ 因此，

$$\mathbb{E}_0\left[(\lambda_{t,s}^1(f_0))^4\right] \leq 8\mu_0(t,s)^4 + 8\|g_0\|_\infty^4\mathbb{E}_0\left[(N^1[0,1]^{d+1})^4\right] < \infty.$$

因此，$I_2 = O(n^{-\alpha/2}) = o(\bar{\epsilon}_n^2)$ 对任意 $\alpha > 1$。因此我们得出结论：当 n 足够大时 h,

$$KL(f, f_0) = n(I_1 + I_2) \leq \kappa n\bar{\epsilon}_n^2(1 + o(1)),$$

这证明了引理 4.2的第一个结论。

现在我们证明第二个陈述。证明依赖于对 $\log L(f_0|N)$ –$\log L(f|N)$ 的方差进行界定，并应用切比雪夫不等式。为界定 $\log L(f_0|N)$–$\log L(f|N)$ 的方差，我们将其分别在 Ω_n 和 Ω_n^c 上进行分解。回顾

$$\log L(f_0|N) - \log L(f|N) = \sum_{i=1}^n\int\log\frac{\lambda_{t,s}^i(f_0)}{\lambda_{t,s}^i(f)}dN_{t,s}^i + \int(\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0))dtds.$$

对于任意 $i \in [n]$，令

$$Z_i := \int\log\frac{\lambda_{t,s}^i(f_0)}{\lambda_{t,s}^i(f)}dN_{t,s}^i + \int(\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0))dtds.$$

注意， $\mathbb{E}_0[Z_i] = KL(f, f_0)$。我们有

$$\mathbb{E}_0[Z_i^2] = \mathbb{E}_0\left[\left(\int \log\frac{\lambda^i_{t,s}(f_0)}{\lambda^i_{t,s}(f)}dN^i_{t,s} + \int(\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_0))dtds\right)^2\right]$$
$$= \mathbb{E}_0\left[\left(\int \log\frac{\lambda^i_{t,s}(f_0)}{\lambda^i_{t,s}(f)}\lambda^i_{t,s}(f_0)dtds + \int \log\frac{\lambda^i_{t,s}(f_0)}{\lambda^i_{t,s}(f)}(dN^i_{t,s}-\lambda^i_{t,s}(f_0)dtds) + \int(\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_0))dtds\right)^2\right]$$
$$= \mathbb{E}_0\left[\left(\int \psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)\lambda^i_{t,s}(f_0)dtds + \int \log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}(dN^i_{t,s}-\lambda^i_{t,s}(f_0))dtds)\right)^2\right]$$
$$\le 2\mathbb{E}_0\left[\left(\int \psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)\lambda^i_{t,s}(f_0)dtds\right)^2\right] + 2\mathbb{E}_0\left[\left(\int \log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}(dN^i_{t,s}-\lambda^i_{t,s}(f_0))dtds)\right)^2\right]$$
$$\le 2\mathbb{E}_0\left[\int \psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2(\lambda^i_{t,s}(f_0))^2dtds\right] + 2\mathbb{E}_0\left[\int(\log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)})^2\lambda^i_{t,s}(f_0)dtds\right].$$

利用柯西-施瓦茨不等式和最后一个不等式中的（38）。我们首先对其进行上界估计

$$\mathbb{E}_0\left[\mathbb{1}_{\Omega_n}\int \psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2(\lambda^i_{t,s}(f_0))^2dtds\right] \quad \text{and} \quad \mathbb{E}_0\left[\mathbb{1}_{\Omega_n}\int(\log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)})^2\lambda^i_{t,s}(f_0)dtds\right]$$

回顾（35），在 Ω_n 上且当 n 足够大时，$\frac{\lambda^1_{t,s}(f)}{\lambda^1_{t,s}(f_0)} \ge \frac{1}{2}$。因此，再次利用对于任意满足 $r = \frac{1}{2}$ 的 $x \ge r > 0$，有 $\psi(x) \le -\log(r)(x-1)^2$，并结合假设 2，可得

$$\psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2(\lambda^i_{t,s}(f_0))^2 \le (\log 2)^2\frac{(\lambda^i_{t,s}(f_0)-\lambda^i_{t,s}(f))^4}{\lambda^i_{t,s}(f_0)^2} \le \frac{(\log 2)^2}{\underline{\mu}^2}(\lambda^i_{t,s}(f_0)-\lambda^i_{t,s}(f))^4.$$

因此，

$$\mathbb{E}_0\left[\mathbb{1}_{\Omega_n}\int \psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2(\lambda^i_{t,s}(f_0))^2dtds\right] \le \frac{(\log 2)^2}{\underline{\mu}^2}\mathbb{E}_0\left[\mathbb{1}_{\Omega_n}\int(\lambda^i_{t,s}(f_0)-\lambda^i_{t,s}(f))^4dtds\right]$$
$$\le \frac{(\log 2)^2}{\underline{\mu}^2}\bar{\epsilon}_n^4(1+c_\alpha\log n)^4 = o(\bar{\epsilon}_n^2),$$

在最后一个不等式中使用（34），并利用假设，有 $\bar{\epsilon}_n = o((\log n)^{-2})$。此外，利用 $|\log(x)| \le -2\log(r)|x-1|$ 对任意 $x \ge r > 0$ 且满足 $r = \frac{1}{2}$，还可得

$$(\log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)})^2\lambda^i_{t,s}(f_0) \le (2\log 2)^2\frac{(\lambda^i_{t,s}(f_0)-\lambda^i_{t,s}(f))^2}{\lambda^i_{t,s}(f_0)} \le \frac{(2\log 2)^2}{\underline{\mu}}(\lambda^i_{t,s}(f_0)-\lambda^i_{t,s}(f))^2,$$

这意味着，利用（41）可得

$$\mathbb{E}_0\left[\mathbb{1}_{\Omega_n}\int(\log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)})^2\lambda^i_{t,s}(f_0)dtds\right] \le \frac{(2\log 2)^2}{\underline{\mu}}\mathbb{E}_0\left[\int(\lambda^i_{t,s}(f_0)-\lambda^i_{t,s}(f))^2dtds\right] \le (\log 2)\kappa\bar{\epsilon}_n^2.$$

现在我们对剩余项进行界定

$$\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int \psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2(\lambda^i_{t,s}(f_0))^2dtds\right] \quad \text{and} \quad \mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int(\log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)})^2\lambda^i_{t,s}(f_0)dtds\right].$$

利用（42）和（43），可得

$$\left|\log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right| = \log\frac{\lambda^i_{t,s}(f)}{\lambda^1_{t,s}(f_0)} \vee \log\frac{\lambda^i_{t,s}(f_0)}{\lambda^1_{t,s}(f)} \leq \log\left(1+\frac{\bar{\epsilon}_n(1+N[0,1]^{d+1})}{\lambda^i_{t,s}(f_0)\wedge\lambda^i_{t,s}(f)}\right) \leq \frac{\bar{\epsilon}_n(1+N[0,1]^{d+1})}{\lambda^i_{t,s}(f_0)\wedge\lambda^i_{t,s}(f)}.$$

