

# 基于深度强化学习与梯度优化的接吻数问题求解研究

李预胜、李君迟、陈长、官天睿

2026 年 1 月 11 日

## 摘要

本文研究  $d$  维欧几里得空间中的经典几何难题——接吻数问题（Kissing Number Problem）。针对高维空间中搜索空间呈指数级增长及局部最优解密集的挑战，提出一种基于 Gram 矩阵重构的深度强化学习求解框架。首先，通过将搜索对象从欧氏坐标空间转化至内积空间，彻底消除旋转冗余性，实现解空间的有效压缩。其次，结合基于 Transformer 的策略与价值网络，利用蒙特卡洛树搜索（MCTS）在半正定锥流形上进行序列化决策。此外，引入“Self-Use”自进化机制，通过精英样本迭代自训练与动态回溯修复策略，显著提升了算法在高维流形中的探索效率。实验结果表明，该框架在  $1 \sim 8$  维空间均能稳定达到已知最优解，并在  $9 \sim 10$  维空间取得了远超梯度下降基线的搜索结果，验证了数据驱动方法在复杂几何堆积问题中的巨大潜力。

**关键词：**接吻数问题，梯度优化，深度强化学习，蒙特卡洛树搜索，几何堆积

## 1 问题阐述

接吻数问题是几何堆积理论中的一个经典问题。该问题旨在寻找在  $d$  维欧几里得空间  $\mathbb{R}^d$  中，能够同时与一个中心单位球体相切且彼此互不重叠的单位球体的最大数量。我们将这个最大数量记为  $\tau_d$ 。

形式化地，设中心球体为  $S_0$ ，其球心位于原点，半径为 1。我们需要寻找一组向量  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ ，其中  $\mathbf{x}_i \in \mathbb{R}^d$ ，满足以下两个条件：

- 接触条件：**所有外部球体与中心球体接触，即  $\|\mathbf{x}_i\| = 2R = 2$ （若考虑球心距离），或者在简化模型中，我们将所有向量归一化，使其位于单位球面  $\mathbb{S}^{d-1}$  上，即  $\|\mathbf{x}_i\| = 1$ 。
- 非重叠条件：**任意两个外部球体  $S_i$  和  $S_j$  之间不发生重叠。

在几何上，这等价于在  $d$  维单位球面  $\mathbb{S}^{d-1}$  上分布  $N$  个点，使得任意两点之间的欧几里得距离至少为 1（假设球半径为 0.5 归一化后的情况）。更直观地，我们需要使得任意两个向量  $\mathbf{x}_i, \mathbf{x}_j$  之间的夹角  $\theta_{ij}$  不小于  $60^\circ$ （即  $\pi/3$ ）。

数学描述为寻找最大的  $N$ ，使得存在点集  $\{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \mathbb{S}^{d-1}$  满足：

$$\mathbf{x}_i \cdot \mathbf{x}_j \leq \cos(60^\circ) = 0.5, \quad \forall i \neq j \quad (1)$$

## 2 背景介绍

接吻数问题的历史可以追溯到 1694 年艾萨克·牛顿与大卫·格里高利关于三维空间接吻数的著名争论。牛顿认为  $\tau_3 = 12$ ，而格里高利认为可能是 13。直到 1953 年，Schütte 和 van der Waerden 才给出了  $\tau_3 = 12$  的严谨证明。

随着维度的增加，问题的复杂度呈指数级上升。目前，只有少数维度的接吻数是确定的：

- $d = 1$ :  $\tau_1 = 2$

- $d = 2$ :  $\tau_2 = 6$
- $d = 3$ :  $\tau_3 = 12$
- $d = 4$ :  $\tau_4 = 24$
- $d = 8$ :  $\tau_8 = 240$  (源于  $E_8$  晶格)
- $d = 24$ :  $\tau_{24} = 196560$  (源于 Leech 晶格)

