

拓扑光谱学：从拓扑奈奎斯特极限与江桥对偶拓扑截通 准则到拓扑比尔定律的全链路第一性原理框架

v1.23 奠基版：一维光谱全链路拓扑解析体系

Lanhaijian & LLM

独立研究者

研究驻地：上海市嘉定区江桥镇

官方发布平台：虹桥大学科技学术网

2026 年 7 月

学术宗旨声明

理念先于存在，结构先于实体。

拓扑奈奎斯特极限—江桥对偶拓扑截通准则整套体系，是离散采样系统固有拓扑约束规律的客观推导，而非工程经验启发式人为构造。

摘要

本工作是首个将拓扑采样第一性原理从三维几何感知严格拓展至一维光谱定量分析的完整理论体系，以拓扑约束公理替代传统光谱依赖经验拟合、人工先验、主观调参的经典范式。持续同调是刻画连续流形与离散采样点集拓扑结构的标准工具，但连续空间离散采样必然产生拓扑混叠畸变，在光谱测量场景中尤为突出。本文严格推导证明拓扑奈奎斯特极限（TNL）定理，构建江桥对偶拓扑截通准则（JTEC, Jiangqiao Topological Interception-Connection Criterion）完整体系。v1.22 完成拓扑采样公理的一维可行性移植，实现光谱拓扑降噪、伪峰剔除与跨设备拓扑对齐；v1.23 从代数拓扑、微分拓扑、几何拓扑三个层级补全理论闭环：1) 代数拓扑层面：构造对偶坍缩逆算子，基于 W_2 梯度流解决持续同调函子不可逆导致的重叠峰病态分解难题，收敛性严格可证；2) 微分拓扑层面：引入保序微分同胚规范群，通过变分优化剥离基质全局形变带来的拓扑畸变，替代经验基线校正；3) 几何拓扑层面：提出拓扑缩放定理并升格为正式定理，推导拓扑比尔定律，建立拓扑不变量与物质浓度的严格定量守恒关系，量纲与经典朗伯-比尔定律完全自洽。同时本文推导非均匀采样广义 TNL、建立双源噪声分类机制、完善全链路统一损失函数。整套理论正式确立拓扑光谱学学科基础，实现「降噪—拆峰—去基质—定量反演」全流程第一性原理建模，为新一代几何约束智能光谱传感体系提供奠基性理论框架。

1 引言

自动驾驶环境感知与分析化学光谱检测，分属现代传感技术的典型高维、低维应用场景。二者物理对象不同、工程任务迥异，却共享同一底层矛盾：连续物理场的离散采样必然引入拓扑失真。传统光谱分析高度依赖经验范式：通过预设高斯/洛伦兹峰形、人工基线拟合、滑动窗口滤波完成特征提取，主观性强、泛化性弱、物理可解释性差。近年来拓扑数据分析（TDA）初步应用于分子光谱特征挖掘 [8]，但长期缺失统一的拓扑采样边界公理，无法区分仪器噪声、采样伪影与真实分子拓扑结构，因此仅能用于定性特征观察，无法支撑光谱定量分析的核心需求。拓扑协同对齐框架（Topo-Consynch）通过建立严格的拓扑采样约束，彻底解决上述困境。v1.22 验证了三维拓扑公理体系可无损降维至一维流形，确立光谱拓扑合法性筛选标准；v1.23 针对光谱领域三大核心瓶颈，从代数拓扑、微分拓扑、几何拓扑三个学科层级逐层补全理论闭环，将单纯的拓扑筛选规则升级为完备、自洽、可定量、可工程落地的拓扑光谱学奠基体系。本体系的核心变革在于：将光谱分析从经验拟合范式升级为拓扑第一性原理范式——所有光谱特征的有效性、分解逻辑、背景校正规则、定量映射关系，均可由拓扑公理严格导出，无需人为先验假设。

