arXiv:2611.32270v1 [math.ST] 24 Nov 2026

用于逻辑斯蒂-正态积分的经认证的极小极大正态尺度混合

李睿杰*

2026年11月24日

摘要

逻辑斯蒂分布函数是正态分布函数的尺度混合，Monahan 和 Stefanski 用 K项混合 $\Lambda(x) \approx \sum_i w_i \Phi(s_i x)$ 对其进行近似，该混合经选择后可使实数轴上的一致误差最小。每个这样的混合都能将逻辑斯蒂-正态积分化为正态分布函数的 K 之和，并具有该混合的一致误差界。在统计学中，此积分是正态分布下逻辑斯蒂 sigmoid 函数的期望；在半导体物理学中，它是高斯–费米积分。Monahan 和 Stefanski 列出了 $K \le 8$ 的常数，并猜想最优误差会在 $2K$ 个点上等振荡；然而，当超过 $K = 8$时，雷梅兹交换算法和局部求解器会停滞于不满足等振荡条件的驻点。我们证明，混合族的切空间是一个切比雪夫系统，因此每个内部极小化子都必然满足 $2K$ 点等振荡，且任何此类解都是唯一的全局极小极大解；我们还证明了 $K \le 11$的内点达到。认证算法采用多精度算术审查候选混合，并以经证明的极值点数量上界作为完备性检验；三条搜索流程达到 $K = 9$，10，11，最后一条得到的认证最坏情况误差为 1.26×10^{-11}。一个带符号的 $K = 22$ 混合在 3.605×10^{-15}处具有松弛认证。在二元结局的因果中介分析中，经认证的求和可替代有偏的泰勒展开和随机积分，只需进行一次闭式求值。

关键词。 极小极大逼近、等振荡、切比雪夫系统、正态尺度混合、逻辑斯蒂-正态积分、因果中介。

数学　　学科分类（2020年）。 65D15, 33F05, 41A50, 41A52, 41A25, 62-08。

1 引言

逻辑斯蒂分布函数 $\Lambda(x) = 1/(1+e^{-x})$ 是正态分布函数 [1, 2]，$\Lambda(x) = \int_0^\infty \Phi(sx)\, q(s)\, ds$ 的精确连续尺度混合，其中 q为柯尔莫哥洛夫–斯米尔诺夫混合密度， Φ 为标准正态分布函数。Monahan 和 Stefanski [3] 将该混合离散为 K 个原子，由此得到逼近族

$$f_K(x; w, s) = \sum_{i=1}^{K} w_i\, \Phi(s_i x), \qquad w_i \ge 0,\ \sum_i w_i = 1,\ s_i > 0, \tag{1}$$

并选择其参数，使实数轴上的最坏情况误差最小化， $\Delta_K^\star = \inf_{(w,s)} \sup_x |\Lambda(x) - f_K(x;\, w, s)|$。

*卫生服务与结果研究，新加坡国立医疗集团，Mandalay Road 1号，新加坡 308205，新加坡。电子邮件： ruijie.li@nhghealth.com.sg 和 ruijie@leefy.net。

Monahan 和 Stefanski 将 $K \le 8$ 的极小极大权重和尺度制成表格，这些权重和尺度是 Remez 交换算法的固定点， [4], 且最坏情况误差为 $\Delta_8^\star \approx 2.1 \times 10^{-9}$。他们留下了两个未解问题。首先，他们提出的等振荡猜想认为，最优的 K项混合的误差在 $(0,\infty)$ 上恰有 $2K$ 个等幅交替极值，但这只是数值观察，尚无证明。经典的切比雪夫交错定理和线性Remez理论 [5] 并不适用，因为混合族关于尺度是非线性的。其次，表格未涵盖超过 $K = 8$的情形。当超过 $K = 8$ 时，极小极大误差会降至 10^{-9}以下，而等振荡方程组会出现不具交替性的驻点、代数鞍点以及狭窄的各向异性吸引域；在这些区域中，Remez迭代和牛顿迭代会停滞或发散。

该近似是为逻辑斯蒂-正态积分构造的

$$I(\eta, v) = \int_{-\infty}^{\infty} \Lambda(\eta + m)\, \frac{1}{\sqrt{2\pi v}}\, \exp\!\left(-\frac{m^2}{2v}\right) dm, \tag{2}$$

即正态分布下逻辑斯蒂 sigmoid 函数的期望，该期望没有初等闭式。由于每个正态项 $\Phi(s_i(\eta + m))$ 与高斯密度相乘后都能以闭式积分求得，因此用 f_K 替换 Λ ，可将 (2) 化为有限和

$$I(\eta, v) \approx \hat{I}(\eta, v) = \sum_{i=1}^{K} w_i\, \Phi\!\left(\frac{s_i \eta}{\sqrt{1 + s_i^2 v}}\right), \qquad |I - \hat{I}| \le \sup_x |\Lambda(x) - f_K(x; w, s)| \tag{3}$$

在 $(\eta,\ v) \in \mathbb{R} \times (0,\ \infty)$ 上一致成立。因此，极小极大混合将 (2) 化为无需分支的闭式求值，并附带无参数误差界。

只要对高斯潜变量取平均时涉及逻辑斯蒂链接函数，就会出现积分 (2)。含协变量测量误差和随机效应的逻辑斯蒂回归促成了最初的近似方法 [3]。该积分也出现在对具有对数正态分布利率的短期利率模型和信用模型进行仿真时，此时它是带有逻辑斯蒂初始条件的热方程的解 [6]。在有机半导体器件仿真中，该积分即高斯–费米积分，表示高斯态密度的载流子密度 [7, 8]。其中， η 是从该态密度中心起算的费米能级， $\sqrt{v}$ 是其宽度，两者均以热能为单位。相关文献通过在非简并区和简并区分别使用解析近似 [7], 、误差函数级数 [9, 10], 或数值求积 [11]来计算该积分。这些器件的漂移-扩散模型通过两个比值使用该积分，即扩散增强因子 $I/\partial_\eta I$ 和简并因子 $e^{-\eta} I$ [8]，因此需要以相对精度计算非简并区的小值。因此，定理 3.2 将该混合与一个已知的指数倾斜恒等式结合起来，从而对所有 η 给出相对误差界。在二元结局和连续中介变量的因果中介分析中，(2) 是计算核。在此，自然间接效应——即经由中介变量传递的结局概率变化——是两个此类积分之差。

当只需要式 (2) 的个别数值时，求值问题便已解决。Pirjol [6]在 η 中推导出一个精确递推，并给出一个收敛的泊松求和级数，可将积分在给定点处（$\eta,\ v$）求值至工作精度所允许的准确度。然而，在仿真、优化和不确定性传播中，所用的积分并非在单个点上求值，而是作为参数空间上的函数来处理。例如，不确定性传播会针对数千个（$\eta,\ v$）抽样反复求值，而对所得区间的任何陈述都承袭了内层求值的误差。对于截断级数，收敛速率和截断误差取决于求值点，因此要在抽样循环中作出一致保证，就必须控制每次抽样的截断阶数。随着 v 增大，固定节点的高斯–埃尔米特求积会失去精度，因为被积函数在 $\eta + m = \pm i\pi$ 处的极点与实轴的距离趋近于 $\pi/\sqrt{2v}$ ，而求积 —

变量。有限极点展开在 $\mathbb{R}$ 两端趋于相同的极限，而 Λ 则分别趋于 0 和 1，因此它无法在 $\mathbb{R}$ 上一致逼近 Λ 。若闭式的最坏情况误差是一个经认证的常数，且在看到任何参数之前就已固定，那么就无需针对每次抽样进行控制。该有限和还具有闭式导数，并且经过进一步的高斯卷积后仍为有限和（第 5.5 节）。在固定点处，这种混合并不能提高逐点计算器的有效位数。它的优势在于闭式、对所有（$\eta,\ v$）都成立的经认证一致界，以及由此得到的精确有限和表达式。第 5.4 节将这种方法与高斯–埃尔米特求积、高斯–费米近似、有理极点展开和泊松求和级数进行比较。

本文解决了 Monahan 和 Stefanski 提出的两个尚未解决的问题，并给出了经认证的范围。该猜想悬而未决三十四年，是因为这一族不属于线性切比雪夫理论的范畴。混合切空间的 Haar 性质提供了缺失的等振荡机制，而且该机制可推广到具有相同切空间结构的其他原子族（定理 2.6）。我们的贡献如下。

1. 我们证明了 Monahan–Stefanski 等振荡猜想对每个内部极小值点都成立（定理 2.2 和定理 2.3）。混合流形的切空间是一个切比雪夫（Haar）系统，通过 Kolmogorov 变分论证，它迫使极小极大误差恰好在 $2K$ 个点上等振荡。同一零点计数工具还导出全局唯一性，因此任何等振荡的内点解都是唯一的全局极小极大解（定理 2.5），因为该族在 Rice 意义下是变可解的。根据经认证的常数，我们证明了 $K = 1$、……、11 的内点达到（定理 2.4），因此该猜想在经认证的范围内无条件成立，完整证明见附录 A 至附录 G。

2. 我们建立了一个切向 Wronskian 恒等式（定理 2.7），并对 $K \le 5$ 给出符号证明，通过 $K = 12$ 进行精确验证，且在每个经认证的解处通过直接 Sturm 求值予以确认。该恒等式提供了一个独立于全局论证的显式局部唯一性证书，并为 $K = 11$ 搜索中的尺度分离惩罚提供指导。

3. 我们给出认证算法（算法 1）；根据定理 E.1 的极值界，该算法对极值点的扫描是完备的，并据此认证 $K = 9$,10, 11 的极大极小常数（$w^\star, s^\star, \Delta_K^\star$）（第 3 节）。每种情况都需要不同的全局搜索流程：协方差矩阵自适应进化策略（CMA-ES） [12] 结合 Remez，并在 $K = 9$ 处进行多精度牛顿精修；在 $K = 10$ 处采用扰动参考迭代 Remez；以及以 Sturm 为种子的 CMA-ES，结合链式 Levenberg–Marquardt（LM） [13, 14] ，并在 $K = 11$处进行最终的 100 位精度精修。在 $K = 11$ 处，认证最坏情况误差为 $\Delta_{11}^\star = 1.258 \times 10^{-11}$。

4. 我们描绘了超越 $K = 11$ （第 4 节）的结构：在 $K = 12,$ 13 处猜想的平台；经认证的混合，否定其延伸至 $K = 14,$ 15 的可能性；贯穿 $K = 17$ 的宽松 Bernstein–Akhiezer [15] 证书；证明 $\Delta_K^\star \to 0$ 如 $K \to \infty$ （定理4.2）；以及带符号类别的 $K = 22$前沿，其中包含经认证的严格有符号30点等振荡。

5. 我们展示其在本研究动机所源自的应用中的影响（第 5 节）。在因果中介分析中，经认证的混合取代了有偏的 Valeri–VanderWeele泰勒展开，以及其他地方采用的随机模拟或固定网格求积，只需一个经认证的闭式求和，每10^4 次抽样耗时不到 15 毫秒，即可得出非对称准贝叶斯区间。

2 经认证的极小极大正态尺度混合

要认证一个候选混合，并将求解器的固定点提升为全局极小极大解，需要具备本节确立的三个结构性质。首先，混合流形的切空间是一个切比雪夫系统（定理 2.2），这迫使极小极大误差在$2K$ 个点上等振荡（定理 2.3）。混合拉普拉斯变差递减界将任意 K 项混合的误差在 $(0,\ \infty)$ 上的局部极值点数量限制在 $2K$ 个以内（定理 E.1）。其次，零点计数差分论证表明，任何 $2K$ 等振荡的内点混合都是唯一的全局极小极大解（定理 2.5）。第三， $K \le 11$ 的内点达到性质（定理 2.4）消除了经认证范围内的内点性假设。另一个切向 Wronskian 恒等式（定理 2.7）提供了显式的有限-K 局部唯一性证书，以及 $K = 11$ 搜索中所用的尺度分离惩罚。

2.1 等振荡、达到与唯一性

令 $\mathcal{F}_K \subset C^\infty(\mathbb{R})$ 表示具有严格正权重和有序正的互异尺度 $0 < s_1 < \cdots < s_K$ 的 K 项正态尺度混合 (1) 所构成的流形，并将其紧化记为 $\mathcal{F}_K$，其中有 $w_i \ge 0$ 和 $s_i \in [0, \infty]$。在任意内点 $f \in \mathcal{F}_K$ 处，切空间为

$$V = T_f\mathcal{F}_K = \operatorname{span}\{\Phi(s_1 z), \dots, \Phi(s_K z),\ z\,\phi(s_1 z), \dots, z\,\phi(s_K z)\}\,, \tag{4}$$

在对由 $\sum_i w_i = 1$ 所诱导的 Φ 系数施加线性约束 $\sum_i a_i = 0$ 后，其维数为 $2K-1$。此处，$\phi = \Phi'$ 是标准正态密度。由于 $\Lambda(-z) = 1 - \Lambda(z)$ 且 $\sum_i w_i = 1$，误差 $e(z) = \Lambda(z) - f(z)$ 在 $\mathbb{R}$ 上为奇函数，并在原点处消失，因此 $\mathbb{R}$ 上的极小极大问题可化为 $(0, \infty)$。切空间收集了权重和尺度发生微小变化时混合的一级变化，因而决定了最坏情况误差能否在局部得到降低。在线性切比雪夫逼近中， n 维逼近空间的 Haar 性质保证最优解处至少有 $n+1$ 次交替。在非线性族 $\mathcal{F}_K$ 中，切空间在每一点都承担这一角色，这正是其 Haar 性质成为以下结果所依据的结构性事实的原因。

本节的每个结果都归结为零点计数，而这些计数以一个经典界为基础。

引理 2.1（广义指数和的零点）。广义指数和 $\sum_{i=1}^{m} P_i(x)\, e^{\lambda_i x}$具有互异的实数 λ_i，且 $\deg P_i = r_i$，其实零点按重数计至多为 $(\sum_i (r_i + 1)) - 1$。特别地， $\sum_{i=1}^{N} c_i e^{\mu_i x}$ 至多有 $N-1$ 个零点，而 $\sum_{i=1}^{K}(c_i + d_i x)\, e^{\lambda_i x}$ 至多有 $2K-1$ 个零点。

这个界是经典结果 [16–18]。乘以 $e^{-\lambda_1 x}$，再求导 $r_1 + 1$ 次以消去第一组项，并应用罗尔定理，便可对 $\sum_i (r_i + 1)$ 使用归纳法得出该界。

定理 2.2（切空间的 Haar 性质）。设 V 如式 (4) 所示。任一非零的 $h \in V$ 在 $(0, \infty)$ 内至多有 $2K-2$ 个互异零点。等价地， V是在 $(0, \infty)$ 上维数为 $2K-1$ 的切比雪夫（*Haar*）系统。

证明。令 $h \in V$非零，并将 $h(z)$ 写为 $= \sum_{i=1}^{K} a_i \Phi(s_i z) + \sum_{i=1}^{K} b_i\, z\, \phi(s_i z)$，其中 $\sum_i a_i = 0$。由于 $\Phi(0) = 1/2$ 且 $0 \cdot \phi(0) = 0$， $h(0) = 0$；又由于当 $z \to \infty$ 时， $\Phi(s_i z) \to 1$ 且 $z\phi(s_i z) \to 0$，$\lim_{z\to\infty} h(z) = 0$。假设 h 在 $(0, \infty)$ 内有 $N \ge 2K-1$ 个互异零点。在边界 0 和 ∞ 处补上零点后，根据罗尔定理， h' 在 $(0, \infty)$ 内至少有 $N+1$ 个零点。作代换 $x = z^2$,$\lambda_i = s_i^2/2$ 后，$h'(z)$ 在 $(0, \infty)$ 内的零点与 $g(x) = \sum_i c_i \exp(-\lambda_i x) + \sum_i d_i\, x \exp(-\lambda_i x)$ 在 $(0, \infty)$ 内的零点一一对应，其中 $(c, d) \ne 0$，因为正尺度下系数映射可逆。由定理 2.1，该 g 在 $(0, \infty)$ 内至多有 $2K-1$ 个零点。等价地，函数族 $\{\exp(-\lambda_i x)\ \ x\exp(-\lambda_i x)\}_{i=1}^{K}$ 在 $[0, \infty)$ $[19,\ 20]$ 上构成一个 $2K$ 维扩展完全 Chebyshev 系统。因此 $N+1 \le 2K-1$ 且 $N \le 2K-2$，矛盾。 □

推论 *2.3*（等振荡）。设 $f^\star \in F_K$ 在某个内点处达到 $\Delta_K^\star = \sup_{z\ge 0} |\Lambda(z) - f^\star(z)|$。则误差 $e^\star = \Lambda - f^\star$ 在 $(0,\infty)$ 内恰有 $2K$ 个交替极值，其中 $|e^\star(z_j)| = \Delta_K^\star$ 对于 $j = 1, \dots, 2K$，且 $\operatorname{sgn} e^\star(z_j) = (-1)^{j+1} \operatorname{sgn} e^\star(z_1)$。由于 Λ 和 F_K 关于 $z = 0$ 对称，区间 $\mathbb{R}$ 上的扩展误差恰有 $4K$ 个这样的极值点。

证明概要。 交替极值点数量的下界可通过非线性切比雪夫逼近的柯尔莫哥洛夫判据 [21, 22]，由定理 2.2 推出。若 $f^\star$ 有 $N \le 2K-1$个交替极值点，则 V的 Haar 性质可得一个切向方向 $h \in V$，使其在每个极值点处都与误差的符号一致。随后，沿实现 h 的一阶参数变动，并控制一致的二阶余项，即可减小误差，这与极小极大性相矛盾。对于任何具有不同正尺度的 K项混合， $2K$的相应上界均成立（定理 E.1）。其误差的导数 e' 是关于 $u = z^2/2$的混合拉普拉斯变换，由一个正混合密度和 K个原子组成。双重剥离罗尔归纳法（即符号正则拉普拉斯核 [19]的变差递减性质所对应的零点计数形式）至多允许此类变换有 $2K$ 个零点，而 $e^\star$ 的每个内部极值点都是 $(e^\star)'$的一个零点。完整的证明链条（预备设定、存在性、Haar 性质、下界、上界、唯一性和对称性，见附录 A 至附录 G；附录 H 给出了受阻的另一种全正性路径，附录 I 则提供了证明依赖索引）见各附录。

□

定理 2.3 针对内部极小值点陈述。下一项结果见附录 B 中的证明，它在经认证的范围内去除了这一假设。达到是构造性的，且没有任何极小化子发生退化。

命题 2.4 （ $K \le 11$的内点达到）。对于 $K = 1, \dots, 11$ 极小极大 $\Delta_K^\star$ 能够达到，并且每个极小化子都是 F_K 的内点，其中没有消失权重，也不存在两个尺度重合或逃逸至 0 或 ∞ 的情形。

证明概要。在紧化$\mathcal{F}_K$ 上，由于参数域的紧性以及上确界范数在逐点极限下的下半连续性，可以达到极小值。极小化子可能通过两种边界机制离开内部，而从 $K = 1$ 基础情形开始对 K 进行归纳，并利用其在 $K-1$ 层级提供的闭包界，可以排除这两种机制。

