通过基于锚点的扩散几何近似，衔接距离与谱位置编码

Zimo Yan 和 Zheng Xie* **以及 Runfan Duan、Chang Liu 和 Wumei Du**国防科技大学，
中国长沙 {yanzimo20,xiezheng81,duanrunfan24,liuchang_,wumeidu}@nudt.edu.cn

arXiv:2601.04517v1 [cs.IT] 8 Jan 2026

摘要

分子图学习受益于能够捕捉局部邻域和全局拓扑的位置编码信号。两类常用方法分别是由拉普拉斯算子或扩散算子导出的谱编码，以及基于最短路径信息构建的锚点距离编码，但二者之间的确切关系尚未得到充分理解。我们将距离编码视为扩散几何的一种低秩替代，并推导出一个显式的三边测量映射，可根据变换后的锚点距离和锚点谱位置重建截断后的扩散坐标；对于随机正则图，该映射具有逐点保证和弗罗贝尼乌斯间隙保证。在使用共享的基于 GNP 的药物-药物相互作用预测骨干网络对 DrugBank分子图进行实验时，距离驱动的 Nyström 方案能够紧密恢复扩散几何；拉普拉斯编码和距离编码的表现均显著优于不使用编码的基线。

1 引言

分子图建模任务——从头图生成到分子推理——需要能够同时捕捉局部化学邻域和全局拓扑的表示（2017年；2020；2020）。一个反复出现的挑战是，如何注入结构化位置信息，使模型能够利用长程依赖关系和全局组织结构。在这项工作中，我们使用图神经过程（GNP）作为灵活的概率骨干网络（2025c）。

人们探索了许多位置与结构信号，包括随机游走/传播统计量、结构角色描述符，以及图 Transformer 中的相对位置偏置（2022；2023；2021）。我们重点关注两类尤为常见的方法。谱编码/扩散编码（例如拉普拉斯特征映射、热核和扩散映射特征）可提供与扩散几何对齐的坐标，但往往依赖代价高昂的特征分解或精心设计的近似方法（2007；2023；2022）。距离/锚点编码以每个节点到一个小型锚点集的最短路径距离来表示该节点（也可选择进行变换），无需显式谱计算，是一种简单且可扩展的替代方案（2025a；2009）。

尽管锚点距离特征得到广泛使用，但关于其何时能够近似扩散几何的明确代数描述仍然缺失；对于恢复的坐标以及诱导的距离矩阵/核矩阵，也缺乏同时具有实际意义的误差度量。在分子场景中，这一缺口尤为重要：扩散几何是一种自然的归纳偏置，而计算预算又受到限制，因此产生了一个实际问题：在药物-药物相互作用（DDI）预测等任务中，距离编码何时能够作为谱编码的有理论依据的替代方案（2018；2018）。

动机。 我们提出的问题是：我们能否构造一个显式的代数映射，将基于锚点的最短路径编码映射为截断的扩散（谱）坐标，并为逐点误差和矩阵级误差提供可证明的保证？这种联系能够将离散距离基元与扩散几何联系起来，并阐明以距离特征替代谱计算背后的准确率-效率权衡。

我们通过理论分析和实验来回答这一问题。在理论方面，我们推导出一个显式的三边测量算子，可根据变换后的锚点距离和锚点谱位置重建截断扩散坐标；并在局部单调距离链接的随机正则图模型下，建立了逐点保证和弗罗贝尼乌斯间隙保证。在实证方面，我们使用距离驱动的 Nyström 方案评估 DrugBank 上的扩散几何恢复，并在共享的 GNP 骨干网络下评估 DrugBank 和 ChCh-Miner 上的药物-药物相互作用预测（2020；

*通讯作者：Zheng Xie（xiezheng81@nudt.edu.cn）。

2020；2023）。总体而言，拉普拉斯编码和距离编码均优于 NoPE，其中拉普拉斯编码带来的提升最为稳定。

我们的主要贡献如下：

1. 我们通过将基于锚点的距离编码形式化为扩散几何的低秩替代，并开发一种显式的三边测量算子，在随机正则图假设下为其提供逐点和 Frobenius 差距保证，从而架起谱代数桥梁。

2. 我们证明，基于距离的 Nyström 近似能够在 DrugBank分子图上紧密还原扩散几何；只需适量锚点，即可获得准确的核与嵌入近似。

3. 我们采用用于药物-药物相互作用预测的通用 GNP 架构，对比 NoPE、距离编码和 LapPE，并在 DrugBank 和 ChCh-Miner 上针对距离变换和锚点数量开展消融实验。

2 文献综述

本节回顾图学习中的结构信息和位置信息，重点关注分子图。我们讨论纯局部计算在结构上的局限性，概述谱／扩散方法与基于锚点的距离编码这两类常见的位置编码，并回顾药物-药物相互作用模型，以明确本文所提出的距离特征与扩散几何之间的谱代数桥梁所要解决的空白。

2.1 图结构学习与局部性限制

许多图学习方法采用基于局部性的计算方式，其中最具代表性的是消息传递：节点表示通过聚合并变换相邻节点的特征进行迭代更新，构成了许多 GNN 架构及相关图编码器的骨干网络（2025c；2017年；2016年）。这类设计在广泛的应用领域中展现出强劲的实证表现，包括分子性质预测。

尽管取得了这些成功，局部性仍带来了结构上的局限。标准消息传递的表达能力受 Weisfeiler-Leman 测试集的上界约束，因此，非同构图和结构上不同的节点仍可能无法区分（2019；2018）。此外，长程依赖关系可能因过平滑和过压缩效应而难以捕捉（2021；2022）。这些问题促使人们在局部邻域聚合之外，为节点特征补充显式的结构性或位置性信号。

2.2 谱位置编码

除谱方法和基于距离的设计外，已有研究还探讨了随机游走／传播统计量、结构角色、Transformer相对位置偏置以及子结构计数描述符（2022；2023；2021）。我们聚焦于谱／扩散方法和锚点距离编码，因为它们提供了自然的几何视角，也是我们明确建立联系的两类方法。

谱位置编码从拉普拉斯算子或扩散算子的特征向量中导出坐标（例如，拉普拉斯特征映射和扩散映射）（2003年；2005年；2006年），并被广泛用作固定的拉普拉斯位置特征，以注入全局拓扑信息并打破对称性（2023）。一些变体通过学习重加权／组合或基于传播的近似来调整或近似这些特征（2022；2023），但由于依赖全局模式，谱方法在大型或不断演变的图上仍可能代价高昂。

2.3 基于距离的位置编码

基于距离的位置编码通过节点到一小组锚节点的最短路径距离（或其简单变换）来表示每个节点（2009；2019）。这类特征依赖最短路径计算，而非特征分解，因此通常易于实现，并且能借助标准图算法程序良好扩展。

这类编码已用于提升模型超越魏斯费勒-莱曼测试的表达能力（2009），并通过最短路径距离及相关结构先验构建具备结构感知能力的图 Transformer（2021）。然而，它们与谱几何/扩散几何之间的联系往往只停留在定性讨论层面，因此有限锚点集何时能够近似扩散几何，以及如何在坐标和核/距离层面量化近似误差，仍有待研究。

2.4 基于图的药物-药物相互作用预测模型

基于图的方法已成为药物-药物相互作用（DDI）预测的核心。早期深度学习方法将分子指纹或序列表示与标准神经网络

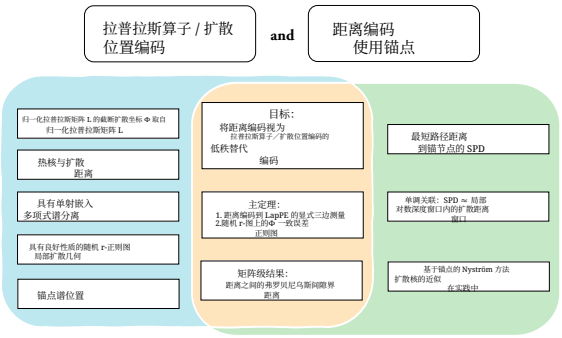


图 1：谱/扩散位置编码与锚点距离位置编码概览，以及本文建立的代数桥梁。

架构（2018），而后续研究利用生物医学知识图谱，并采用图神经网络或知识图谱神经网络对异构网络中的相互作用进行建模（2018；2020）。多模态框架进一步整合分子图、靶标、通路和文本，丰富了药物表示（2020）。

较新的架构通过双图设计和协同注意力机制，显式利用分子结构来捕捉药物对之间的子结构层级交互（2025b；2025年）。这些研究结果一再凸显了结构信息对于准确预测药物-药物相互作用的重要性；然而，大多数模型采用固定的位置/结构特征设计，并侧重于架构或数据集成选择，而在统一的理论视角下系统地关联基于距离的编码与谱扩散几何方面，相关研究相对有限。

2.5 总结与尚存差距

谱编码能提供具有几何意义的坐标，但全局计算成本高；距离编码则易于扩展且简单，却通常仅有经验性依据，也缺乏统一的代数理论来说明它们何时可作为扩散几何的低秩替代。为弥合这一差距，我们提出重构算子并给出误差界，以建立距离编码与谱编码之间的联系；此外，我们在统一的基于 GNP 的框架中评估了由此产生的准确率与效率权衡，并将其用于药物-药物相互作用预测（图 1 和图 2）。

3 预备知识

我们汇总分析中使用的符号和基本几何对象，并形式化两类位置编码：基于锚点的最短路径距离编码和截断拉普拉斯谱坐标，以及第4节研究的代数比较问题。

3.1 符号约定

设 $G=(V,E)$ 是一个有限连通无向图，且具有 $|V|=n$。对于 $u,v\in V$，令 $\mathrm{SPD}(u,v)$ 表示最短路径距离。对于 $u\in V$ 和 $R\in\mathbb{N}$，定义半径为 R 的球

$$B_G(u,R)=\{\,v\in V:\mathrm{SPD}(u,v)\le R\,\}. \quad (1)$$

令 $A\in\{0,1\}^{n\times n}$ 为邻接矩阵，D 为对角度矩阵。在理论部分，我们关注 r-正则图，因此 $D=rI$，并使用

$$L:=I-\frac{1}{r}A. \quad (2)$$

设 $\{(\lambda_j,\varphi_j)\}_{j=1}^{n}$ 为 L 的特征值对，其中 $\{\varphi_j\}$ 在 $\ell^2(V)$ 中正交归一，且 $0=\lambda_1\le\lambda_2\le\cdots\le\lambda_n$。对于 $t>0$，定义热半群

$$K_t=e^{-tL}, \quad (3)$$

其元素为 $k_t(u,v)=(K_t)_{uv}$。时刻 t 的扩散距离等价地表示为

$$\begin{aligned} d_t(u,v)^2&=\sum_{j=1}^{n}e^{-2t\lambda_j}\big(\varphi_j(u)-\varphi_j(v)\big)^2\\ &=k_{2t}(u,u)+k_{2t}(v,v)-2k_{2t}(u,v). \end{aligned} \quad (4)$$

对于整数 $m\ge 1$，定义截断扩散映射嵌入

$$\Phi_t^{(m)}(v)=\big(e^{-t\lambda_{j+1}}\varphi_{j+1}(v)\big)_{j=1}^{m}\in\mathbb{R}^m, \quad (5)$$