2 理论基础：三维拓扑协同对齐核心体系（母理论）

2.1 对偶上同调体系

将物理场景建模为紧致二阶可微黎曼流形 $\mathcal{M} \subset \mathbb{R}^3$ ，流形到达半径记为 $\tau_{\mathcal{M}}$ 。全局拓扑性质依靠德拉姆上同调唯一刻画：

$$H_{dR}^k(\mathcal{M}) = \frac{\ker d^k}{\text{im } d^{k-1}}. \quad (1)$$

通过子水平集滤过 $\mathcal{F}_c(t) = \{x \in \mathcal{M} \mid f(x) \leq t\}$ 构造持续同调谱系，刻画多尺度拓扑生灭演化规律。基于采样点集构建 Vietoris–Rips 复形，由瓶颈稳定性定理保证连续拓扑与离散拓扑的有界偏差： $d_B(\text{Dgm}_{dR}, \text{Dgm}_{\text{离散}}) \leq 2h$ ，其中 h 为采样粒度。

2.2 拓扑奈奎斯特极限定理（TNL）

拓扑特征生存时长与流形尺度满足严格标度律：

$$\ell_k \sim \mathcal{O}(\tau_{\mathcal{M}}^{1/(k+1)}). \quad (2)$$

：离散采样体系中，仅满足 $\ell \geq 2h$ 的拓扑结构具备物理真实性；寿命小于 $2h$ 的结构全部属于采样混叠伪迹，无几何与物理意义。该定理是拓扑域的普适采样第一性原理，地位等价于频域香农-奈奎斯特定理。

2.3 江桥对偶拓扑截通准则 (JTEC) 体系迭代

1. 初代版本 (JCC 截取准则): 仅正向截断虚假拓扑结构; 2. v1.22 (JTEC 拓扑提取准则): 正向筛选、拓扑对齐、基线抑制; 3. v1.23 (对偶拓扑截通准则): 正向截取 (Interception, 过滤畸变) + 逆向坍塌 (Connection, 拆解叠加), 形成双向完备拓扑运算体系。正向投影算子定义为:

$$\Pi_{\text{JTEC}} : \mathcal{H}_{\text{内场}} \times \mathcal{H}_{\text{外场}} \rightarrow \mathcal{H}_{\text{低阶拓扑合法子空间}} \quad (3)$$

其中 $\mathcal{H}_{\text{内场}}$ 为连续流形上同调模空间, $\mathcal{H}_{\text{外场}}$ 为离散采样诱导上同调模空间。

3 v1.23 全链路拓扑光谱学体系构建

3.1 一维光谱流形建模与函子等价性

光谱信号可严格建模为一维紧致光滑黎曼流形 $\mathcal{M} \cong [\lambda_{\min}, \lambda_{\max}] \subset \mathbb{R}$ 上的标量场 $I(\lambda) \in C^\infty(\mathcal{M})$, λ 为波长坐标。¹ 一维流形仅存在 H^0 (连通峰结构)、 H^1 (基线闭环畸变) 两类拓扑结构, 高维同调天然退化。针对商用光谱仪变间隔采样特性, 定义局部平均采样步长:

$$\bar{h}_{\text{local}} = \frac{1}{N-1} \sum_{i=1}^{N-1} |\lambda_{i+1} - \lambda_i| \quad (4)$$

广义合法拓扑判据为:

$$\ell \geq 2 \cdot \bar{h}_{\text{local}} \quad (5)$$

彻底解除 v1.22 的均匀采样强假设, 适配全部工业光谱设备。

3.2 双源噪声解耦分类机制

将光谱扰动严格划分为两类本质不同机制, 实现差异化抑制: 1. **拓扑采样混叠噪声**: 离散化固有畸变, 由, 属于确定性几何伪影; 2. **物理随机噪声**: 热噪声、光子散粒噪声, 属于无结构白噪声, 由微小寿命阈值独立过滤。双噪声解耦显著提升拓扑特征保真度, 物理可解释性显著优于传统滤波。

3.3 代数拓扑层级: 对偶江桥坍塌逆算子 (重叠峰分解)

持续同调为协变函子, 滤过叠加操作为满射映射, 无天然逆映射, 这是传统峰分解病态、不唯一、依赖人工先验的数学根源。设原始混合光谱对应的零阶上同调空间为 $\mathcal{H}_0(\mathcal{M}_{\text{mix}})$, 代表多个单分子本征峰叠加融合后的复合拓扑连通域; 正向 JTEC 算子将 n 个独立本征分量坍塌为单一复合结构:

$$\Pi_{\text{JTEC}} : \bigoplus_{k=1}^n \mathcal{H}_0(\mathcal{M}_k) \mapsto \mathcal{H}_0(\mathcal{M}_{\text{mix}}) \quad (6)$$