首先，某个权重可能消失，或者两个尺度可能重合。对于 $K \ge 2$ 而言，这样一个减少计数的极小化子将位于$\overline{\mathcal{F}_{K-1}}$ 中，并产生至少为 $\Delta_{K-1}^\star > \Delta_K^\star$ 的误差。这个严格不等式不需要达到假设，因为第 3 节中给出的 K-混合从上方界定了 $\Delta_K^\star$ ，而经审计的（$K-1$）-混合则通过德拉瓦莱·普桑不等式（定理 F.2）从下方界定了 $\Delta_{K-1}^\star$。在 $K = 1$ 处不存在这样的极小化子，因为 $\sum_i w_i = 1$迫使单个权重等于 1。

其次，质量可能逃逸到边界尺度 $s = 0$ 和 $s = \infty$，在那里它会产生一个常数项或原点处的阶跃函数，并导致上确界误差等于该质量的一半。对于 $K \ge 2$ ，与 $\Delta_{K-1}^\star$的比较可在 $\Delta_K^\star / \Delta_{K-1}^\star < (1 - 4\Delta_K^\star)/2$时排除这种情况。经认证的比值不超过 0.19，而阈值始终大于 0.498。在 $K = 1$ 处，唯一的边界点是常数 $1/2$ 和单位阶跃，各自的误差均为 $1/2$。

□

推论 2.5 (全局唯一性). 对于 $K = 1, \dots, 11$，从 F_K 对 Λ 进行的一致最佳逼近是唯一的；任何在 $2K$ 个点处误差等振荡的内部候选解，都是该唯一的全局极小极大解。

证明思路。两个不同的实数集$\mathcal{F}_K$ 中的混合之差，是至多由 $2K$ 个不同的Φ原子构成、系数和为零的有符号线性组合。这样的组合在 $(0,\infty)$ 上至多有 $2K-2$ 个零点。在极小极大水平上，两个混合之一的误差在 $2K$ 个点处交替变化，因此至少有 $2K-1$ 个零点；这正是德拉瓦莱·普桑论证的等号情形（附录 F）。经典地，同样的

结论可由变差可解性 [22, 定理 7-4], [20, III.4.6] 得出。结合定理 2.3 和定理 2.4 即得所述结论。 □

定理 2.3 和定理 2.4 共同无条件地解决了$K \le 11$的 Monahan–Stefanski 猜想，而定理 2.5 将第 3 节中的每个经认证的等振荡极小值提升为唯一的严格全局极小值，因此任何 Remez 型不动点都获得了终止证书。图 1 展示了由此产生的等振荡。

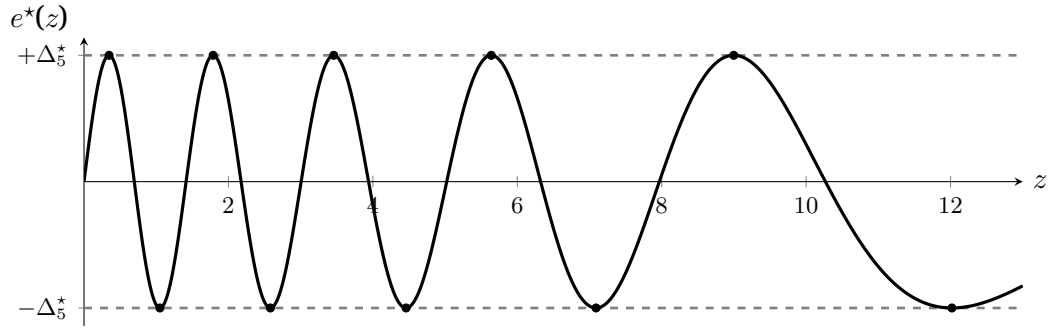


图 1：在 $(0,\ \infty)$ 上经认证的 $K = 5$ 极小极大误差 $e^\star(z) = \Lambda(z) - \sum_{i=1}^{5} w_i^\star \Phi(s_i^\star z)$，由表 1 中经认证的常数计算得到。与定理 2.3 一致，它在恰好 $2K = 10$ 个点处达到 $\pm\Delta_5^\star$（虚线，$\Delta_5^\star = 5.953 \times 10^{-7}$），这些点的符号严格交替（已标记），并因奇性在原点处为零。随着 z 增大，各极值点逐渐分开，最外侧的极值点接近 $z = 12$

注 2.6（其他原子族）。 定理 2.2 的证明仅使用了这样一个事实：切空间的导数在代换 $x = z^2$ 下映射到一个扩展切比雪夫空间，这一性质并非高斯原子所特有。它可以推广到任意光滑原子族$\{G(s_i z)\}$，其中具有偶密度 G'，且导数块 $\{G'(s_i z)\ \ z\,G''(s_i z)\}$ 在 $x = z^2$ 下变换为一个已知的扩展切比雪夫空间。该导数块只有在 $G''(u)$ 是 $u\,G'(u)$ 的常数倍时，才会退化为高斯形式$\{G'(s_i z)\ \ z^2 G'(s_i z)\}$，而这一条件刻画了高斯分布，因此其他任何原子都由一般导数块所支配。该类中的每个 G 都继承了切空间的 Haar 性质，并由此得到至少 $2K$ 次交替的下界。相应的上界和唯一性论证则需要更多结构，因为精确计数依赖于连续的G尺度混合及其符号正则导数核，而唯一性则依赖于带符号组合 $2K$ 原子的 $2K - 2$零点界。逻辑函数在 $G'' = G'(1 - 2G)$ 的情形下是一个合理的成员，但现有证明尚未解决这一情形。

2.2 朗斯基行列式局部唯一性

上述唯一性论证通过整个函数族的零点计数结构发挥作用。独立验证还需要进行代数检验，该检验可直接在每个经过认证的尺度向量 $s^\star$ 处求值，其尺度因子则可在全局搜索过程中衡量与尺度合并的接近程度（第 3.2 节）。切向 Wronskian 恒等式同时满足这两项要求。切向 Wronskian 的一个余子式若在有序正锥上处处非零，便可证明切空间基在每个候选点处均非退化，而无需依赖等振荡定理。令 $Q(z; s)$ 表示基函数组 $\{\Phi(s_i z)\ \ z\,\phi(s_i z)\}_{i=1}^{K}$ 的 $(2K-1) \times (2K-1)$ 切向 Wronskian 余子式；该基函数组通过删除 0 阶导数行和最大索引列 $z\,\phi(s_K z)$ 得到。各行按导数阶数升序排列，剩余各列则先按 Φ 块、再按 $z\,\phi$ 块排列；这一顺序确定了附录 F 中闭式的符号。

定理 2.7（结构性纯单项式恒等式，$K \le 5$）。对于 $2 \le K \le 5$ 和有序、互异的正数 $s_1 < s_2 < \cdots < s_K$，约化切向 *Wronskian* 余子式满足 $Q(z;\ s) = c_K(s)\ z^{(2K-1)(K-1)}$，其中 $c_K(s)$ 是由 *2* 的幂、各尺度的幂以及平方尺度两两差的偶次幂构成的带符号乘积，具体形式见附录 *F*；它在有序正锥上非零。因此，*Wronskian* 在 F_K 的每个内点均非奇异。

该证明是一次完整的符号计算。对 $K \le 4$进行 Bareiss 无分式消元 [23]，并对 $K = 5$ 在精确整数算术中进行分阶段后缀子式展开，得到约化行列式这一精确的多元多项式；该多项式按所述方式进行因式分解（附录 F）。等波动候选解是严格局部极小值，这是一个独立结论，在附录 F 中由唯一性推论得出；该推论依据 Haar 和 de la Vallée Poussin链得出严格性，而非依据 Q。

猜想 2.8（一般情形的纯单项式恒等式 K）。 定理 *2.7* 的恒等式对每个 $K \ge 2$均成立。

定理 2.8 在整数尺度 $s =$（1、2、.....、K）上对 $6 \le K \le 12$得到了精确验证。Bareiss 行列式与闭式计算结果相同，且在 $K = 12$处超过 400 位数字。该恒等式是一个可检验的证书，而不是一条新的唯一性路径。对于第 3 节中的每个经认证的混合，在经认证的 $s^\star$ 处 Q 的不消失性还通过在精确有理数算术中直接进行 Sturm实根计数得到确认，因此其中的局部唯一性证书并不依赖于定理 2.8。

3 $K = 1$、.....、11的经认证的解

第 2 节关于达到性和唯一性的结果，基于构造出其一致误差随 K 严格递减并在 $2K$ 个点处等波动的混合。本节报告 $K = 1$、.....、11 的经认证常数，说明认证与评估算法，详细介绍求解 $K = 9$、10、11 的搜索流程，并记录每项结果的复现方法。

表 1 报告了 $K = 1$、.....、11 的经认证的极小极大常数。对于 $K \le 8$，这些是 Monahan–Stefanski 的情形，[3], 我们使用 100 位精度运算重新求解了这些情形。在 $K = 8$这一最大情形下，我们的常数与其发表值在八位有效数字内一致。按原样计算时，发表的常数所得上确界相对于 $\Delta_8^\star$ 在相对量上超出 4.9×10^{-7}，或者在去除其权重超过 1 所产生的1.9×10^{-7}后，超出 10^{-15}。对于 $K = 9$、10、11，这些常数是新的。每个条目均在 100 位有效数字运算下满足定理 2.3。对于 $K = 9$、10、11，所有 $2K$ 极值点的绝对值均相等，其相对离差低于 10^{-50}，且符号严格交替。

比值 $\Delta_{K+1}^\star / \Delta_K^\star$ 在 $K = 1 \to 2, \dots, 10 \to 11$ 时分别为 0.054、0.085、0.108、0.126、0.141、0.153、0.164、0.173、0.182 和 0.189。该序列单调递增并趋近于 1，但直到 $K = 11$ 时仍严格小于 1。对 $K = 1, \dots, 10$ 范围内 $\log \Delta_K^\star$ 关于 $\{1, K, K^2, \log K, K \log K\}$进行最小二乘拟合，并向前外推一个阶跃，预测值为 $\Delta_{11}^\star \approx 1.256 \times 10^{-11}$；认证值与其相差不超过 0.19%，这为该认证提供了独立的一致性检验。图 2 绘出了全部十一条认证误差曲线，以及由这些常数确定的包络线。

注 3.1（证书保证的内容）。 认证值 $\Delta_K^\star$ 是对所部署的特定权重和尺度，在 100 位精度运算中审计得到的 $|\Lambda(z) - \sum_i w_i \Phi(s_i z)|$ 的上确界，也是经验证所部署混合能够满足的上界。对于式 (3) 中的 $\hat{}I$，无论该混合是否为全局极小极大，精度保证 $|I - \hat{}I| \le \Delta_K^\star$ 都对该混合成立。全局最优性，即不存在任何 K 项混合能够达到更小的一致误差，则是由定理 2.3 和 2.5 针对 $K \le 11$ 证明的独立且更强的结论。部署误差界不依赖于此。

表 1： $K = 1$至11 的经认证极大极小常数 $\Delta_K^\star$，以及完成各案例验证的搜索流程。M&S 表示 Monahan 和 Stefanski 的案例 [3]„ 这些案例在 100位审计下重新求解，所得结果与他们发表的常数在 $K = 8$（最大案例）下精确到八位有效数字。CMA-ES 表示协方差矩阵自适应进化策略，LM 表示 Levenberg–Marquardt，MPFR 表示 GNU 多精度浮点库。所显示的数值均四舍五入至图示位数。经审计的全精度权重和尺度可在配套代码库中找到，位于 `inst/certs/minimax_k01.csv` 至 `minimax_k11.csv`；经审计的界本身则位于 `inst/certs/deltas.csv`。

K	经认证（100位审计） $\Delta_K^\star$	极值点	流程
1	9.4573×10^{-3}	2	M&S
2	5.1013×10^{-4}	4	M&S
3	4.3559×10^{-5}	6	M&S
4	4.7167×10^{-6}	8	M&S
5	5.9529×10^{-7}	10	M&S
6	8.3872×10^{-8}	12	M&S
7	1.2857×10^{-8}	14	M&S
8	2.1086×10^{-9}	16	M&S
9	$3.656450583166549933 \times 10^{-10}$	18	CMA-ES + Remez +MPFR 牛顿法
10	$6.646923022720961975 \times 10^{-11}$	20	扰动参考雷梅兹算法 + MPFR 牛顿法
11	$1.258437696716086980 \times 10^{-11}$	22	Sturm 种子 CMA-ES +链式 Levenberg–Marquardt 算法 + 精修

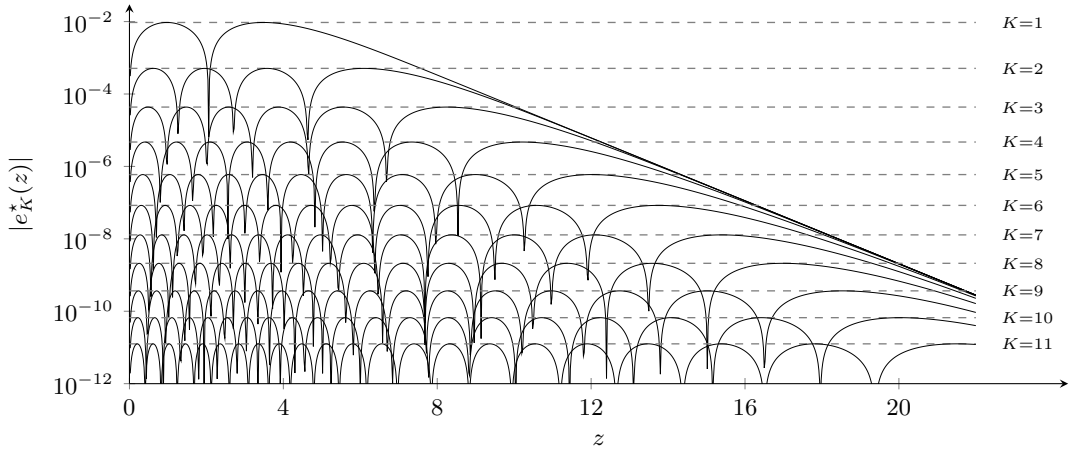


图 2：使用表 1 中的常数，以 256 位算术计算得到的整个经认证范围 $K = 1$，……，11 内经认证的极小极大误差幅值 $|e_K^\star(z)| = |\Lambda(z) - \sum_{i=1}^{K} w_i^\star \Phi(s_i^\star z)|$。灰色虚线标出经认证的 $\Delta_K^\star$。与定理 2.3 一致，曲线 K 恰好在 $2K$ 个点处与其包络线相切，且这些点全部位于窗口内；最外侧的极值点从 $z = 3.48$ 处向外移动，在 $K = 1$ 处达到 $z = 21.50$，并在 $K = 11$处达到另一位置。向下的尖峰是 $e_K^\star$的零点交叉，在对数坐标轴上被截断。越过各自最外侧的极值点后，每条曲线都会趋于共同的渐近线$1 - \Lambda(z) \sim e^{-z}$，因为此时高斯尾部已衰减至低于逻辑函数尾部，因此各曲线按顺序与该渐近线相接， $K = 1$ 最先， $K = 11$ 最后。包络线按文中给出的比率呈几何级数下降，因此这十一个证书跨越了近九个数量级

反之，对于这一族函数，通过初等差分论证证明的德拉瓦莱·普桑不等式（附录 F）表明，在经审计值相对离差为 ε 以内的水平上验证交替性，可将该值乘以（$1-\varepsilon$）作为下界。因此，每个经认证的常数都会在给定离差范围内，对真实极小极大值给出双侧包围。

3.1 认证与求值算法

搜索会返回一个候选解；只有在计算出其误差的上确界并验证其交替性后，候选解才会成为证书。算法 1 就是执行这一步。对于每个 K 和每种搜索流程，它都相同；它将候选解的权重与尺度作为精确小数输入，因此返回的界适用于实际发布的常数。

由于发布的权重以有限小数存储，令 $d = 1 - \sum_i w_i$ 表示它们的单位和缺陷；对于精确常数，该值为零，而对于此处发布的每一组权重，其绝对值都小于 10^{-39} 。于是 e（$-z$）$= d - e$（z），因此 e' 为偶函数，负半轴上的临界点与正半轴上的临界点互为镜像，并且当 $z \to -\infty$ 时， e（z）趋于 0；当 $z \to +\infty$时，则趋于 d 。因此， $|e|$在实数轴上的上确界必在（0， ∞）上的某个临界点处，或在其镜像点（$-\infty$，0）处取得，除非该上确界至多为 $|d|$。

为找出 $(0,\infty)$ 上的所有临界点，算法 1 在有限网格 e' 上扫描$G \subset (0, Z]$，以寻找符号变化。对于一般函数，这种扫描并不完整，因为落在两个相邻网格节点之间或超出 Z 的临界点对不会被检测到，从而低估上确界。定理 E.1 消除了这一族函数中的这个障碍。具有不同正尺度的任意 K 项混合，其误差在 $(0,\infty)$ 上至多有 $2K$ 个局部极值，因此，若网格扫描括定了 $2K$ 个 e' 的符号变化，就能证明已找到每一个局部极值。

算法 1 K项混合的认证
要求： 实权重 w_i 和不同的尺度 $s_i > 0$，均以精确小数形式给出，精度为 $d = 1 - \sum_i w_i$，以及一个网格 $G \subset (0, Z]$，以及两个以有效数字表示的工作精度 $p_{\text{scan}} < p$
确保： 界 $\Delta \geq \sup_{z\in\mathbb{R}} |e(z)|$ ，其中 $e(z) = \Lambda(z) - \sum_i w_i \Phi(s_i z)$，以及其交替性的跨度 ε
1: 以精度 p_{scan} 在 G 上计算 $e'(z) = \Lambda'(z) - \sum_i w_i s_i \phi(s_i z)$，并保留每对相邻使 e' 取相反符号的网格点 ▷ 括区间
2: 将每个括区间二分至精度 p ，得到一个临界点 ▷ $z_1 < \cdots < z_n$
3: 要求 $n = 2K$ 且 $\operatorname{sgn} e(z_{j+1}) = -\operatorname{sgn} e(z_j)$ 对于 $j = 1, \ldots, 2K-1$ ▷ 根据定理 E.1 完成
4: $\Delta \leftarrow$ 最大j 最大$\{|e(z_j)|,\ |d - e(z_j)|,\ |d|\}$，按报告的最后一位向上取整 ▷ 两个半直线
5: $\varepsilon \leftarrow 1 - \min_j |e(z_j)|/\Delta$ ▷ $(1-\varepsilon)\Delta - |d| \leq \Delta_K^\star \leq \Delta + |d|$

第 1 阶段对 e'的零点进行括区间定位，第 2 阶段细化临界点，第 3 阶段检查完备性和交替性，第 4 阶段在两个半直线上界定 $|e|$ ，第 5 阶段为正权重构造双侧包围。将权重除以其总和，得到 $\mathcal{F}_K$ 中的一个元素，其误差与 e相差至多 $|d|$。包围的上界成立，因为该元素属于$\mathcal{F}_K$；下界则由定理 F.2 的德拉瓦莱·普桑不等式给出，该不等式应用于其 $2K$ 交替取值。随常数一同提供的界由算法 1 重新计算，采用了由 16，000 个在（$0, 80$] 上等间距的点和 3，001 个在 $[10^{-4}, 1]$上对数间距的点组成的并集，精度为 $p_{\text{scan}} = 40$和 $p = 140$。所用算术是多精度浮点运算，而非区间算术，因此该证书是在所述精度下进行的验证（第 3.3 节）。第 4 节的松弛认证所涉及的临界点少于 $2K$ 个，具体而言，在 $K = 14$、15、16、17、22 处分别有 22、22、21、21 和 31 个。对于这些混合，第 3 和第 5 阶段不适用，扫描的完备性依赖于 G 的稠密性。

算法 2 是构建该证书所针对的算法。它将式 (3) 重述为一个过程 .