因此，我们得到

$$\begin{aligned}
&\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\left(\log\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2\lambda^i_{t,s}(f_0)dtds\right] \leq \mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\bar{\epsilon}_n^2(1+N[0,1]^{d+1})^2(\lambda^1_{t,s}(f_0))^{-1}(1\vee\frac{\lambda^1_{t,s}(f_0)}{\lambda^1_{t,s}(f)})^2\right]\\
&\leq \frac{4}{\underline{\mu}}\bar{\epsilon}_n^2\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}(N[0,1]^{d+1})^2(1+\frac{2\bar{\epsilon}_n(1+N[0,1]^{d+1})}{\underline{\mu}})^2\right]\\
&\lesssim \bar{\epsilon}_n^2\sqrt{\mathbb{P}_0(\Omega_n^c)\mathbb{E}_0\left[(N[0,1]^{d+1})^8\right]} = o(\bar{\epsilon}_n^2),
\end{aligned} \tag{44}$$

u再次得到（40），并且 $\mathbb{P}_0(\Omega_n^c) = o(1)$。类似地，利用 $(a+b)^2 \leq 2a^2+2b^2$，

$$\begin{aligned}
&\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\psi\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2(\lambda^i_{t,s}(f_0))^2dtds\right]\\
&\leq 2\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int\log\left(\frac{\lambda^i_{t,s}(f)}{\lambda^i_{t,s}(f_0)}\right)^2(\lambda^i_{t,s}(f_0))^2dtds\right] + 2\mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\int(\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_0))^2dtds\right]\\
&\lesssim \mathbb{E}_0\left[\mathbb{1}_{\Omega_n^c}\bar{\epsilon}_n^2(1+N[0,1]^{d+1})^2(1+\frac{\bar{\epsilon}_n(1+N[0,1]^{d+1})}{\underline{\mu}})\right]\\
&\quad+\sqrt{\mathbb{P}_0(\mathbb{1}_{\Omega_n^c})\mathbb{E}_0\left[\left(\int(\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_0))^2dtds\right)^2\right]} = o(\bar{\epsilon}_n^2),
\end{aligned}$$

其中，在最后一个等式中，我们对右端的第一项使用了（44），并且利用了

$$\begin{aligned}
\mathbb{E}_0\left[\left(\int(\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_0))^2dtds\right)^2\right] &\leq \mathbb{E}_0\left[\int(\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_0))^4dtds\right]\\
&\leq 8\|\mu-\mu_0\|_\infty^4 + 8\|g-g_0\|_\infty^4\mathbb{E}_0\left[(N^i[0,1]^{d+1})^4\right] \lesssim \bar{\epsilon}_n^4 = o(\bar{\epsilon}_n^2),
\end{aligned}$$

利用（34）。因此我们可以得出

$$\mathbb{V}ar[Z_i^2] = \mathbb{E}_0[Z_i^2] - \mathbb{E}_0^2[Z_i] \leq \mathbb{E}_0[Z_i^2] \leq \epsilon_n^2(\kappa(\log 2)+o(1)) \leq 2\kappa\epsilon_n^2,$$

当 n 足够大时。接下来应用切比雪夫不等式：对于任意 $x>0$,

$$\mathbb{P}_0\left[\log L(f_0|N) - \log L(f|N) - KL(f,f_0) > x\right] \leq \frac{n\mathbb{E}_0[Z_i^2]}{x^2}$$

因此，对于任意足够大的 $\epsilon>0$ 和 n，

$$\mathbb{P}_0\left[\log L(f_0|N) - \log L(f|N) > \kappa(1+\epsilon)n\bar{\epsilon}_n^2 + x\right] \leq \frac{2\kappa n\bar{\epsilon}_n^2}{x^2},$$

并且对于任意足够大的 $\epsilon>0$ 和 n，结合 $x = x_1 n\bar{\epsilon}_n^2$ 和 $x_1>0$，我们得到，

$$\begin{aligned}
&\mathbb{P}_0\left[\log L(f_0|N) - \log L(f|N) > (\kappa(1+\bar{\epsilon})+x_1)n\bar{\epsilon}_n^2\right] \leq \frac{2\kappa}{x_1^2 n\bar{\epsilon}_n^2}\\
\iff\ &\mathbb{P}_0\left[\log L(f_0|N) - \log L(f|N) > b_1 n\bar{\epsilon}_n^2\right] \leq \frac{b_2}{n\bar{\epsilon}_n^2},
\end{aligned}$$

其中 $b_1 = \kappa(1+\epsilon)+x_1 > \kappa$ 和 $b_2 = \frac{2\kappa}{x_1^2}$。通过选择 $x_1 = \kappa\epsilon$，对于某个固定的 $\epsilon>0$，我们得到$b_1 = \kappa(1+2\epsilon)$ $b_2 = \frac{2}{\kappa\epsilon^2}$，由此完成本引理的证明。

□

B.3 引理 4.3 的证明

引理 B.3 (检验)。 在假设 *4* 和 *2*下，存在一个检验函数 $\phi := \phi(N, \epsilon_n)$ 使得

$$\mathbb{E}_0[\phi \mathbb{1}_{\Omega_n}] = o(1)$$
$$\sup_{f \in \mathcal{F}_n} \mathbb{E}_f[(1-\phi)\mathbb{1}_{\Omega_n}\mathbb{1}_{f \in A_n}] \le e^{-b_2 n\epsilon_n^2}$$

其中 $b_2 > c_1$， A_n 定义见 (12)，且Ω_n 定义见引理*4.1*。

证明。 主要思路是构造单独的检验函数 $\phi(f_1) = \phi(f_1, N, \epsilon_n)$，用于检验参数 $f_1 \in A_n$ 是否符合 f_0:

$$\phi(f_1) := \mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i}) - \Lambda^{i,0}(S_{1,i}) > v_n} \vee \mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i}^c) - \Lambda^{i,0}(S_{1,i}^c) > v_n} \tag{45}$$

其中 $S_{1,\ i} := \{(t,\ s) \in [0,1]^{d+1} :\ \lambda_{f_1}^i(t,\ s) \ge \lambda_{f_0}^i(t,\ s)\}$， $v_n > 0$ 是一个序列，且

$$\Lambda^{i,0}(S_{1,i}) = \int_{S_{1,i}} \lambda_{f_0}^i(t,s)dtds.$$

注意， $\Lambda^{i,0}(S_{1,i})$ 是 N^i 在 S_1 上、在 $\mathbb{P}_0$ 条件下的补偿子，因此，在 $\mathbb{P}_0$ 条件下，$\frac{1}{n}\sum_i N^i(S_{1,i}) - \Lambda^{i,0}(S_{1,i})$ 和 $\frac{1}{n}\sum_i N^i(S_{1,i}^c) - \Lambda^{i,0}(S_{1,i}^c)$ 都集中在 0 附近。因此，直观地说，在 $\mathbb{P}_0$ （原假设）下，只要 v_n 不太小， $\phi(f_1)$ 就应趋于 0。在 $\mathbb{P}_{f_1}$ （备择假设）下，平均而言， $\lambda_{f_1}^i(t,s)$ 大多要么大于（情况 1），要么小于（情况 2） $\lambda_0^i(t,s)$。在情况 1 下， $|S_{1,i}| > |S_{1,i}^c|$，并注意到 $\Lambda^{i,1}(S_{1,i}) > \Lambda^{i,0}(S_{1,i})$。因此，在 $\mathbb{P}_{f_1}$ 条件下，在这种情况下，$\frac{1}{n}\sum_i N^i(S_{1,i}) - \Lambda^{i,0}(S_{1,i}) = \frac{1}{n}\sum_i N^i(S_{1,i}) - \Lambda^{i,1}(S_{1,i}) + \Lambda^{i,1}(S_{1,i}) - \Lambda^{i,0}(S_{1,i})$ 将集中于 $\lim_{n\to\infty}\frac{1}{n}\sum_i \Lambda^{i,1}(S_{1,i}) - \Lambda^{i,0}(S_{1,i}) > 0$，这意味着 $\phi(f_1)$ 应趋于 1。类似地，对于情况 2，可用 $S_{1,i}^c$ 替代 $S_{1,i}$ 进行推理。