对于其他维度（如  $d = 5, 6, \dots$ ），目前数学界仅能给出下界和上界。

传统上，寻找接吻数下界主要依赖于代数构造或随机几何分析。然而，在高维空间中构造满足约束的构型极具挑战性。近年来，随着深度学习和强化学习在组合优化问题中的成功应用（例如 AlphaGo, AlphaTensor, AlphaEvolve），利用数据驱动的方法在巨大的解空间中进行启发式搜索成为可能。

通过将接吻数问题建模为优化问题或序列决策过程，机器学习算法可以在不依赖复杂代数先验的情况下，自动发现接近或达到已知最优解的几何构型，甚至可能发现新的下界。

### 3 基线方法：基于梯度的势能最小化

为了评估后续强化学习方法的有效性，我们首先构建了一个基于梯度的连续优化方法作为基线。该方法将接吻数问题视为物理系统中的**势能最小化问题**。

#### 3.1 核心思想

我们在  $d$  维单位球面上随机初始化  $N$  个点。如果任意两点之间的夹角小于  $60^\circ$ ，则认为它们之间存在“排斥力”。我们的目标是通过移动这些点的位置，使得系统的总 Violation Energy 降为零。如果最终能量为零，则说明找到了一个大小为  $N$  的有效接吻构型。

#### 3.2 算法实现

##### 3.2.1 初始化

基线算法采用 PyTorch 框架实现，我们生成  $N$  个服从标准正态分布的随机向量，并将其归一化到单位球面上：

$$\mathbf{x}_i \leftarrow \frac{\mathbf{z}_i}{\|\mathbf{z}_i\|}, \quad \text{其中 } \mathbf{z}_i \sim \mathcal{N}(0, I_d) \quad (2)$$

并将这些向量设定为可求导参数，以便利用自动微分计算梯度。

##### 3.2.2 损失函数定义

我们定义损失函数为所有违反角度约束的点对的余弦相似度误差之和。首先计算 Gram 矩阵  $G = XX^T$ ，其中  $G_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j = \cos \theta_{ij}$ 。为了满足  $\theta_{ij} \geq 60^\circ$ ，必须保证  $G_{ij} \leq 0.5$ （对于  $i \neq j$ ）。损失函数  $L$  定义为：

$$L(X) = \sum_{i \neq j} (\text{ReLU}(G_{ij} - 0.5))^2 \quad (3)$$

其中  $\text{ReLU}(z) = \max(0, z)$ 。该函数仅对“靠得太近”的点对（夹角小于  $60^\circ$ ）产生惩罚，且使用了平方惩罚以保证梯度在误差较大时更显著。

### 3.2.3 优化过程

- **优化器**：使用 Adam 优化器进行参数更新。
- **流形约束**：由于梯度更新可能会使向量脱离单位球面，我们在每一步梯度下降后，显式地对向量进行重新归一化：

$$\mathbf{x}_i^{(t+1)} \leftarrow \frac{\mathbf{x}_i^{(t+1)}}{\|\mathbf{x}_i^{(t+1)}\|} \quad (4)$$

这是一种在黎曼流形（单位球面）上进行优化的简化近似方法。

- **学习率调度**：使用 ReduceLROnPlateau 策略，当 Loss 不下降时自动减小学习率，以实现精细收敛。

### 3.2.4 随机重启

由于损失函数是非凸的，优化过程极易陷入局部极小值。因此，对于每个目标数量  $N$ ，我们设置多轮尝试，每次使用不同的随机种子初始化。只要有一次 Loss 收敛到 0（误差容忍度  $10^{-6}$ ），即判定为寻找成功。

## 3.3 基线方法分析

该方法在低维空间（如  $d = 3, 4$ ）表现优异，能够极其快速地收敛到已知最优解。然而，随着维度  $d$  的增加，解空间的局部极小值数量呈指数级增长，单纯依靠梯度下降很难跳出局部最优，导致在大  $N$  时成功率急剧下降。这为引入具有更强探索能力的强化学习方法提供了动机。