算法 2 利用经认证的混合计算逻辑斯蒂-正态积分**要求：** $\eta \in \mathbb{R}$, $v \ge 0$, 以及一个带有界 Δ 的经认证的混合 (w, s)**确保：** $\hat{I}$，其中 $|I(\eta, v) - \hat{I}| \le \Delta_1$: **对于** $i = 1, \dots, K$ **执行**2 $t_i \leftarrow s_i \eta / \sqrt{1 + s_i^2 v}$: : **结束循环**4: $\hat{I} \leftarrow \sum_{i=1}^{K} w_i \Phi(t_i)$

3 ———

计算成本为 K Φ 的值。它没有分支、没有迭代，也没有调节参数，并且其误差界在每个（η, v）处都是相同的常数，这正是第 5 节的采样循环所依赖的。

注 3.2（小值处的相对精度）。 算法 2 的界是绝对误差界。当 $\eta \to -\infty$积分衰减为 $e^{\eta + v/2}$ 时，而 $\hat{I}$ 以高斯速率衰减，因此 $\hat{I}$ 的相对误差在此处趋近于 1。因此，对于需要在小值处达到固定相对精度的应用，例如非简并半导体的载流子密度，仅使用 Δ 是不够的。一个已知的精确恒等式可以补足这一点。由于 $\Lambda(x) = e^x \Lambda(-x)$，对正态变量 $Y \sim N(\eta, v)$ 进行指数倾斜 $E[e^Y g(Y)] = e^{\eta + v/2} E[g(Y + v)]$可得

$$I(\eta, v) = e^{\eta + v/2}\, I(-\eta - v, v). \tag{5}$$

Crouch 和 Spiegelman [24] 将其列为对称关系之一，而它是 Pirjol 递推 [6] 与反射 $I(-\eta, v) = 1 - I(\eta, v)$ 的结合。在 $\eta = -v$ 处的特例，即精确值 $I(-v, v) = e^{-v/2}/2$，标志着半导体文献中从非简并区过渡到简并状态 [7]。Gory、Craigmile 和 MacEachern [25] 在八项混合之前应用该递推，并通过数值计算评估精度增益。借助经过认证的 Δ，一次应用即可得到相对误差的界。对于 $\eta < -v/2$，使用算法 2 计算右端项。乘积的误差至多为 $e^{\eta + v/2}\Delta$。变量 $-\eta - v$ 超过 $-v/2$，且 I随其第一个自变量增加，因此 $I(\eta, v) \ge e^{\eta + v/2} I(-v/2, v)$，相对误差至多为$\Delta / I(-v/2, v)$。对于 $\eta \ge -v/2$，直接计算达到相同的界，因此该界在每个 η处都成立，并且仅取决于 v。在 $K = 11$ 处，由于 $I(-v/2, v)$ 在这两个方差之间从 0.40 降至 2.1×10^{-3}，它在 $v = 1$ 处为 3.2×10^{-11}，在 $v = 36$处为 6.1×10^{-9}，因此保证会随v 增大而减弱。算法 2 的直接求和仍然是默认方法。对于和为 1 的权重，它恒等地满足 $\hat{I}(\eta, v) + \hat{I}(-\eta, v) = 1$，而倾斜计算则放弃了这一性质；第 5 节的效应表达式也是针对该方法给出的。

3.2 流水线与失效模式

这三个新案例面临同一个难题，因此需要采用三种不同的补救措施。在候选极值点$0 < z_1 < \cdots < z_{2K}$处，对于 $\sigma = \operatorname{sgn} e(z_1)$，等振荡方程组要求 $e(z_j; w, s) = (-1)^{j+1} \sigma \Delta$对 $j = 1, \dots, 2K$ 成立，且满足 $\sum_i w_i = 1$；在完整系统精修中，还需结合 $e'(z_j; w, s) = 0$。记 F为其关于权重、尺度和 Δ的残差，记 J为 F的雅可比矩阵。在每个 $K \ge 9$ 处，平方残差 $\|F\|^2$ 都存在鞍点和非等振荡局部极小值；牛顿法、Levenberg– Marquardt 和内尔德–米德法很容易到达这些点，却难以脱身。局部几何无法提供一阶信号，引导优化走向定理 2.3 所保证的等振荡全局极小值。每个 K 所采用的补救措施，都会注入一个非局部移动，以适配其 K的吸引域几何（表 2）。

$K = 9$**：全局逃逸** 标准的局部和全局搜索均收敛于 $\Delta \approx 6.72 \times 10^{-10}$ 处的非等振荡驻点。这些搜索包括牛顿法、内尔德–米德法 [26], 微分进化、Sobol 与极值范围播种、在 $K = 9$个样本上的随机盆地跳跃，以及从对数间隔字典中枚举 个九尺度子集。

在该点，误差仅在 18 个极值点中的 12 个处达到等振荡所要求的上确界的 99%。每个失败变体的协方差或阶跃函数大小在对数尺度空间中都是各向同性的，而目标盆地则具有很强的各向异性（经认证的尺度比为 $s^\star_{\max}/s^\star_{\min} \approx 6.3$）。在 9 维对数尺度空间上运行协方差矩阵自适应进化策略（CMA-ES） [12] 并进行 50 次重启后，找到了正确的盆地，因为其自适应协方差打破了这种对称性。以双精度进行 Remez 热启动后，在相对离差为 1。1% 时，误差降至 $\Delta \approx 3.67 \times 10^{-10}$。随后，MPFR 中的 256 位牛顿精修经过 13 次迭代收敛，而 115 位数字的全系统牛顿求解则生成了表 1 中的 100 位数字常数。

$K = 10$：离散度降低 $K = 9$ 流程在 $\Delta = 6.7095 \times 10^{-11}$到达了正确的吸引域，得到 20 个交替极值点，但离散度为 2.7%；此时牛顿法发散，而 Levenberg–Marquardt 则停滞于 $J^\top F = 0$，这是代数鞍点的典型特征。更大范围的 CMA-ES 扫描在 6.6952×10^{-11}找到同一结构的另一个驻点，但数值略有不同，因此已无更好的吸引域可寻。扰动参考迭代Remez策略将极值点视为控制参数。该策略将每个 $z_j^\star \to z_j^\star(1+\delta)$ 扰动 $\delta \in \pm\{1, 2, 3, 5, 8, 10, 20, 30\}\%$，并对六个幅值最小的极值点中的点对进行联合扰动，同时在z_j 固定的条件下求解牛顿等振荡方程组。经过三轮迭代，离散度从 2.7% 依次降至 0.33% 和 0.016%，之后 MPFR 牛顿法和 115 位精度的全系统牛顿求解均实现了二次收敛。

$K = 11$：链式 Levenberg–Marquardt $K = 11$ 问题同时包含这两种失效模式。其对数尺度上的极小极大盆地宽度小于 5×10^{-5}，从下文所述的 22 极值点候选解出发执行一次完整牛顿步便会离开该盆地，$\|F\|$ 从 1.4×10^{-12} 升至 1.7×10^{-7}。最小极小极大权重从 $w^{(10)\star}_{\text{最小}} \approx 2.45 \times 10^{-4}$ 降至 $w^{(11)\star}_{\text{最小}} \approx 1.03 \times 10^{-4}$，因此，基于此前解校准的权重下限不再适用。在成功的流程出现之前，有 29 种策略均告失败。它们分为三类：交替惩罚全局搜索、在坐标空间和对数尺度空间中进行局部牛顿精修，以及 $K = 10 + 1$ 原子插入。第一类只有在相对离差达到数量级为 1 时，或根本无法达到 22 次交替。唯一的例外是 $\|F\|^2$ 的一个鞍点，位于 $\Delta = 1.303 \times 10^{-11}$ 处（离差为 7.4%）；从该点出发，第二类方法将离差提高到 13.5% 和 11.7%。第三类中的每次尝试最终都坍缩到一个 $K = 10$ 吸引子。

有两个关键步骤完成了认证。第一步是以 Sturm 为种子的 CMA-ES 运行；当 $\log |c_K(s)|$（定理 2.8 的切向 Wronskian 系数）低于目标值时，目标函数便对试探解施加惩罚，从而引导搜索趋向彼此间隔较大的尺度。随后以 100 位精度运算重新审查该运行结果，发现有 22 个极值点，而线性规划网格验证器（一个双精度例程，在网格上针对固定尺度计算最佳权重）报告了 21 个。网格将一个幅度较小的极值点与其相邻极值点合并了。第二步是在 100 位精度运算中采用带自适应阻尼 λ 的链式 Levenberg–Marquardt。它在 $500, 500, 500, 10^4,$ 的五个区段中穿过狭窄的吸引域而未脱离，并最多进行了 2×10^5 次迭代，在 $\|F\| < 10^{-30}$时停止。最后六次迭代显示出牛顿二次收敛，此时 $\|F\|$ 从 4.2×10^{-20} 降至2.95×10^{-34}。从此处开始，对完整等振荡系统进行的 115 位精度牛顿求解收敛至残差 5.6×10^{-98}，发布的 $K = 11$ 常数即为其结果。

表 2 总结了这一模式。K 的三个取值产生了三种盆地几何结构和三种补救方案，没有任何单一的全局搜索流程能在这三种情况下都取得成功。

表 2： $K = 9$，10、11 处的吸引域几何、陷阱与补救措施。每种情况都需要根据其几何特征，对搜索动态进行非局部修改。CMA-ES 表示协方差矩阵自适应进化策略，LM 表示 Levenberg–Marquardt，F 是等振荡残差，J 是其雅可比矩阵。

K	吸引域几何	Trap	补救措施
9	错误吸引域占主导	非等振荡驻点	全局逃逸，CMA-ES
10	右侧吸引域，代数鞍点	$J^\top F = 0$	扰动参考雷梅兹算法
11	右侧吸引域，狭窄	完整牛顿步离开吸引域	链式 Levenberg–Marquardt 算法，100 位精度

3.3 可复现性

表 1 中每个条目的经认证常数均可在 `https://github.com/li-ruijie/LogisNormIntegral`公开获取，具体形式为每个K 对应的文件 `inst/certs/minimax_k01.csv` 至`minimax_k11.csv` （包含权重和尺度），以及包含界的 `inst/certs/deltas.csv` 。 $K = 14$、.....、17 处的松弛 Bernstein–Akhiezer 证书和带符号的 $K = 22$ 前沿位于同一目录中， `minimax_k14_BA_cert.csv` 至 `minimax_k17_BA_cert.csv` 以及`minimax_k22_signed_BA_cert.csv`也在其中，其经审计的界位于同一 `deltas.csv`中。求解过程中使用了三种工作精度。等振荡审计和 $K = 11$ 链式 Levenberg–Marquardt 运行采用 100 位有效数字的算术（Python `mpmath` [27]）完成，而对 $K \le 11$ 常数的牛顿法求解使用了 115 位工作数字。 $K = 9$ 和 $K = 10$的牛顿精修，以及第 3.2 节中 $K = 11$ 坐标空间和对数空间的精修，采用 256 位 MPFR 算术 [28] （R `Rmpfr` [29]）。全局搜索采用双精度运行。随后，使用算法 1 以 140 位有效数字重新计算 `deltas.csv` 中的界，同样采用 `mpmath`。这些证书是在所述精度下进行的浮点验证，而不是区间包络。在 50 位与 100 位有效数字下进行的所有交叉检查均至少有 50 位数字一致，这表明舍入效应远低于所报告的数字精度。

在准备本文期间，作者使用 Claude Opus 5（Anthropic）进行语言编辑，并将其用作编程助手。使用该工具后，作者对内容进行了必要的审阅和编辑，并对文章内容承担全部责任。

4 超越 $K = 11$ 的结构

在 $K = 11$ 以上的范围内，搜索均未找到严格的 $2K$ 点等振荡，因此问题变成了在此范围内仍能证明和认证什么。本节确立非递增界，提出 $K = 12$，13 时的平台猜想，反驳 $K = 14$，15 时的相同等式，通过 $K = 22$ 映射松弛认证，证明 $\Delta_K^\star \to 0$ (定理4.2)，并认证带符号的 $K = 22$ 前沿。

4.1 非递增界与有限平台

定理 2.3 的严格 $2K$ 交替性情形适用于 $K = 1, \ldots, 11$。超过 $K = 11$ 后，经认证的极小极大值非递增，$\Delta_K^\star \le \Delta_{11}^\star$，因为将 11 原子最优解补入 $K - 11$ 个额外原子（总权重为 ε ）后，其误差至多改变 ε，且 ε 可以任意小。

猜想 4.1 (有限平台)。 $\Delta_{12}^\star = \Delta_{13}^\star = \Delta_{11}^\star$。

目前只证明了 $\le$ 方向，而猜想的平台不可能延伸到 $K = 13$之后。在 $K = 14$，15 处，通过展示经 100 位数字

精度运算审计的混合，推翻了相应的等式。这些混合是在放宽后的 Bernstein–Akhiezer 判据下得到认证的，该判据下文记作 BA-(K+1)，要求至少有 $K+1$ 个临界点，而非 $2K$个临界点，其误差水平至少达到经审计的上确界误差的 0.9 倍（表 3）。这两个见证分别将 $\Delta_{14}^{\star} \le 7.233 \times 10^{-12}$ 和 $\Delta_{15}^{\star} \le 7.211 \times 10^{-12}$置于严格低于 $\Delta_{11}^{\star}$的位置。在 $K = 12$，13 处，三条独立的、各含 3，000 个种子的 Sobol 多起点流均未找到低于 $\Delta_{11}^{\star}$的误差；以填充后的 $K = 11$ 最优解为初始值进行的运行也回到了该解。在每个 $K = 14$，15 见证所对应的 13 尺度子集上固定尺度并重新优化权重，同样未能得到低于 $\Delta_{11}^{\star}$的误差。这三种工具（以种子初始化的多起点搜索、约束 Remez 权重再优化和尺度子集选择）正是产生 $K = 14$，15 见证的同一套工具。通过计数零点的自然途径无法得出结论，因为定理 2.1 的尖锐节点界允许相关的带符号混合拥有至少与 $K = 11$证书的等振荡所强制要求的一样多的内部零点。

4.2 认证层级与跨-K 记录

表 3 汇总了从 $K = 11$ 到 $K = 17$ 以及在 $K = 22$处得到认证的结果，即 $K = 11$处的严格证书、其上的宽松证书，以及两项 $K = 22$ 记录。如式 (1) 所示，权重为正的混合构成正类；权重可正可负且总和为 1 的混合则构成第 4.4 节的有符号类。

表 3： $K = 11$、...、17 的正类认证层级、正类 $K = 22$ 记录（经更高精度审计，但未提供证书文件），以及已知最佳的有符号 $K = 22$前沿，扩展了表 1。 $K = 12$、13 处猜想的平台对应定理 4.1 的等式 $\Delta_K^{\star} = \Delta_{11}^{\star}$ （仅证明了 $\le$ ）。Bernstein–Akhiezer BA-(K+1) 证书将严格的$2K$-等振荡放宽到 K+1 个临界点，而 BA-K 证书则放宽到 K；两者的水平都至少达到所展示混合的经审计上确界误差的 0.9 倍。 $K = 11$、14、15、16、17 和 22 项均通过 100 位精度运算验证。宽松证书的数值向上舍入到所示的最后一位，因此每个印出的数值本身都是 $\Delta_K^{\star}$的认证上界。

K	经认证的误差（100位审计）	证书	来源
11	1.258×10^{-11}	严格的 $2K$-等振荡	表 1
12, 13	$\le \Delta_{11}^{\star}$	平台猜想（$\le$ 已证明）	本文
14	7.233×10^{-12}	BA-(K+1) 放宽	本文
15	7.211×10^{-12}	BA-(K+1) 放宽	本文
16	4.066×10^{-12}	BA-(K+1) 松弛	本文
17	2.544×10^{-12}	BA-K 松弛	本文
22	2.811×10^{-14}	BA-(K+1) 松弛，正类	本文
22	3.605×10^{-15}	带符号的 BA-(K+1) 松弛形式	本文

宽松证书保留了严格保证的上界部分，即在所述精度下，所展示混合的经审计的上确界误差从上方界定 $\Delta_K^{\star}$，但不包含 de la Vallée Poussin 下界。放宽仅涉及所展示的等振荡结构，并未对 $K \ge 12$声称其具有最优性。在 BA-(K+1) 判据下，反驳平台等式的 $K = 14$ 和$K = 15$ 混合的经审计上确界误差分别为 7.23290×10^{-12} 和7.21007×10^{-12}，而 $K = 16$ 达到 4.065×10^{-12}。在 $K = 17$处， K点 BA-K 判据认证了2.544×10^{-12}。超过 $K = 17$ 后，在双精度线性规划网格流水线下，Bernstein–Akhiezer 判据不再适用于正类；该流程围绕基于网格的切比雪夫权重步骤优化尺度（第 3.2 节），其权重步骤在接近 2.4×10^{-13}处出现下限。该下限是算法伪影，而非极小极大界。对该平台的位势论解读是：在分岔为双区间平衡测度之前，先出现单区间饱和；其改进的起点为

$K=14$，详见附录 K。

表 4 记录了严格情形如何退化。对于 $K=1,\dots,11$，每个经认证的极小化解都将所有 $2K$ 临界点置于上确界处。在 $K=12$ 临界点计数 $n_{0.99}$ 中，距离上确界 1% 以内的数量骤减，到 $K=19$时最低降至一个，而原始计数 n_{crit} 仍维持在接近 K数量级。该转变并非秩亏。双精度奇异值分解中的秩上限起初表明有效维数约为 22。关于权重、尺度和极小极大水平的等振荡方程组具有一个 $2K\times(2K+1)$ 雅可比矩阵。在 60 位精度运算下，该矩阵在 $K=16$ $(32/32)$、$K=17$ $(34/34)$ 以及 $K=22$ 正类记录（$44/44$，最小奇异值为 $\approx 1.3\times10^{-21}$）处均具有满行秩。因此，严格 $2K$ 方程组在数值上是适定的，障碍在于找到全局吸引域。

表 4：正类流水线记录的跨 K 临界点图。列 n_{crit} 统计表列配置在 $0<z\le 200$ 上的误差临界点数量，而 $n_{0.99}$ 统计其中距上确界不超过 1% 的临界点；严格等振荡要求 $n_{0.99}=2K$。超过该窗口后，逻辑函数的指数尾部会压过所有高斯尾部，因此误差在该处单调，除 $K=19$、20、21 外；这些位置各自还有一个极值点，其位置约为 $z=486$、294 和 387，幅度均小于 10^{-128}。$K\ge12$ 行是双精度线性规划网格流水线的记录。其在 $K=12$、……、15 处的直接非退化解，上确界误差高于正文的填充界 $\Delta_K^\star\le\Delta_{11}^\star$，也高于表 3 中经认证的 $K=14$、15 界，因此该图记录的是临界点结构，而非最佳界。在 $K=16$ 处，上确界误差即为审计认证值，而两个计数列给出了涵盖近最优流水线配置的范围。正类中的 $K=22$ 行已被表 3 中精度更高的条目（2.811×10^{-14}）取代；有符号类中的该行则已被正文的前沿取代（在 3.604×10^{-15} 处有 30 个主导交错点）。