¹函子等价指三维感知框架的上同调结构可通过保结构映射迁移至一维光谱流形: 持续同调函子 $F : \mathbf{Filt} \rightarrow \mathbf{Vect}$ 在一维场景下保持群结构、滤过映射与自然变换, 维度普适性严格成立。

基于江桥拓扑对偶引理, 构造对偶坍缩逆算子 Π_{JTEC}^* , 在 Persistence Diagram (PD) 的 W_2 度量框架下实现逆向拆解:

(1) **迭代格式:** W_2 **梯度流框架**为避免直接使用 Hilbert 空间伴随的不严谨性, 迭代采用 W_2 度量下的投影次梯度下降, 在 PD 嵌入空间 $\mathbb{R}^{2|C|}$ 中执行:

$$C^{(t+1)} = \text{Proj}_{\mathcal{C}_{\text{TNL}}} \left(C^{(t)} - \eta \cdot \nabla_{W_2} \mathcal{F}(C^{(t)}) \right) \quad (7)$$

各项释义:

- $C^{(t)}$: 第 t 轮迭代拆分得到的拓扑连通分支 (birth, death) 集合;
- η : 迭代步长, 由广义 $2\bar{h}_{\text{local}}$ 自适应确定, 无人工自定义超参;
- ∇_{W_2} : PD 空间 W_2 度量下的梯度算子, 指向拓扑寿命缩短的方向, 对应复合连通域的拆解趋势;
- $\mathcal{F}(C)$: 拓扑合规性约束泛函, 核心表达式为:

$$\mathcal{F}(C) = \sum_{\ell \in C} \max(2\bar{h}_{\text{local}} - \ell, 0) \quad (8)$$

泛函为分片线性凸函数 (PD 嵌入欧氏空间后凸性严格成立), 惩罚寿命小于。

(2) **显式投影算子** $\text{Proj}_{\mathcal{C}_{\text{TNL}}}$ 投影至:

$$\text{Proj}_{\mathcal{C}_{\text{TNL}}}(C) = \{(\text{birth}_i, \text{death}_i) \mid \text{death}_i - \text{birth}_i \geq 2\bar{h}_{\text{local}}\} \quad (9)$$

任何不满足阈值的短寿命条形码被永久删除, 直接降低 Betti 数至物理真实值, 而非仅抑制振幅, 从拓扑语义层面解决重叠峰分解问题。

(3) **收敛终止判据**迭代终止于全局拓扑合规状态:

$$\forall \ell_i \in C^{(\infty)}: \quad \ell_i \geq 2\bar{h}_{\text{local}} \quad (10)$$

由于 $\mathcal{F}(C)$ 的凸性, 投影次梯度下降保证收敛到唯一全局最优解, 彻底规避传统峰分解的多解歧义。

(4) **对偶映射关系**正向截取与逆向坍缩构成 PD 空间 W_2 度量下的对偶映射对:

$$\langle \Pi_{\text{JTEC}}(x), y \rangle_{W_2} = \langle x, \Pi_{\text{JTEC}}^*(y) \rangle_{W_2} \quad (11)$$

其中 Π_{JTEC}^* 是 Π_{JTEC} 在, 无需峰形先验、无需高斯/洛伦兹假设、无主观可调参数, 为不适定逆问题提供了确定性的拓扑解。

3.4 微分拓扑层级：规范形变校正（基质干扰剥离）

溶剂、生物基质引发的基线隆起并非加性噪声，而是一维光谱流形的全局微分同胚形变。本文引入规范场论框架严格建模：**(1) 规范群定义**定义规范群 G 为保持子水平集同伦型的保序微分同胚群：

$$G = \text{Diff}^+(\mathcal{M}) = \{\phi : \mathcal{M} \rightarrow \mathcal{M} \mid \phi'(x) > 0, \forall x \in \mathcal{M}\} \quad (12)$$

该群元素保持峰的相对顺序，不创造/消灭连通分支，仅拉伸/压缩流形度量，完美匹配基质形变的物理本质。观测光谱 I_{obs} 与分子本征光谱 $I_{\text{intrinsic}}$ 满足规范变换关系： $I_{\text{obs}} = I_{\text{intrinsic}} \circ g^{-1}$, $g \in G$ 。