以及相应的截断扩散距离

$$d_t^{(m)}(u,v)=\|\Phi_t^{(m)}(u)-\Phi_t^{(m)}(v)\|_2. \quad (6)$$

给定锚点 $a_1,\dots,a_{m+1}\in V$，记

$$p_i=\Phi_t^{(m)}(a_i)\in\mathbb{R}^m,\quad i=1,\dots,m+1. \quad (7)$$

3.2 问题陈述

给定一个图 $G=(V,E)$ 和一个锚点集 $S=\{a_1,\dots,a_{m+1}\}\subset V$。对于任意节点 $v\in V$，定义（变换后的）距离编码

$$\tilde\zeta(v\mid S)=\big(\psi(\mathrm{SPD}(a_1,v)),\dots,\psi(\mathrm{SPD}(a_{m+1},v))\big)^\top\in\mathbb{R}^{m+1}, \quad (8)$$

其中 $\psi:\ \mathbb{R}_{\ge0}\to\mathbb{R}_{\ge0}$ 是第4节中指定的单调标量变换；原始距离编码对应于 $\psi=\mathrm{id}$。目标拉普拉斯谱编码是截断坐标

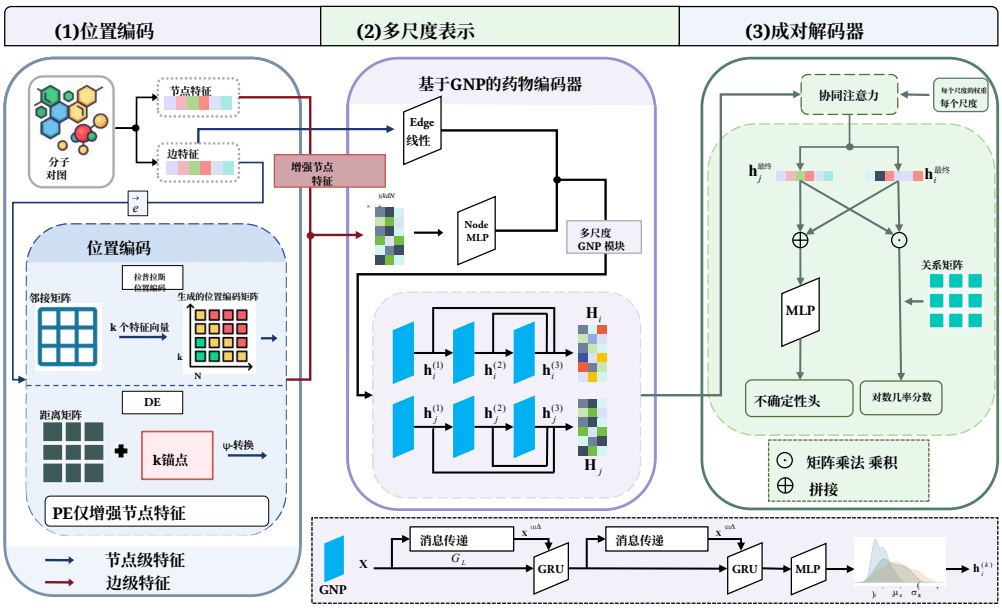


图 2：提出的药物-药物相互作用框架：位置编码增强分子图，随后由共享的多尺度 GNP 进行编码，并通过协同注意力和关系感知评分进行解码。

$\Phi_t^{(m)}(v) \in \mathbb{R}^m$ （或其符号/基不变变体）。

我们的核心问题是：

在对 G 施加适当的结构条件并采用适定的锚点配置 S后，能否构造一个显式的代数映射，将 $\zeta(v \mid S)$作为输入，并输出 $\Phi_t^{(m)}(v)$，同时对误差进行定量控制，并给出诱导的距离/核量之间的矩阵级差异界（例如 *Frobenius* 范数）？

在第4节中，我们将这些条件形式化，并证明这种链接确实存在；其误差由结构失配参数 δ控制（精确常数除外），且适用于对数邻域尺度范围，与先前分析中常见的局部尺度一致。

4 理论分析

本节分析 r-正则图上最短路径距离编码 (DE) 与截断拉普拉斯谱坐标 (LapPE) 之间的代数关系。所有证明均推迟至附录 A.4。本节中，我们始终使用 $L = I - \frac{1}{r}A$ 形式的对称归一化拉普拉斯算子，使扩散算子为 $K_t = e^{-tL}$，截断扩散映射嵌入为 $\Phi_t^{(m)}$，定义见第3节。

4.1 假设

假设 1。 在理论分析中，我们考虑从随机正则模型 $G \sim \mathcal{G}_{n,\ r}$ 抽取的图，即在固定度为 $r \geq 3$ 的条件下， $[n]$ 上所有带标号 r-正则图的均匀分布。

假设 2. 存在 $m \in \mathbb{N}$ 和 $\alpha > 0$，使得对于 $G \sim \mathcal{G}_{n,r}$，以高概率而言，截断嵌入 $\Phi_t^{(m)}$：$V \to \mathbb{R}^m$ 是单射并满足

$$\min_{u \neq v} \left\| \Phi_t^{(m)}(u) - \Phi_t^{(m)}(v) \right\|_2 \geq n^{-\alpha}. \quad (9)$$

对于$G \sim \mathcal{G}_{n,r}$，在分析拉普拉斯算子和基于距离的位置编码时常用的相同对数尺度范围内，这些假设以高概率成立。

4.2 主要理论结果

我们首先定义一个显式的三边测量算子，将距离特征映射到一个 m 维的谱坐标。

定义 1。固定 $t > 0$ 和 $m \in \mathbb{N}$，并令 a_1、…、$a_{m+1} \in V$ 为锚点，其中

$$p_i = \Phi_t^{(m)}(a_i) \in \mathbb{R}^m, \quad i = 1, \ldots, m+1. \quad (10)$$

定义

$$A = 2 \begin{pmatrix} (p_1 - p_{m+1})^\top \\ \vdots \\ (p_m - p_{m+1})^\top \end{pmatrix} \in \mathbb{R}^{m \times m}. \quad (11)$$

对于任意 $r=(r_1,\dots,r_{m+1})^\top\in\mathbb{R}^{m+1}$，定义

$$b(r)=\begin{pmatrix}\|p_1\|_2^2-\|p_{m+1}\|_2^2+r_{m+1}^2-r_1^2\\ \vdots\\ \|p_m\|_2^2-\|p_{m+1}\|_2^2+r_{m+1}^2-r_m^2\end{pmatrix}\in\mathbb{R}^m. \tag{12}$$

令 $\psi:\ [0,R]\to\mathbb{R}_+$ 为定理 *2* 中严格递增的标量函数，并将其逐元素作用于 $y\in\mathbb{R}^{m+1}$ 记作 $\psi_*(y)$，即

$$\big(\psi_*(y)\big)_i:=\psi(y_i),\quad i=1,\dots,m+1. \tag{13}$$

回顾到锚点 $S=\{a_1,\dots,a_{m+1}\}$ 的（原始）距离编码：

$$\zeta(v\mid S)=\big(\mathrm{SPD}(a_1,v),\dots,\mathrm{SPD}(a_{m+1},v)\big)^\top\in\mathbb{R}^{m+1}. \tag{14}$$

当 A可逆时，定义

$$T(v):=A^{-1}b\big(\psi_*(\zeta(v\mid S))\big)\in\mathbb{R}^m. \tag{15}$$

辅助量。令 $\{(\lambda_j,\varphi_j)\}_{j=1}^n$ 为 L 的正交归一特征分解。对于 $m\in\mathbb{N}$ 和 $t>0$，定义截断尾项

$$\mathrm{Tail}_t^{(m)}(u,v)^2:=\sum_{j=m+2}^{n}e^{-2t\lambda_j}\big(\varphi_j(u)-\varphi_j(v)\big)^2. \tag{16}$$

令 $p:\ V\to\mathbb{R}^m$，且令 $V_1,\dots,V_{m+1}\overset{\text{i.i.d.}}{\sim}\mathrm{Unif}(V)$。定义退化概率

$$\eta_n:=\mathbb{P}\big(\det[p(V_1)-p(V_{m+1})\ \cdots\ p(V_m)-p(V_{m+1})]=0\big). \tag{17}$$

定理 2。固定 $t>0$ 和 $m\in\mathbb{N}$。设 $G=(V,E)$ 为有限连通图，且令 $R\ge1$。设 $\psi:\ [0,R]\to\mathbb{R}_+$ 为严格递增函数。定义

$$\Delta_{\mathrm{geom}}(G;R,t,\psi):=\sup_{\mathrm{SPD}(u,v)\le R}\big|d_t(u,v)-\psi(\mathrm{SPD}(u,v))\big| \tag{18}$$

and

$$\Delta_{\mathrm{tail}}(G;R,t,m):=\sup_{\mathrm{SPD}(u,v)\le R}\mathrm{Tail}_t^{(m)}(u,v). \tag{19}$$

那么，对于所有的 $u,v\in V$，满足 $\mathrm{SPD}(u,v)\le R$，

$$\big|d_t^{(m)}(u,v)-\psi(\mathrm{SPD}(u,v))\big|\le\Delta_{\mathrm{geom}}(G;R,t,\psi)+\Delta_{\mathrm{tail}}(G;R,t,m). \tag{20}$$

为便于后续使用，定义局部连结/截断误差水平

$$\delta_n^L:=\Delta_{\mathrm{geom}}(G;R,t,\psi)+\Delta_{\mathrm{tail}}(G;R,t,m). \tag{21}$$

命题 3。固定 $m\ge1$ 并令 $p:\ V\to\mathbb{R}^m$ 为截断扩散嵌入 $\Phi_t^{(m)}$ v $a_1\ \dots\ a_{m+1}$ $\overset{\text{iid}}{\sim}$ ()：

()。抽取锚点，，$\overset{\text{iid}}{\sim}\mathrm{Unif}(V)$ 并令 $p_i:=p(a_i)$。定义

$$M:=\big[p_1-p_{m+1}\ \cdots\ p_m-p_{m+1}\big]\in\mathbb{R}^{m\times m}. \tag{22}$$

当且仅当 $\det(M)\neq0$时，$p_1,\dots,p_{m+1}$ 在 $\mathbb{R}^m$ 中仿射无关。此外，定义 *1* 中的三边测量矩阵满足 $A=2M^\top$。

假设 (17) 中的退化概率 η_n 满足 $\eta_n=o(1)$。那么 $\mathbb{P}\big(\det(M)\neq0\big)=1-\eta_n=1-o(1)$。

注4 (抖动锚点几乎必然是一般的**)**。固定 $m\ge1$，并设 $p_1,\dots,p_{m+1}\in\mathbb{R}^m$ 为任意点。设 $\varepsilon>0$，且令 $\xi_1,\dots,\xi_{m+1}\in\mathbb{R}^m$ 为随机向量，其联合分布关于勒贝格测度绝对连续。令$\tilde p_i:=p_i+\varepsilon\xi_i$ 并

$$\tilde M:=\big[\tilde p_1-\tilde p_{m+1}\ \cdots\ \tilde p_m-\tilde p_{m+1}\big],$$