K	n_{crit}	$n_{0.99}$	上确界误差
1–11	$2K$	$2K$	表 1，严格
12	16	15	5.04×10^{-10}
13	20	15	1.57×10^{-10}
14	18	15	6.30×10^{-11}
15	19	13	1.72×10^{-11}
16	21–25	4–16	4.07×10^{-12}
17	21	3	2.54×10^{-12}
18	22	2	4.23×10^{-13}
19	25	1	3.30×10^{-13}
20	21	2	2.97×10^{-13}
21	20	1	2.52×10^{-13}
22	23	1	2.41×10^{-13}

4.3 极小极大误差在极限中的消失

上述填充界将每个 $K\ge12$ 处的 $\Delta_K^\star$ 限制在 $\Delta_{11}^\star$，但仍无法确定该序列是否会一直下降到零。一个将混合密度离散化的显式构造解决了这一问题。

命题 4.2（消失的极小极大误差）。极限 $\lim_{K\to\infty}\Delta_K^\star=0$。

证明。定义尺度不变模量 $D(\rho)=\sup_{u>0}|\Phi(\rho u)-\Phi(u)|$。对于 $\rho\ge1$，由 ϕ 在 $(0,\infty)$ 上的单调性可得 $|\Phi(\rho u)-\Phi(u)|=\int_u^{\rho u}\phi\le(\rho-1)\,u\phi(u)\le(\rho-1)/\sqrt{2\pi e}$，因为 $\max_{u>0}u\phi(u)=\phi(1)=1/\sqrt{2\pi e}$，且情形 $\rho\le1$ 对称，因此 $D(\rho)=\sup_{r\in[1/\rho,\rho]}D(r)\to0$ 当 $\rho\downarrow1$ 时。

固定 K，并将 $[1/K, K]$ 按对数尺度划分为 K 个公比为 $\rho_K = K^{2/K} \to 1$ 的区间 I_j。放置一个权重为 $w_j = \int_{I_j} q\,ds$ 的原子 $s_j \in I_j$，将尾部质量 $\tau_K = \int_{(0,1/K)\cup(K,\infty)} q\,ds$ 加到两端的原子上，使得 $\sum_j w_j = 1$，且混合属于 $\mathcal{F}_K$。对于 $s \in I_j$ 和 $z > 0$，令 $u = s_j z$ 和 $\rho = s/s_j$，可得 $|\Phi(sz) - \Phi(s_j z)| \le D(\rho_K)$。根据第 1 节的尺度混合表示 $\Lambda(z) = \int_0^\infty \Phi(sz)\; q(s)\; ds$，

$$\Big|\Lambda(z) - \sum_{j=1}^{K} w_j \Phi(s_j z)\Big| \le \sum_j \int_{I_j} |\Phi(sz) - \Phi(s_j z)|\, q\, ds + 2\tau_K \le \widetilde{D}(\rho_K) + 2\tau_K,$$

其中 $2\tau_K$ 界定了两项尾部贡献。根据 $\Lambda(-z) = 1 - \Lambda(z)$ 和 $\sum_j w_j = 1$，误差关于 z 为奇函数，因此该界在整个 $\mathbb{R}$ 上均成立。当 $K \to \infty$ 时，$D(\rho_K) \to 0$ 且 $\tau_K \to 0$，因此 $\Delta_K^\star \to 0$。 □

自然的极限对象就是混合密度 q，其闭式为 $q(s) = (2/s^3) \sum_{n \ge 1} (-1)^{n-1} n^2 \exp(-n^2/(2s^2))$ [2]。其低阶矩可通过狄利克雷 η 函数写成闭式，详见附录 J；经认证的原子已经实现了其中两个。二阶矩 $\sum_i w_i^\star (s_i^\star)^2$ 趋近于 $7\zeta(3)/(2\pi^2) = 0.42627\ldots$（精确到七位数字，由 $K = 10$ 给出），而对数尺度矩 $\sum_i w_i^\star \log\, s_i^\star$ 趋近于 $(\gamma_{\mathrm{EM}} + \log\, 2 - 2\log\, \pi)/2 = -0.50954\ldots$（精确到七位数字），其中 γ_{EM} 为欧拉–马歇罗尼常数。

这两个极限最初都是搜索认证常数 $\Delta_{11}^\star$ 闭式时得到的副产收敛值。整数关系搜索（PSLQ）[30]在三轮搜索中均未返回任何关系。第一轮针对相关莫德尔积分 [31] 的模变换常数，搜索系数高度不超过 10^{50}的情形。第二轮针对九个族中的认证权重和尺度的 81 项统计量，以及按秩排序的尺度的 21 个幂律拟合，搜索高度不超过 10^{25}的情形。第三轮针对以 K收敛的统计量的极限值，其中包括上述两个极限，搜索高度不超过 10^{60}的情形。因此，在所测试的精度下，该认证常数不满足低高度的任何整数关系，而这两个收敛值是极限混合密度 q 的矩，并非 $K = 11$所特有的常数。

4.4 更高精度下的有符号前沿

采用更高精度计算可以消除双精度下限。在正类中进行尺度搜索时，以交替次数为目标函数，依据 BA-(K+1) 准则，在 $K = 22$处达到 2.810×10^{-14}，这就是表 3 中的相应条目。在有符号类中使用相同目标函数运行协方差矩阵自适应进化策略（CMA-ES）（$\sum_i w_i = 1$，全部 $w_i \ne 0$，符号不受限制），在 $K = 22$处达到 $\Delta = 3.604 \times 10^{-15}$，其误差在$\ge 0.99\,\|e\|_\infty$处具有 30 个严格交替极值。该界约为双精度单位舍入误差 2^{-53}的三十三分之一。

以正规方程形式求解指定交替性的牛顿系统时，在接近 4.5×10^{-20} 处停滞，这是将雅可比条件数平方所造成的伪象（$\kappa \approx 1.8 \times 10^{15}$ 变为 $\kappa^2 \approx 3 \times 10^{30}$，耗尽了该求解阶段的工作精度）。相反，直接对 J 的奇异值分解（SVD）进行阶跃函数计算，并结合逐方向阻尼和测地加速 [32, 33]，可将残差降至 1.05×10^{-90}，并证明严格有符号30点等振荡。该前沿处的误差有31个临界点，即30个主导交错点，以及另一个低于该水平的临界点。

若要获得双精度下精确的证书，使误差低于双精度单位舍入误差 2^{-53}，就需要满足完整的严格 44 点等振荡。无论尝试何种路径，都无法从这一前沿达到该点数，各项诊断也都指向同一机制。从显式的 44 瓣等波纹种子出发进行函数同伦会受条件数阻碍，因为该种子带有$\kappa \approx 2 \times 10^{97}$。将原子数增加到 $K = 23$ 和 $K = 24$ 要么会回到前沿，要么

会到达误差高出六个数量级的吸引域，且权重彼此抵消。在缺失的临界点插入后，于 $K = 22$ 吸引域内为基于奇异值分解的求解器提供初始值，求解会在 $\|F\| \approx 3.6 \times 10^{-14}$到 5.9×10^{-14} 处停滞，并回到该前沿的 31 个临界点。初始条件数（$\kappa \approx 10^{28}$ 到10^{35}）意味着在该求解器的工作精度下，可达到的误差下限约为 10^{-45} 到 10^{-52}，因此这种停滞源于结构，而非精度限制。要生成缺失的 13 个临界点，必须经历一个带有 $e' = e'' = 0$的退化构型分岔；在此过程中，相应的雅可比矩阵行会变得线性相关，而局部牛顿流无法越过这样的折叠。

带符号混合与正权重混合一样，利用三角不等式给出积分界，$|I(\eta,\ v) - \hat{}I(\eta,\ v)| \le \Delta$ 其中 $\hat{}I$ 如式（3）所示，无论权重的符号如何。中介分析应用并不依赖这些更高阶的K 性质，因为经过认证的 $K = 11$ 混合已将误差控制在比第 5 节的泰勒余项低十个数量级的范围内。

5 因果中介的应用及与其他求值方法的比较

第 3 节和第 4 节认证了混合常数，并给出了算法 2，用于以固定的一致误差界计算逻辑斯蒂-正态积分（2）。本节将该求值方法应用于二元结局的因果中介分析，此时自然间接效应可化为两个逻辑斯蒂-正态积分之差。用经过认证的求和式取代标准泰勒展开，可以消除其截断偏差；本节还构造准贝叶斯区间（第 5.3 节）。随后，本节将该混合方法与高斯–埃尔米特求积、高斯–费米近似、有理极点展开以及 Pirjol 的泊松求和级数进行基准比较（第 5.4 节），并展示该闭式可用于顺序或并行的k 正态中介变量组合，且只需进行一次求值（第 5.5 节）。

5.1 因果中介中的积分

考虑一个针对二元结局 $Y \in \{0, 1\}$ 和连续中介变量 M 的两阶段模型，其中暴露为 z，协变量为 c。结局遵循逻辑斯蒂回归 $P(Y = 1 \mid z, m, c) = \Lambda(\theta_0 + \theta_1 z + \theta_2 m + \theta_3 z m + \theta_4^\top c)$，中介变量遵循同方差正态回归$M \mid z, c \sim N(\gamma_0 + \gamma_1 z + \gamma_2^\top c, \sigma_M^2)$。在标准识别假设和跨世界独立性 [34], 下，当直接暴露固定为 z 、中介变量按暴露 z' 时的分布抽取，预期结局等于 $\Lambda(\theta_0 + \theta_1 z + (\theta_2 + \theta_3 z)M + \theta_4^\top c)$ 对 $M \sim N(\gamma_0 + \gamma_1 z' + \gamma_2^\top c, \sigma_M^2)$ 的期望。该期望就是逻辑斯蒂-正态积分 (2)，其中$\eta(z, z', c) = \theta_0 + \theta_1 z + (\theta_2 + \theta_3 z)(\gamma_0 + \gamma_1 z' + \gamma_2^\top c) + \theta_4^\top c$，且 $v(z) = (\theta_2 + \theta_3 z)^2 \sigma_M^2$。暴露改变 $0 \to z'$ 对中介变量的自然间接效应，是反事实对比 $\mathrm{NIE}(z, z', c) = \mathbb{E}[Y(z, M(z')) - Y(z, M(0)) \mid c] = I(\eta(z, z', c)\ v(z)) - I(\eta(z, 0, c)\ v(z))$；令 $z = z' = 1$ 即可得到总自然间接效应。M 的正态性使每个潜在结果均值都化为单个积分 (2)，而同方差性使 v 独立于 z' 和 c。

Valeri 和 VanderWeele（V&VW）的估计量 [35], 通过其 SAS 和 SPSS 宏成为应用因果中介分析中的默认方法，它用围绕 $m = 0$ 的二阶泰勒截断 $\hat{}I_{\mathrm{V\&VW}} = \Lambda(\eta) + \frac{1}{2} v \Lambda''(\eta)$ 替代 $I(\eta,\ v)$。由于 $\Lambda(\eta + m)$ 在复平面上的 $\eta + m = \pm i\pi$ 处有极点，其关于 m 的泰勒级数收敛半径有限，为 $\sqrt{\eta^2 + \pi^2}$，因此当 $\eta \neq 0$ 时，对于任意 $v > 0$，在高斯密度 $(-\infty,\ \infty)$ 上逐项求期望都会发散。对于二阶截断，奇数阶高斯矩为零，而四阶矩 $3v^2$ 提供因子 $3/4! = \frac{1}{8}$，从而将余项界定为

$$|R_{\mathrm{V\&VW}}(\eta, v)| \le \tfrac{1}{8}\, v^2 \sup_m |\Lambda^{(4)}(\eta + m)| \le \tfrac{v^2}{8}\, C_4, \qquad C_4 = \sup_x |\Lambda^{(4)}(x)| \approx 0.128. \tag{6}$$

该界对任意 $(\eta,\ v)$ 都是有限的，但会随 v 呈二次增长。一个具有代表性的情形 $\sigma_M = 1$,$\theta_2 = 1.5$, $\theta_3 = 0.5$, $z = 1$ 给出 $v = 4$，余项界约为 0.26。该界与表 L.1 中范围为 0.24 至 0.41 的间接效应相当，足以掩盖关注的效应。因此，泰勒方法取决于具体情形：在 v 较小且交互作用较弱时准确，但在实践中常见的中等-v 情形下存在偏差。经认证的混合 (3) 则具有无参数界 $\Delta_K^\star$。

在部署软件中，同一个积分通过三种方法之一进行求值，而经认证的混合分别取代了这些方法。准贝叶斯模拟（如 `mediation` 软件包 [36], 所示）对中介变量进行抽样，并对拟合的结果概率取平均，同时会产生蒙特卡罗误差，还需承担内部抽样循环的成本。在 `regmedint` [37]中实现的解析泰勒展开虽为闭式形式，但继承了无界余项（6）。固定网格数值积分是 Mplus [38, 第 8.1 节 3], 采用的路径，它在标准化范围 $[-5, 5]$ 上的固定 101 点网格上对被积函数进行求值；这是一种确定性规则，其精度由网格间距决定。附录 L 量化了该混合方法与前两种路径的比较；对于第三种路径，它则用更短且经认证的求和取代 101 节点网格。

5.2 正确性不对称性与一个 worked example

表 5 根据应用中介分析所关注的标准，对比了两种求值方法。经认证的混合适用于整个参数域 $(\eta,\ v) \in \mathbb{R} \times (0,\ \infty)$。没有任何情景需要使用泰勒展开，而该展开在中等-$v$ 情形下会失效。

表 5：Valeri 和 VanderWeele (V&VW) 的泰勒展开 [35] 与自然间接效应 (NIE) 的经认证的极小极大混合方法之间的正确性不对称性。

	V&VW 泰勒展开	经认证的混合
误差界	$O(\sigma_M^4)$，在 θ_3 中无界	$\Delta_K^\star$ 一致地
认证	启发式，未经证明	已证明的全局极小极大（$K \le 11$）
参数依赖性	是（随 v 增大而减弱）	none
闭式自然间接效应	近似	在 $\mathcal{F}_K$ 中精确
非对称区间	仅对称德尔塔法	准贝叶斯

考虑一个合成实例，其中 $(\theta_0,\ \ldots\ldots,\ \theta_4) = (-1.0,\ 0.5,\ 1.5,\ 0.5,\ -0.2)$、$(\gamma_0,\ \gamma_1,\ \gamma_2) = (0.5,\ 1.2,\ 0.3)$，以及 $\sigma_M = 0.8$，在 $z = z' = 1$ 和 $c = 0$ 处进行评估，使得 $v = 2.56$。在此实例中，经过认证的 $K = 11$ 混合给出 NIE $= 0.30464$，其绝对误差至多为 $2\Delta_{11}^\star < 2.52 \times 10^{-11}$；由于该效应是两个积分之差，这一数值是单积分界的两倍。泰勒展开在相同系数下给出 0.34238，绝对误差为 0.0377，约为真实效应的 12%。

配套代码库中的默认混合是经认证的 $K = 22$ 有符号前沿，其最坏情况误差为 3.605×10^{-15}。此处和表 L.1 中展示的是正类 $K = 11$ 混合；在相同的参数抽样下，两者精确到小数点后十位均一致，因此采用哪种经认证的混合对报告值没有影响。

5.3 准贝叶斯区间

由于极大似然估计量具有渐近正态性，从其拟合抽样分布中抽取 $R = 10^4$ 个参数向量，并在每次抽样时计算式 (3)，即可得到非对称准贝叶斯区间 [39]（根据伯恩斯坦–冯·米塞斯对应关系，这些区间是渐近有效的可信限）。每次抽样需计算两个积分，每个积分调用 K 次 Φ，因此 $R = 10^4$ 次抽样需要 $2KR$ 个 Φ 值，并可在 15 毫秒内完成。

附录 L 比较了三种结局患病率区间下、 $R = 10^4$ 次外层抽样时的四种内积分策略。经认证的混合、内层蒙特卡洛以及第 5.4 节的逐点计算器在所有区间内均达成一致。仅使用泰勒展开会产生偏差，分别为 -13、 $+9$ 和 $+25$ 个概率尺度百分点；该偏差会随发生率改变符号，并且在高发生率下，其区间达到 1.06，超过自然概率界。按代码库默认值 $K = 22$ 计时，经认证的混合在 $M_{\mathrm{inner}} = 10^4$ 时比内层蒙特卡洛快 380 多倍，比逐点路线快 13 倍以上。重复抽样实验确认，混合方法区间在所有区间内均保持名义 95% 水平，而泰勒区间则未达到该水平，且偏离方向会随发生率改变。

因此，经认证的混合以一次闭式求值取代泰勒展开、内层蒙特卡洛和固定网格求积，并附带经认证的统一误差界，代价是需要使用德尔塔法区间。

5.4 替代求值方法与逐点路线

除中介分析软件采用的随机方法和泰勒方法外，数值分析和物理学文献还通过求积规则、高斯–费米近似、有理极点展开，以及 Pirjol 的逐点递推和泊松求和级数来求值 (2) [6] （表 6）。在求积规则中，Crouch 和 Spiegelman [24] 针对这类积分设计了一种规则，应用对象是逻辑斯蒂-正态模型；该梯形法则根据显式误差界来选择步长。标准高斯–埃尔米特求积直接针对正态权函数进行积分，但在求积变量中，被积函数的极点 $\eta + m = \pm i\pi$ 距实轴为 $\pi/\sqrt{2v}$，因此随着 v 增大，收敛速度会下降。使用 11 个节点时，其函数求值次数与 $K = 11$ 混合相同；在 $|\eta| \le 30$ 上的最大实测误差，在 $v = 1$ 处为 1.4×10^{-7}，在 $v = 9$处为 2.1×10^{-3}。即使使用 100 个节点，在 $v = 9$ 处的误差仍为 2.3×10^{-10}；相比之下，混合在每个 v处的认证界均为 1.26×10^{-11}。

第1节引用的半导体文献以高斯–费米名称讨论了同一个积分。其双分支近似 [7] 是一个闭式，作者认为其误差对于 $-v < \eta < 0$低于 10% 到 15%。以对 $|\eta| \le 30$ 和 $v \in \{4,\ 9,\ 16,\ 25,\ 36\}$ 进行自适应求积的结果为基准，在文章绘制的宽度范围内，其最大绝对误差为 4.5×10^{-3}，最大相对误差为 18%。误差函数级数 [9, 10] 在极限情况下是精确的，但收敛缓慢。在六个自变量 $\eta \in \{1,\ 3\}$、 $v \in \{1,\ 9,\ 25\}$处，要达到 10^{-6} 的误差，需要 56 到 492 项。Selvaggi 将该级数作为数值方法的检验，并指出数值积分是更高效的求值途径。反函数关系也有解析近似，即以载流子密度为自变量的费米能级 [40]。