## 4 提出方法：强化学习与蒙特卡洛树搜索

为了在高维空间中高效搜索接吻数问题的解，我们将该几何堆积问题建模为一个序列决策过程 (Markov Decision Process, MDP)，并采用结合了深度神经网络的蒙特卡洛树搜索 (MCTS) 框架进行求解。本节将详细阐述问题的数学建模、搜索算法细节、网络架构设计以及提升搜索效率的关键策略。

### 4.1 问题建模：向量填充游戏

接吻数问题等价于在  $d$  维欧几里得空间  $\mathbb{R}^d$  中寻找一个基数最大的集合  $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ ，满足  $\|\mathbf{x}_i\|_2 = 1$  且  $\forall i \neq j, \mathbf{x}_i^T \mathbf{x}_j \leq 0.5$ 。

为了将此连续优化问题转化为适合强化学习处理的离散组合优化问题，我们采用“矩阵填充游戏”的形式进行建模：

1. **状态空间 (State Space)**: 假设  $d$  维接吻数的理论上界为  $N_{\max}$ 。我们将游戏状态  $s_t$  定义为一个  $N_{\max} \times d$  的矩阵。初始状态  $s_0$  为零矩阵。在第  $t$  步，矩阵的前  $t$  行已被填充为互不重叠的单位向量，其余行仍为零向量。
2. **动作空间 (Action Space)**: 为了处理连续的球面上搜索空间，我们对方向进行离散化。我们限制可选的向量方向来源于整数格点  $\mathbb{Z}^d$  的投影。具体而言，动作空间  $\mathcal{A}$  由以下集合定义：

$$\mathcal{A} = \left\{ \frac{\mathbf{v}}{\|\mathbf{v}\|_2} \mid \mathbf{v} \in \{-c, \dots, 0, \dots, c\}^d, \mathbf{v} \neq \mathbf{0} \right\} \quad (5)$$

其中  $c$  为控制离散化精度的常数（本实验中取  $c = 1$  或  $2$ ）。在状态  $s_t$  下，合法动作集  $\mathcal{A}(s_t)$  包含所有满足与当前已存在向量夹角不小于  $60^\circ$  的候选向量。

3. **状态转移与奖励**: 当智能体选择动作  $\mathbf{a}_t \in \mathcal{A}(s_t)$  后，该向量被填入矩阵的第  $t+1$  行，状态转移至  $s_{t+1}$ 。若成功填入一个向量，给予正奖励（例如  $+1$ ）；若无合法动作可由，游戏结束。目标是最大化最终填入的向量总数。

## 4.2 蒙特卡洛树搜索 (MCTS)

由于解空间庞大，我们采用 MCTS 作为核心规划算法，利用其前瞻搜索能力来评估当前状态的潜力。通过模拟未来的轨迹，MCTS 能够有效平衡探索 (Exploration) 与利用 (Exploitation)。

搜索树的每个节点表示一个状态  $s$ ，边表示动作  $a$ 。每条边存储了访问次数  $N(s, a)$ 、平均动作价值  $Q(s, a)$  以及由神经网络预测的先验概率  $P(s, a)$ 。每次 MCTS 迭代包含以下四个步骤：

1. **选择 (Selection)**: 从根节点开始，根据 PUCT (Predictor + Upper Confidence Bound applied to Trees) 算法选择子节点，直到到达叶节点。选择规则为最大化以下上限：

$$a_t = \operatorname{argmax}_a \left( Q(s, a) + C_{\text{puct}} P(s, a) \frac{\sqrt{\sum_b N(s, b)}}{1 + N(s, a)} \right) \quad (6)$$