我们有 $\mathbb{P}\big(\det(\tilde M)\neq0\big)=1$。

接下来的两个定理分别给出逐点重构误差 $\|\Phi_t^{(m)}(v)-T(v)\|_2$ 以及由此产生的矩阵级偏差的界。

定理 5。设 $G\sim\mathcal{G}_{n,r}$ 具有固定的 $r\ge3$，并固定 $t>0$ 和 $m\in\mathbb{N}$。设 $a_1,\dots,a_{m+1}$ 为满足命题 *3*（以及在需要时满足上文抖动一般性注释）的锚点，并令 $\Phi_t^{(m)}$ 为截断谱嵌入。假设假设 *1–2* 成立，且定理 *2* 对 $d_t^{(m)}$ 成立，误差为 δ_n^L。

对于满足 $\mathrm{SPD}(a_i,v)\le R$（对所有 i均成立）的节点 $v\in V$，定义

$$r_i:=\psi\big(\mathrm{SPD}(a_i,v)\big),\quad r_i^\star:=d_t^{(m)}(a_i,v). \tag{23}$$

那么，以概率 $1-o(1)$（关于 G），我们有

$$\big\|\Phi_t^{(m)}(v)-T(v)\big\|_2\ \le\ \|A^{-1}\|_{\mathrm{op}}\sqrt{m}\,\Big(4\,\rho_R\,\delta_n^L+2(\delta_n^L)^2\Big), \tag{24}$$

其中

$$\rho_R:=\max_{0\le d\le R}\psi(d)\ +\ \delta_n^L. \tag{25}$$

特别是，对于 $\delta_n^L\le1$，

$$\big\|\Phi_t^{(m)}(v)-T(v)\big\|_2\ \le\ \|A^{-1}\|_{\mathrm{op}}\sqrt{m}\,\big(4\rho_R+2\big)\,\delta_n^L. \tag{26}$$

定理 6. 在与定理 *5* 相同的设定和假设下，定义

$$(D_{\mathrm{SPD}})_{v,i}=\mathrm{SPD}(v,a_i),(D_{\mathrm{diff}}^{(m)})_{v,i}=d_t^{(m)}(v,a_i), \tag{27}$$

并令 ψ_*（D_{SPD}）表示对 ψ 逐元素应用。则以概率 $1-o$（1），

$$\left\|D_{\mathrm{diff}}^{(m)}-\psi_*(D_{\mathrm{SPD}})\right\|_F \le \delta_n^L\sqrt{n(m+1)}. \quad (28)$$

具体而言，每个条目的平均偏差至多为 δ_n^L。

证明见附录 A.4。定理 5 将局部单调链接（定理 2）与定义 1 中线性系统的扰动分析结合起来；定理 6 则通过将该链接一致地应用于所有节点-锚点对，并对所得误差求和得出。

4.3 相对于 NoPE 的表达能力影响

我们回顾一个标准结论：距离编码对随机正则图上的消息传递图神经网络有何作用。

定理 7。设 $G\sim\mathcal{G}_{n,\,r}$ 取自随机 r-正则图模型，其中 $r\ge 3$固定；设 $\mathcal{F}_{\mathrm{MP}}$ 表示不使用位置编码（*NoPE*）的 T层消息传递图神经网络类，其区分能力受 *1-WL* 测试（*2019*；*2018*）的上界约束。令 $\mathcal{F}_{\mathrm{DE}}$ 表示增强了基于到 $k=\Theta$ 个锚点的最短路径距离的距离编码的 T 层消息传递图神经网络类，如（*2009*）所述。则存在一族定义在 $\mathcal{G}_{n,\,r}$ 上的节点级分类任务，使得对于 G，以高概率，

1. 在 $\mathcal{F}_{\mathrm{MP}}$ 中，不存在 *NoPE* 消息传递图神经网络能够实现目标标签（因为 *1-WL* 无法区分相关节点）；但 *2.* 在 $\mathcal{F}_{\mathrm{DE}}$ 中，存在某种 *DE* 增强图神经网络能够精确区分所有标签类别。

特别是在随机正则图上，函数类 $\mathcal{F}_{\mathrm{DE}}$ 的表达能力严格强于 *NoPE* 类 $\mathcal{F}_{\mathrm{MP}}$。

推论 1. 在假设 *1* 和 *2* 下（并采用上述相同的随机正则图／扩散几何设定），考虑我们骨干架构的 *NoPE* 和距离编码（*DE*）变体：*(i)* $\mathcal{F}_{\mathrm{NoPE}}$，不含位置编码的骨干网络；*(ii)* $\mathcal{F}_{\mathrm{DE}}$，通过 $\zeta(\cdot\mid S)$ 并使用$k=\Theta(\log n)$ 个锚点增强的同一骨干网络。那么，以高概率而言，对于 $G\sim\mathcal{G}_{n,\,r}$ 和随机锚点 S，$\mathcal{F}_{\mathrm{DE}}$的表达能力严格强于 $\mathcal{F}_{\mathrm{NoPE}}$；具体而言，存在一些节点分类任务，某个经距离编码增强的实例可以实现，而任何 *NoPE* 实例都无法实现。

推论 1 由定理 7 以及 NoPE 消息传递图神经网络与 1-WL 测试集等价这一结论推出（2019；2018）。结合定理 5–6，可得随机正则图上的表达能力链

$$\text{NoPE}\ \subsetneq\text{DE 增强}\ \approx\text{LapPE 增强}, \quad (29)$$

其中，$\approx$ 的含义与上述重构界和弗罗贝尼乌斯间隙界一致。

5 实验设置

我们在统一的、基于 GNP 的药物-药物相互作用预测骨干网络下，对位置编码进行受控比较。除非另有说明，所有实现和超参数细节均遵循（2025b），详见附录 B。

数据集与协议。 我们在 Drug- Bank 和 ChCh-Miner 上进行评估，其中节点是药物，边表示已知的相互作用。我们使用 RDKit 根据 SMILES 构建分子图。我们采用基线方法中的归纳式链接预测协议，将药物对划分为训练集/验证集/测试集。

基线方法。 为单独考察位置编码的影响，我们固定骨干网络，并比较三个主要变体：NoPE、距离编码（DE）和拉普拉斯位置编码（LapPE）。此外，在相同的骨干网络和训练协议下，我们还报告 RWSE 和热核特征（HKS），作为参考性的位置/结构基线方法。

模型配置。 所有模型均使用 PyTorch 和 PyTorch Geometric 实现，并采用标准的原子/键特征。LapPE 使用归一化拉普拉斯矩阵的前 m 个非平凡特征向量；DE 使用到 k 锚点的最短路径距离，并应用径向变换 $\psi(\cdot)$。RWSE 使用步数为 $\{1,2,4,8,16\}$ 时的返回概率，HKS 则使用截断拉普拉斯谱中的扩散时间 $\{0.1,0.5,1,2,5\}$。架构、归一化和优化设置见附录 B。

评估。药物-药物相互作用预测被视为二分类任务；我们报告留出测试集上的 AUROC 和 F1 分数，其中 F1 阈值根据验证集选定。主要结果（NoPE/DE/LapPE）为三个随机种子的平均值；RWSE/HKS 则根据可用的运行结果报告。

消融实验。 我们通过在固定锚点数量的情况下改变 DrugBank 上的$\psi(\cdot)$，以及在固定 $\psi(\cdot)$的情况下改变 ChCh-Miner 上的锚点数量 k，来探究距离编码的设计。所有消融实验均沿用

表1：随机 r-正则图上的验证（r=6、t=1.0、 m=8；三个随机种子）。

（a）链接和矩阵级偏差。

n	R	$\hat{\delta}_L^n$（均值±标准差）	Frob-差距（均值±标准差）
256	6	0.1580 ± 0.0035	0.0350 ± 0.0009
512	7	0.1099 ± 0.0051	0.0290 ± 0.0028
1024	7	0.0855 ± 0.0040	0.0217 ± 0.0014
2048	8	0.0664 ± 0.0024	0.0156 ± 0.0002

（b）逐点重构和条件数。

n	R	$\|\Phi - T\|_2$（中位数）	cond(A)（中位数）
256	6	0.449	46.89
512	7	0.757	112.41
1024	7	0.181	63.05
2048	8	0.0845	30.00

主要的预处理和训练设置；详细结果见第 6.4 节。

6 实验结果与分析

我们从三个角度评估所提出的 DE-LapPE 桥接方法：(i) 在随机 r-正则图上进行理论对齐验证（第4节）；(ii) 在真实 DrugBank 分子图上进行受控谱近似；以及 (iii) 在固定 GNP 骨干网络（NoPE/DE/LapPE）下进行下游药物-药物相互作用预测，并针对 $\psi(\cdot)$ 和锚点数量 k开展有针对性的 DE 消融实验。

6.1 随机 r-正则图上的理论对齐验证

我们在 $G \sim \mathcal{G}_{n,r}$ 上开展受控研究，使用 $r = 6$ 和 $n \in \{256, 512, 1024, 2048\}$（每个 n 使用三个随机种子）。对于每个图，我们构造 $L = I - \frac{1}{r}A$，并计算截断扩散坐标 $\Phi_t^{(m)}$，其中使用 $t = 1.0$ 和 $m = 8$。我们将局部性半径设为 $R = \lceil \log n \rceil$，并使用保序回归，根据 SPD$(u, v) \le R$ 的点对，拟合从 SPD 到 $d_t^{(m)}$ 的单调映射 ψ。

我们报告：(i) 经验链接误差 $\hat{}\delta_L^n := \max_{\mathrm{SPD}(u,v)\le R} |d_t^{(m)}(u, v) - \psi(\mathrm{SPD}(u, v))|$；(ii) 节点-锚点矩阵上的归一化 Frobenius 差距 $\|D^{(m)}{}_{\mathrm{diff}} - \psi_*(D_{\mathrm{SPD}})\|_F / \sqrt{n(m+1)}$；以及 (iii) 使用 $m+1$ 个锚点通过 $T(v)$进行三边测量重建，并以逐点误差中位数 $\|\Phi_t^{(m)}(v) - T(v)\|_2$和 mediancond($A$)进行汇总。

$\hat{}$ 表1显示，随着 n 增大， δ_L^n 和归一化弗罗贝尼乌斯间隙均减小，而三边测量重建误差在典型情况下仍然较小。

6.2 DrugBank 上扩散几何的谱近似

我们分析了80个DrugBank分子图（15–200个节点）。对于每个图，我们计算扩散核（$t = 1$）和扩散映射（前 $m = 8$个分量），并使用 $k = 32$个锚点（最远点采样）和蒂霍诺夫正则化，通过基于锚点的 DE Nyström 方案近似该核。

Nyström-DE 精度较高：核的相对 Frobenius 误差 0.024 ± 0.021（中位数 0。020），Procrustes对齐后的坐标 MSE 3.9×10^{-4}，以及近似扩散距离与精确扩散距离之间的平均绝对皮尔逊相关系数为 0。988。

图 3 进一步展示了拟合的局部保序链接函数 ψ^*（含 $R = \lceil \log n \rceil$）及条件数诊断结果，显示链接值/间隙值较小，而条件数呈重尾分布，这也说明了正则化的必要性。