为了控制我们在事件 Ω_n 上检验的第一类错误和第二类错误，我们采用伯恩斯坦集中不等式，并将Hansen 等人 [2015] 命题 2 中的方法调整到引理A.4所述的时空情境:

$$\mathbb{P}_0\left(\frac{1}{n}\sum_{i=1}^n N^i(S_i) - \Lambda^i(S_i) \ge x\right) \le e^{-\frac{nx^2}{2(\bar{\sigma}^2 + \bar{b}x)}},$$

对任意 $x > 0$，且满足 $\bar{\sigma}$， b 常数与子集 $(S_i)_i$ 无关。

对于第一类错误 $\mathbb{E}_0[\phi]$，我们将利用 L_1 的一个最小覆盖网， $\mathcal{F}_n$ （定义见假设 4），其由半径为 $\zeta j\epsilon_n$ 的球构成，其中 $\zeta > 0$ 为一个常数，其取值稍后确定，记为 $\mathcal{N}_j$。然后，在假设4下， $\mathcal{N}_j$ 的基数（即覆盖数）有如下界:

$$|\mathcal{N}_j| = C(\zeta j\epsilon_n, \mathcal{F}_j, \|\cdot\|_1) \le C(\zeta_0\epsilon_n, \mathcal{F}_n, \|\cdot\|_1) =: |\mathcal{N}_0| \le e^{c_3 n\epsilon_n^2},$$

如果 $j \ge \zeta_0/\zeta$，则此结论成立；而这在 $M \ge \zeta_0/\zeta$ 时成立，并且存在 $\mathcal{N}_0$ 一个最小的 L_1 覆盖网，其覆盖 $\mathcal{F}_n$ 所用的球半径为 $\zeta_0\epsilon_n$ （回顾一下， ζ_0 在假设 4中定义）。对于任意 $f_j \in \mathcal{N}_j$，我们定义检验函数$\phi(f_j)$，如式 (45) 所示，其中序列 v_n 依赖于 j，其定义如下。接着，我们定义全局检验函数如下:

$$\phi = \max_{j \ge M}\max_{f_j \in \mathcal{N}_j} \phi(f_j).$$

对于第二类错误 $\sup_{f\in\mathcal{F}_n}\mathbb{E}_f[(1-\phi)\mathbb{1}_{f\in A_n}]$，我们采用 A_n （定义见（12））的切片方法。对于任意 $j \ge M$，我们定义“切片”

$$\mathcal{F}_j := \{f \in \mathcal{F}_n : j\epsilon_n \le d_S(f, f_0) \le (j+1)\epsilon_n\},$$

从而可以将 A_n 重新表示为（不失一般性地假设 M 为整数）

$$A_n = \bigcup_{j=M}^{\infty} \mathcal{F}_j.$$

由于 $\mathcal{F}_j \subseteq \mathcal{F}_n \subseteq \cup_{f_j \in \mathcal{N}_j}\{f \in \mathcal{F}_n : \ \|f - f_1\|_1 \leq \zeta j\epsilon_n\}$，因此

$$\begin{aligned}\sup_{f\in\mathcal{F}_n} \mathbb{E}_f[(1-\phi)\mathbb{1}_{f\in A_n}] &\leq \sum_{j\geq M} \sup_{f\in\mathcal{F}_n} \mathbb{E}_f[(1-\phi)\mathbb{1}_{f\in\mathcal{F}_j}] \\ &\leq \sum_{j\geq M} \sup_{f_j\in\mathcal{N}_j} \sup_{f\in\mathcal{F}_n:\|f-f_j\|_1\leq\zeta j\epsilon_n} \mathbb{E}_f[(1-\phi(f_j))\mathbb{1}_{f\in\mathcal{F}_j}]\end{aligned}$$

现在我们指定 v_n 在 $\phi(f_j)$ 中的序列，并将 $\phi(f_j)$ 分解为两个子检验，即对于任意 $f_j \in \mathcal{N}_j$，我们定义

$$\begin{aligned}\phi(f_j) &= \phi^+(f_j) \vee \phi^-(f_j) \\ \phi^+(f_j) &= \mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i}) - \Lambda^{i,0}(S_{1,i}) > v_n} \\ \phi^-(f_j) &= \mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S^c_{1,i}) - \Lambda^{i,0}(S^c_{1,i}) > v_n}.\end{aligned}$$

我们定义 $v_n = x_1 j\epsilon_n$ ($x_1 > 0$，这是一个稍后确定取值的常数)，并应用引理 A.4，其中 $x = x_1 j\epsilon_n = v_n$、 $x_1 > 0$ 和 $S_i = S_{1,\ i}$:

$$\mathbb{E}_0[\phi^+(f_j)] = \mathbb{P}_0\left[\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i}) - \Lambda^{i,0}(S_{1,i}) > x_1 j\epsilon_n\right] \leq e^{-\frac{x_1^2 n j^2 \epsilon_n^2}{2(\bar{\sigma}^2 + \bar{b}x_1 j\epsilon_n)}}.$$

我们可以对 $S_i = S^c_{1,\ i}$ 应用相同的不等式，并得到对于检验函数

$$\phi^-(f_j) = \mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S^c_{1,i}) - \Lambda^{i,0}(S^c_{1,i}) > v_n},$$

that

$$\mathbb{E}_0[\phi^-(f_j)] \leq e^{-\frac{x_1^2 n j^2 \epsilon_n^2}{2(\bar{\sigma}^2 + \bar{b}x_1 j\epsilon_n)}}.$$

因此，我们得到:

$$\mathbb{E}_0[\phi(f_j)] = \mathbb{E}_0[\phi^+(f_j) \vee \phi^-(f_j)] \ \leq \mathbb{E}_0[\phi^+(f_j)] + \mathbb{E}_0[\phi^-(f_j)] \leq 2e^{-\frac{x_1^2 n j^2 \epsilon_n^2}{2(\bar{\sigma}^2 + \bar{b}x_1 j\epsilon_n)}}.$$

我们分两种情况讨论:

- $j\epsilon_n > \bar{\sigma}^2/(\bar{b}x_1)$。那么上面的右端为

$$\leq 2e^{-\frac{x_1^2 n j^2 \epsilon_n^2}{4\bar{b}x_1 j\epsilon_n}} = 2e^{-\frac{x_1 n j\epsilon_n}{4b}}.$$

- $j\epsilon_n \leq \bar{\sigma}^2/(\bar{b}x_1)$。那么上面的右端为

$$\leq 2e^{-\frac{x_1^2 n j^2 \epsilon_n^2}{4\bar{\sigma}^2}}.$$

回想一下，我们的全局检验函数定义为

$$\phi := \max_{j\geq M} \max_{f_1\in\mathcal{N}_j} \phi(f_j) \leq \sum_{j\geq M} \sum_{f_1\in\mathcal{N}_j} \phi(f_j).$$

注意，前述不等式右端第二个求和中的项数不超过$|\mathcal{N}_0|$。

第一类错误。 我们可以将 $\mathbb{E}_0[\phi]$上界估计为

$$\begin{aligned}\mathbb{E}_0[\phi] &\le \sum_{j\ge M}\sum_{f_1\in\mathcal{N}_j}\mathbb{E}_0[\phi_{1,j}] \le \sum_{j>\bar\sigma^2/(\bar b x_1\epsilon_n)} 2|\mathcal{N}_0|e^{-\frac{x_1 nj\epsilon_n}{4\bar b}} + \sum_{M\le j\le\bar\sigma^2/(\bar b x_1\epsilon_n)} 2|\mathcal{N}_0|e^{-\frac{x_1^2 nj^2\epsilon_n^2}{4\bar\sigma^2}}\\ &\le \sum_{j>\bar\sigma^2/(\bar b x_1\epsilon_n)} 2e^{c_3 n\epsilon_n^2}e^{-\frac{x_1 nj\epsilon_n}{4\bar b}} + \sum_{M\le j\le\bar\sigma^2/(\bar b x_1\epsilon_n)} 2e^{c_3 n\epsilon_n^2}e^{-\frac{x_1^2 nj^2\epsilon_n^2}{4\bar\sigma^2}}\\ &\le 2e^{c_3 n\epsilon_n^2}\sum_{j\ge M} e^{-\frac{\min(x_1,x_1^2)nj\epsilon_n^2}{4\max(\bar b,\bar\sigma^2)}}\\ &\le 4e^{-\frac{\min(x_1,x_1^2)Mn\epsilon_n^2}{8\max(\bar b,\bar\sigma^2)}} = o(1).\end{aligned}$$