另一类方法是在积分之前先逼近西格玛函数。在电子结构计算中，反射逻辑斯蒂函数 $\Lambda(-x) = 1/(1 + e^{x})$ 就是费米–狄拉克分布，通常用具有简单极点的有理函数来逼近。Lin、Lu、Ying 和 E [41] 构造了极点展开式，其计算成本随逼近区间宽度的对数增长。Moussa [42]以双精度列出了半直线上的极小极大有理逼近函数，并指出当极点数量较大时，传统的雷梅兹算法无法计算该函数。由于有限个简单极点之和在无穷远处趋于零，而 $\Lambda(-x)$ 在 $x \to -\infty$ 时趋于一，因此该类函数的误差在此处趋于一，且其中没有任何一个函数能在整个实数轴上对 $\Lambda(-x)$ 作一致逼近。正态尺度混合具有与 Λ 相同的两个渐近极限，因此有可能得到一致误差界。在表 6 的闭式行中，Probit 缩放 $\Phi(c\eta/\sqrt{1 + c^2 v})$ 是式 (3) 中的 $K = 1$ 成员。~~Ga~~Gaynor、Schwartz 和 Lin [43] 将其用于共同结局下的闭式中介效应点估计，精度达到 $\Delta_1^{\star}$ 所明确界定的两位数精度下限。Gory、Craigmile 和 MacEachern [25] 采用 Monahan 和 Stefanski 的八项混合计算该积分，并首先应用 Pirjol递推，

从而使混合在原点附近进行计算。在这些闭式行中，只有正态尺度混合族兼具（$\eta,\ v$）的闭式表达和无需参数的认证界。

表 6：第 5.3 节和附录 L 中方法之外，对式 (2) 的其他求值方式。Probit 行是混合族 (1) 中的 $K=1$ 成员，其优化常数达到表 1 的 $\Delta_1^\star$。拉普拉斯收缩阈值由 $\sup_x \Lambda'(x) = 1/4$ 得出；帕德范围则使 Sigmoid 自变量在正态变量两个标准差范围内保持在精度区域 $|x| \lesssim 3$。高斯–埃尔米特数值是对 $|\eta| \le 30$ 测得的最大误差。双分支近似和误差函数级数的数值是在正文所给自变量上测得的。正文讨论了逐点路线。

路线	闭式	一致保证	局限
Probit 缩放 $\Phi\left(c\eta/\sqrt{1+c^2 v}\right)$	yes	2.3×10^{-2} 在 $c=\sqrt{3}/\pi$，$\Delta_1^\star$ 优化	两位数精度下限
泰勒（矩）展开	yes	无，余项 (6)	对所有……都发散 $v>0$ 当 $\eta\neq0$ 时，半径 $\sqrt{\eta^2+\pi^2}$
拉普拉斯	隐式求解模式	none	中等偏度误差 v, 收缩性需要 $v<4$
双曲正切中的帕德	yes	none	尾部错误，需要 $\lvert\eta\rvert+2\sqrt{v}\lesssim3$
变分子式下界 [44, 45]	yes	单侧	下界而非求值
高斯–埃尔米特求积，n 节点	yes	无，误差随 v增大	2.1×10^{-3} 在 $n=11$处，$v=9$
双分支近似 [7]	yes	none	4.5×10^{-3} 绝对误差，18% 相对误差
误差函数级数 [9, 10]	无穷级数	none, 无截断界	针对 10^{-6}，从 56 项到 492 项
Pirjol递推和泊松级数 [6]	逐点例程	逐点截断控制	在（$\eta,\ v$）中没有闭式
经认证的混合（3）	yes	$\Delta_K^\star$，1.26×10^{-11} 处 $K=11$	固定近似每 K的下限

Pirjol 的逐点路线 [6] 与所有闭式行截然不同。精确递推和泊松求和级数以工作精度计算式 (2) 的单个值。

这一比较基于测量，脚本 `PirjolTiming.R` 随附捕获的输出，并发布于`inst/reproduction`。在覆盖 $v\le40$ 的八点网格上，其 R 实现在截断阶数 30 时，复现高斯–克朗罗德参考值的误差至多为 2.3×10^{-16}，每次调用耗时 $21.5\,\mu$ 秒；截断阶数 10 时，误差至多为 1.6×10^{-12}，每次调用耗时 $8.7\,\mu$ 秒。认证的 $K=11$ 求和计算相同点时，观测误差至多为 2.4×10^{-14}，处于其认证界内，每次调用耗时 $1.7\,\mu$ 秒。因此，截断阶数足够高时，逐点路线的单点精度更优，而两种路线的计算时间都在微秒量级。

逐点路线无法提供的是中介分析应用所需的对象。每次调用返回的是一个数值，而不是一个函数，因此无法获得效应的闭式表达式、其导数，以及在第 5.3 节的 10^4 准贝叶斯抽样中进行数组求值的功能。此外，截断阶数必须在每次抽样时加以控制，因为级数收敛速度会随 $(\eta,\ v)$ 而变化。后续处理在各种可能的展开式中筛选出一个最优级数，并提供尾部误差控制，从而在不改变这种

逐点保证特征的情况下，改进了逐次抽样管理。对相关 logit-正态矩的同期处理从另一侧说明了同样的权衡。该方法用截断至十一项的级数近似逻辑函数，并配有拟合的近零修正、针对指数项的溢出保护，以及用于在不同分支之间进行选择的区域阈值。它将准确性报告为实测范围，而不是界。经认证的混合在每个点都提供先验的一致保证 $\Delta_K^\star$，且该保证在进行任何抽样之前就已固定。因此，这两种路线互为补充。对于已知参数下的单个数值，Pirjol 的求值器更为适用；对于作为估计量、采样循环和证明中函数而被使用的积分，经认证的混合更为适用。

在准贝叶斯循环中，逐点路线可用作内层引擎。在低发生率和中等发生率区域，它能将经认证的混合复现到脚本保留的六位小数；截断阶数为 10 时，每 10^4 次抽取耗时 0.17 秒，阶数为 30 时耗时 0.43 秒。尽管如此，它仍需在每次抽取时管理截断阶数，并通过数值微分计算导数；而且在其定义域的两端都会返回非有限值，高患病率区域正好会触及这两端。附录 L 对这些成本逐一进行了测量。两种路线的数值误差至少都比统计噪声低八个数量级，因此，在此应用中，混合的优势在于其保证的稳健性和闭式表达，而非数值精度。

5.5 多个中介变量：成本不变的组合

闭式表达可以组合使用。嵌套多重中介反事实的识别需要满足超出第 5 [34], 节单中介假设集的条件；本小节的结论仅涉及计算，并且仅在嵌套均值可识别为所述高斯泛函时成立。考虑两个序贯中介变量，其结局模型为 logit $P(Y=1\mid z,\ m_1,\ m_2)=\theta_0+\theta_1 z+\theta_2 m_1+\theta_3 m_2$。由于每个潜在结局中 z 保持固定，因此允许存在暴露-中介交互；正态中介模型为 $M_1\mid z''\sim N(\gamma_{10}+\gamma_{11}z'',\ \sigma_1^2)$ 和 $M_2\mid z',\ m_1\sim N(\gamma_{20}+\gamma_{21}z'+\gamma_{22}m_1,\ \sigma_2^2)$。嵌套的二维潜在结果均值可精确化简为，

$$\mathbb{E}\big[Y\big(z,\,M_1(z''),\,M_2(z',M_1(z''))\big)\big]=I(\eta_{\mathrm{tot}},\,v_{\mathrm{tot}}),\tag{7}$$

其中，$c_1=\theta_2+\theta_3\gamma_{22}$ 是 m_1 在 logit 尺度上的总系数，$\eta_{\mathrm{tot}}=\theta_0+\theta_1 z+\theta_3(\gamma_{20}+\gamma_{21}z')+c_1(\gamma_{10}+\gamma_{11}z'')$，以及 $v_{\mathrm{tot}}=\theta_3^2\sigma_2^2+c_1^2\sigma_1^2$。其论证采用热半群，即 Pirjol [6] 所发展的对式 (2) 的解读。作为 (η,v) 的函数，积分 $I(\eta,v)=(\Lambda*\phi_v)(\eta)$（其中 ϕ_v是 $N(0,v)$ 的稠密性），就是 Λ 在时刻 v的高斯平滑。因此，再对 η 进行一次高斯积分，就会推进时间变量，而方差沿路径相加。通过归纳法，可将式 (7) 扩展至 k 个序贯中介变量，且总共只需进行一次经认证的求值。加入中介变量之间的交互作用 $\theta_{23}m_1m_2$ 后，内层方差 $(\theta_3+\theta_{23}m_1)^2\sigma_2^2$ 会依赖于 m_1。此时时间参数不再在整个卷积中保持恒定，因此化简后需要对经认证的一维核心进行 $(k-1)$ 维求积。对于通过 $a+b^\top M$进入结局模型、且服从联合正态分布 $M=(M_1,\dots,M_k)^\top\sim N(\mu,\Sigma)$ 的 k 个并行中介变量，由于线性组合仍服从正态分布，故有 $\eta=a+b^\top\mu$ 和 $v=b^\top\Sigma\,b$，因此经认证的求值成本不依赖于 k。在每一层中，一致界在组合运算下均得以保持，因为绝对值不超过 $\Delta_K^\star$ 的误差，其期望本身也不超过 $\Delta_K^\star$；而本小节中的每个等式都是固定闭式 (3) 的恒等式。

6 结论

我们证明了逻辑斯蒂 sigmoid 函数的正态尺度混合近似在每个内部极小值点处都满足 Monahan–Stefanski 等振荡猜想，这是基于

混合切空间的 Haar 性质。随后，一个初等的零点计数差分论证将每个经认证的等振荡极小值提升为唯一的全局极小极大值。对于 $K \le 11$，内点达到（定理 2.4）及由此得出的猜想均无条件成立。我们认证了 $K = 9$，10，11 的极大极小常数，得到最坏情况误差 1.26×10^{-11}。在 $K = 11$ 之后，我们认证了上确界误差严格小于 $\Delta_{11}^{\star}$、取值为 $K = 14$，15 的混合，证明极大极小误差在极限下趋于零（定理4.2），并认证了在 $K = 22$ 处严格有符号30点等振荡。对于逻辑斯蒂-正态积分，这带来了具有无参数误差界的闭式求值；在因果中介分析中，它以毫秒级的耗时取代了有偏的泰勒展开以及随机积分或基于网格的积分。该求值可复合，因此 k 个串联或并联的正态中介变量可归结为对同一闭式的一次经认证的求值（第 5.5 节）。一项历经三十四年的猜想如今已在经认证的范围内成为定理，而它所涉及的积分也有了经认证的闭式，其误差约为双精度单位舍入误差的三十三倍。仍有四个问题悬而未决。第一个是证明或反驳 $K = 12$，13 平台（定理 4.1）。第二个是在 $K > 11$ 处给出严格证书，这似乎需要全局或拓扑存在性工具，而非局部求解器。第三个是对松弛的伯恩斯坦–阿赫泽尔证书前沿作出严格刻画；这是经典逼近理论中的问题，超出了本文的范围。第四个是一般的K 切向 Wronskian 恒等式（定理 2.8）。

声明

资助 本文获得新加坡国家医学研究理事会资助，资助编号为 MOH-SEEDFD24jul-0001。

利益冲突 作者声明不存在与本文内容相关的利益冲突。

数据可用性 本文未收集新数据，所有计算均使用第 5 节所述的合成数据。

代码可用性 经认证的常数和复现脚本可在 `https://github.com/li-ruijie/LogisNormIntegral` 获取，永久 DOI 将在最终录用前注册。

作者贡献 R.L. 是本文唯一作者，负责理论分析、算法设计、实现、数值实验和稿件撰写。

附录 A 预备设定：族、切空间与边界值

本附录规定了后续每个证明都会使用的记号、切空间以及误差的边界行为。固定 $K \ge 1$。 K项正态尺度混合族为

$$\mathcal{F}_K = \Big\{ f(\,\cdot\,; w, s) = \sum_{i=1}^{K} w_i \, \Phi(s_i \,\cdot\,) \;:\; w_i > 0,\ \textstyle\sum_{i=1}^{K} w_i = 1, \\ 0 < s_1 < s_2 < \cdots < s_K \Big\}. \tag{A.1}$$

其闭包 $\mathcal{F}_K$ 允许 $w_i \geq 0$ 以及重合或边界尺度 $s_i \in [0,\infty]$。$\mathcal{F}_K$ 的内部点是指所有 $w_i > 0$ 且尺度互不相同、为正且有限的点。约束 $\sum_i w_i = 1$ 去除了一个权重自由度，因此 $\mathcal{F}_K$ 是一个维数为 $2K-1$的光滑流形。

在内部点 $f \in \mathcal{F}_K$，切空间由关于自由参数的 f 的偏导数张成，

$$V = T_f\mathcal{F}_K = \operatorname{span}\{\, \Phi(s_1 z), \dots, \Phi(s_K z),\ z\,\phi(s_1 z), \dots, z\,\phi(s_K z) \,\}, \tag{A.2}$$

其中 $\partial f/\partial w_i = \Phi(s_i z)$ 且 $\partial f/\partial s_i = w_i\, z\,\phi(s_i z)$。单纯形约束 $\sum_i w_i = 1$ 对 Φ 系数施加了唯一的线性关系 $\sum_i a_i = 0$，因此 $\dim V = 2K-1$。

引理 A.1 (切向量的边界值)。每个 $h \in V$，写作 $h(z) = \sum_{i=1}^{K} a_i\,\Phi(s_i z) + \sum_{i=1}^{K} b_i\, z\,\phi(s_i z)$，其中 $\sum_i a_i = 0$，均满足 $h(0) = 0$，且 $\lim_{z\to\infty} h(z) = 0$。

证明。由于 $\Phi(0) = \frac{1}{2}$，且 $0\cdot\phi(0) = 0$，因此 $h(0) = \frac{1}{2}\sum_i a_i = 0$。由于 $z \to \infty$，$\Phi(s_i z) \to 1$，且 $z\,\phi(s_i z) \to 0$，所以 $\lim_{z\to\infty} h(z) = \sum_i a_i = 0$。 □

这两个边界零点在附录 C 和 F 中发挥作用：它们通过罗尔定理，将 h 的内点零点计数转化为 h' 的零点计数。

附录 B 极小极大问题的存在性与内点达到

仅凭参数紧致性，就能在紧化后的参数集合中找到一个误差不超过 $\Delta_K^\star$ 的点。经认证的严格递减排除了减少计数的极限，经认证的比率界则通过定理 B.3 排除了极限面；从 $K = 1$ 基础情形出发进行归纳，可将极小化子置于 $K \leq 11$ 的开族中。

命题 B.1 (存在性). 对于每个 $K \geq 1$，都存在某个 $f^\star \in F_K$ 满足 上确界$_z\,|\Lambda - f^\star| \leq \Delta_K^\star$。对于$K = 1$，……，11，下确界还可以在 F_K 本身中达到。某个内点 $f^\star$ 满足 上确界$_z\,|\Lambda - f^\star| = \Delta_K^\star$，每个极小化子都位于 F_K 中，而 F_K 中的每个点的一致误差至少为 $\Delta_K^\star$。

证明。令 $(f_n) \subset \mathrm{F}_K$ 为一个极小化序列，$f_n = \sum_i w_i^{(n)}\Phi(s_i^{(n)} z)$，且上确界为$_z\,|\Lambda - f_n| \to \Delta_K^\star$。权重位于紧单纯形中，因此取子序列后 $w_i^{(n)} \to w_i^\star \in [0,1]$ 满足$\sum_i w_i^\star = 1$，且每个尺度序列都位于紧集 $[0,\infty]$中，因此进一步取子序列可得$s_i^{(n)} \to s_i^\star \in [0,\infty]$。参数收敛迫使 $f_n \to f^\star = \sum_i w_i^\star \Phi(s_i^\star z) \in \mathrm{F}_K$ 在每个 $z \neq 0$ 处逐点收敛，约定为 $\Phi(0\cdot z) = \frac{1}{2}$ 以及 $\Phi(\infty\cdot z) = \mathbf{1}\{z > 0\}$。除了阶跃函数极限的单个点 $z = 0$ 外，每个求和项关于其参数均连续。因此，对每个 $z \neq 0$，有 $|\Lambda(z) - f^\star(z)| = \lim_n |\Lambda(z) - f_n(z)| \leq \Delta_K^\star$；在原点处，$|\Lambda(0) - f^\star(0)| = m_\infty/2 = \lim_{z\downarrow 0} |\Lambda(z) - f^\star(z)| \leq \Delta_K^\star$，其中 m_∞是 $s = \infty$处的极限质量。因此上确界为$_z\,|\Lambda - f^\star| \leq \Delta_K^\star$。

这个上确界是否等于 $\Delta_K^\star$ 是另一个问题。原则上，$\mathcal{F}_K$ 可能包含误差严格低于开族上确界的点，因为阶跃函数无法在原点附近由连续混合一致逼近，而只有经认证的常数才能弥合这一差距。对于 $K \leq 11$，通过对 K使用归纳法即可证明这一点。假设$\mathcal{F}_{K-1}$ 中的每个点的误差都至少为 $\Delta_{K-1}^\star$（对于 $K = 1$，此假设显然成立）。减小计数的极限（即消失权重或重合尺度的情形）属于$\mathcal{F}_{K-1}$，因此对于~~$K \geq$~~ 2，根据经认证的严格递减（表 1），其误差至少为 $\Delta_{K-1}^\star > \Delta_K^\star$；而在 $K = 1$ 处不存在这类极限，因为 $\sum_i w_i = 1$ 迫使唯一的权重为 1。

对于 $K = 1$，可直接排除 $s \in \{0,\infty\}$ 处的边界质量，因为这类点只有常数函数 1 2 和单位阶跃函数两种，它们的误差均为 1 2；对于 $K \geq 2$，则可依据定理 B.2 和 B.3 排除，其闭包假设就是第 $K-1$ 层相同的归纳界。因此，$\mathcal{F}_K \setminus \mathcal{F}_K$ 中每个点的误差都严格大于 $\Delta_K^\star$，极限~~—~~$f^\star$在值为 $\Delta_K^\star$ 处属于 $\mathcal{F}_K$，而 $\mathcal{F}_K$ 中每个点的误差至少为 $\Delta_K^\star$，由此推进了归纳过程。 □

等振荡定理针对一个内点极小化子陈述。内点达到取决于极小极大序列的严格单调性。

命题 B.2（严格单调性下的内点达到）。固定 $K \geq 2$，其中 $\Delta_K^\star < \Delta_{K-1}^\star$，并假设 F_{K-1} 中的每个点的一致误差都至少为 $\Delta_{K-1}^\star$。那么，没有任何极小化子 $f^\star$ 会有消失权重或两个重合尺度。此外，如果每个极限面上的极小极大误差（至多叠加 $K-1$ 个真正的高斯分量的常数、单位阶跃或二者）都严格大于 $\Delta_K^\star$，那么每个极小化子都位于内点。对于 $K = 2, \dots, 11$，这三个假设均成立：严格不等式来自经认证的常数（表 *1*），闭包界来自定理 *B.1* 的归纳法，而面条件来自定理 *B.3*。

证明。降计数退化（即权重消失或两个有限正尺度重合）会留下至多含 $K-1$ 个不同高斯分量的混合，因此 $f^\star \in \mathcal{F}_{K-1}$，根据闭包假设和严格递减性，其误差至少为 $\Delta_{K-1}^\star > \Delta_K^\star$。这样的点无法达到最小值，由此证明第一个结论。