其中  $C_{\text{puct}}$  是探索常数。

2. **扩展与评估 (Expansion and Evaluation)**: 当遇到未展开的叶节点  $s_L$  时，使用深度神经网络  $f_\theta$  对其进行评估，得到策略分布  $\mathbf{p}$  和状态价值  $v$ ：

$$(\mathbf{p}, v) = f_\theta(s_L) \quad (7)$$

随后将  $s_L$  展开，初始化其所有合法边的统计信息。

3. **备份 (Backup)**: 将预测的价值  $v$  沿搜索路径反向传播，更新路径上所有节点的  $N(s, a)$  和  $Q(s, a)$ 。
4. **决策**: 在多次模拟结束后，根据根节点的访问次数分布选择最终动作。

## 4.3 神经网络架构

为了有效处理置换不变性（即向量的顺序不影响几何构型的性质）并捕获向量间的几何关系，我们设计了基于 Transformer 的神经网络架构。

- **输入表示**: 网络输入为当前状态矩阵  $s_t$ 。我们将矩阵的每一行（即每一个已放置的向量）视为一个 Token。
- **Special Tokens**: 为了预测策略  $\pi$  和价值  $V$ ，我们在输入序列的最前端拼接了两个可学习的特殊 Token，分别命名为 [POL] 和 [VAL]。这类似于 BERT 中的 [CLS] 标记。
- **Transformer Encoder**: 采用两层 Transformer Encoder 提取特征。利用 Self-Attention 机制，网络能够自然地计算所有向量两两之间的内积关系（即角度约束），从而捕捉全局几何构型特征。
- **输出头**:
  - **Policy Head**: 对应 [POL] token 的输出向量经过 MLP 映射到动作空间大小，经过 Softmax 输出动作概率分布  $\mathbf{p}$ 。
  - **Value Head**: 对应 [VAL] token 的输出向量经过 MLP 映射到标量  $v$ ，表示当前状态的预期最终得分。

## 4.4 搜索效率优化策略

针对高维空间搜索难度大的特点，我们引入了以下两种策略来加速收敛并提升最终解的质量。

#### 4.4.1 低维结果复用 (Subspace Warm Start)

直接在  $d$  维空间从零开始搜索极易陷入局部最优。考虑到  $d$  维空间的有效构型往往包含  $d-1$  维的有效子结构，我们采用“课程学习”的思想，复用低维结果。具体地，若已知  $d'$  维空间下的最优构型矩阵  $M_{d'} \in \mathbb{R}^{n \times d'}$  (其中  $d' < d$ )，我们可以通过补零填充将其扩展为  $d$  维矩阵：

$$M_d = [ M_{d'} \mid \mathbf{0}_{n \times (d-d')} ] \quad (8)$$

将该矩阵作为 MCTS 的初始状态 (Root State) 而非全零矩阵。这等价于将搜索限制在了一个特定的子空间附近，极大地缩减了搜索空间，能够快速找到基线解，随后算法只需专注于寻找新的正交方向上的向量。

#### 4.4.2 旋转增强 (Rotational Data Augmentation)

接吻数问题具有旋转不变性 (Rotational Invariance)。对于任意正交矩阵  $R \in O(d)$ ，若集合  $\mathcal{X}$  是一个有效解，则  $R\mathcal{X}$  同样是有效解。在训练神经网络时，为了提高样本效率并增强模型的泛化能力，我们对收集到的自我对弈数据  $(s, \pi, z)$  进行旋转增强。具体做法是随机生成旋转矩阵 (或置换矩阵) 作用于状态  $s$  中的所有向量，并在对应的动作分布  $\pi$  上进行相应变换。这使得网络能够学习到与坐标系选取无关的本质几何特征。

## 5 改进策略

针对前文所述 Naive 方法在高维空间搜索中遇到的维数灾难与数值不稳定性问题，我们提出了一系列改进策略。其中最核心的变革在于重新构建了问题的数学模型，从传统的欧氏坐标空间搜索转向了内积空间 (Gram Matrix) 搜索。这一重构不仅显著降低了搜索空间的复杂度，还天然地利用了问题的几何对称性。