其中，在第三个不等式中，我们利用了 $\max\left(e^{-\frac{x_1 nj\epsilon_n}{4\bar b}}, e^{-\frac{x_1^2 nj^2\epsilon_n^2}{4\bar\sigma^2}}\right) \le e^{-\frac{\min(x_1,x_1^2)nj\epsilon_n^2}{4\bar b\,\bar\sigma^2}\max(\bar b,\bar\sigma^2)}$；在最后一个不等式中，我们利用了 $\min\left(\frac{x_1}{8\bar b}, \frac{x_1^2}{\bar\sigma^2}, M\right)\max(\bar b,\bar\sigma^2) > c_3$ 对于足够大的 M 和 n 成立。

第二类错误。 现在我们对 $\sup_{f\in\mathcal{F}_n}\mathbb{E}_f[(1-\phi)\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in A_n}]$进行上界估计。首先注意，对于任意 $j'\ge M$ 和 $f_1'\in\mathcal{N}_{j'}$,

$$1-\phi = 1-\max_{j\ge M}\max_{f_j\in\mathcal{N}_j}\phi(f_j) \le 1-\phi(f_{j'}).$$

回忆一下，由于 $A_n=\bigcup_{j\ge M}\mathcal{F}_j$，我们有

$$\begin{aligned}\sup_{f\in\mathcal{F}_n}\mathbb{E}_f[(1-\phi)\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in A_n}] &\le \sum_{j\ge M}\sup_{f\in\mathcal{F}_n}\mathbb{E}_f[(1-\phi)\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in\mathcal{F}_j}]\\ &\le \sum_{j\ge M}\sup_{f_j\in\mathcal{N}_j}\sup_{f\in\mathcal{F}_n,\|f-f_1\|\le\zeta j\epsilon_n}\mathbb{E}_f[(1-\phi(f_j))\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in\mathcal{F}_j}],\end{aligned}$$

利用对于任意 j、$\mathcal{F}_n\subset\bigcup_{f_j\in\mathcal{N}_j}\{f\in\mathcal{F}_n$ 都有 $\|f-f_j\|_1\le\zeta j\epsilon_n\}$。令 $f_j\in\mathcal{N}_j$ 和 $f\in\mathcal{F}_n$ 满足$\|f-f_j\|\le\zeta j\epsilon_n$。我们有

$$\begin{aligned}\mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi(f_j))] &= \mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^+(f_j)\vee\phi^-(f_j))]\\ &\le \mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^+(f_j))]\wedge\mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^-(f_j))] \qquad (46)\\ \mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^+(f_j))] &= \mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,0}(S_{1,i})<v_n}]\\ &= \mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})+(\Lambda^{i,f}(S_{1,i})-\Lambda^{i,0}(S_{1,i}))<v_n}],\end{aligned}$$

其中 $S_{1,i}:=\{(t,s)\in[0,1]^d:\ \lambda^i_{f_j}(t,s)\ge\lambda^i_{f_0}(t,s)\}$, $v_n>0$。我们可以将 $\frac{1}{n}\sum_i\Lambda^{i,f}(S_{1,i})-\Lambda^{i,0}(S_{1,i})$ 的下界估计为

$$\begin{aligned}\frac{1}{n}\sum_i\Lambda^{i,f}(S_{1,i})-\Lambda^{i,0}(S_{1,i}) &= \frac{1}{n}\sum_i\Lambda^{i,f}(S_{1,i})-\Lambda^{i,f_1}(S_{1,i})+\Lambda^{i,f_1}(S_{1,i})-\Lambda^{i,0}(S_{1,i})\\ &= \frac{1}{n}\sum_i\int_{S_{1,i}}(\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_j)+\lambda^i_{t,s}(f_j)-\lambda^i_{t,s}(f_0))dtds\\ &\ge -\frac{1}{n}\sum_i\int_{S_{1,i}}|\lambda^i_{t,s}(f)-\lambda^i_{t,s}(f_j)|dtds+\frac{1}{n}\sum_i\int_{S_{1,i}}(\lambda^i_{t,s}(f_j)-\lambda^i_{t,s}(f_0))dtds\\ &\ge -d_S(f,f_j)+\frac{1}{n}\sum_i\int_{S_{1,i}}(\lambda^i_{t,s}(f_j)-\lambda^i_{t,s}(f_0))dtds. \qquad (47)\end{aligned}$$

利用引理A.1，在 Ω_n 上，对于任意 $f \in \mathcal{F}_j$,

$$d_S(f, f_j) \leq N_0 \|f - f_j\|_1 \leq N_0 \zeta j \epsilon_n. \tag{48}$$

令 $f \in \mathcal{F}_j$。现在我们考虑两种情况：

- **情况 1**:

$$\frac{1}{n} \sum_i \int_{S_{1,i}} (\lambda^i_{t,s}(f_j) - \lambda^i_{t,s}(f_0)) dtds > \frac{1}{n} \sum_i \int_{S^c_{1,i}} (\lambda^i_{t,s}(f_0) - \lambda^i_{t,s}(f_j)) dtds.$$

- **情况 2**:

$$\frac{1}{n} \sum_i \int_{S_{1,i}} (\lambda^i_{t,s}(f_j) - \lambda^i_{t,s}(f_0)) dtds \leq \frac{1}{n} \sum_i \int_{S^c_{1,i}} (\lambda^i_{t,s}(f_0) - \lambda^i_{t,s}(f_j)) dtds.$$

首先假设**情况 1**成立。那么

$$\frac{1}{n} \sum_i \int_{S_{1,i}} (\lambda^i_{t,s}(f_j) - \lambda^i_{t,s}(f_0)) dtds > \frac{d_S(f_0, f_j)}{2}.$$

此外，

$$\begin{aligned} d_S(f_0, f_j) = \frac{1}{n} \sum_{i=1}^n \int |\lambda^i_{t,s}(f_j) - \lambda^i_{t,s}(f_0)| dtds &\geq \frac{1}{n} \sum_{i=1}^n \int |\lambda^i_{t,s}(f) - \lambda^i_{t,s}(f_0)| dtds \\ &\quad - \frac{1}{n} \sum_{i=1}^n \int |\lambda^i_{t,s}(f_j) - \lambda^i_{t,s}(f)| dtds \\ &\geq j\epsilon_n - \frac{1}{n} \sum_{i=1}^n \int |\lambda^i_{t,s}(f_j) - \lambda^i_{t,s}(f)| dtds = j\epsilon_n - d_S(f, f_j), \end{aligned}$$

因为 $f \in \mathcal{F}_j$。此外，利用引理 A.1，在 Ω_n，我们有

$$d_S(f, f_j) \leq N_0 \|f - f_j\|_1 \leq N_0 \zeta j \epsilon_n,$$

因此，

$$d_S(f_0, f_j) \geq j\epsilon_n (1 - N_0 \zeta) > 0,$$

如果 $\zeta < N_0^{-1}$。

由（47）和（48），可得

$$\begin{aligned} \frac{1}{n} \sum_i \Lambda^{i,f}(S_{1,i}) - \Lambda^{i,0}(S_{1,i}) &\geq -d_S(f, f_j) + \frac{d_S(f_0, f_j)}{2} \\ &\geq -N_0 \zeta j \epsilon_n + \frac{j\epsilon_n (1 - N_0 \zeta)}{2} \\ &\geq (1/2 - 3N_0 \zeta/2) j \epsilon_n \geq j\epsilon_n/4, \end{aligned}$$