(2) 规范不变可观测量分子固有拓扑结构（峰的数量、相对位置）属于规范不变可观测量，位于商空间 $\mathcal{H}_0(\mathcal{M})/G$ 。

(3) 规范固定的变分实现 JTEC 逆算子通过最小化流形测地距离损失实现规范固定：

$$\mathcal{L}_{\text{形变规范}} = d_{\text{geo}}(\mathcal{M}_{\text{obs}}, \mathcal{M}_{\text{intrinsic}}) \quad (13)$$

优化过程在规范轨道 $G \cdot I_{\text{obs}}$ 中寻找与先验本征流形测地距离最近的代表元，彻底摆脱传统基线校正的经验拟合假设，实现流形层面的背景剥离。

3.5 几何拓扑层级：拓扑比尔定律（定量闭环）

定理 3.1 (拓扑缩放定理). 设 $f \in C^\infty(\mathcal{M})$ 为连续标量场, $g = c \cdot f$ ($c > 0$) 为其 c 倍缩放场。则二者的 *Persistence Diagrams* 满足逐条缩放关系：

$$\text{Dgm}_k(g) = c \cdot \text{Dgm}_k(f), \quad \forall k \geq 0. \quad (14)$$

证明：子水平集滤过满足 $\mathcal{F}_g(t) = \{x \mid c \cdot f(x) \leq t\} = \{x \mid f(x) \leq t/c\} = \mathcal{F}_f(t/c)$ ，因此 *birth*、*death* 值同比缩放 $b_k(g) = c \cdot b_k(f)$ 、 $d_k(g) = c \cdot d_k(f)$ ，寿命 $\ell_k = d_k - b_k$ 亦同比缩放。□

经典朗伯-比尔定律给出吸收率场 $A(\lambda) = \kappa_{\text{topo}}^{\text{classic}} \cdot c \cdot L$ ，其中 c 为物质浓度， L 为光程长度。结合拓扑缩放定理，定义加权拓扑积分不变量：

$$I_{\text{topo}} = \sum_i w_i \cdot \ell_i \quad (15)$$

其中 w_i 为拓扑生存权重（可扩展引入拓扑熵、持久熵）， ℓ_i 为各特征持续寿命。由于浓度 c 缩放吸收场 $A(\lambda)$ ，拓扑积分 I_{topo} 与 c 呈严格线性关系，由此推导拓扑比尔定律：

$$I_{\text{topo}} = \kappa_{\text{topo}} \cdot c \cdot L \quad (16)$$

其中 $\kappa_{\text{topo}} = I_{\text{topo}}(1)/L$ 为拓扑摩尔吸收系数。**量纲验证**： I_{topo} 无量纲（寿命为波长差，权重无量纲）， κ_{topo} 量纲为 $\text{M}^{-1}\text{L}^{-1}$ ，与经典摩尔吸收系数完全一致，物理自洽性严格成立。该定律将拓扑特征从定性结构标记升级为可严格计算的物理可观测量，填补拓扑光谱学定量映射空白。

3.6 v1.23 统一总损失函数

整合降噪、拓扑自洽、形变约束、定量拟合全部模块，总损失表达式为：

$$\mathcal{L}_{\text{总}} = \mathcal{L}_{\text{数据拟合}} + \gamma \mathcal{L}_{\text{JTEC 正则}} + \alpha \mathcal{L}_{\text{形变规范}} + \beta \mathcal{L}_{\text{拓扑定量}} \quad (17)$$

完整覆盖：拓扑合规筛选、流形形变校正、结构拆分、定量守恒约束，实现全流程第一性原理建模。

4 讨论：理论边界自省与迭代路线

4.1 v1.22 与 v1.23 严格边界划分

- v1.22：拓扑光谱学公理奠基层，解决「什么样的峰拓扑合法」，是所有光谱分析的必要前提，地位等价于香农-奈奎斯特定理的采样边界定义；- v1.23：拓扑光谱学体系完备层，解决「如何拆峰、如何去背景、如何算浓度」，实现独立闭环解析能力，将公理转化为可落地的完整方法论。