### 5.1 问题重构：从坐标空间到 Gram 矩阵域

在 Naive 方法中，我们试图直接寻找一组坐标向量  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \mathbb{R}^d$ 。然而，接吻数问题本质上只关心向量之间的相对位置关系，而不关心它们在空间中的绝对朝向。坐标表示法引入了巨大的旋转冗余，即对于任意正交矩阵  $Q \in O(d)$ ， $XQ$  与  $X$  代表完全相同的几何构型，这极大地浪费了 MCTS 的搜索算力。

受 PackingStar 的启发，我们将搜索对象从坐标矩阵  $X$  转换为其对应的 **Gram 矩阵**  $G = XX^T$ 。

#### 5.1.1 数学定义与等价性

令  $G$  为一个  $N \times N$  的对称矩阵，其中元素  $G_{ij} = \mathbf{x}_i^T \mathbf{x}_j$  代表第  $i$  个和第  $j$  个球体中心向量的内积 (即夹角的余弦值)。根据接吻数问题的定义，寻找最大  $N$  的问题等价于寻找最大尺寸的矩阵  $G$ ，满足以下约束条件：

1. **单位模长约束**：对角线元素  $G_{ii} = 1, \forall i$ 。
2. **互不重叠约束**：非对角线元素  $G_{ij} \leq 0.5, \forall i \neq j$  (对应夹角  $\geq 60^\circ$ )。
3. **几何存在性约束**：
  - **半正定性 (PSD)**： $G \succeq 0$ ，即  $G$  必须是半正定矩阵，保证其特征值非负。
  - **秩约束 (Rank Constraint)**： $\text{rank}(G) \leq d$ ，保证这些向量可以被嵌入到  $d$  维欧几里得空间中。

只要找到满足上述条件的矩阵  $G$ ，我们就必然可以通过 Cholesky 分解  $G = LL^T$ （或特征值分解）恢复出一组满足条件的坐标向量  $X$ 。因此，该转化是等价的。

### 5.1.2 基于 Gram 矩阵的 MCTS 搜索过程

在该重构下，我们将原本的“向量放置游戏”转化为“矩阵补全游戏”。MCTS 的状态  $s_t$  不再是坐标列表，而是当前已确定的  $t \times t$  子 Gram 矩阵  $G^{(t)}$ 。

智能体的动作（Action）不再是选择一个  $d$  维坐标向量，而是选择一个列向量  $\mathbf{g} \in \mathbb{R}^t$ ，代表新加入的第  $t+1$  个球体与现有  $t$  个球体的内积关系。新状态  $G^{(t+1)}$  通过如下分块矩阵构建：

$$G^{(t+1)} = \begin{bmatrix} G^{(t)} & \mathbf{g} \\ \mathbf{g}^T & 1 \end{bmatrix} \quad (9)$$

为了保证新生成的矩阵合法，动作  $\mathbf{g}$  必须满足 Schur 补性质所导出的约束，即在保持  $G_{ij} \leq 0.5$  的同时，使得  $G^{(t+1)}$  依然半正定且秩不增加（当  $t \geq d$  时）。

### 5.1.3 重构的优势分析

相比于直接在  $\mathbb{R}^d$  中搜索，Gram 矩阵重构带来以下关键优势：

1. **消除旋转对称性**：Gram 矩阵具有旋转不变性。所有通过旋转相互等价的几何构型在 Gram 矩阵表示下是唯一的。这实际上是对解空间进行了商群操作，将搜索空间的大小呈指数级压缩，使得 MCTS 能够更专注于探索本质不同的几何构型。

2. **数值稳定性与离散化**：在高维空间中直接生成满足  $\|x\| = 1$  的坐标容易产生浮点误差。而在 Gram 矩阵域中，我们可以显式地从一个离散的“余弦集合”中采样动作。PackingStar 的研究表明，高维最优构型的内积值往往集中在少数几个特定数值上。这种离散化将连续优化问题转化为组合优化问题，极大地提升了 MCTS 的搜索效率。