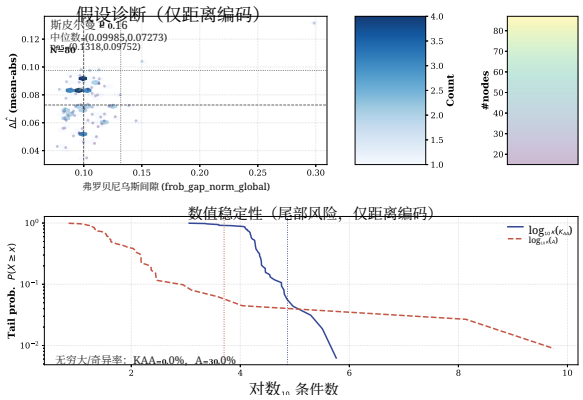


图 3：**DrugBank 扩散几何诊断（80 个图）**。**上：** 归一化弗罗贝尼乌斯间隙与单调性残差 ΔL（中位数/p95 和斯皮尔曼 ρ）。**下：** $\log_{10}\kappa(K_{AA})$ 和 $\log_{10}\kappa(A)$ 的尾部互补累积分布函数。

6.3 位置编码对药物-药物相互作用预测的影响

在相同的 GNP-DDI 骨干网络、训练计划和数据划分下，我们比较 NoPE、DE、LapPE 和两种参考基线方法（RWSE、HKS）；唯一变化的是位置编码模块。表 2 报告了测试集 AUROC 和 F1 分数（三次运行的均值±标准差）。

LapPE 整体表现最佳。在 DrugBank 上，位置信息至关重要：NoPE 的表现为 0.890/0.820（AUROC/F1），距离编码提升至 0.976/0.927，而 LapPE 则进一步提升至 0.980/0.934。在 ChCh-Miner 上，整体表现已相当出色，但与 NoPE（0.938/0.870）和距离编码（0.938/0.869）相比，LapPE 仍带来稳定提升（0.946/0.879）。R WSE 的表现

表 2：在相同的 GNP-DDI 骨干网络和训练方案下，DrugBank 和 ChCh-Miner 上的测试集 AUROC 和 F1 分数（三次运行的均值 ± 标准差）。仅改变位置编码。

数据集	方法	测试集 AUROC	测试集F1
DrugBank	NoPE	0.890 ± 0.002	0.820 ± 0.003
	距离编码	0.976 ± 0.002	0.927 ± 0.004
	LapPE	0.980 ± 0.003	0.934 ± 0.006
	RWSE	0.892 ± 0.008	0.812 ± 0.003
	HKS	0.863 ± 0.013	0.787 ± 0.007
ChCh-Miner	NoPE	0.938 ± 0.003	0.870 ± 0.002
	DE	0.938 ± 0.006	0.869 ± 0.004
	LapPE	0.946 ± 0.002	0.879 ± 0.003
	RWSE	0.944 ± 0.003	0.876 ± 0.002
	HKS	0.941 ± 0.002	0.870 ± 0.004

表 3：在 DrugBank 和 ChCh-Miner 上对距离编码（距离编码）的消融研究。每个数据集的最佳配置以阴影标出。

(a) DrugBank（$k = 16$，变化的 ψ（·））

ψ	Par.	验证集AUC	验证集F1	测试集AUC	测试集F1
恒等函数	1.168	0.9430	0.8795	0.9451	0.8835
exp	1.168	0.9762	0.9279	0.9760	0.9281
log1p	1.168	0.9723	0.9196	0.9721	0.9207

(b) ChCh-Miner（$\psi(d)$ = 指数函数（$-d$），变化的 k）

k	Par.	验证集AUC	验证集F1	测试集AUC	测试集F1
4	0.123	0.9415	0.8715	0.9410	0.8718
8	0.123	0.9411	0.8715	0.9411	0.8727
16	0.123	0.9406	0.8709	0.9412	0.8722
32		0.9455	0.8778	0.9453	0.8784

在 ChCh-Miner 上具有竞争力，而在此插件设置下，HKS 在 DrugBank 上的表现较弱。[1]

6.4 距离编码消融实验结果

在 DrugBank（表3a）上，使用 $k = 16$ 个锚点时，ψ（·）的选择至关重要：恒等映射表现较差（测试集 0.9451/0.8835 AUROC/F1 分数），而 $\psi =$ 指数函数表现最佳（测试集 0.9760/0.9281），log(1+d) 略逊一筹。参数数量相同（1.168M），因此性能提升来自距离塑形，而非容量。

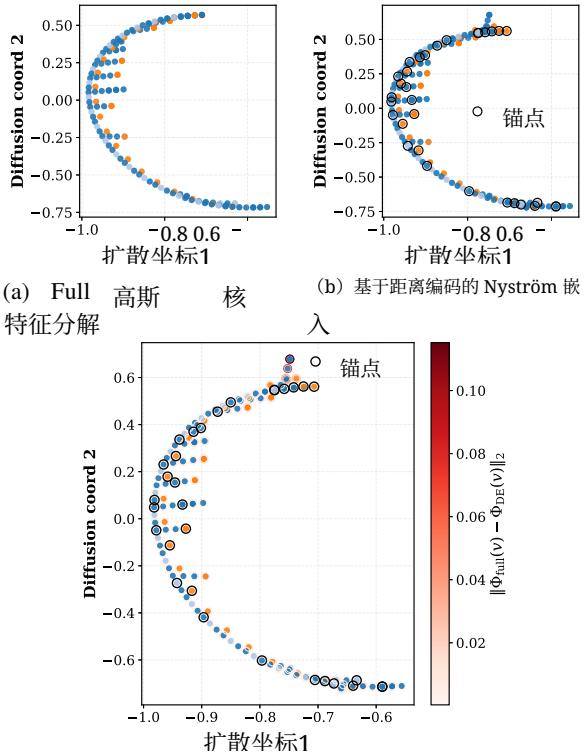
在 ChCh-Miner（表3b）上，固定 $\psi(d) = \exp(-d)$ 后，$k \in \{4, 8, 16\}$（测试集 $\approx 0.941/0.872$）的性能保持稳定，而在 $k = 32$（测试集 0.9453/0.8784）时有所提升，参数量仅略有增加（0.123M→0.124M）。

[1]RWSE 使用步数为 $\{1, 2, 4, 8, 16\}$ 时的返回概率；HKS 使用扩散时间 $\{0.1, 0.5, 1, 2, 5\}$，截断谱维度为 k=32（归一化拉普拉斯矩阵）。

6.5 DrugBank分子图的定性案例研究

我们选取一个 DrugBank 分子（DB00006，含 155 个原子）进行可视化，对比完整高斯核特征分解得到的参考扩散映射嵌入与基于 DE 的 Nyström 嵌入，后者使用 $k = 32$ 个锚点（经过 Procrustes 对齐）。如图 4(a-b) 所示，DE Nyström 能够很好地匹配全局扩散几何。

图 4(c) 报告了节点级误差 $\|\Phi_{\text{full}}(v) - \Phi_{\text{DE}}(v)\|_2$：大多数节点的误差处于 10^{-2} 量级或更低，均值为 7.8×10^{-3} 和最大值 1.15×10^{-1}。结合第6.2小节的汇总结果，这表明在实际应用中，数量适中的锚点集能够高保真地恢复主要扩散坐标。



(a) Full 高斯 核特征分解

(b) 基于距离编码的 Nyström 嵌入

(c) 距离编码 Nyström 的节点级嵌入误差

图 4：单个 DrugBank分子图（DB00006）的基于扩散的嵌入定性比较。

局限性

我们指出了理论和实验中的一些局限性与实际考量。

理论范围。 我们的保证是在随机正则图模型和扩散与最短路径距离之间的局部

链接下建立的。这些保证并未直接涵盖实践中常见的具有显著度异质性、明显社区结构、边权重或方向性的图。

近似设计选择。 误差对截断级别、锚点放置方式和径向变换的依赖关系目前仅得到部分刻画。除针对性的消融实验外，我们未分析数据驱动的锚点选择、学习得到的距离度量，或学习得到的（可能非单调的）距离变换。

经验普适性。 我们在两个药物-药物相互作用数据集和一种多尺度图神经过程骨干网络上进行了评估；结果可能无法完全推广至其他领域、标签或整理机制不同的数据集，或其他架构。

数据与评估注意事项。 药物-药物相互作用资源可能并不完整，且容易受到报告偏差和整理偏差的影响；被当作负例的未观测药物对中也可能包含假阴性。因此，标准数据划分下的性能可能无法充分反映真实世界的药物警戒场景。

潜在风险。 我们的药物-药物相互作用预测实验旨在用于研究和基准评测，而非临床决策。若误用或过度依赖预测结果（假阳性/假阴性），在未经专家审查和外部验证的情况下，可能会得出不恰当的结论。

未来方向。 位置编码被视为固定的预计算特征；将锚点、径向映射和谱正则化项联合进行端到端学习，我们留待未来工作。

人工智能助手的使用。 我们使用人工智能助手支持代码开发（例如调试和编写样板代码）。所有实验结果、分析和主张均由作者完成并验证。

致谢

作者谨向所有匿名审稿人致以诚挚谢意，感谢他们认真审阅并提出富有见地的建议。

参考文献

Uri Alon 和 Eran Yahav。2021。论图神经网络的瓶颈及其实际影响。发表于国际学习表征会议。

Mikhail Belkin 和 Partha Niyogi。2003年。用于降维和数据表征的拉普拉斯特征映射。*Neural Computation*, 15(6):1373–1396。

Michael M Bronstein、Joan Bruna、Yann LeCun、Arthur Szlam 和 Pierre Vandergheynst。2017年。几何深度学习：超越欧几里得数据。*IEEESignal Processing Magazine*，34(4):18–42。

Mengjie Chen、Ming Zhang 和 Cunquan Qu。2025年。迈向可解释的药物-药物相互作用预测：一种结合分子层面和网络层面解释的基于图的方法。载于第*31*届 *ACM SIGKDD* 知识发现与数据挖掘会议论文集（第*2*卷），第203–214页。Gautam Chinta、Jay Jorgenson 和 Anders Karlsson。2015年。正则图上的热核与广义 Ihara ζ 函数公式。*Monatshefte fürMathematik*，178(2):171–190。

Fan Chung 和丘成桐。1999年。覆盖、热核与生成树。《电子组合学杂志》，6(1):R12。

Ronald R Coifman 和 Stéphane Lafon。2006年。扩散映射。《应用与计算调和分析》，21(1):5–30。

Ronald R Coifman、Stephane Lafon、Ann B Lee、Mauro Maggioni、Boaz Nadler、Frederick Warner 和 Steven W Zucker。2005年。几何扩散作为调和分析和数据结构定义的工具：扩散映射。《美国国家科学院院刊》，102(21):7426–7431。

Yifan Deng、Xinran Xu、Yang Qiu、Jingbo Xia、Wen Zhang 和 Shichao Liu。2020。用于预测药物-药物相互作用事件的多模态深度学习框架。《生物信息学》，36(15):4316–4322。

VijayPrakash Dwivedi、ChaitanyaK Joshi、Anh Tuan Luu、Thomas Laurent、Yoshua Bengio 和 Xavier Bresson。2023年。图神经网络基准测试。《机器学习研究杂志》，24(43)：1–48。

Moshe Eliasof、Fabrizio Frasca、Beatrice Bevilacqua、Eran Treister、GalChechik 和 HaggaiMaron。2023年。通过随机特征传播实现图位置编码。发表于国际机器学习会议，第 9202–9223 页。PMLR。

Matthias Fey 和 Jan Eric Lenssen。2019。使用 PyTorch Geometric 进行快速图表示学习。*ICLR 2019* 图与流形上的表征学习研讨会论文集。