通过选择

$$(1/2 - 3N_0 \zeta/2) \geq \frac{1}{4} \iff \zeta \leq \frac{N_0}{6}.$$

因此，取 $v_n = nj\epsilon_n/8$，可得

$$\begin{aligned}\mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^+_{1,j})] &\le \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<v_n-j\epsilon_n/4}]\\ &= \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<(1/8-1/4)j\epsilon_n}]\\ &= \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<-j\epsilon_n/8}]. \end{aligned} \tag{49}$$

T为了界定右端，我们使用引理 A 中陈述的伯恩斯坦不等式的另一种形式 .3:

$$\mathbb{P}_f\left(\sum_{i=1}^n N^i(S_{1,i})-\Lambda^i(S_{1,i}) \le -\sqrt{2vx}-\frac{x}{3}\right)\le e^{-x},$$

对任意 $x>0$，且当 $v \ge \sum_i \int_{S_{1,\,i}} \lambda^i_{t,\,s}(f)dtds = \sum_i \Lambda^{i,\,f}(S_{1,\,i})$ 时。我们首先求出这样的 v。我们有

$$\sum_i \Lambda^{i,f}(S_{1,i}) = \sum_i \int_{S_{1,i}} \lambda^i_{t,s}(f)dtds \le \sum_i \int \lambda^i_{t,s}(f)dtds \le n\|\mu\|_1 + \|g\|_1 \sum_i N^i[0,1]^{d+1}.$$

此外，由于 $f\in\mathcal{F}_j$ 且在 Ω_n 上，利用引理A.2，我们有

$$\begin{aligned}\|\mu\|_1 + \|g\|_1 \frac{1}{n}\sum_i N^i[0,1]^{d+1} &\le \|\mu_0\|_1 + \|g_0\|_1 \frac{1}{n}\sum_i N^i[0,1]^{d+1} + d_S(f,f_0)\\ &\le \|\mu_0\|_1 + e_0\|g_0\|_1 + \|g_0\|_1 + (j+1)\epsilon_n. \end{aligned} \tag{50}$$

其中 $e_0 = \frac{\bar{\mu}}{1-\|g_0\|_1}$。因此，令 $C_0 = \|\mu_0\|_1 + \|g_0\|_1(e_0+1)$，得到

$$\sum_i \Lambda^{i,f}(S_{1,i}) \le nC_0 + (j+1)n\epsilon_n =: v.$$

因此，

$$\mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^+(f_j))] = \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<-\sqrt{2vx}-\frac{x}{3}}] \tag{51}$$

$$\le \mathbb{P}_f\left[\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i}) < -\sqrt{2vx}-\frac{x}{3}\right] \le e^{-x}. \tag{52}$$

我们区分两种情况：

- $j\epsilon_n \le C_0+1$。然后，我们将（52）应用于 $x = x_1 nj^2\epsilon_n^2$，其中 $0 \le x_1 \le \frac{3}{56(C_0+1)}$。特别地，$x_1(C_0+1) \le 1$ 因此 $\sqrt{x_1}(C_0+1) \le x_1(C_0+1)$。接着，

$$\begin{aligned}\sqrt{2vx}+\frac{x}{3} &= \sqrt{2x_1 vn}j\epsilon_n + \frac{x_1 nj^2\epsilon_n^2}{3}\\ &= \left(\sqrt{2x_1(C_0+(j+1)\epsilon_n)} + \frac{x_1 j\epsilon_n}{3}\right)nj\epsilon_n\\ &\le \left(2\sqrt{x_1(C_0+1)} + \frac{x_1(C_0+1)}{3}\right)nj\epsilon_n\\ &\le \frac{7x_1(C_0+1)}{3}nj\epsilon_n \le \frac{nj\epsilon_n}{8}. \end{aligned}$$

因此，结合（52），我们得到

$$\mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<-j\epsilon_n/8}] \le \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<-\sqrt{2vx}-\frac{x}{3}}] \le e^{-nx_1j^2\epsilon_n^2}.$$

• $j\epsilon_n > C_0 + 1$。然后，我们将（52）与 $x = x_0 nj\epsilon_n$结合，其中 $x_0 \le \min(\frac{C_0+1}{32^2}, \frac{3}{16})$。特别地$2\sqrt{\frac{x_0}{C_0+1}} \le \frac{1}{16}$。于是我们有

$$\begin{aligned}\sqrt{2vx} + \frac{x}{3} &= \sqrt{2x_0 vnj\epsilon_n} + \frac{x_0 nj\epsilon_n}{3} \\ &= \left(\frac{\sqrt{2x_0(C_0 + (j+1)\epsilon_n)}}{j\epsilon_n} + \frac{x_0}{3}\right) nj\epsilon_n \\ &\le \left(2\frac{\sqrt{x_0}}{\sqrt{j\epsilon_n}} + \frac{x_0}{3}\right) nj\epsilon_n \\ &\le \left(2\frac{\sqrt{x_0}}{\sqrt{C_0+1}} + \frac{x_0}{3}\right) j\epsilon_n \le \frac{nj\epsilon_n}{8}.\end{aligned}$$

因此，结合（49）和（52），我们得到

$$\begin{aligned}\mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^+(f_j))] &\le \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\frac{1}{n}\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<-j\epsilon_n/8}] \\ &\le \mathbb{E}_f[\mathbb{1}_{\Omega_n}\mathbb{1}_{\sum_{i=1}^n N^i(S_{1,i})-\Lambda^{i,f}(S_{1,i})<-\sqrt{2vx}-\frac{x}{3}}] \le e^{-x_0 nj\epsilon_n}.\end{aligned}$$

注意，上述界与 f 和 f_1无关。因此，我们已经证明

$$\sup_{f_1\in\mathcal{N}_j}\sup_{f\in\mathcal{F}_n,\|f-f_1\|\le\zeta j\epsilon_n}\mathbb{E}_f[(1-\phi^+(f_j))\mathbb{1}_{\Omega_n}\mathbb{1}_{f\in\mathcal{F}_j}] \le \begin{cases} e^{-x_1 nj^2\epsilon_n^2} & \text{if } j\epsilon_n \le C_0+1 \\ e^{-x_0 nj\epsilon_n} & \text{if } j\epsilon_n > C_0+1\end{cases}.$$

现在假设**情况 2** 成立。那么

$$\frac{1}{n}\sum_i \int_{S^c_{1,i}} (\lambda^i_{t,s}(f_1) - \lambda^i_{t,s}(f_0))dtds \ge \frac{d_S(f_0,f_1)}{2}.$$

将 $S_{1,i}$ 替换为 $S^c_{1,i}$，并进行相同的计算，可得

$$\sup_{f_1\in\mathcal{N}_j}\sup_{f\in\mathcal{F}_n,\|f-f_1\|\le\zeta j\epsilon_n}\mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}\mathbb{1}_{\Omega_n}(1-\phi^-(f_j))] \le \begin{cases} e^{-x_1 nj^2\epsilon_n^2} & \text{if } j\epsilon_n \le C_0+1 \\ e^{-x_0 nj\epsilon_n} & \text{if } j\epsilon_n > C_0+1\end{cases}$$