极限退化会将质量置于尺度 $s = 0$ 上，此时 $\Phi(0 \cdot z) \equiv \frac{1}{2}$ 贡献一个常数；也会置于尺度 $s = \infty$ 上，此时 $\Phi(\infty \cdot z) = \mathbf{1}\{z > 0\}$ 在原点贡献一个单位阶跃。令 m_0 和 m_∞ 分别表示这两个尺度上的质量。当 $z \to +\infty$ 有限原子饱和至其权重时，$s = 0$ 原子达到其权重的一半，而 $s = \infty$ 原子达到其权重，从而得到 $f^\star(+\infty) = 1 - m_0/2$ 与 $\Lambda(+\infty) = 1$ 相差一个尾部误差 $m_0/2$。在 $z = 0$ 处，阶跃原子跨越连续值 $\Lambda(0) = \frac{1}{2}$ 跳跃 m_∞，因此误差至少为 $m_\infty/2$。于是 $\sup_z |\Lambda - f^\star| \geq \frac{1}{2}\max(m_0, m_\infty)$，所以任何极限面极小化子所携带的质量都趋于零，至多只剩下 $K-1$ 个真正的高斯分量，其上方的质量为 $\max(m_0, m_\infty) \leq 2\Delta_K^\star$。每个极限面的值都大于 $\Delta_K^\star$ 这一假设排除了这种残差，因此每个极小化子都位于内部。 □

引理 B.3（极限面的误差高于内部值）。固定 $K \geq 2$，其中 $\Delta_K^\star / \Delta_{K-1}^\star < (1 - 4\Delta_K^\star)/2$，并假设 F_{K-1} 中的每个点的一致误差至少为 $\Delta_{K-1}^\star$。那么，F_K 的每个极限面上的极小极大误差都严格大于 $\Delta_K^\star$。

证明。设 f 是满足 $\|\Lambda - f\| \leq \Delta_K^\star$ 的任一极限面上的点，并以 m_0, m_∞ 表示其在 $s = 0$ 和 $s = \infty$ 处的质量。尾部界 $\sup_z |\Lambda - f| \geq \frac{1}{2}\max(m_0, m_\infty)$ 给出 $\max(m_0, m_\infty) \leq 2\Delta_K^\star$，因此极限质量 $\mu = m_0 + m_\infty \leq 4\Delta_K^\star$。将至多 $K-1$ 个真正的高斯分量重新归一化，得到 $g = (f - \frac{1}{2}m_0 - m_\infty \mathbf{1}\{z > 0\})/(1-\mu) \in F_{K-1}$，故根据闭包假设有 $\|\Lambda - g\| \geq \Delta_{K-1}^\star$。令 $h = \frac{1}{2}m_0 + m_\infty \mathbf{1}\{z > 0\} - \mu\Lambda$，逐项检查 $z \leq 0$ 和 $z > 0$ 可得 $\|h\| \leq \frac{1}{2}\max(m_0, m_\infty) \leq \Delta_K^\star$，而恒等式 $(1-\mu)(\Lambda - g) = (\Lambda - f) + h$ yields

$$\|\Lambda - f\| \geq (1-\mu)\,\Delta_{K-1}^\star - \Delta_K^\star \geq (1 - 4\Delta_K^\star)\,\Delta_{K-1}^\star - \Delta_K^\star.$$

该假设 使右端严格大于 $\Delta_K^\star$，这与 $\|\Lambda - f\| \leq \Delta_K^\star$ 矛盾。 □

对于 $K = 2$、……、11，经认证的解严格满足 $\Delta_K^\star < \Delta_{K-1}^\star$，从而符合定理 B.2 的严格递减假设。经认证的比率 $\Delta_K^\star / \Delta_{K-1}^\star$ 在整个 $2 \leq K \leq 11$ 范围内至多为 0。19，而阈值 $(1 - 4\Delta_K^\star)/2$ 大于 0。498（因为 $\Delta_K^\star \leq \Delta_2^\star < 6 \times 10^{-4}$），因此定理 B.3 的比率假设在每个层级都成立。

它与定理 B.2 共享的闭包假设，由定理 B.1 的归纳法加以验证：从 $K = 1$ 基例（常数 1 2 和单位阶跃，两者的误差均为 $\frac{1}{2} > \Delta_1^\star$）开始，逐级推进。在每一级，引理排除了各个面，严格递减排除了减小计数的极限，而闭包界则由此导出，供下一级使用。因此，对于 $K = 2, \dots, 11$，定理 B.2 的全部三个假设均成立；又由于 $K = 1$ 情形已在定理 B.1 的证明中直接解决，故对于 $K = 1, \dots, 11$，极小化子位于内点。关于 $K = 11$ 之后的严格性和比值条件仍未解决，因此下文的等振荡定理以内点达到为前提。

附录 C 切空间的 Haar 性质

零点计数工具是以下关于广义指数和实零点的经典界。

引理 C.1（广义指数和的零点）。设 g 为广义指数和$g(x) = \sum_{i=1}^{m} P_i(x)\, e^{\lambda_i x}$，其中实数 λ_i 互不相同，多项式 $P_i \not\equiv 0$ 的次数为 r_i。那么 g 至多有 $\left(\sum_{i=1}^{m}(r_i+1)\right)-1$ 个实零点（按重数计）。特别地，含有 N 项的普通指数和至多有 $N-1$ 个零点，而重数为二的和 $\sum_{i=1}^{K}(c_i + d_i x)\, e^{\lambda_i x}$ 至多有 $2K-1$个零点。

证明。这些界是经典结果 [16–18]。对总维数 $M = \sum_i (r_i+1)$ 作归纳。若 $M = 1$，则 $g = c\,e^{\lambda x}$ 处处非零。对于 $M > 1$，乘以 $e^{-\lambda_1 x}$ 不会改变零点，因此令$\lambda_1 = 0$ 并写作 $g = P_1 + \sum_{i\ge 2} P_i\, e^{\lambda_i x}$。若 $m = 1$，则 $g = P_1$ 是次数为 $r_1 = M-1$的多项式，界显然成立，因此假设 $m \ge 2$。求导 r_1+1 次会消去 P_1，而其余各项都变为 $\tilde{P}_i\, e^{\lambda_i x}$，其次数为 $\tilde{P}_i = r_i$。因此 $g^{(r_1+1)}$ 是同一形式的和，总维数为 $M-(r_1+1)$，故由归纳法可知，它至多有 $M-(r_1+1)-1$ 个零点。应用罗尔定理 r_1+1 次，可得 g至多有 $M-1$ 个零点。 □

定理 C.2（切空间的 Haar 性质）。 V 的切空间（*A.2*）是定义在$(0,\infty)$上的一个维数为 $2K-1$ 的切比雪夫（*Haar*）系统，也就是说，每个非零的 $h \in V$ 在$(0,\infty)$内至多有 $2K-2$ 个互异零点。

证明。设 $h \in V$ 非零，按定理 A.1 的形式写为 $h(0) = 0$，且 $h(z) \to 0$，其中 $z \to \infty$。反设 h 在 $(0,\infty)$ 中有 $N \ge 2K-1$ 个互异零点。再加上边界零点 0 和 ∞，由罗尔定理可得 $\#\{z \in (0,\infty) : h'(z) = 0\} \ge N+1$。求导并利用 $\phi'(t) = -t\,\phi(t)$，

$$h'(z) = \sum_{i=1}^{K} (a_i s_i + b_i)\,\phi(s_i z) - \sum_{i=1}^{K} b_i s_i^2\, z^2\, \phi(s_i z).$$

代入 $\phi(s_i z) = (2\pi)^{-1/2} e^{-s_i^2 z^2/2}$，令 $x = z^2$ 和 $\lambda_i = s_i^2/2 > 0$（均互不相同），则 h' 在 $(0,\ \infty)$ 上的零点与在 $(0,\ \infty)$ 上的零点一一对应。

$$g(x) = \sum_{i=1}^{K} c_i\, e^{-\lambda_i x} + \sum_{i=1}^{K} d_i\, x\, e^{-\lambda_i x}$$

对于依赖于 (a_i, b_i) 的标量 c_i, d_i。这些标量并非全为零，因为映射 $(a,b) \mapsto (c,d)$ 对正尺度可逆，因此 $h \not\equiv 0$ 迫使 $g \not\equiv 0$。根据定理 C.1 的重数为二的情形，这种形式的每个非零 g至多有 $2K-1$ 个零点。等价地，族 $\{e^{-\lambda_i x},\, x\,e^{-\lambda_i x}\}_{i=1}^{K}$是 $[0,\infty)$ [19, 20]上一个$2K$维的扩展完全 Chebyshev 系统。因此 $N+1 \le 2K-1$，也就是说 $N \le 2K-2$，与 $N \ge 2K-1$矛盾。 □

附录 D 下界：至少 $2K$ 交替次数

柯尔莫戈洛夫变分判据结合附录 C 中证明的 Haar 性质，给出了内部极小化子必须具有的交替次数下界。

推论 D.1（等振荡，下界）。设 $f^\star \in F_K$ 在内点处达到极小极大 $\Delta_K^\star$（定理 *B.1* 和 *B.2*）。那么误差 $e^\star = \Lambda - f^\star$ 在（$0,\ \infty$）内至少有 $2K$ 个交替极值，也就是说，存在 $0 < z_1 < \cdots < z_{2K}$ 满足 $|e^\star(z_j)| = \Delta_K^\star$，且$\operatorname{sgn} e^\star(z_j) = (-1)^{j+1} \operatorname{sgn} e^\star(z_1)$。

证明。 首先注意到 $\Delta_K^\star > 0$。误差若为零，则 Λ 会成为有限正态尺度混合，这是不可能的，因为 $1-\Lambda(z)$以 e^{-z} 的速率衰减，而任何混合的尾部都以其最小尺度对应的高斯速率衰减。由于 Λ 和高斯函数均为实解析函数，且 $e^\star \to 0$ 在 $+\infty$处，因此极值集合 $E = \{z \in (0,\ \infty) :\ |e^\star(z)| = \Delta_K^\star\}$ 是有限的。将这些点排序，并按 $\operatorname{sgn}\ e^\star$恒定划分为极大的连续段，令 m 为连续段的数量。假设 $m \le 2K-1$，并在每对相邻且符号相反的连续段之间选取一个分隔节点，从而得到 $t_1 < \cdots < t_{m-1}$，这些节点位于满足 $|e^\star| < \Delta_K^\star$的间隙中。根据定理 C.2，$V$是维数为 $2K-1 \ge m$的 Haar 空间，因此根据标准切比雪夫构造，其中存在一个 h，其唯一的符号变化发生在 $m-1 \le 2K-2$ 个节点处，且在每个连续段 $[21, 22]$上均有 $\operatorname{sgn}\ h \equiv \operatorname{sgn}\ e^\star$。于是，在整个 E上都有 $e^\star\, h > 0$。

切向方向可以通过参数变动实现，这正是内点性发挥作用之处。写作$h(z) = \sum_i a_i \Phi(s_i^\star z) + \sum_i b_i\, z\, \phi(s_i^\star z)$，其中 $\sum_i a_i = 0$，曲线 $f_\varepsilon = f(\cdot;\ w_i^\star + \varepsilon a_i,\quad s_i^\star + \varepsilon b_i / w_i^\star)$ 对所有充分小的 $\varepsilon \ge 0$ 都保持在$\mathcal{F}_K$ 中，因为权重始终为正且总和为 1，尺度也保持有序，正如族方程 (A.1) 所要求的那样。该曲线满足 $f_\varepsilon = f^\star + \varepsilon h + R_\varepsilon$，且 $\sup_{z\ge 0} |R_\varepsilon(z)| = O(\varepsilon^2)$。由于混合的每个二阶参数导数在 $[0,\infty)$ 上都有界，余项关于 z 是一致的。

取 U的一个闭邻域，其中 $e^\star h \ge c > 0$。在 U上，$|e^\star - \varepsilon h|^2 \le (\Delta_K^\star)^2 - 2\varepsilon c + O(\varepsilon^2)$，因此当 $\varepsilon > 0$ 充分小时，$|\Lambda - f_\varepsilon| \le \Delta_K^\star - \varepsilon c/\Delta_K^\star + O(\varepsilon^2) < \Delta_K^\star$。在 U之外，连续函数 $|e^\star|$ 对于某个 $\delta_0 > 0$仍严格低于$\Delta_K^\star - \delta_0$，因为其极值集合位于 U 内，且 $e^\star$ 在无穷远处趋于零；而 $O(\varepsilon)$ 的变动保持了这一间隙。因此，对所有充分小的 $\varepsilon > 0$，都有 $\sup_z |\Lambda - f_\varepsilon| < \Delta_K^\star$，这与极小极大性矛盾。这种参数空间扰动就是切向 Haar 流形 $[21, 22]$, 的柯尔莫哥洛夫判据的显式推导。因此 $m \ge 2K$，且从每一段中取一个极值点，便得到$0 < z_1 < \cdots < z_{2K}$，其幅值为 $\Delta_K^\star$，符号严格交替。 □

附录 E 上界：至多 $2K$ 个极值点，基于全正性

本节通过计算微分误差作为混合拉普拉斯变换时的零点，直接证明了匹配的 $2K$ 个极值点上界。证明不依赖权重的正性或最优性，因此该界适用于任意带符号混合。

定理 E.1 (等振荡、上界与精确计数)。 设 $f(z) = \sum_{i=1}^K w_i \Phi(s_i z)$ 具有实权重 w_i 和不同的正尺度 s_i。则 $e = \Lambda - f$ 在$(0,\infty)$内至多有 $2K$ 个局部极值。特别地，在定理 *D.1* 的假设下，极小极大误差 $e^\star$ 至多有 $2K$ 个内部极值点；因此结合定理 *D.1* 可知，它恰好在 $(0,\infty)$的 $2K$ 个点处等振荡，并且（根据附录 *G*）在 $\mathbb{R}$的 $4K$ 个点处等振荡。此计数解决了正态尺度混合族在每个内部极小化子处的 *Monahan–Stefanski* 猜想，因此对于 $K \le 11$，该猜想无条件成立（定理 *B.1*）。

证明。 根据逻辑斯蒂 sigmoid 函数的精确正态尺度混合表示 $[1, 2]$ $[3,$ 式 18.4.6，该表示以尺度 $\sigma = 1/s]$ 表示，

$$\Lambda(z) = \int_0^\infty \Phi(sz)\, q(s)\, ds, \qquad q > 0, \quad \int_0^\infty q(s)\, ds = 1, \tag{E.1}$$

其中 q 为柯尔莫哥洛夫–斯米尔诺夫混合密度（文献中称为 Stefanski 混合密度）。界 $\int_0^\infty s\, q(s)\, ds < \infty$ 使积分号下求导合理，并且对于 $z > 0$,

$$e'(z) = \int_0^\infty s\, \phi(sz)\, q(s)\, ds - \sum_{i=1}^K w_i s_i\, \phi(s_i z). \tag{E.2}$$

在 $u = z^2/2$ 和 $\lambda = s^2$ 下，高斯密度变为拉普拉斯核，$\phi(sz) = (2\pi)^{-1/2} e^{-\lambda u}$，而 $\sqrt{2\pi}\, e'$ 则变为混合拉普拉斯变换

$$F(u) = \int_0^\infty \rho(\lambda)\, e^{-\lambda u}\, d\lambda - \sum_{i=1}^{K} c_i\, e^{-\lambda_i u}, \qquad \rho(\lambda) = \tfrac{1}{2}\, q(\sqrt{\lambda}\,), \quad c_i = w_i s_i, \tag{E.3}$$

其中 $\lambda_i = s_i^2$ 和 $\rho > 0$。由于指数因子能够吸收这些项，ρ 关于每个 $e^{-\lambda u}$, $u > 0$的多项式矩均有限。根据下文的定理 E.2，无论 c_i的符号如何，F 在 $(0,\ \infty)$ 中至多有 $2K$ 个零点。这个界是符号正则拉普拉斯核 [19, （第 I 章和第 V 章）的变差递减性质的零点计数形式。由于 $u = z^2/2$ 是 $(0,\ \infty)$ 到自身的双射，e' 在那里至多有 $2K$ 个零点，而 e 的每个内部局部极值都是 e'的一个零点，因此 e 在 $(0,\ \infty)$ 中至多有 $2K$ 个局部极值。

在内部极小值点处，极值集合 $\{z > 0 :\ |e^\star(z)| = \Delta_K^\star\}$ 中的每一点都是 $e^\star$的全局极值点，因而也是局部极值点，所以该集合至多包含 $2K$ 个点。结合定理 D.1 的下界，可知点数恰为 $2K$。 □

注记 E.2（双重剥离零点计数）。 混合拉普拉斯变换的界（E.3）是定理 C.1 中罗尔定理机制从有限和推广到密度加原子变换后的结果。设$F(u) = \int_0^\infty \rho(\lambda)\ e^{-\lambda u}\, d\lambda + \sum_{i=1}^{K} c_i\, e^{-\lambda_i u}$ 具有几乎处处为正的密度，以及 K 个符号任意的原子。乘以 $e^{\lambda_1 u}$ 不会改变零点；根据罗尔定理，求导至多损失一个零点，同时将$(\rho(\lambda)\ c_i)$替换为$(\rho(\lambda)(\lambda_1 - \lambda)\ \ c_i(\lambda_1 - \lambda_i))$。因此，第一个原子被消去，但代价是密度因子现在在 λ_1处变号。再次进行相同的剥离，使密度再次乘以$(\lambda_1 - \lambda)$，从而恢复几乎处处为正的密度$\rho(\lambda)(\lambda_1 - \lambda)^2$，并使第一个原子继续保持消去状态。两次罗尔步骤因此消去一个原子，所以 K个原子需要 $2K$ 个零点；而无原子的基例是一个严格为正且没有零点的积分。因此，F 在$(0,\ \infty)$中至多有 $2K$ 个零点。

附录 x F 唯一性与切向 Wronskian 恒等式

唯一性证明是初等且自足的，只需将附录 C 的零点计数工具应用于两个竞争混合模型之差即可。按照经典术语，由于每个内点处的满维 Haar 切向量使该族成为次数为 $2K - 1$ 的变可解族，也可由变差可解性得出相同结论。对于变可解族，柯尔莫哥洛夫判据是充要条件，且最佳逼近唯一 [22, 定理 7-4], [20, III.4.6]。下文的直接论证使这些附录不必涉及这套理论。

引理 F.1 (差函数零点界)。 设 $f, g \in F_K$ 互不相同。则 $d = f - g$ 不恒等于零，且在$(0, \infty)$内至多有 $2K - 2$ 个零点，按重数计。

证明。 合并两个参数列表中的重合尺度，$d(z) = \sum_j c_j\, \Phi(\sigma_j z)$ 是由 $\sum_j c_j = 1 - 1 = 0$ 个 $m \le 2K$ 不同的 Φ-原子组成的组合。并非每个 c_j 都会消失，因为若所有系数都消失，就会迫使两个混合共享尺度和权重。因此 $d \not\equiv 0$；对于 $d \equiv 0$，这将迫使 $d' \equiv 0$；而 $d'(z) = \sum_j c_j \sigma_j\, \phi(\sigma_j z)$ 在 $x = z^2$ 下映射为一个非平凡的普通指数和，根据定理 C.1，该指数和只有有限个零点。如定理 A.1 中所述，$d(0) = 0$，且 $d(z) \to 0$ 作为 $z \to \infty$。如果 d 在 $(0, \infty)$ 中按重数计有 $2K - 1$ 个零点，那么对这些零点及两个边界零点应用罗尔定理，将得到 d' 在 $(0, \infty)$ 中按重数计至少有 $2K$ 个零点。 d' 的指数和像含有 $m \le 2K$ 项，因此根据定理 C.1，按重数计至多有 $m - 1 \le 2K - 1$ 个零点。 □