Justin Gilmer、Samuel S Schoenholz、Patrick F Riley、Oriol Vinyals 和 George E Dahl。2017年。用于量子化学的神经网络消息传递。发表于国际机器学习会议，第 1263–1272 页。Pmlr。

Chao He 和其他 1 位作者。2022。基于图神经网络的多类型特征融合，用于药物相互作用预测。《*BMC*生物信息学》，23(224)。

TN Kipf。2016年。使用图卷积网络进行半监督分类。*arXiv* 预印本 *arXiv:1609.02907*。

Pan Li、Yanbang Wang、Hongwei Wang 和 Jure Leskovec。2009。距离编码设计：用于结构表示学习的、更强大的图神经网络。*CoRR*，第 00142–00142 页。

Xuan Lin、Zhe Quan、Zhi-Jie Wang、Tengfei Ma 和 Xiangxiang Zeng。2020。KGNN：用于药物-药物相互作用预测的知识图神经网络。载于*IJCAI*，第 380 卷，第 2739–2745 页。

Ilya Loshchilov 和 Frank Hutter。2019。解耦权重衰减正则化。国际学习表征会议。ICLR。

Mei Ma 和 Xiujuan Lei。2023。一种基于分子结构和相互作用的药物-药物相互作用预测双图神经网络。*PLOS* 计算生物学，19(1):e1010812。

Sohir Maskey、Ali Parviz、Maximilian Thiessen、Hannes Stärk、Ylli Sadikaj 和 Haggai Maron。2022。用于图表示学习的广义拉普拉斯位置编码。*arXiv* 预印本 *arXiv:2210.15956*。

Christopher Morris、Martin Ritzert、Matthias Fey、William L Hamilton、Jan Eric Lenssen、Gaurav Rat-tan 和 Martin Grohe。2019 年。Weisfeiler和 Leman 走向神经网络：高阶图神经网络。载于*AAAI* 人工智能会议论文集，第 33 卷，第 4602–4609 页。

Ding Niu、Shuai Wang、Chunyu Li 和其他 1 位作者。2024 年。DAS-DDI：一种结合药物关联性与相似性的双视图框架，用于药物-药物相互作用预测。《生物医学信息学杂志》。即将刊出。

Adam Paszke、Sam Gross、Francisco Massa、Ada m Lerer、James Bradbury、Gregory Chanan、Trevor Killeen、Zeming Lin、Natalia Gimelshein、Luca Antiga 和其他 1 位作者。2019 年。PyTorch：一种命令式风格的高性能深度学习库。神经信息处理系统进展，第 32 卷。

Ladislav Rampášek、Mikhail Galkin、Vijay Prakash Dwivedi、Anh Tuan Luu、Guy Wolf 和 Dominique Beaini。2022。通用、强大且可扩展的图变换器方案。载于神经信息处理系统进展。

Jae Yong Ryu、Hyun Uk Kim 和 Sang Yup Lee。2018。深度学习提高了对药物-药物及药物-食物相互作用的预测能力。《美国国家科学院院刊》，115(18)：E4304–E4311。

Jian Sun、Maks Ovsjanikov 和 Leonidas J. Guibas。2009。基于热扩散的简洁且经严格证明具有信息性的多尺度特征。载于《计算机图形论坛》，第28卷，第1383–1392页。

James Topping、Francesco Di Giovanni、Benjamin P Chamberlain、Xiaowen Dong 和 Michael M Bronstein。2022。通过曲率理解图上的过压缩与瓶颈问题。载于 国际学习表征会议。

Ulrike Von Luxburg。2007。谱聚类教程。*Statistics and computing*，17(4):395–416。

Jiarui Wang 等人。2024。Structnet-ddi：基于分子结构表征的药物-药物相互作用预测表示学习。*Molecules*，29(20)：4829。

Zonghan Wu、Shirui Pan、Fengwen Chen、Guodong Long、Chengqi Zhang 和 Philip S Yu。2020。《图神经网络综合综述》。*IEEE* 神经网络与学习系统汇刊，32(1):4–24。

Keyulu Xu、Weihua Hu、Jure Leskovec 和 Stefanie Jegelka。2018。图神经网络的表达能力有多强？*arXiv* 预印本 *arXiv:1810.00826*。

Zimo Yan、Zheng Xie、Chang Liu 和 Yuan Wang。2025a。通过拉普拉斯谱编码解决图神经过程中的节点可识别性问题。*arXiv* 预印本 *arXiv:2511.19037*。

Zimo Yan、Zheng Xie、Jie Zhang、Yiping Song 和 Hao Li。2025b。一种用于药物-药物相互作用预测的跨药物协同注意力多尺度图神经过程。*arXiv* 预印本 *arXiv:2509.15256*。

Zimo Yan、Jie Zhang、Zheng Xie、Chang Liu、Yizhen Liu 和 Yiping Song。2025c。Metamolgen：用于从头分子设计的神经图模体生成模型。*arXiv* 预印本 *arXiv:2504.15587*。

Zhitao Ying、Tianjun Cai、Shengjie Luo、Shuxin Zheng、Guolin Ke、Di He、Yanming Shen 和 Tie-Yan Liu。2021。Transformer 在图表征任务中的表现真的很差吗？载于神经信息处理系统进展，第 34 卷，第 28877–28888 页。

Jiaxuan You、Rex Ying、Xiang Ren、William L Hamilton 和 Jure Leskovec。2019。用于图级任务的位置感知图神经网络。载于第 *36* 届国际机器学习会议论文集，第 7134–7143 页。PMLR。位置感知图神经网络的图级变体。

Jie Zhou、Ganqu Cui、Shengding Hu、Zhengyan Zhang、Cheng Yang、Zhiyuan Liu、Lifeng Wang、Changcheng Li 和 Maosong Sun。2020。图神经网络：方法与应用综述。*AI open*，1：57–81。

MarinkaZitnik、Monica Agrawal 和 Jure Leskovec。2018。利用图卷积网络建模多重用药副作用。《生物信息学》，34(13):i457–i466。

A 距离编码与拉普拉斯谱坐标之间的代数关系

在本附录中，我们对第4节中陈述的结果给出详细证明。这些结果通过一个显式的代数映射，将基于最短路径距离的距离编码（DE）与拉普拉斯算子谱坐标联系起来。

全文均使用第3节和第4节中引入的记号与基本假设。特别地，$G=(V,E)$ 是一个有限的、简单的、连通的无向图，L 是第3节中定义的归一化拉普拉斯矩阵，截断谱嵌入 $\Phi_t^{(m)}$ 和截断扩散距离 $d_t^{(m)}$ 均已在其中给出，并且随机正则模型、单调关联、锚点一般位置和谱单射性假设即为假设1–2。

我们还使用定义 1 中引入的 DE-to-LapPE 三边测量算子 $T(\cdot)$，现为方便起见简要回顾如下。

A.1 三边测量算子与距离矩阵回顾

固定 $t>0$ 和 $m\in\mathbb{N}$，令 a_1、……、$a_{m+1}\in V$ 为满足以下条件的锚点：

$$p_i=\Phi_t^{(m)}(a_i)\in\mathbb{R}^m,\quad i=1,\dots,m+1. \tag{30}$$

$m\times m$ 矩阵 A 以及映射 $b\colon\mathbb{R}^{m+1}\to\mathbb{R}^m$ 为

$$A=2\begin{pmatrix}(p_1-p_{m+1})^\top\\ \vdots\\ (p_m-p_{m+1})^\top\end{pmatrix}, \tag{31}$$

$$b(r)=\begin{pmatrix}\|p_1\|_2^2-\|p_{m+1}\|_2^2+r_{m+1}^2-r_1^2\\ \vdots\\ \|p_m\|_2^2-\|p_{m+1}\|_2^2+r_{m+1}^2-r_m^2\end{pmatrix}, \tag{32}$$

对于任意 $r=(r_1,\dots,r_{m+1})^\top\in\mathbb{R}^{m+1}$。

设 $\psi\colon[0,R]\to\mathbb{R}_+$ 为定理 2 中的严格递增函数。对于向量 $y=(y_1,\dots,y_{m+1})^\top\in\mathbb{R}^{m+1}$，我们逐元素应用 ψ，并记为

$$\big(\psi_*(y)\big)_i=\psi(y_i),\quad i=1,\dots,m+1. \tag{33}$$

对于节点 $v\in V$ 和锚点集 $S=\{a_1,\dots,a_{m+1}\}$，回顾距离编码

$$\zeta(v\mid S)=\big(\mathrm{SPD}(a_1,v),\dots,\mathrm{SPD}(a_{m+1},v)\big)^\top. \tag{34}$$

当且仅当 A 可逆时，定义 1 规定

$$T(v)=A^{-1}b\big(\psi_*(\zeta(v\mid S))\big)\in\mathbb{R}^m. \tag{35}$$

此外，节点到锚点的最短路径距离矩阵和截断扩散距离矩阵为

$$\big(D_{\mathrm{SPD}}\big)_{v,i}=\mathrm{SPD}(v,a_i), \tag{36}$$

$$\big(D_{\mathrm{diff}}^{(m)}\big)_{v,i}=d_t^{(m)}(v,a_i)=\big\|\Phi_t^{(m)}(v)-\Phi_t^{(m)}(a_i)\big\|_2, \tag{37}$$

我们将逐元素应用 ψ 得到的矩阵记作 $\psi_*(D_{\mathrm{SPD}})$。

A.2 定理 2 的证明

我们首先回顾随机正则图上的一个标准对数深度探索窗口。Li-Wang-Wang-Leskovec（2009）在距离编码分析中证明并将此结果作为关键输入；此处我们将其作为黑箱结论引入，以确定一个具体的 $R=\Theta(\log n)$ 情形，在该情形下，局部邻域呈现类树扩张。

引理 8(Li 等人（2009）)。固定一个整数 $r\geq 3$，并令 $G\sim\mathcal{G}_{n,\,r}$。存在一个充分小的常数 $\epsilon>0$，使得以概率 $1-o(n^{-3/2})$，以下结论成立。选取一个根 $u\in V$，并定义 *BFS* 层集合

$$Q_k:=\{v\in V:\ \mathrm{SPD}(u,v)=k\}. \tag{38}$$

令 p_k 表示在配置模型的揭示过程中，第 $(k+1)$ 个 *BFS* 步骤开始时，与 Q_k 相接且仍未匹配的边界半边数量。那么，对于每个整数

$$k\in\Big(\frac{\epsilon}{5}\cdot\frac{\log n}{\log(r-1)}+1,\ \Big(\frac{2}{3}-\epsilon\Big)\cdot\frac{\log n}{\log(r-1)}\Big), \tag{39}$$

有

$$|Q_k|\geq(r-1-\epsilon)^{k-1}\quad\textit{and}\quad p_k\geq(r-1-\epsilon)\,|Q_k|. \tag{40}$$