结合（46），总体而言，这意味着

$$\begin{aligned}\sum_{j\ge M}\sup_{f_1\in\mathcal{N}_j}\sup_{f\in\mathcal{F}_n,\|f-f_1\|\le\zeta j\epsilon_n}\mathbb{E}_f[\mathbb{1}_{f\in\mathcal{F}_j}(1-\phi(f_j))\mathbb{1}_{\Omega_n}] &\le \sum_{M\le j\le(C_0+1)\epsilon_n^{-1}} e^{-x_1^2 nj^2\epsilon_n^2} + \sum_{j\ge(C_0+1)\epsilon_n^2} e^{-x_0 nj\epsilon_n} \\ &\le e^{-\min(x_0,x_1)nM\epsilon_n^2/2} = e^{-b_2 n\epsilon_n^2}.\end{aligned}$$

当 M 足够大时，$b_2 = \frac{\min(x_0,\ x_1\ M)}{2}$和 $b_2 > c_1$。

□

B.4 引理证明 4.4

引理 B.4. 对于任意 $f \in A_n^c$，存在一个常数 $p_0 > 0$，使得在 Ω_n 上，

$$\mathbb{E}_f[Z_1] \ge p_0\|f - f_0\|_1.$$

证明. 回顾

$$Z_1 := \int_0^{t_2^1} \int_{[0,1]^d} |\lambda_{t,s}^i(f) - \lambda_{t,s}^i(f_0)| dtds,$$

其中 t^1_2为 t 为第二个事件发生的时间；如果事件少于 2 个，则为 $t_2^1 = 1$。定义事件

$$\Omega_0 := \{N^1[0,1]^{d+1} = 0\},$$

即，Ω_0 表示序列中没有事件的事件。于是，在 Ω_0 上， $t_2^1 = 1$以及

$$\mathbb{E}_f[Z_1 \mathbb{1}_{\Omega_0}] = \mathbb{E}_f \left[\mathbb{1}_{\Omega_0} \int_0^1 \int_{[0,1]^d} |\mu(t,s) - \mu_0(t,s)| dtds \right] = \|\mu - \mu_0\|_1 \mathbb{P}_f[\Omega_0].$$

此外，令 $\mathbb{Q}$ 表示在 $[0,1]^{d+1}$ 上具有单位强度的齐次泊松点过程的测度，则有

$$\mathbb{E}_f[\mathbb{1}_{\Omega_0}] = \mathbb{E}_{\mathbb{Q}}[\mathcal{L}_f \mathbb{1}_{\Omega_0}],$$

其中 $\mathcal{L}_f$ 是定义在 $[0,1]^{d+1}$ 上的似然过程，定义为

$$\mathcal{L}_f = \exp\left(1 - \int \lambda_{t,s}^1(f) dtds + \int \log(\lambda_{t,s}^1(f)) dN_{t,s}\right)$$

在 Ω_0, 上

$$\mathcal{L}_f = \exp\left(1 - \int \mu(t,s) dtds\right) \geq \exp\left(1 - \|\mu\|_1\right).$$

由式（19），在Ω_n,

$$\|\mu\|_1 \leq \|\mu_0\|_1 + \frac{3e_0}{2}\|g_0\|_1 + d_S(f, f_0) \leq \|\mu_0\|_1 + \frac{3e_0}{2}\|g_0\|_1 + M\epsilon_n$$

由于 $f \in A_n^c$。因此，对于足够大的 n，在 $\Omega_0 \cap \Omega_n$ 上,

$$\mathcal{L}_f \geq \exp\left(-\|\mu_0\|_1 - \frac{3e_0}{2}\|g_0\|_1\right) =: \mathcal{L}_0,$$

因此,

$$\mathbb{E}_f[Z_1 \mathbb{1}_{\Omega_0}] \geq \mathbb{E}_f[Z_1 \mathbb{1}_{\Omega_0 \cap \Omega_n}] \geq \mathcal{L}_0 \|\mu - \mu_0\|_1 \mathbb{Q}(\Omega_0). \tag{53}$$

由于 $\mathbb{Q}$ 是强度为 1 的齐次泊松过程,

$$\mathbb{Q}(\Omega_0) = e^{-1}.$$

因此，我们可以得出结论:

$$\mathbb{E}_f[Z_1 \mathbb{1}_{\Omega_0}] \geq \mathcal{L}_0 e^{-1} \|\mu - \mu_0\|_1.$$

现在令 $\tau \in (0, 1-2a)$，并定义子空间 $S_\tau^- = [0,\tau] \times [0,1]^d$、 $S_\tau = [\tau, \tau + a] \times [0,1]^d$、$S_\tau^+ = [\tau + a, \tau + 2a] \times [0,1]^d$ 以及事件:

$$\Omega_\tau = \{N^1(S_\tau^-) = 0, N^1(S_\tau) = 1, N^1(S_\tau^+) = 0\}.$$

注意，在 Ω_τ 上，$t_2^1 > t_1^1 + a_\circ$ 于是，

$$\begin{aligned}\mathbb{E}_f[Z_1\mathbb{1}_{\Omega_\tau}] &\geq \mathbb{E}_f\left[\mathbb{1}_{\Omega_\tau}\int_0^\tau\int_{[0,1]^d}|\mu(t,s)-\mu_0(t,s)|dtds\right]\\ &+\mathbb{E}_f\left[\mathbb{1}_{\Omega_\tau}\int_{t_1^1}^{t_1^1+a}\int_{[0,1]^d}|\mu(t,s)+g(t-t_1^1,s-s_1^1)-\mu_0(t,s)-g_0(t-t_1^1,s-s_1^1)|dtds\right]\\ &\geq \mathbb{E}_f\left[\mathbb{1}_{\Omega_\tau}\int_{t_1^1}^{t_1^1+a}\int_{[0,1]^d}\left||\mu(t,s)-\mu_0(t,s)|-|g(t-t_1^1,s-s_1^1)-g_0(t-t_1^1,s-s_1^1)|\right|dtds\right]\end{aligned} \tag{54}$$

首先假设

$$\begin{aligned}&\|\mu-\mu_0\|_1 \geq \frac{\mathcal{L}_0 a e^{-2a-\tau}}{2}\|g-g_0\|_1\\ \Longrightarrow\ &\|\mu-\mu_0\|_1+\frac{\mathcal{L}_0 a e^{-2a-\tau}}{2}\|\mu-\mu_0\|_1 \geq \frac{\mathcal{L}_0 a e^{-2a-\tau}}{2}\|f-f_0\|_1\\ \Longrightarrow\ &\|\mu-\mu_0\|_1 \geq \frac{\mathcal{L}_0 a e^{-2a-\tau}}{2+\mathcal{L}_0 a e^{-2a-\tau}}\|f-f_0\|_1.\end{aligned}$$

那么由（53）可得，

$$\mathbb{E}_f[Z_1] \geq \mathbb{E}_f[Z_1\mathbb{1}_{\Omega_0}] \geq \frac{\mathcal{L}_0^2 a e^{-2a-\tau-1}}{2+\mathcal{L}_0 a e^{-2a-\tau}}\|f-f_0\|_1.$$

在备择假设情况下

$$\begin{aligned}&\|\mu-\mu_0\|_1 < \frac{\mathcal{L}_0 a e^{-2a-\tau}}{2}\|g-g_0\|_1\\ \Longrightarrow\ &\|f-f_0\|_1 < \|g-g_0\|_1+\frac{\mathcal{L}_0 a e^{-2a-\tau}}{2}\|g-g_0\|_1\\ \Longrightarrow\ &\|g-g_0\|_1 \geq \frac{2}{2+\mathcal{L}_0 a e^{-2a-\tau}}\|f-f_0\|_1,\end{aligned}$$

then

$$\begin{aligned}\mathbb{E}_f[Z_1\mathbb{1}_{\Omega_\tau}] &\geq \mathbb{E}_f\left[\mathbb{1}_{\Omega_\tau}\int_{t_1^1}^{t_1^1+a}\int_{[0,1]^d}|g(t-t_1^1,s-s_1^1)-g_0(t-t_1^1,s-s_1^1)|-|\mu(t,s)-\mu_0(t,s)|dtds\right]\\ &\geq \mathbb{E}_f\left[\mathbb{1}_{\Omega_\tau}\int_{t_1^1}^{t_1^1+a}\int_{[0,1]^d}|g(t-t_1^1,s-s_1^1)-g_0(t-t_1^1,s-s_1^1)|dtds\right]-\|\mu-\mu_0\|_1\mathbb{E}_f\left[\mathbb{1}_{\Omega_\tau}\right].\end{aligned}$$