命题 F.2（德拉瓦莱·普桑界及其等号情形）。设 $f \in F_K$ 在 $2K$ 个点 $0 < z_1 < \cdots < z_{2K}$ 处具有交替的误差$e_f = \Lambda - f$，即存在某些 $\delta > 0$ 和 $\epsilon = \pm 1$，使得对于每个 j，$(-1)^j \epsilon\, e_f(z_j) \geq \delta$。则 $\Delta_K^\star \geq \delta$，且每个满足 $\sup_z |\Lambda - g| \leq \delta$ 等于 f 的 $g \in F_K$ 都是如此。

证明。设 $g \in \mathcal{F}_K$ 满足 $\sup_z |\Lambda - g| \leq \delta$，并假设 $g \neq f$。在每个交错点处，$d(z_j) = (f-g)(z_j) = e_g(z_j) - e_f(z_j)$，其中 $|e_g(z_j)| \leq \delta \leq (-1)^j \epsilon\, e_f(z_j)$，因此 $(-1)^j \epsilon\, d(z_j) \leq 0$，且 d 的符号相对于 $2K$ 点弱交替。一个不恒为零、解析且其符号相对于 $2K$ 个有序点弱交替的函数，在 $[z_1, z_{2K}]$中至少有 $2K-1$ 个零点（按重数计）。每一对相邻点都会贡献一个变号零点或一个端点零点。若相邻两对都将同一个端点计作零点，那么该零点要么发生变号，此时两个相邻开区间之一还包含另一个零点；要么不变号，此时其重数为偶数，计为两个零点。这一计数与定理 F.1 的界 $2K-2$ 矛盾，因此 $g = f$。因此 $\Delta_K^\star \geq \delta$，因为一个满足 $\sup_z |\Lambda - g| < \delta$ 的混合 g，根据上述结论将等于 f，而 f 的误差在模意义下达到 δ。 □

推论 F.3 (全局唯一性)。对于每个 K，只要极小极大在内点取得（特别是 $K = 1, \ldots, 11$，根据定理 *B.1*），$\Delta_K^\star$在 F_K中的极小化子便是唯一的；任何误差在 $2K$ 个点处以 $\Delta_K^\star$ 为水平交替的内点混合，都是该极小化子。

证明。一个内部极小值点 $f^\star$ 在 $2K$ 个点处以 $\Delta_K^\star$ 为水平线呈现等波动（定理 D.1）。任何其他极小化子 g都具有上确界$_z |\Lambda - g| = \Delta_K^\star$，因此，根据定理 F.2 的等号情形，结合 $\delta = \Delta_K^\star$，可得 $g = f^\star$。□

推论 F.4 (严格局部唯一性)。在定理 *F.3* 的假设下，极小化子$(w^\star, s^\star)$ 是严格的，即 F_K 中的任何其他混合的一致误差都严格高于 $\Delta_K^\star$。特别地，它是极小极大目标函数的一个严格且局部孤立的最小值，这正是经典雷梅兹-雅可比路线通过隐函数定理 *[21, 22]*所得到的结论。

证明。根据定理 F.3，最小值仅在 $f^\star$ 处取得，因此每个 $g \neq f^\star$ 都有 $\sup_z |\Lambda - g| > \Delta_K^\star$。 □

严格局部唯一性的另一种显式证书，不依赖上述等振荡论证，且在有限 K下可用，是关于合流切向 Wronskian 的一个结构恒等式。令 $Q(z; s)$ 表示基函数组$\{\Phi(s_i z),\ z\,\phi(s_i z)\}_{i=1}^{K}$ 删去 0 阶导数行和最大索引列 $z\,\phi(s_K z)$ 后所得的约化（$2K-1$）$\times$ ($2K-1$)切向 Wronskian 余子式。各行按导数阶数从小到大排列，即 $1, \ldots, 2K-1$；保留的列则先排列为 Φ 块 $\Phi(s_1 z), \ldots, \Phi(s_K z)$，再排列为 $z\,\phi$ 块 $z\,\phi(s_1 z), \ldots, z\,\phi(s_{K-1} z)$。下文闭式的符号取决于这一排列。对行或列进行置换只会使 Q 整体符号改变，因此非零结论与排列无关。

定理 F.5 (结构性纯单项式恒等式，$K \leq 5$)。对于 $2 \leq K \leq 5$ 和有序、互异的正数 $s_1 < s_2 < \cdots < s_K$，约化切向朗斯基行列式余子式可分解为 $Q(z;\ s) = c_K(s)\, z^{(2K-1)(K-1)}$，其中 $c_K(s)$ 是由 *2* 的幂、尺度的幂以及平方尺度的两两差的偶数次幂构成的带符号乘积，在有序正锥上非零，因此 $Q(z;\ s) \neq 0$ 在 $(0, \infty)$ 上成立。

证明。对于每个 $K \in \{2、3、4、5\}$，约化的 $(2K-1) \times (2K-1)$ 切向矩阵的元素均属于$\mathbb{Q}[z, s_1, \ldots, s_K]$。采用 Bareiss 无分式消元 [23] ($K \leq 4$)，并通过精确整数算术进行分阶段后缀子式展开，再在多个点和素数上进行模交叉校验（$K = 5$），得到

行列式的精确多元多项式。该多项式可因式分解为 $c_K(s)\ z^{(2K-1)(K-1)}$，其中

$$c_K(s) = (-1)^{\lfloor (K-1)/2 \rfloor}\, 2^{K-1}\, s_K \prod_{i=1}^{K-1} s_i^3 \prod_{1 \le i < j \le K-1} (s_i^2 - s_j^2)^4 \prod_{i=1}^{K-1} (s_i^2 - s_K^2)^2 .$$

对于按序排列且互不相同的正数 s，每个因子均非零，因此 $c_K(s) \neq 0$ 和 $Q(z;\, s) \neq 0$ 在 $(0, \infty)$ 上成立。 □

$c_K(s)$ 中的指数符合节点 $\lambda_i = s_i^2$ 的合流范德蒙德计数。每个尺度 $i < K$ 对应一个大小为二的块（其两个切向方向 $\Phi(s_i z)$ 和 $z\,\phi(s_i z)$），被删去列所对应的尺度 s_K 则对应一个大小为一的块。每对节点出现时的指数等于其块大小的乘积：在前 $K-1$ 个尺度内为 $2 \cdot 2 = 4$，与 s_K 相比则为 $2 \cdot 1 = 2$。猜想中的一般-K 形式由此计数外推而来。其尚待证明的内容是 z 中的纯单项式因式分解。

猜想 F.6 （一般 K下的纯单项式恒等式）。 **</style}定理 F.5 的因式分解成立；对于每个，** （s） 的闭式都相同。

对于 $6 \le K \le 12$，该公式在整数尺度 $s = (1,\ 2,\ \ldots\ldots,\ K)$ 上逐点验证成立。Bareiss 行列式与该公式求得相同的整数；在 $K = 8$ 处，该整数达到 140 位，在 $K = 12$ 处达到 412 位。一般-K 符号证明尚未解决，但此处并不需要。上述唯一性链仅依赖附录 C 的零点计数工具；此外，在每个经过认证的$s^\star$ 处，还通过对 Q 在 $\mathbb{Q}$ 上进行 Sturm 根计数，直接确认了 Q 不为零。

附录 G 对称性：从 $(0,\ \infty)$ 到实数集 $\mathbb{R}$

目前为止得到的计数是在 $(0,\ \infty)$ 上的。误差的奇性将这些计数推广到整条实数轴上，并使每个计数加倍。

引理 G.1 (极小极大误差是奇函数)。 对于任意 $f \in F_K$，误差 $e = \Lambda - f$ 满足 $e(-z) = -e(z)$ 对所有 $z \in \mathbb{R}$均成立。

证明。 逻辑斯蒂 sigmoid 函数满足 $\Lambda(-z) = 1 - \Lambda(z)$，且 $\Phi(-u) = 1 - \Phi(u)$，因此 $\Phi(-s_i z) = 1 - \Phi(s_i z)$。利用 $\sum_i w_i = 1$，$f(-z) = \sum_i w_i (1 - \Phi(s_i z)) = 1 - f(z)$。因此 $e(-z) = (1 - \Lambda(z)) - (1 - f(z)) = -e(z)$。 □

$e^\star$ 的奇性迫使 $e^\star(0) = 0$，且极值点成镜像对出现。定理 E.1 在 $(0,\ \infty)$ 上的 $2K$ 个等振荡点映射为 $(-\infty,\ 0)$ 上的 $2K$ 个点，因此扩展误差在 $\mathbb{R}$上恰有 $4K$ 个等振荡点，且原点处没有极值点。

附录 H 一种替代的（受阻）全正性路径

一种经典的替代方法，是直接对累积核 $\Phi(sz)$应用全正性理论，从而读出等振荡计数，避免附录 E 中的求导。本节说明了这条路径受阻的原因。累积核是对数凹的，但并非全正。

引理 H.1 (正态分布函数的对数凹性)。 **</style}映射 对数** $\Phi(z)$ 在 $\mathbb{R}$ 上严格凹，且逆风险率 $z \mapsto \phi(z)/\Phi(z)$ 严格递减。

证明。 由 $\phi' = -z\,\phi$，$(\log \Phi)''(z) = -\phi(z)[z\,\Phi(z) + \phi(z)]/\Phi(z)^2$, 因此其符号与$-[z\,\Phi(z) + \phi(z)]$相同。令 $g(z) = z\,\Phi(z) + \phi(z)$。则 $g'(z) = \Phi(z) > 0$ 且 $g(z) \to 0$，因为 $z \to -\infty$，所以$g > 0$，从而得到 $(\log \Phi)'' < 0$。同一不等式也使 $(\phi/\Phi)' = -g\,\phi/\Phi^2 < 0$成立。 □

命题 H.2 (累积核并非全正)。 核 $\Phi(sz)$ 在 $(0,\infty)\times\mathbb{R}$上不是 *2* 阶全正的。对于 $0<s_1<s_2$，子式 $D(z_1,z_2)=\Phi(s_1z_1)\,\Phi(s_2z_2)-\Phi(s_1z_2)\,\Phi(s_2z_1)$ 在某些有序自变量处严格为负，而在另一些有序自变量处严格为正。

证明。 当 $(s_1,s_2)=(1,2)$ 时，在 $(z_1,z_2)=(1,2)$ 处， $D=\Phi(1)\Phi(4)-\Phi(2)^2=0.84132-0.95502=-0.11370<0$；而在 $(z_1,z_2)=(-2,-1)$ 处， $D=\Phi(-2)^2-\Phi(-1)\Phi(-4)=0.000518-0.000005=+0.000513>0$。二阶全正核的每个 2×2 子式都非负，因此 $\Phi(sz)$ 不是全正的。

□

稠密性核不存在这一障碍。在 $u=z^2/2$, $\lambda=s^2$ 下，高斯密度 $\phi(sz)$ 变为拉普拉斯核 $e^{-\lambda u}$，它在第 [19, 章 I] 中对所有阶数都是符号正则的。因此，附录 E 中的变差递减论证可通过微分误差 (E.2) 应用于这一稠密性核，并由此得到上界。

附录 I 证明依赖图

表 I.1 列出了主要结果。预备设定（定理 A.1）提供了定理 C.2 和 F.1 所需的边界值。Haar 性质（定理 C.2）通过零点计数引理（定理 C.1）导出下界（定理 D.1），并通过差函数零点界（定理 F.1）和德拉瓦莱普桑命题（定理 F.2）导出唯一性链（定理 F.3 和 F.4）。切向 Wronskian 恒等式（定理 F.5）通过符号计算独立证明。上界（定理 E.1）独立于哈爾路径。它仅基于尺度混合方程（E.1）以及高斯密度核的双重剥离零点计数（定理 E.2）。下界和上界结合得到在 $(0,\infty)$ 上的精确 $2K$ 计数，随后由对称性引理（定理 G.1）将其扩展到 $4K$ 上的 $\mathbb{R}$。存在性（定理 B.1 和 B.2）保证对于 $K\le11$存在一个内部极小值点。该分支是图中唯一表面上的循环论证。定理 B.1 调用定理 B.2 和 B.3 以排除极限面，而定理 B.2 本身又依赖于定理 B.3。反过来，二者都假设了水平 $K-1$ 处的闭包界，而这正是定理 B.1 在低一个层级所给出的结果。因此，三者通过从 $K=1$ 基点开始对 K 进行归纳法共同得到解决，正如附录 B 所述，整个链条不存在循环论证。

附录J 极限混合密度的闭式矩

K原子极小极大族的自然极限对象是第1节中的柯尔莫哥洛夫–斯米尔诺夫混合密度 q。其低阶矩具有闭式，现记录于此以支持第 4 节 4.3。将尺度混合表示 (E.1) 写作 $Z=NS$，其中 $Z\sim$ Logistic(0,1) 与 $N\sim N(0,1)$ 相互独立，且 $\Lambda(z)=\mathbb{E}[\Phi(z/S)],$，因此 q 是核尺度倒数的分布 $1/S$。逻辑变量 Z 的梅林变换为 $\mathbb{E}[|Z|^s]=2\,\Gamma(s+1)\;\eta_{\mathrm{D}}(s)$，其中 $\eta_{\mathrm{D}}(s)=(1-2^{1-s})\;\zeta(s)$ 是狄利克雷 η 函数。除以半正态矩 $\mathbb{E}[|N|^s]$ 并按奇偶性分类，可得对于 $k\ge1$,

$$\begin{aligned}\mathbb{E}[S^{2k}]&=2^{k+1}\,k!\,\eta_{\mathrm{D}}(2k)=\frac{(4^k-2)\,|B_{2k}|\,\pi^{2k}}{(2k-1)!!},\\ \mathbb{E}[S^{2k+1}]&=(2k+1)!!\,\eta_{\mathrm{D}}(2k+1)\,\sqrt{2\pi},\end{aligned}\tag{J.1}$$

因此，偶数阶矩是通过伯努利数 B_{2k} 得到的 π^{2k} 的有理倍数（例如 $\mathbb{E}[S^2]=\pi^2/3$）。奇数阶矩则包含奇数阶 ζ 值（例如 $\mathbb{E}[S^3]=\frac{9}{4}\sqrt{2\pi}\,\zeta(3)$）。Korolev 和 Shevtsova [48] 用黎曼 ζ 函数和欧拉伽马函数表示了柯尔莫哥洛夫分布所有实数阶的矩，并据此给出了绝对矩

表 I.1：附录主要结果及其依赖关系索引。

结果	陈述	依赖于
定理 A.1	边界值 $h(0) = h(\infty) = 0$	预备设定
定理 B.1	在 $\mathcal{F}_K$ 上达到，对 $1 \le K \le 11$ 误差至少为 $\Delta^\star_K$ 在整个 $\mathcal{F}_K$ 上	紧性，经认证的严格递减，定理 B.3，and 定理 B.2
定理 B.2	内点达到（$2 \le K \le 11$）	严格单调性，定理 B.3，以及定理 B.1 在 $K-1$
定理 B.3	极限面超过 $\Delta^\star_K$	经认证的比率界和定理 B.1 在 $K-1$
定理 C.1	指数和零点计数 $\le 2K-1$	经典
定理 C.2	V Haar 维数为 $2K-1$	定理 A.1 和 C.1
定理 D.1	$\ge 2K$ 交替次数	定理 C.2
定理 E.1	$\le 2K$ 极值点，精确的 $2K/4K$	尺度混合方程 (E.1)，定理 E.2，定理 D.1，and 定理 G.1
定理 E.2	双重剥离零点计数 $\le 2K$ 在几乎处处为正的密度且具有有限多项式矩	罗尔定理
定理 F.1	的差有 $\le 2K-2$ 个零点	定理 A.1 和 C.1
定理 F.2	de la Vallée Poussin 界，等号情形	定理 F.1
定理 F.3	全局唯一性	定理 D.1 和 F.2
定理 F.4	严格局部唯一性	定理 F.3
定理 F.5	$Q(z; s) \ne 0$ $(K \le 5)$	符号的已验证的 $K \le 12$
定理 F.6	对一般 K 的恒等式	
定理 G.1	的误差为奇数，$4K$ 在 $\mathbb{R}$ 上	预备设定
定理 H.1	ϕ/Φ 严格递减	初等的
定理 H.2	$\Phi(sz)$不是全正的	数值反例

逻辑斯蒂分布的阶数为 $p > -1$。由于在表示式 (E.1) 中， S 是柯尔莫哥洛夫变量的两倍，因此矩 (J.1) 及其负阶对应形式是该结果的特例。

核尺度 $1/S$ 通过以下公式反演矩公式 (J.1)： $\mathbb{E}_q[s^k] = \mathbb{E}[S^{-k}]$，其对数尺度分布的均值和方差具有闭式表达式

$$\begin{aligned} \operatorname{mean}_q(\log s) &= \tfrac{1}{2}\bigl(\gamma + \log 2 - 2\log\pi\bigr) = -0.50954\ldots, \\ \operatorname{Var}_q(\log s) &= \frac{\pi^2}{24} - 2(\log 2)^2 - 4\log 2\,\log\pi - 2\,\zeta''(0) - \log^2(\pi/2) = 0.08524\ldots \end{aligned} \tag{J.2}$$

这里， γ 是欧拉–马歇罗尼常数，而 $\zeta''(0)$ 是黎曼ζ函数在原点处的二阶导数。

产生对数尺度统计量 (J.2) 的反演，也通过将梅林关系 $\mathbb{E}[|Z|^s] = 2\,\Gamma(s+1)\;\eta_{\mathrm{D}}(s)$ 延拓至负参数，得到了 q的二阶矩。在 $s = -2$ 处，零点 $\eta_{\mathrm{D}}(-2) = 0$ 抵消了 $\Gamma(s+1)$的极点；将剩余的导数项除以延拓后的半正态矩，再通过 $\zeta'(-2) = -\zeta(3)/(4\pi^2)$计算，得到$\eta_{\mathrm{D}}'(-2) = 7\zeta(3)/(4\pi^2)$，从而有

$$\mathbb{E}_q[s^2] = 2\,\eta_{\mathrm{D}}'(-2) = \frac{7\zeta(3)}{2\pi^2} = 0.42627\ldots, \tag{J.3}$$

经认证的原子的二阶矩趋近的常数（第 4.3 节）。随附的 $K = 10$证书将值 (J.3) 精确复现到七位有效数字。

附录　　K 平台的双区间平衡测度解读

第 4 节中猜想的有限平台是当 $\Delta_K^\star = \Delta_{11}^\star$ 且 $K = 12,13$ 时取等号；而在$K = 14, \ldots, 17$ 时，依据放宽后的 Bernstein–Akhiezer 判据认证的混合，其经审计的上确界误差低于 $\Delta_{11}^\star$。这个猜想的平台可以从位势论角度加以解读，将尚未解决的平台问题重新表述为附录 J 中极限混合密度的问题，但并未解决这些问题。将原子数 K 视为逼近次数，将最优对数尺度位置 $\{x_i^\star = \log\, s_i^\star\}$ 的闭包视为极限密度 ρ_∞的支集，并将渐近原子密度视为平衡测度。在单一区间上，该测度服从反正弦律，极小极大误差随逼近次数单调下降。在两个区间上，极值问题存在于一条亏格一超椭圆曲线上，平衡测度在每个区间上都有一个分量；误差可能停滞不变，直到逼近次数足以使交替次数分布到两个分量上 [15, 49, 50]。