引理 9。设 T_r 为无限 r-正则树，且设 $t>0$。令 L_{T_r} 表示 T_r 上的（归一化）图拉普拉斯算子，$K_t:=e^{-tL_{T_r}}$ 表示其热半群，其热核为 $k_t(x,y):=(K_t)^{xy}$。那么存在一个函数 $h_t:\mathbb{N}_0\to(0,\infty)$，使得

$$k_t(x,y)=h_t(d(x,y)),\quad d(x,y):=\mathrm{SPD}_{T_r}(x,y). \tag{41}$$

若在 T_r 上定义扩散距离为

$$d_t(x,y)^2 := k_{2t}(x,x) + k_{2t}(y,y) - 2k_{2t}(x,y), \tag{42}$$

则 $d_t(x,y) = \psi_{\text{tree}}(d(x,y))$ 与

$$\psi_{\text{tree}}(d) := \sqrt{2\big(h_{2t}(0) - h_{2t}(d)\big)}. \tag{43}$$

此外，对于固定的 $t>0$， $h_{2t}(d)$ 关于 d严格递减，因此 ψ_{tree} 严格递增。

证明。由于 T_r 是距离传递的，对于满足 $d(x,y) = d(x',y')$ 的任意两对点 (x,y) 和 (x',y')，存在 T_r 的一个自同构 π，使得$\pi(x) = x'$ 且 $\pi(y) = y'$。拉普拉斯算子L_{T_r} 在自同构下保持不变，因此热半群 $K_t = e^{-tL_{T_r}}$ 与 π可交换。将这一不变性写到核的层面可得$k_t(x,y) = k_t(\pi(x)\ \pi(y)) = k_t(x',y')$，这意味着存在一个径向函数 h_t，使得$k_t(x,y) = h_t(d(x,y))$。正则树上 h_t 的完整推导和显式公式见（2015年）；基于覆盖的方法对无限正则树与有限正则图上的热核之间的关系进行了处理，见（1999年）。

于是，扩散距离恒等式可通过代数运算得到。由于 T_r 是顶点传递的，$k_{2t}(x,\ x) =k_{2t}(y,\ y) = h_{2t}(0)$，而径向性给出 $k_{2t}(x,\ y) =h_{2t}(d(x,\ y))$。代入定义可得

$$\begin{aligned} d_t(x,y)^2 &= h_{2t}(0) + h_{2t}(0) - 2h_{2t}(d(x,y)) \\ &= 2\big(h_{2t}(0) - h_{2t}(d(x,y))\big), \end{aligned} \tag{44}$$

因此，$d_t(x,\ y) = \psi_{\text{tree}}(d(x,\ y))$，其中 ψ_{tree} 如前所述。对于固定的 d，$h_{2t}(d)$ 关于 $t>0$ 的严格单调性已在上文引用的正则树热核分析中得到证明，这意味着 ψ_{tree} 严格递增。□

现在我们证明定理 2。除了定理中所述的几何比较假设外，唯一非平凡的步骤是量化将扩散距离截断至前 m 个非平凡特征模态所产生的影响。我们将这一截断计算直接纳入证明。

定理 *2* 的证明。固定 $t>0$ 和 $m \in \mathbb{N}$。设 $G=(V,E)$ 是一个有限连通图，其归一化拉普拉斯矩阵 L 为对称矩阵。设 $(\lambda_j,\ \phi_j)_{j=1}^n$ 是一个标准正交特征基，其中 $0=\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$。对于 $t>0$，定义完整扩散距离

$$d_t(u,v)^2 := \sum_{j=2}^{n} e^{-2t\lambda_j}\big(\phi_j(u) - \phi_j(v)\big)^2, \tag{45}$$

以及截断扩散距离

$$d_t^{(m)}(u,v)^2 := \sum_{j=2}^{m+1} e^{-2t\lambda_j}\big(\phi_j(u) - \phi_j(v)\big)^2. \tag{46}$$

定义谱尾能量

$$\text{Tail}_t^{(m)}(u,v) := \left(\sum_{j=m+2}^{n} e^{-2t\lambda_j}\big(\phi_j(u) - \phi_j(v)\big)^2\right)^{1/2}. \tag{47}$$

任取一对 $u,v \in V$，使得 $\text{SPD}(u,v) \le R$。我们首先通过一个精确分解，将 $d_t^{(m)}(u,v)$ 与 $d_t(u,v)$ 联系起来。将定义 $d_t(u,v)^2$ 的求和在索引 $m+1$ 处分开，得到

$$\begin{aligned} d_t(u,v)^2 = &\sum_{j=2}^{m+1} e^{-2t\lambda_j}\big(\phi_j(u) - \phi_j(v)\big)^2 \\ &+ \sum_{j=m+2}^{n} e^{-2t\lambda_j}\big(\phi_j(u) - \phi_j(v)\big)^2. \end{aligned} \tag{48}$$

根据定义，第一个求和等于 $d_t^{(m)}(u,v)^2$，第二个求和等于 $\text{Tail}_t^{(m)}(u,v)^2$,，因此

$$d_t(u,v)^2 = d_t^{(m)}(u,v)^2 + \text{Tail}_t^{(m)}(u,v)^2. \tag{49}$$

令 $a := d_t^{(m)}(u,v)^2 \ge 0$，且 $b := \text{Tail}_t^{(m)}(u,v)^2 \ge 0$。那么 $d_t(u,v) = \sqrt{a+b}$，且$d_t^{(m)}(u,v) = \sqrt{a}$，所以

$$\begin{aligned} d_t(u,v) - d_t^{(m)}(u,v) &= \sqrt{a+b} - \sqrt{a} \\ &= \frac{(a+b)-a}{\sqrt{a+b}+\sqrt{a}} = \frac{b}{\sqrt{a+b}+\sqrt{a}}. \end{aligned} \tag{50}$$

由于 $\sqrt{a+b}+\sqrt{a} \ge \sqrt{a+b} \ge \sqrt{b}$, 我们得到

$$\begin{aligned} 0 \le d_t(u,v) - d_t^{(m)}(u,v) &= \frac{b}{\sqrt{a+b}+\sqrt{a}} \le \frac{b}{\sqrt{b}} \\ &= \sqrt{b} = \text{Tail}_t^{(m)}(u,v), \end{aligned} \tag{51}$$

这意味着

$$\big|d_t^{(m)}(u,v) - d_t(u,v)\big| \le \text{Tail}_t^{(m)}(u,v). \tag{52}$$

现在，我们将 $d_t^{(m)}(u,\ v)$ 与 $\psi(\text{SPD}(u,\ v))$ 进行比较。根据三角不等式，

$$\begin{aligned} \big|d_t^{(m)}(u,v) - \psi(\text{SPD}(u,v))\big| \le &\big|d_t^{(m)}(u,v) - d_t(u,v)\big| \\ &+ \big|d_t(u,v) - \psi(\text{SPD}(u,v))\big|. \end{aligned} \tag{53}$$

在定理陈述中的截断假设所给出的高概率事件上，第一项满足

$$\begin{aligned}\left|d_t^{(m)}(u,v)-d_t(u,v)\right| &\le \mathrm{Tail}_t^{(m)}(u,v)\\ &\le \sup_{\mathrm{SPD}(x,y)\le R}\mathrm{Tail}_t^{(m)}(x,y)\le \delta_n^{\mathrm{trunc}}.\end{aligned} \tag{54}$$

在定理陈述中的几何比较假设所给出的高概率事件上，第二项满足

$$\left|d_t(u,v)-\psi(\mathrm{SPD}(u,v))\right|\le \delta_n^{\mathrm{geom}}. \tag{55}$$

取两个事件的交集并结合两个不等式，可得

$$\left|d_t^{(m)}(u,v)-\psi(\mathrm{SPD}(u,v))\right|\le \delta_n^{\mathrm{trunc}}+\delta_n^{\mathrm{geom}}=\delta_n^{L}. \tag{56}$$

由于这些界在同一个交集事件上对所有满足 $\mathrm{SPD}(u,\ v)\ \le\ R$ 的点对都一致，因此结论以高概率在$G\sim\mathcal{G}_{n,\ r}$上成立。 □

我们将在下一个重构定理的证明中使用定理 2。

A.3 命题 3 的证明

命题 *3* 的证明。设 $a_1,\ \ldots,\ a_{m+1}$ 是在 V 上独立同分布的均匀随机变量，并令 $p_i := \Phi_t^{(m)}(a_i)\in\mathbb{R}^m$。定义 $M=[p_1-p_{m+1}\ \cdots\ p_m-p_{m+1}]\in\mathbb{R}^{m\times m}$。

由于 $(a_1,\ldots,a_{m+1})$ 和 $(V_1,\ldots,V_{m+1})$ 具有相同的分布，根据定义 **(??)** 可得

$$\mathbb{P}\big(\det(M)=0\big)=\eta_n. \tag{57}$$

因此 $\mathbb{P}(\det(M)\ \neq 0)\ =1-\eta_n=1-o(1)$。

接下来假设采用抖动构造。固定 G 和锚点，从而固定 $p_1,\ldots,p_{m+1}$。令 $\varepsilon>0$，并令 $\xi_1,\ldots,\ \xi_{m+1}\in\mathbb{R}^m$ 为具有联合~密度的独立同分布随机变量，定义 $\tilde{p}_i := p_i+\varepsilon\xi_i$。令 $M=[\tilde{p}_1-\tilde{p}_{m+1}\ \cdots\ \tilde{p}_m-\tilde{p}_{m+1}]$。那么

$$\tilde{p}_i-\tilde{p}_{m+1}=(p_i-p_{m+1})+\varepsilon(\xi_i-\xi_{m+1}), \tag{58}$$

so

$$\tilde{M}=M+\varepsilon X,\quad X:=[\xi_1-\xi_{m+1}\ \cdots\ \xi_m-\xi_{m+1}]. \tag{59}$$

~因此，$\det(M)\ =\det(M+\varepsilon X)$ 是 $(\xi_1、\ldots\ldots、\xi_{m+1})$ 的各个条目所构成的多项式。该多项式并非恒等于零：事实上，选择一个确定性实现，使得 $\xi_{m+1}\ =\ 0$并令 $\xi_i\ =\ e_i$ 为 $i\ =$ 1、……、m，其中 $\{e_i\}$ 是 $\mathbb{R}^m$的标准基。于是 $X=I_m$ 并且

$$\det(M+\varepsilon X)=\det(M+\varepsilon I_m). \tag{60}$$

作为关于 ε 的多项式，$\det(M+\varepsilon I_m)$ 的首项为 ε^m，因此它不是零多项式，所以存在一些噪声值，使得 $\det(M+\varepsilon X)\ \neq 0$。因此，零集

$$Z:=\{(\xi_1,\ldots,\xi_{m+1})\in(\mathbb{R}^m)^{m+1}:\ \det(M+\varepsilon X)=0\} \tag{61}$$

是一个真代数簇，且其勒贝格测度为零。由于 $(\xi_1,\ldots,\xi_{m+1})$ 具有联合密度，

$$\mathbb{P}\big((\xi_1,\ldots,\xi_{m+1})\in Z\mid G,a_1,\ldots,a_{m+1}\big)=0, \tag{62}$$

~这等价于 $\mathbb{P}(\det(M)\neq\ \ 0\ \ \mid G,a_1,\ldots,\ a_{m+1})\ =1$。这证明了抖动下的仿射独立性。 □

A.4 定理 5 的证明

定理 *5* 的证明。固定 $G\ \ \sim\ \ \mathcal{G}_{n,r}$ 个锚点$a_1,\ldots,\ a_{m+1}$，以及一个节点 $v\ \ \in\ \ V$，使得对所有 i，$\mathrm{SPD}(a_i,v)\ \le R$。记

$$z=\Phi_t^{(m)}(v)\in\mathbb{R}^m, p_i=\Phi_t^{(m)}(a_i)\in\mathbb{R}^m, \tag{63}$$