此外，

$$\begin{aligned}&\mathbb{E}_f\left[\mathbb{1}_{\Omega_\tau}\int_{t_1^1}^{t_1^1+a}\int_{[0,1]^d}|g(t-t_1^1,s-s_1^1)-g_0(t-t_1^1,s-s_1^1)|dtds\right]\\ &=\mathbb{E}_{\mathbb{Q}}\left[\mathcal{L}_f\mathbb{1}_{\Omega_\tau}\int_{t_1^1}^{t_1^1+a}\int_{[0,1]^d}|g(t-t_1^1,s-s_1^1)-g_0(t-t_1^1,s-s_1^1)|dtds\right]\\ &\geq \mathcal{L}_0\mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_{\Omega_\tau}\int_0^a\int_{[0,1]^d}|g(u,s-s_1^1)-g_0(u,s-s_1^1)|duds\right]=\mathcal{L}_0\mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_{\Omega_\tau}\right]\|g-g_0\|_1.\end{aligned}$$

我们有

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[\mathbb{1}_{\Omega_\tau}\right] &= \mathbb{P}_{\mathbb{Q}}\left[N^1(S_\tau^-)=0\right]\mathbb{P}_{\mathbb{Q}}\left[N^1(S_\tau)=1\right]\mathbb{P}_{\mathbb{Q}}\left[N^1(S_\tau^+)=0\right]\\ &= e^{-\tau}\times ae^{-a}\times e^{-a} = ae^{-2a-\tau}.\end{aligned}$$

由此我们得到

$$\begin{aligned}\mathbb{E}_f[Z_1\mathbb{1}_{\Omega_\tau}] &\geq \mathcal{L}_0 ae^{-2a-\tau}\|g-g_0\|_1 - \|\mu-\mu_0\|_1\\ &\geq \frac{\mathcal{L}_0 ae^{-2a-\tau}}{2}\|g-g_0\|_1\\ &\geq \frac{\mathcal{L}_0 ae^{-2a-\tau}}{2+\mathcal{L}_0 ae^{-2a-\tau}}\|f-f_0\|_1.\end{aligned}$$

由于 Ω_0 和 Ω_τ 是互不相交的事件，我们可以得出

$$\begin{aligned}\mathbb{E}_f[Z_1] &\geq \mathbb{E}_f[Z_1\mathbb{1}_{\Omega_0}] + \mathbb{E}_f[Z_1\mathbb{1}_{\Omega_\tau}]\\ &\geq \min\left(\frac{\mathcal{L}_0 ae^{-2a-\tau}}{2+\mathcal{L}_0 ae^{-2a-\tau}}, \frac{\mathcal{L}_0^2 ae^{-2a-\tau-1}}{2+\mathcal{L}_0 ae^{-2a-\tau}}\right)\|f-f_0\|_1\\ &= \frac{\mathcal{L}_0 ae^{-2a-\tau}}{2+\mathcal{L}_0 ae^{-2a-\tau}}\min\left(1,\mathcal{L}_0 e^{-1}\right)\|f-f_0\|_1 = \frac{\mathcal{L}_0^2 ae^{-2a-\tau-1}}{2+\mathcal{L}_0 ae^{-2a-\tau}}\|f-f_0\|_1,\end{aligned}$$

因为 $\mathcal{L}_0 < 1$。因此，令

$$p_0 := \frac{\mathcal{L}_0^2 ae^{-2a-\tau-1}}{2+\mathcal{L}_0 ae^{-2a-\tau}}, \tag{55}$$

对于任意 $\tau \in (0, 1-2a)$，我们得到引理 4.4 的结论。 □

参考文献

Ryan Prescott Adams、Iain Murray 和 David JC MacKay。具有高斯过程强度的泊松过程中可处理的非参数贝叶斯推断。载于第 *26* 届国际机器学习年会论文集，第 9–16 页，2009年。

Alba Bernabeu、Jiancang Zhuang 和 Jorge Mateu。《时空霍克斯点过程：综述》。《农业、生物与环境统计学杂志》，30(1)：89–119，2025年。doi：10.1007/s13253-024-00653-7。网址https://doi.org/10.1007/s13253-024-00653-7。

Pierre Br´emaud。《点过程与排队论，鞅动力学》。施普林格出版社，纽约，1981年。

Biao Cai、Jingfei Zhang 和 Yongtao Guan。基于高维多元点过程的潜在网络结构学习。《美国统计学会杂志》，第1–14页，2022年。

Feng Chen 和 Peter Hall。非平稳自激点过程的推断及其在超高频金融数据建模中的应用。《应用概率杂志》，50(4)：1006–1024，2013年。

Simon Clinet 和 Nakahiro Yoshida。基于高频观测的多变量 Hawkes 过程统计推断。《随机过程及其应用》，127(6):1800–1839，2017。doi: 10.1016/j.spa.2016.09.017。

Riley Crane 和 Didier Sornette。网络上传染过程的稳健动态建模。美国国家科学院院刊，105(41):15649–15653，2008。doi: 10.1073/ pnas.0803685105。Daryl J. Daley 和 David Vere-Jones。点过程理论导论：第二卷：一般理论与结构。施普林格，纽约，第 2 版，2003。Patric Dolmeta 和 Matteo Giordano。用于基于协变量的强度估计的高斯过程方法。载于国际函数与算子统计研讨会，第 185–192 页。施普林格，2025年a。

Patric Dolmeta和Matteo Giordano。协变量驱动的泊松过程独立同分布观测的非参数贝叶斯分析。*arXiv* 预印本 *arXiv:2509.02299*，2025b。

Sophie Donnet、Vincent Rivoirard 和 Judith Rousseau。具有 Aalen 乘性强度的计数过程的后验集中率。伯努利，20(4):1990–2028，2014。doi: 10.3150/ 13-BEJ547。网址 https://doi.org/10.3150/13-BEJ547。

Sophie Donnet、Vincent Rivoirard 和 Judith Rousseau。多元 Hawkes 过程的非参数贝叶斯估计。统计年鉴，48(5):2698–2727，2020年。

Michael Eichler、Rainer Dahlhaus 和 Julian Dueck。采用非参数链接函数的多元 Hawkes 过程图模型。时间序列分析杂志，38(2):225–242，2017。doi:10.1111/jtsa.12174。

Subhashis Ghosal 和 Aad W Van der Vaart。《非参数贝叶斯推断基础》，第 44 卷。剑桥大学出版社，2017。

Matteo Giordano、Alisa Kirichenko 和 Judith Rousseau。针对协变量驱动的非齐次点过程进行非参数贝叶斯强度估计。《伯努利》，2025年。即将出版。

Nicolas R. Hansen、Patricia Reynaud-Bouret 和 Vincent Rivoirard。多元点过程的套索方法与概率不等式。《style id='1'>伯努利》，21(1)：83–143，2015年。doi：10.3150/13-BEJ561。