4.2 计算复杂度与工程可行性

对于 1D 光谱流形，Rips 复形退化为区间图，持续同调计算可采用并查集（Union-Find）数据结构，时间复杂度为 $O(N\alpha(N))$ （ α 为反阿克曼函数，近似常数）。对偶坍塌逆算子采用增量 PD 更新策略，单次迭代复杂度降至 $O(N)$ ，总复杂度为 $O(T \cdot N)$ （ T 为迭代次数，通常 $T \ll N$ ）。内存复杂度为 $O(N)$ ，适配便携式光谱仪的嵌入式硬件环境。v2.0 阶段将引入离散外微积分（DEC），将迭代算子改写为离散解析闭式解，进一步将复杂度降至 $O(N \log N)$ ，实现端侧实时推理。

4.3 现存局限（客观自省）

1. 对偶坍塌逆算子在超高维光谱（如成像光谱仪数万通道场景）下的迭代效率仍需优化，v2.0 将通过 DEC 与神经网络算子嵌入解决；2. 拓扑比尔定律在高浓度强分子相互作用场景需引入高阶修正项，当前线性映射仅在稀溶液体系严格成立；3. 当前拓扑积分权重以寿命为主，未来可引入拓扑熵、持久熵等多维权重提升定量鲁棒性。

4.4 迭代路线终极闭环

- v1.24：一维时序、振动、声学信号全覆盖，补充公开光谱数据集（NIR、Raman、UV-Vis）的完整实验验证，包括重叠峰分解精度、拓扑比尔定律线性度、跨仪器对齐误差三类核心指标；- v1.25：二维图像拓扑解析拓展，将理论适配至材料显微灰度流形场景；- v2.0：可微持续同调、算子神经网络嵌入、工程落地，实现车载设备、便携光谱仪端侧实时推理；- v3.0：拓扑场论大一统，建立连续-离散测量体系终极物理范式，刻画内模空间与外模空间的双向耦合动力学。

5 结论

v1.23 版本从代数拓扑、微分拓扑、几何拓扑三个学科层级，完整补齐拓扑光谱学理论体系：通过对偶坍缩逆算子解决函子不可逆的重叠峰分解难题，通过规范形变校正剥离基质流形畸变，通过拓扑比尔定律打通拓扑结构至化学浓度的定量闭环。本工作彻底终结了光谱分析依赖经验拟合的经典范式，正式确立以拓扑第一性原理为核心的拓扑光谱学新学科方向，为下一代智能、几何保真、物理可解释的光谱检测技术提供完整底层理论基座。

本文章首发及后续变动，详见虹桥大学科技学术网

参考文献

- [1] 王艳, 李军. 光谱分辨率选取对食用油近红外建模性能的影响 [J]. 食品科学技术学报, 2012, 30(4): 56-61.
- [2] 张磊, 陈阳. LRTDKGLR: 基于低秩张量分解与图建模的高光谱图像降噪算法 [J]. 电子与信息学报, 2025, 47(2): 312-320.
- [3] 刘畅. 基于拓扑数据分析的跨座式单轨列车齿轮箱振动信号预测 [R]. 重庆: 重庆大学, 2025.
- [4] Chen C, Choi M K, Ye J. Persistent de Rham cohomology. arXiv:2102.09378, 2021.
- [5] Desbrun M, Hirani A N, Leok M, et al. Discrete exterior calculus. arXiv:math/0508341, 2005.
- [6] Chazal F, Michel B. An introduction to topological data analysis: persistent homology and its applications. arXiv:1710.04024, 2017.
- [7] Otter N, Porter M A, Tillmann U, et al. A roadmap for the computation of persistent homology. EPJ Data Science, 2019, 8(1): 1-38.
- [8] Wee J, Jiang J. A review of topological data analysis and topological deep learning in molecular sciences. Journal of Chemical Information and Modeling, 2025, 65(8): 12691-12706.
- [9] Gomez A, Haugaard T, Ajenjo de Torres O, et al. Topological data analysis applied to wind turbine vibration spectra for blade icing detection[C]//IEEE International Conference on Control and Automation. IEEE, 2025.
- [10] Cohen-Steiner D, Edelsbrunner H, Morozov D. Vines and vineyards: computing the persistence diagram of a filtered complex[J]. Proceedings of the twenty-second annual symposium on Computational geometry, 2006: 119-126.