以及其相对于 S 的距离编码为

$$y_i=\mathrm{SPD}(a_i,v),\quad y=\zeta(v\mid S)=(y_1,\ldots,y_{m+1})^\top. \tag{64}$$

步骤 1：从最短路径距离到近似半径。令 $y_i:=\mathrm{SPD}(a_i,v)$。根据定理 2，存在一个严格递增函数 ψ 和一个序列 $\delta_n^L\ =\ o(1)$，使得以高概率，对所有满足 $y_i\le R$ 的 i，

$$\left|d_t^{(m)}(a_i,v)-\psi(y_i)\right|\le\delta_n^L. \tag{65}$$

定义

$$r_i:=\psi(y_i),\quad r:=(r_1,\ldots,r_{m+1})^\top. \tag{66}$$

为便于后续使用，定义精确截断半径

$$r_i^\star:=d_t^{(m)}(a_i,v),\quad i=1,\ldots,m+1, \tag{67}$$

以及相应的误差

$$\zeta_{i,n}:=r_i^\star-r_i, \tag{68}$$

从而

$$r_i^\star=r_i+\zeta_{i,n},\quad |\zeta_{i,n}|\le\delta_n^L. \tag{69}$$

步骤 2：使用 $\{r_i^\star\}$根据 $d_t^{(m)}$ 的定义，对于每个锚点 a_i，我们有

$$d_t^{(m)}(a_i, v) = \left\|\Phi_t^{(m)}(a_i) - \Phi_t^{(m)}(v)\right\|_2 = \left\|p_i - z\right\|_2. \tag{70}$$

将两边平方，得到

$$\left\|z - p_i\right\|_2^2 = \left(r_i^\star\right)^2. \tag{71}$$

利用恒等函数

$$\left\|z - p_i\right\|_2^2 = \|z\|_2^2 - 2\langle z, p_i\rangle + \|p_i\|_2^2, \tag{72}$$

这就变为

$$\|z\|_2^2 - 2\langle z, p_i\rangle + \|p_i\|_2^2 = \left(r_i^\star\right)^2. \tag{73}$$

这些是关于未知量 $z \in \mathbb{R}^m$ 和标量 $\|z\|_2^2$ 的 $m+1$ 个方程。

为消去 $\|z\|_2^2$，我们用对应一般索引 $i \in \{1, \ldots, m\}$ 的方程减去对应索引 $m+1$ 的方程。对于每个这样的 i，我们考虑

$$\begin{aligned}&\left(\|z\|_2^2 - 2\langle z, p_i\rangle + \|p_i\|_2^2\right) - \left(\|z\|_2^2 - 2\langle z, p_{m+1}\rangle + \|p_{m+1}\|_2^2\right)\\ &= \left(r_i^\star\right)^2 - \left(r_{m+1}^\star\right)^2.\end{aligned} \tag{74}$$

在左侧，我们逐项计算

$$\begin{aligned}&\|z\|_2^2 - 2\langle z, p_i\rangle + \|p_i\|_2^2 - \|z\|_2^2 + 2\langle z, p_{m+1}\rangle - \|p_{m+1}\|_2^2\\ &= -2\langle z, p_i\rangle + 2\langle z, p_{m+1}\rangle + \|p_i\|_2^2 - \|p_{m+1}\|_2^2.\end{aligned} \tag{75}$$

将内积因式分解后得到

$$-2\langle z, p_i\rangle + 2\langle z, p_{m+1}\rangle = 2\langle z, p_{m+1} - p_i\rangle = -2\langle z, p_i - p_{m+1}\rangle. \tag{76}$$

因此

$$-2\langle z, p_i - p_{m+1}\rangle + \|p_i\|_2^2 - \|p_{m+1}\|_2^2 = \left(r_i^\star\right)^2 - \left(r_{m+1}^\star\right)^2. \tag{77}$$

两边乘以 -1，得到

$$2\langle z, p_i - p_{m+1}\rangle = \|p_i\|_2^2 - \|p_{m+1}\|_2^2 + \left(r_{m+1}^\star\right)^2 - \left(r_i^\star\right)^2. \tag{78}$$

对于每个 $i = 1, \ldots, m$，这都是关于 z各分量的线性方程。

现在，我们将关于 $i = 1, \ldots, m$ 的这些方程叠放在一起。左侧的第 i 个元素为

$$2(p_i - p_{m+1})^\top z. \tag{79}$$

定义

$$A = 2\begin{pmatrix}(p_1 - p_{m+1})^\top\\ \vdots\\ (p_m - p_{m+1})^\top\end{pmatrix}. \tag{80}$$

那么，堆叠后的左侧恰好为 Az。在右侧，定义

$$b^\star = \begin{pmatrix}\|p_1\|_2^2 - \|p_{m+1}\|_2^2 + (r_{m+1}^\star)^2 - (r_1^\star)^2\\ \vdots\\ \|p_m\|_2^2 - \|p_{m+1}\|_2^2 + (r_{m+1}^\star)^2 - (r_m^\star)^2\end{pmatrix}. \tag{81}$$

m 个线性方程可以简洁地写为

$$Az = b^\star. \tag{82}$$

根据命题 3-**??**，向量 $p_1 - p_{m+1}, \ldots, p_m - p_{m+1}$ 在$\mathbb{R}^m$中线性无关。因此，$A/2$ 的各行线性无关，故 A 满秩 m 且可逆。因此存在唯一解

$$z = A^{-1}b^\star. \tag{83}$$

由于 $z = \Phi_t^{(m)}(v)$，这便将真实的谱坐标表示为精确截断半径 $r_i^\star$的函数。

步骤 3：关联精确右端项与近似右端项。实践中，我们使用 r_i 而非 $r_i^\star$来构造 $T(v)$。对应的右端项为

$$b(r) = \begin{pmatrix}\|p_1\|_2^2 - \|p_{m+1}\|_2^2 + r_{m+1}^2 - r_1^2\\ \vdots\\ \|p_m\|_2^2 - \|p_{m+1}\|_2^2 + r_{m+1}^2 - r_m^2\end{pmatrix}. \tag{84}$$

根据定义 1，基于 DE 的近似为

$$T(v) = A^{-1}b(r). \tag{85}$$

现在我们逐项比较 $b^\star$ 和 $b(r)$。对于$i = 1, \ldots, m$，有

$$\begin{aligned}b_i^\star - b(r)_i &= \left(\|p_i\|_2^2 - \|p_{m+1}\|_2^2 + (r_{m+1}^\star)^2 - (r_i^\star)^2\right)\\ &\quad - \left(\|p_i\|_2^2 - \|p_{m+1}\|_2^2 + r_{m+1}^2 - r_i^2\right)\\ &= (r_{m+1}^\star)^2 - (r_i^\star)^2 - r_{m+1}^2 + r_i^2.\end{aligned} \tag{86}$$

利用 $r_i^\star = r_i + \zeta_{i,n}$，我们计算得

$$\begin{aligned}(r_{m+1}^\star)^2 - r_{m+1}^2 &= (r_{m+1} + \zeta_{m+1,n})^2 - r_{m+1}^2\\ &= r_{m+1}^2 + 2r_{m+1}\zeta_{m+1,n} + \zeta_{m+1,n}^2 - r_{m+1}^2\\ &= 2r_{m+1}\zeta_{m+1,n} + \zeta_{m+1,n}^2,\end{aligned} \tag{87}$$

同理

$$\begin{aligned}(r_i^\star)^2 - r_i^2 &= (r_i + \zeta_{i,n})^2 - r_i^2\\ &= 2r_i\zeta_{i,n} + \zeta_{i,n}^2.\end{aligned} \tag{88}$$

代入这些表达式，我们得到

$$\begin{aligned}&(r^\star_{m+1})^2-(r^\star_i)^2-r^2_{m+1}+r^2_i\\&=\big((r^\star_{m+1})^2-r^2_{m+1}\big)-\big((r^\star_i)^2-r^2_i\big)\\&=\big(2r_{m+1}\zeta_{m+1,n}+\zeta^2_{m+1,n}\big)-\big(2r_i\zeta_{i,n}+\zeta^2_{i,n}\big).\end{aligned}\tag{89}$$

因此

$$b^\star_i-b(r)_i=2r_{m+1}\zeta_{m+1,n}+\zeta^2_{m+1,n}-2r_i\zeta_{i,n}-\zeta^2_{i,n}.\tag{90}$$

现在利用以下不等式对 $|b^\star_i-b(r)_i|$ 进行界定

$$|\zeta_{i,n}|\le\delta^L_n,\quad |\zeta_{m+1,n}|\le\delta^L_n.\tag{91}$$

由于 m 固定且嵌入维数有限，存在一个常数 $B_t>0$，使得

$$\|\Phi^{(m)}_t(u)\|_2\le B_t\tag{92}$$

对所有 $u\in V$ 均成立。因此，对所有 i,

$$r^\star_i=d^{(m)}_t(a_i,v)=\|\Phi^{(m)}_t(a_i)-\Phi^{(m)}_t(v)\|_2\le 2B_t,\tag{93}$$

并根据 $r_i=r^\star_i-\zeta_{i,\ n}$，我们有

$$|r_i|\le|r^\star_i|+|\zeta_{i,n}|\le 2B_t+\delta^L_n.\tag{94}$$

因此，对所有充分大的 n $\delta^L_n\le$（使得 1），存在一个常数 $\widetilde{B}_t>0$（仅取决于 $t,\ m,\ r$），使得

$$|r_i|\le\widetilde{B}_t,\quad |r_{m+1}|\le\widetilde{B}_t\quad\text{for all } i=1,\dots,m.\tag{95}$$

Using the triangle inequality, we obtain

$$\begin{aligned}|b^\star_i-b(r)_i|&\le 2|r_{m+1}|\,|\zeta_{m+1,n}|+\zeta^2_{m+1,n}+2|r_i|\,|\zeta_{i,n}|+\zeta^2_{i,n}\\&\le 2\widetilde{B}_t\delta^L_n+(\delta^L_n)^2+2\widetilde{B}_t\delta^L_n+(\delta^L_n)^2\\&=4\widetilde{B}_t\delta^L_n+2(\delta^L_n)^2.\end{aligned}\tag{96}$$

当 n 足够大时，$\delta^L_n\le 1$ 推出 $(\delta^L_n)^2\le\delta^L_n$，因此存在一个常数 $C_1>0$（仅取决于 $\widetilde{}$ 和 B_t），使得

$$|b^\star_i-b(r)_i|\le C_1\,\delta^L_n,\quad i=1,\dots,m.\tag{97}$$

合并这些 m 不等式可得

$$\|b^\star-b(r)\|_2\le\sqrt{m}\,C_1\,\delta^L_n.\tag{98}$$

步骤 4：线性系统的稳定性与结论。真实的谱坐标与基于距离编码的坐标满足

$$Az=b^\star,\quad AT(v)=b(r).\tag{99}$$

相减可得

$$A\big(z-T(v)\big)=b^\star-b(r).\tag{100}$$

由于 A 可逆，我们有

$$z-T(v)=A^{-1}\big(b^\star-b(r)\big).\tag{101}$$

取范数可得

$$\big\|z-T(v)\big\|_2\le\|A^{-1}\|_{\mathrm{op}}\,\big\|b^\star-b(r)\big\|_2.\tag{102}$$