Alan G Hawkes。若干自激和相互激发点过程的谱。《style id='1'>生物统计学，58 (1)：83–90，1971年。

Anthony Hills、Talia Tseriotou、Xenia Miscouridou、Adam Tsakalidis 和 Maria Liakata。激发情绪变化：用于变化检测建模的时间感知层级 Transformer。载于 Lun-Wei Ku、Andre Martins 和 Vivek Srikumar 编，计算语言学协会论文集：*ACL 2024*，第 12526–12537 页，泰国曼谷，2024年8月。计算语言学协会。doi：10.18653/v1/2024.findings-acl.744。网址：https://aclanthology.org/2024.findings-acl.744/。

Bowen Huang、Issei Sato、RyoNakano 和 KenjiFukumizu。大型语言模型可以学习时间点过程。*arXiv* 预印本 *arXiv:2401.03436*，2024。网址https://arxiv.org/abs/2401.03436。

Mikyoung Jun 和 Scott Cook。恐怖主义的灵活多变量时空霍克斯过程模型。《应用统计年鉴》，18(2):1378–1403，2024。

Alan Karr。点过程及其统计推断。Routledge，2017。

Alisa Kirichenko 和 Harry Van Zanten。基于高斯过程的泊松过程强度学习的最优性。《机器学习研究期刊》，16(1):2909–2919，2015年1月。ISSN 1532-4435。

Alisa Kirichenko、Harry Van Zanten 等。基于高斯过程的泊松过程强度学习的最优性。《机器学习研究期刊》，16(1):2909–2919，2015年。Jonathan Koh、Franc¸ois Pimont、Jean-Luc Dupuy 和 Thomas Opitz。通过带有中等和极端标记的点过程进行时空野火建模。《应用统计年鉴》，17(1):560–582，2023年。Tsz-Kit Jeffrey Kwan、Feng Chen 和 William TM Dunsmuir。非平稳霍克斯过程的备择渐近推断理论。《统计规划与推断杂志》，227: 75–90，2023年。Thomas J. Liniger。多元 *Hawkes* 过程。博士论文，苏黎世联邦理工学院，2009年。网址：https://doi.org/10.3929/ethz-a-005940267。苏黎世联邦理工学院博士论文第18403号。Chris Lloyd、Tom Gunter、Michael Osborne 和 Stephen Roberts。高斯过程调制泊松过程的变分推断。载于国际机器学习会议，第1814–1822页。机器学习研究论文集，2015年。Noa Malem-Shinitski、C´esar Ojeda 和 Manfred Opper。利用高斯过程自效应对非线性霍克斯过程进行变分贝叶斯推断。《熵》，24(3):356，2022年。Hongyuan Mei 和 Jason Eisner。神经霍克斯过程：神经自调制的多元点过程。载于神经信息处理系统进展，第30卷，2017年。Xenia Miscouridou、Samir Bhatt、George Mohler、Seth Flaxman 和 Swapnil Mishra。Cox-Hawkes过程：双重随机时空泊松过程。载于*2023*年*IMS*国际统计与数据科学会议（*ICSDS*），第564页，2023年。George O. Mohler、Martin B. Short、Sean Malinowski、Mark Johnson、George Tita、Andrea L. Bertozzi 和 P. Jeffrey Brantingham。犯罪的自激点过程建模。《美国统计学会杂志》，106(493):100–108，2011年。doi: 10.1198/jasa.2011.ap09546。Tin Lok J Ng 和 Thomas B Murphy。带变点的非齐次泊松过程强度函数估计。《加拿大统计学杂志》，47(4):604–618，2019年。Yosihiko Ogata。平稳点过程最大似然估计量的渐近行为。《统计数学研究所年报》，30(1):243–261，1978年。doi: 10.1007/ BF02480216。Yosihiko Ogata。地震发生的统计模型及点过程的残差分析。《美国统计学会杂志》，83(401):9–27，1988年。doi: 10.1080/ 01621459.1988.10478560。Tohru Ozaki。霍克斯自激点过程的最大似然估计。《统计数学研究所年报》，31(1):145–155，1979年。doi: 10.1007/BF02480272。Julia A Palacios 和 Vladimir N Minin。利用基因谱系对种群规模轨迹进行基于高斯过程的贝叶斯非参数推断。《生物统计学》，69(1):8–18，2013年。Carl Edward Rasmussen 和 Christopher KI Williams。《机器学习中的高斯过程》。麻省理工学院出版社，2005年。Alex Reinhart。自激时空点过程及其应用综述。《统计科学》，33(3):299–318，2018年。

Patricia Reynaud-Bouret 和 Emmanuel Roy。关于霍克斯过程的一些非渐近尾部估计。比利时数学学会-西蒙·斯特芬会刊，13(5):883–896，2007。Patricia Reynaud-Bouret、Vincent Rivoirard 和 Christine Tuleau-Malot。脉冲序列分析的拟合优度检验与非参数自适应估计。神经信息处理系统进展，27，2014。Marian-Andrei Rizoiu、Swapnil Mishra、Qinxun Kong、Mark J. Carman 和 Lexing Xie。Sir-hawkes：将流行病模型与霍克斯过程相结合，对有限人群中的扩散过程建模。*2018*年万维网大会论文集，第419–428页，2018。doi：10.1145/3178876.3186108。Emilia Siviero、Guillaume Staerman、St´ephan Cl´emenc¸on 和 Thomas Moreau。时空霍克斯过程的灵活参数推断。*arXiv* 预印本 *arXiv:2406.06849*，2024。网址：https://arxiv.org/abs/2406.06849。预印本。Deborah Sulem、Vincent Rivoirard 和 Judith Rousseau。适用于霍克斯过程的可扩展自适应变分贝叶斯方法。*arXiv* 预印本 *arXiv:2212.00293*，2022。Deborah Sulem、Vincent Rivoirard 和 J udith Rousseau。非线性霍克斯过程的贝叶斯估计。伯努利，30(2):1257–1286，2024。

Wilson Truccolo、Uri T. Eden、Matthew R. Fellows、John P. Donoghue 和 Emery N. Brown。用于关联神经元尖峰活动与尖峰历史、神经元集群及外部协变量效应的点过程框架。《神经生理学杂志》，93(2):1074–1089，2005年。doi: 10.1152/ jn.00697.2004。

Aad van der Vaart 和 JH van Zanten。使用带有逆伽马带宽的高斯随机场进行自适应贝叶斯估计。《统计年刊》，37(5B):2655–2675，2009年。Hongteng Xu 和 Hongyuan Zha。从短的双重删失事件序列中学习霍克斯过程。载于第*33*届国际机器学习会议论文集，2016年。Rui Zhang、Christian Walder 和 Marian-Andrei Rizoiu。稀疏高斯过程调制霍克斯过程的变分推断。*AAAI*人工智能大会论文集，34(04):6803–6810，2020年4月。ISSN 2159-5399。doi: 10.1609/aaai.v34i04.6160。URL http://dx.doi.org/10.1609/aaai.v34i04.6160。Tianyu Zhang 和 Noah Simon。利用筛法在张量积空间中进行回归。《电子统计学杂志》，17(2):3660，2023年。Qingyuan Zhao、Murat A. Erdogdu、Hera Y. He、Anand Rajaraman 和 Jure Leskovec。Seismic：一种用于预测推文流行度的自激点过程模型。载于第*21*届*ACM SIGKDD*知识发现与数据挖掘国际会议论文集，第1513–1522页，2015年。doi: 10.1145/2783258.2783401。Feng Zhou、Zhidong Li、Xuhui Fan、Yang Wang、Arcot Sowmya 和 Fang Chen。利用辅助潜变量对非参数霍克斯过程进行高效推断。《机器学习研究杂志》，21(241):1–31，2020年。URLhttp://jmlr.org/papers/v21/19-930.html。Simiao Zuo、Haoming Jiang、Zichong Li、Tuo Zhao 和 Hongyuan Zha。Transformer霍克斯过程。*arXiv*预印本 *arXiv:2002.09291*，2020年。URLhttps://arxiv.org/abs/2002.09291。