因此，平台可解读为在 $K^\star \le K < K^{\star\star}$ 上达到单区间饱和，其中 $K^\star = 11$ 是携带严格证书的最后一个次数。经认证的 $K = 14,15$ 项改进将下降的起点移至 $K^{\star\star} = 14$，而支撑边缘并未指出唯一的分岔次数。已知最佳对数尺度支撑的上边界，对于 $K = 14$ 见证位于 $+0.48$ ，对于 $K = 16$ 证书则位于 $+0.69$ 。 $K = 15$ 见证在 $+0.78$ 处有一个权重为 5.6×10^{-10} 的原子，其余质量的终点为 $+0.48$。随着次数增长，质量是否通过双区间机制迁移越过单区间边缘，并由此终结所猜想的 $K = 12,13$ 平台，目前仍未有定论。

这种对未决问题的解读并非证明。这里需要提醒一下术语的用法：“Bernstein–Akhiezer”既指第 4 节的宽松证书判据，也指经典的双区间平衡测度，而此处仅指后者。

附录L 因果中介分析中的方法比较

本附录完整呈现了第 5.3 节总结的比较，并列出了第 5.4 节所述抽样循环中逐点路线的实测成本。

表 L.1 比较了三种结局发生率区间下、$R = 10^4$ 外层抽样中的四种内积分策略。其标题说明了经认证的混合、内层蒙特卡洛与逐点计算器之间的一致性、泰勒展开随区间而异的偏差以及运行时间，第 5.3 节则给出了结论。其中覆盖率主张所依据的重复抽样实验，使用的是代码仓库中的脚本 `inst/reproduction/Coverage.R`。

逐点计算器对应的行有一项限定。在高发生率下，截断阶数为 10 时，计算器在 10^4 次抽样中的三次返回非有限值；阶数为 30 时，则有四次返回非有限值。表 L.1 对此作了标记，并在计算分位数前剔除了这些抽样。下面的说明将这些失败追溯到计算器定义域的两个边界。

在中介分析流程中，逐点路线是可用的内层引擎；表 L.1 对其进行了测量，而不是想当然地加以假设，并且在两种路线一致的地方，各路线都对另一条路线提供了佐证。但仍有三项成本。第一，必须针对每次抽样管理截断阶数，因此循环范围误差陈述是经验性的，而不是预先固定的经认证常数。第二，德尔塔法区间和敏感性分析所需的导数，必须通过对过程式例程进行数值微分来获得。第三，也是抽样循环特有的问题，是计算器在其定义域的两端都会返回非有限值。下边界位于 $v = 0.0265$ 附近，并且不随截断阶数变化。上边界既取决于内层正态分布的均值，也取决于 v 在两个阶数下的取值。在阶数为 10 时，当均值为零，它位于 $v = 142$ 附近，高于该区间的所有抽样；而在阶数为 30 时，它落在抽样范围内，因此不存在能够描述这些失败的单一 v 阈值。在表 L.1 中，阶数为 10 时，高发生率抽样的 10^4 次中有三次逐点路线未定义；阶数为 30 时则有四次。阶数为 10 的三次失败，以及阶数为 30 的四次失败中的三次，都低于下边界，此时中介变量总系数 $\theta_2 + \theta_3 z$ 接近于零。该区间中抽样到的最小绝对值为 0.075，而另外两个区间分别为 1.10 和 0.76。剩余的一次阶数为 30 的失败对应于该区间抽样方差的最大值 $v = 56.5$，而另外两处的最大值分别为 9.8 和 13.2。因此，提高阶数以改善精度会从上方收窄定义域；与此同时，与阶数无关的下边界位于间接效应接近零的位置，而未加防护的循环会将第一次出现的此类抽样传播到区间中。

参考文献

[1] D. F. Andrews 和 C. L. Mallows. 正态分布的尺度混合. 《皇家统计学会会刊 *B*辑》, 36(1):99–102, 1974. ISSN 0035-9246. doi: 10.1111/j.2517-6161.1974.tb00989.x.[2] Leonard A. Stefanski. 逻辑斯蒂分布的正态尺度混合表示. 《统计与概率快报》, 11(1):69–70, 1991. ISSN 0167-7152. doi: 10.1016/0167-7152(91)90181-P.[3] John F. Monahan 和 Leonard A. Stefanski. F*(z) 的正态尺度混合近似及逻辑斯蒂-正态积分的计算. 载于 N. Balakrishnan 编, 《逻辑斯蒂分布手册》, 第 529–540 页. Marcel Dekker, 纽约州纽约市, 1992. ISBN 978-0-8247-8587-1.[4] Evgeny Ya. Remez. 关于确定给定次数的逼近多项式.《哈尔科夫数学学会通讯》 *Sér. 4*, 10:41–63, 1934.[5] G. A. Watson. 赋范线性空间中的逼近. 《计算与应用数学杂志》, 121(1-2):1–36, 2000. ISSN 0377-0427. doi: 10.1016/S0377-0427(00)00333-2.

表 L.1：自然间接效应（NIE）及其 95% 准贝叶斯（QB）置信区间（CI）在 $R = 10^4$ 次外层抽样（$n = 500$）下，跨三个结局患病率区间的方法比较。极小极大混合、内层蒙特卡洛（MC）以及截断阶数为 10 和 30 的 Pirjol 逐点泊松求和评估器，均能准确计算内层积分，且结果始终一致；在低患病率和中等患病率区间，混合方法与逐点路线的结果在脚本保留的六位小数上完全一致。高患病率区间的汇总值在小数点后第四位有所差异，因为该区间的逐点分位数是基于缩减后的抽样集计算的。Valeri–VanderWeele（V&VW）泰勒展开呈现出随区间而异的偏差，分别为 -13、$+9$和 $+25$ 个概率尺度百分点。匕首符号表示逐点结果中剔除了非有限抽样值：截断阶数为 10 时有 10^4 个，阶数为 30 时有 4 个，且这些情况均出现在高患病率区间。其中六个值低于评估器的下域边界，此处抽得的中介变量总系数最接近零；剩余的一个阶数为 30 的抽样值则超出了该阶数的上边界。可通过 `inst/reproduction/Example1.R` 复现；其扫描还会运行代码库默认方法 $K = 22$，该方法的结果与所展示的 $K = 11$ 个值在所示精度下相符。逐点结果需要上游源代码，这些代码属于第三方且未随文提供；脚本即使没有这些代码也能运行。参数抽样由 Cholesky 因子生成；对于正定协方差矩阵，该因子是唯一的，因此该扫描在不同 BLAS 和 LAPACK 构建版本下均可复现。我们通过分别在 OpenBLAS 和参考 BLAS 下运行扫描验证了这种可复现性，所有 54 条数值记录均完全一致。本文报告的所有墙钟时间均在一台 16 核工作站上测得。

情景	方法	$\widehat{\mathrm{NIE}}_R$	95% QB 置信区间	Bias	墙钟时间
低（$P_Y = 0.38$）	极小极大 $K = 11$	0.4112	(0.3491, 0.4695)	-0.004	≤ 0.01 秒
	V&VW 泰勒展开	0.2827	(0.1973, 0.3473)	-0.133	≤ 0.01 秒
	内部蒙特卡洛，$M = 10^4$	0.4112	(0.3488, 0.4695)	-0.004	4.8 秒
	Pirjol，$N = 10$	0.4112	(0.3491, 0.4695)	-0.004	0.17 秒
	Pirjol，$N = 30$	0.4112	(0.3491, 0.4695)	-0.004	0。43 秒
中等 ($P_Y = 0.69$)	极小极大 $K = 11$	0.3178	(0.2227, 0.4004)	-0.005	≤ 0.01 s
	V&VW 泰勒展开	0.4174	(0.2648, 0.5572)	$+0.095$	≤ 0.01 s
	内部蒙特卡洛，$M = 10^4$	0.3178	(0.2230, 0.4006)	-0.005	4.8 秒
	Pirjol，$N = 10$	0.3178	(0.2227, 0.4004)	-0.005	0.17 秒
	Pirjol，$N = 30$	0.3178	(0.2227, 0.4004)	-0.005	0.43 秒
高（$P_Y = 0.82$）	极小极大 $K = 11$	0.2438	(0.1436, 0.3311)	-0.004	≤ 0.01 秒
	V&VW 泰勒展开	0.4933	(0.1556, 1.0617)	$+0.245$	≤ 0.01 秒
	内部蒙特卡洛，$M = 10^4$	0.2438	(0.1438, 0.3308)	-0.004	4.8 秒
	Pirjol，$N = 10^\dagger$	0.2439	(0.1438, 0.3311)	-0.004	0.17 秒
	Pirjol，$N = 30^\dagger$	0.2439	(0.1437, 0.3311)	-0.004	0.43 秒

[6] Dan Pirjol. 逻辑斯蒂-正态积分及其推广。《计算与应用数学杂志》, 237 (1):460–469, 2013. ISSN 0377-0427. doi: 10.1016/j.cam.2012.06.016.[7] Gernot Paasch 和 Susanne Scheinert. 高斯态密度下有机物的载流子密度：高斯-费米积分的解析近似。*J. Appl. Phys.*, 107(10): 104501, 2010. ISSN 0021-8979. doi: 10.1063/1.3374475.[8] Patricio Farrell、Matteo Patriarca、Jürgen Fuhrmann 和 Thomas Koprucki. 对费米-狄拉克统计和高斯-费米统计下热力学一致的载流子通量离散格式进行比较。*Opt. Quantum Electron.*, 50(2):101, 2018. ISSN 0306-8919. doi: 10.1007/ s11082-018-1349-8.[9] Tural Mehmetoğlu. 高斯态密度下有机材料载流子密度的解析求值。*J. Comput. Electron.*, 13(4):960–964, 2014. ISSN 1569-8025. doi: 10.1007/s10825-014-0615-7.[10] Jerry P. Selvaggi. 重新探讨高斯态密度下有机材料载流子密度的解析求值。*J. Comput. Electron.*, 17(1):61–67, 2018. ISSN 1569-8025. doi: 10.1007/s10825-017-1113-5.[11]Matteo Patriarca、Patricio Farrell、Jürgen Fuhrmann 和 Thomas Koprucki. 用于简并半导体电荷输运的高精度、基于求积的 Scharfetter-Gummel 格式。*Comput. Phys. Commun.*, 235:40–49, 2019. ISSN 0010-4655. doi: 10.1016/j.cpc.2018.10.004.[12] Nikolaus Hansen 和 Andreas Ostermeier. 进化策略中的完全去随机化自适应。*Evol. Comput.*, 9(2):159–195, 2001. doi: 10.1162/106365601750190398.[13] Kenneth Levenberg. 一种求解某些非线性最小二乘问题的方法。*Q. Appl. Math.*, 2(2):164–168, 1944. ISSN 1552-4485, 0033-569X. doi: 10.1090/qam/10666.[14] Donald W. Marquardt. 一种用于非线性参数最小二乘估计的算法。*J.Soc. Ind. Appl. Math.*, 11(2):431–441, 1963. ISSN 0368-4245, 2168-3484. doi: 10.1137/0111030.[15] Naum I. Akhiezer. 逼近论. Frederick Ungar, 纽约州纽约市, 1956.[16] Georg Pólya. 关于用实根多项式进行逼近。*Rend. Circ. Mat.Palermo*, 36(1):279–295, 1913. ISSN 0009-725X. doi: 10.1007/BF03016033.[17] I. J. Schoenberg. 关于 Pólya 频率函数 I：全正函数及其拉普拉斯变换。*J. Anal. Math.*, 1(1):331–374, 1951. ISSN 1565-8538. doi: 10.1007/BF02790092.[18] Samuel Karlin 和 William J. Studden. 切比雪夫系统：及其在分析学与统计学中的应用. Interscience Publishers, 纽约州纽约市, 1966.[19] Samuel Karlin. 全正性, 第一卷. Stanford University Press, 斯坦福, CA, 1968.[20] Dietrich Braess. 非线性逼近论, 施普林格计算数学丛书第 7 卷。Springer-Verlag, 柏林, 1986.[21] Elliott W. Cheney. 逼近论导论. 切尔西, 纽约州纽约市, 第 2 版, 1982. ISBN 978-0-8284-0317-7.

[22] John R. Rice. 函数逼近，第二卷：非线性与多变量理论. Addison-Wesley, 马萨诸塞州雷丁, 1969.[23] Erwin H. Bareiss. Sylvester 恒等式与多步整数保持型高斯消去法。《数学计算》, 22(103):565–578, 1968. ISSN 1088-6842, 0025-5718. doi: 10.1090/S0025-5718-1968-0226829-0.[24] Edmund A. C. Crouch 和 Donna Spiegelman. 对形如$\int_{-\infty}^{+\infty} f(t) \exp(-t^2)\ dt$ 的积分求值：逻辑斯蒂 - 正态模型的应用。《美国统计学会杂志》, 85(410): 464–469, 1990. ISSN 0162-1459. doi: 10.1080/01621459.1990.10476222.[25] Jeffrey J. Gory、Peter F. Craigmile 和 Steven N. MacEachern. 一类针对边际可解释性进行调整的广义线性混合模型。《统计医学》, 40(2):427–440, 2021. ISSN 0277-6715. doi: 10.1002/sim.8782.[26] John Ashworth Nelder 和 R. Mead. 一种用于函数极小化的单纯形法。《计算机学报》, 7(4):308–313, 1965. ISSN 0010-4620. doi: 10.1093/comjnl/7.4.308.[27] Fredrik Johansson. mpmath：用于任意精度浮点运算的 Python 库。Python 软件包版本 1.3.0，https://mpmath.org/，2023.[28] Laurent Fousse、Guillaume Hanrot、Vincent Lefèvre、Patrick Pélissier 和 Paul Zimmermann. MPFR：支持正确舍入的多精度二进制浮点库。《*ACM* 数学软件汇刊》, 33(2):13, 2007. ISSN 0098-3500, 1557-7295. doi: 10.1145/1236463.1236468.[29] Martin Mächler. Rmpfr：用于可靠多精度浮点运算的 R MPFR。R 软件包版本 1.1-2, https://CRAN.R-project.org/package=Rmpfr，2025.[30] Helaman R. P. Ferguson、David H. Bailey 和 Steve Arno. PSLQ 分析：一种整数关系发现算法。《数学计算》, 68(225):351–369, 1999. doi: 10.1090/S0025-5718-99-00995-3.[31] Sander P. Zwegers. 模拟 *theta* 函数. 博士论文, 乌得勒支大学，乌得勒支，荷兰，2002.[32] Mark K. Transtrum、Benjamin B. Machta 和 James P. Sethna. 为什么非线性数据拟合如此困难？《物理评论快报》, 104(6):060201, 2010. doi: 10.1103/PhysRevLett.104.060201.[33] Mark K. Transtrum、Benjamin B. Machta 和 James P. Sethna. 非线性最小二乘的几何及其在松散模型与优化中的应用。《物理评论 *E*》, 83(3):036701, 2011. doi: 10.1103/PhysRevE.83.036701.[34] Tyler J. VanderWeele. 因果推断中的解释：中介作用与交互作用的方法. 牛津大学出版社, 纽约州纽约市, 2015. ISBN 978-0-19-932587-0.[35] Linda Valeri 和 Tyler J. VanderWeele. 允许暴露-中介交互并作因果解释的中介分析：理论假设及其与 SAS 和 SPSS 宏的实现。《心理学方法》, 18(2):137–150, 2013. ISSN 1939-1463, 1082-989X. doi: 10.1037/a0031034.[36] Dustin Tingley、Teppei Yamamoto、Kentaro Hirose、Luke J. Keele 和 Kosuke Imai. mediation：用于因果中介分析的 R 软件包。《统计软件杂志》, 59(5):1–38, 2014. ISSN 1548-7660. doi: 10.18637/jss.v059.i05.

[37] Kazuki Yoshida 和 Yi Li。regmedint：基于回归的因果中介分析，包含交互作用项和效应修饰项。R 软件包版本 1.0.2，https://CRAN.R- project.org/package=regmedint，2026年。[38] Bengt O. Muthén、Linda K. Muthén 和 Tihomir Asparouhov。使用 *Mplus* 进行回归与中介分析。Muthén & Muthén，美国加利福尼亚州洛杉矶，2016年。ISBN 978-0-9829983-1-1。[39] Gary King、Michael Tomz 和 Jason Wittenberg。充分利用统计分析：改进解释与呈现。《美国政治科学杂志》，44(2):347–361，2000年。ISSN 00925853。doi：10.2307/2669316。[40] Abhinav Sharma 和 Michael Sheinman。具有高斯态密度的有机材料中化学势的解析近似。《物理学杂志 *D*：应用物理》，46(12):125106，2013年。ISSN 0022-3727。doi：10.1088/0022-3727/46/12/125106。[41] Lin Lin、Jianfeng Lu、Lexing Ying 和 Weinan E。费米–狄拉克函数的基于极点的近似。《中国数学年刊（*B*辑）》，30(6):729–742，2009年。ISSN 0252-9599。doi：10.1007/ s11401-009-0201-7。[42] Jonathan E. Moussa。费米–狄拉克分布的极小极大有理逼近。《化学物理杂志》，145(16):164108，2016年。ISSN 0021-9606。doi：10.1063/1.4965886。[43] Sheila M. Gaynor、Joel D. Schwartz 和 Xihong Lin。常见二元结局的中介分析。《统计医学》，38(4):512–529，2019年。ISSN 02776715。doi：10.1002/sim.7945。[44] Tommi S. Jaakkola 和 Michael I. Jordan。通过变分方法进行贝叶斯参数估计。《统计计算》，10(1):25–37，2000年。ISSN 1573-1375。doi：10.1023/A:1008932416310。[45] Niccolò Anceschi、Cristian Castiglione、Tommaso Rigon、Giacomo Zanella 和 Daniele Durante。逻辑斯蒂对数似然的最优且计算上可行的下界。《生物计量学》，2026年。ISSN 0006-3444, 1464-3510。doi：10.1093/biomet/asag060。[46] Dan Pirjol。逻辑斯蒂-正态积分及逻辑斯蒂-正态分布的矩，2026年。[47] John Holmes、Ness Arps 和 Marco Reale。逻辑比值－正态分布的矩的推断应用，2026年。[48] V. Yu. Korolev 和 I. G. Shevtsova。关于柯尔莫哥洛夫分布和逻辑斯蒂分布的矩。《莫斯科大学计算数学与控制论学报》，50(S1):S62–S67，2026年。ISSN 0278-6419。doi：10.3103/ s0278641926700226。[49] Harold Widom。复平面曲线系相关的极值多项式。《数学进展》，3:127–232，1969年。doi：10.1016/0001-8708(69)90005-X。[50] Edward B. Saff 和 Vilmos Totik。带外场的对数势，数学科学基础丛书第316卷。Springer，查姆，2024年。ISBN 978-3-031- 65133-5。doi：10.1007/978-3-031-65133-5。