结合步骤 3 中关于 $\|b^\star-b(r)\|_2$ 的界即可证明该结论。回顾 $z=\Phi^{(m)}_t\ v(\ )$，证明完毕。

□

A.5 定理 6 的证明

定理 6 的证明。我们证明截断扩散距离矩阵 $D^{(m)}{}_{\mathrm{diff}}$ 与变换后的最短路径矩阵 $\psi_*(D_{\mathrm{SPD}})$ 之间差异的 Frobenius 范数界。

回顾

$$\begin{aligned}(D_{\mathrm{SPD}})_{v,i}&=\mathrm{SPD}(v,a_i),\\(D^{(m)}_{\mathrm{diff}})_{v,i}&=d^{(m)}_t(v,a_i),\\(\psi_*(D_{\mathrm{SPD}}))_{v,i}&=\psi(\mathrm{SPD}(v,a_i)).\end{aligned}\tag{103}$$

根据定理 2，以高概率，对于所有节点 $v\in V$ 和锚点 a_i，当 $\mathrm{SPD}(v,a_i)\le R$ 时，

$$\Big|(D^{(m)}_{\mathrm{diff}})_{v,i}-(\psi_*(D_{\mathrm{SPD}}))_{v,i}\Big|\le\delta^L_n.\tag{104}$$

因此，

$$\begin{aligned}&\big\|D^{(m)}_{\mathrm{diff}}-\psi_*(D_{\mathrm{SPD}})\big\|^2_F\\&=\sum_{v\in V}\sum_{i=1}^{m+1}\Big((D^{(m)}_{\mathrm{diff}})_{v,i}-(\psi_*(D_{\mathrm{SPD}}))_{v,i}\Big)^2\\&\le\sum_{v\in V}\sum_{i=1}^{m+1}(\delta^L_n)^2=n(m+1)(\delta^L_n)^2.\end{aligned}\tag{105}$$

开平方得到

$$\big\|D^{(m)}_{\mathrm{diff}}-\psi_*(D_{\mathrm{SPD}})\big\|_F\le\delta^L_n\sqrt{n(m+1)}.\tag{106}$$

这就证明了定理 6。 □

B 实验细节

B.1 数据集

我们在两个广泛使用的药物-药物相互作用预测基准数据集 DrugBank 和 ChCh-Miner 上评估所有方法。两种情况下，我们均遵循先前 DDI 研究和 MPNP-DDI 框架（2025b）采用的归纳式链接预测协议。

DrugBank。DrugBank相互作用网络源自DrugBank 5.0数据库，并已广泛用于近期DDI 研究（2018；2024；2024）。图节点对应小分子药物，边表示有文献记录的相互作用。遵循标准预处理流程（2024；2024），我们得到一个包含约 1，700 种药物和 190，000 个带标签药物对的数据集，这些药物对在原始资源中被划分为 86 种相互作用类型。在本研究中，我们关注二元链接预测设定，并将每对药物视为存在或不存在相互作用。对于每种药物，我们使用 RDKit 根据其 SMILES字符串构建分子图，其中原子作为具有类别型和数值型属性（如原子类型、度、形式电荷和芳香性）的节点，化学键则作为具有键类型特征的边。

ChCh-Miner。ChCh-Miner数据集是一个中等规模的药物-药物相互作用网络，基于已获批准的药物整理而成，并作为 BioSNAP 数据集的一部分发布（2018）。节点代表药物，边表示已知的相互作用。该网络包含1,514 种药物和 48,514 条有记录的药物-药物相互作用链接（2024;2024）。与 DrugBank 一样，我们将药物-药物相互作用预测视为该药物-药物图上的二元链接预测问题。个别药物的分子图以与 DrugBank 相同的方式，根据 SMILES 构建。

数据划分。我们采用 MPNP-DDI（2025b）中的划分策略。所有已知相互作用均按药物对随机划分为训练集、验证集和测试集。负样本通过均匀抽样未观测到的药物对生成，并在各划分中保持固定的正负样本比例，这符合药物-药物相互作用链接预测的常见做法（2024;2022）。我们采用归纳式设置，其中一部分药物仅出现在验证集或测试集中，因此模型必须泛化到未见分子，而不是记忆特定药物对。

B.2 基线方法与模型变体

由于我们的重点是理解固定架构中位置编码的作用，因此在所有设置中保持骨干模型不变，仅改变位置编码模块。具体而言，我们考察三种主要变体（NoPE/DE/LapPE）以进行对照比较，并额外报告两种广泛使用的 PE 基线方法（RWSE 和 HKS），作为采用相同骨干模型和训练流程的合理性检查参考。

•**NoPE**。原始 MPNP-DDI 模型（2025b）不使用任何显式位置编码。节点特征仅由 RDKit 提取的原子级描述符构成，模型依靠其多尺度消息传递方案推断结构信息。

• **距离编码**。通过添加距离编码增强的骨干模型。对于每个分子图，我们采样 k 个锚点原子，并计算每个节点到这些锚点的最短路径距离。对每个距离应用径向变换 $\psi(\cdot)$，并将所得向量与原始节点特征拼接。所有实验共用同一个距离编码模块，其设计将在消融实验中详细研究。

• **LapPE**。通过添加拉普拉斯位置编码增强的骨干模型。对于每个分子图，我们计算归一化拉普拉斯算子的前 m 个非平凡特征向量，并按照谱图神经网络中的标准做法（2023），将其与节点特征拼接。我们选择 m，以匹配 DE 特征的维度，使所有变体具有相当的参数数量。

• **RWSE（参考基线方法）**。通过添加随机游走结构编码（RWSE）增强的骨干模型，其中为每个节点分配一个由 K步随机游走返回概率构成的向量，并按照图 Transformer 和位置编码基准测试中的常见做法（2022），将所得编码与节点特征拼接。我们使用的步数为$\mathcal{T} = \{1, 2, 4, 8, 16\}$。

• **HKS（参考基线）**。在骨干模型中加入热核特征（HKS）（2009），这些特征由归一化拉普拉斯矩阵的截断特征空间构造而成。具体而言，对于每个节点，我们使用时间 $\mathcal{S} = \{0.1, 0.5, 1, 2, 5\}$，基于前$m$ 个特征值对计算扩散时间嵌入，并将其与节点特征拼接。我们将截断特征空间的维度设为 m=32。

为确保比较公平，所有位置编码变体均作为节点特征的插件式增强，并使用相同的骨干网络、优化器和训练预算进行训练。在适用情况下，我们选择 PE

的维度，使其与 DE/LapPE 特征增强的规模相当。

所有模型均使用 PyTorch（2019）实现，并使用 PyTorch Geometric 高效执行图操作（2019）。除非另有说明，我们遵循 MPNP-DDI（2025b）的主要架构与优化方案。

B.2.1 特征归一化。
对于分子图，我们使用 RDKit 构建标准的原子特征和键特征，包括原子类型、度、价态、芳香性、杂化和键类型。类别属性采用独热向量编码，并与数值描述符拼接。对于所有连续特征，我们在训练集上进行逐特征标准化（均值为零、方差为一），并在验证集和测试集上沿用相同的统计量。

对于拉普拉斯位置编码，我们在每个分子图内对原始特征向量进行归一化，使每个特征向量的均值为零、方差为一。为降低对全局符号翻转的敏感性，我们会在训练过程中选择性地加入少量高斯噪声。对于距离编码，首先根据每个图内的非零距离中位数对最短路径距离进行缩放，然后将其输入径向映射 $\psi(\cdot)$。我们的实验考察了三种选择：$\psi(d) = d$、$\psi(d) = \exp(-d)$，以及$\psi(d) = \log(1+d)$。随后，在训练集上对变换后的距离进一步进行标准化。

B.2.2 架构设计。
骨干架构遵循 MPNP-DDI（2025b）的多尺度图神经过程设计。每种药物同时由其原始分子图及对应的线图表示，后者捕获键级交互。节点和边特征（包括存在时的位置编码）均被投影到 64 维的隐藏维度。模型堆叠三个图神经过程模块；每个模块在分子图和线图上分别执行两轮消息传递，随后进行聚合与跨尺度融合，从局部子结构到全局拓扑形成层次化表示。

对于一对药物，分块表示会被输入跨药物协同注意力模块，该模块根据每种药物的配对药物计算其上下文感知嵌入。随后，这些成对嵌入会经过一个多层感知机，以预测发生相互作用的概率。DE 和 LapPE 变体仅在额外的节点级输入上有所不同；包括协同注意力和解码器在内的所有后续处理均为共享。

B.2.3 优化设置。
我们使用 AdamW优化器（2019）训练所有模型，并按照（2025b）采用余弦退火学习率调度器，训练 50 个训练轮次。初始学习率和权重衰减通过在验证集上进行网格搜索确定，搜索范围为标准范围（例如，学习率介于10^{-4} 和 10^{-3} 之间，权重衰减介于 10^{-5}和 10^{-3}之间）。我们对预测的交互概率使用二元交叉熵损失。

根据验证集 AUROC 进行早停：如果连续 10 个训练轮次均未观察到改进，则终止训练并保留最佳检查点。所有实验均使用三个随机种子重复进行，并报告各次运行结果的均值和标准差。

B.2.4 批大小与梯度方差。
由于药物对数量庞大，且多尺度骨干网络的内存开销较高，我们采用包含 32 个药物对的迷你批次，并将梯度累积 4 步，从而得到 128 的有效批大小，与（2025b）中的做法一致。这可以降低随机梯度的方差，同时不超出 GPU 内存限制。我们还采用固定的最大范数进行梯度裁剪，以防止高度连接的分子图偶尔导致梯度爆炸。

B.3 评估指标

我们将药物-药物相互作用预测视为针对药物对的二元分类问题。对于 DrugBank 和 ChCh-Miner，我们使用受试者工作特征曲线下面积（AUROC）和 F1 分数评估模型，这些是药物-药物相互作用预测中的标准指标（2024；2022；2024）。AUROC 衡量模型在所有阈值下的排序质量，而 F1 分数则概括特定阈值下精确率与召回率之间的权衡。

在训练和模型选择期间，我们监测验证集上的 AUROC。对于 F1 分数，我们选择使验证集 F1 分数达到最大值的阈值，并将同一阈值应用于留出测试集。所有报告的数值均使用验证集 AUROC 最佳的检查点，在测试划分上计算，并取三次独立运行的平均值。

B.4 消融研究详情

我们开展了一项有针对性的消融研究，以单独考察距离编码模块中关键设计选择的影响。具体而言，我们考察了：(i) 在 DrugBank 上应用于最短路径距离的径向变换 $\psi(\cdot)$ 的选择，同时固定锚点数量；(ii) 在 ChCh-Miner 上的锚点数量 k，同时固定 $\psi(\cdot)$。

在 DrugBank 上，我们固定锚点数量，并比较三种径向函数：$\psi(d) = d$，$\psi(d) = \exp(-d)$，以及 $\psi(d) = \log(1 + d)$。在 ChCh-Miner 上，我们固定 $\psi(d) = \exp(-d)$ 并改变 $k \in \{4, 8, 16, 32\}$，以研究近似质量与模型复杂度之间的权衡。所有消融实验均采用与主要实验相同的数据预处理、模型架构、优化器、批大小和早停标准。详细的数值结果及进一步分析见第 6.4 节。有关更多实现细节（数据集、基线方法和超参数），请参阅

h

ttps://anonymous.4open.science/r/Bridging

-距离 -and-谱 -位置 编码-E48B。