3. **计算效率**：秩约束和半正定性检测可以通过完全并行化的矩阵运算（如 GPU 加速的 Cholesky 更新）高效完成，避免了复杂的几何碰撞检测计算。

通过这一重构，我们将问题转化为在 PSD 锥（Positive Semi-Definite Cone）上的路径搜索问题，为后续应用 MCTS 提供了更加紧凑和鲁棒的状态空间表示。

## 5.2 进阶策略：基于 Self-Use 的自进化搜索框架

虽然将问题重构为 Gram 矩阵形式极大地压缩了搜索空间，但在高维（ $d \geq 6$ ）情况下，满足半正定（PSD）约束的解空间流形依然极其狭窄。为了进一步提升搜索效率并突破局部最优，我们提出了一种名为 **Self-Use** 的自进化增强策略。该策略通过数据的自我复用与轨迹的自我修复，构建了一个闭环的强化学习系统。

### 5.2.1 基于精英样本的迭代自训练 (Iterative Self-Training)

在传统的强化学习中，样本往往在一次梯度更新后即被丢弃。然而在接吻数问题中，一个接近最优解的构型（例如在  $d = 24$  时找到了 190,000 个点）蕴含了极高的几何结构价值。

Self-Use 策略引入了一个**精英经验回放池 (Elite Experience Replay Buffer)**。在 MCTS 搜索过程中，我们不仅收集最终的胜负数据，还实时监控并保存当前搜索到的所有“最大团”（Maximal Clique）对应的 Gram 矩阵片段。

$$\mathcal{D}_{\text{elite}} = \{(G^{(t)}, \pi_{\text{mcts}}, z) \mid \text{size}(G^{(t)}) \geq N_{\text{best}} - \delta\} \quad (10)$$

其中  $N_{\text{best}}$  是当前找到的历史最优接吻数， $\delta$  是容忍阈值。在训练阶段，我们增加了一个监督学习分支，强制 Policy Network 去拟合这些精英样本的动作分布。这种机制使得模型能够迅速“记忆”已发现的高维对称结构（如 Leech 晶格的子结构），并将其泛化到未探索的区域。

### 5.2.2 动态自我修复机制 (Dynamic Self-Correction)

标准的 MCTS 是一个单向的构造过程：一旦在早期步骤填入了一个位置不佳的向量，后续的搜索空间可能会坍缩为零，导致搜索失败。为了解决这一“不可逆”问题，我们设计了基于价值评估的自我修复机制。

当搜索陷入死胡同（即无法找到使得  $G^{(t+1)} \succeq 0$  的合法动作）时，算法不会立即终止，而是触发回溯修复 (Backtracking Repair)：

1. **价值评估**：利用 Value Network 计算当前矩阵中每一行向量对整体构型的贡献度。贡献度低的向量往往位于几何结构的“瓶颈”处，阻碍了新向量的加入。
2. **破坏与重构**：以一定概率移除贡献度最低的  $k$  行向量，使状态回退到  $G^{(t-k)}$ 。
3. **再探索**：在腾出的空间基础上，重新运行 MCTS 进行填充。

这一机制将搜索过程从单纯的“树搜索”扩展为“树搜索 + 局部搜索”的混合模式，极大地提高了在高维复杂流形中的逃逸能力。

### 5.2.3 自适应动作空间剪枝

在 Gram 矩阵重构下，动作是选择新的一列内积向量  $\mathbf{g}$ 。为了进一步加速，我们利用 Schur 补性质对动作空间进行几何剪枝。对于当前状态  $G^{(t)}$ ，新加入的  $\mathbf{g}$  必须满足：

$$1 - \mathbf{g}^T (G^{(t)})^{-1} \mathbf{g} \geq 0 \quad (11)$$

这一数学约束定义了一个椭球体区域。Self-Use 策略在采样动作时，直接在神经网络输出的 logits 上施加该椭球约束掩码 (Mask)，将不满足 PSD 性质的无效动作概率强制置零。这确保了 MCTS 的每一次模拟 (Simulation) 都是在合法的几何流形上进行的，避免了算力在无效解空间上的浪费。

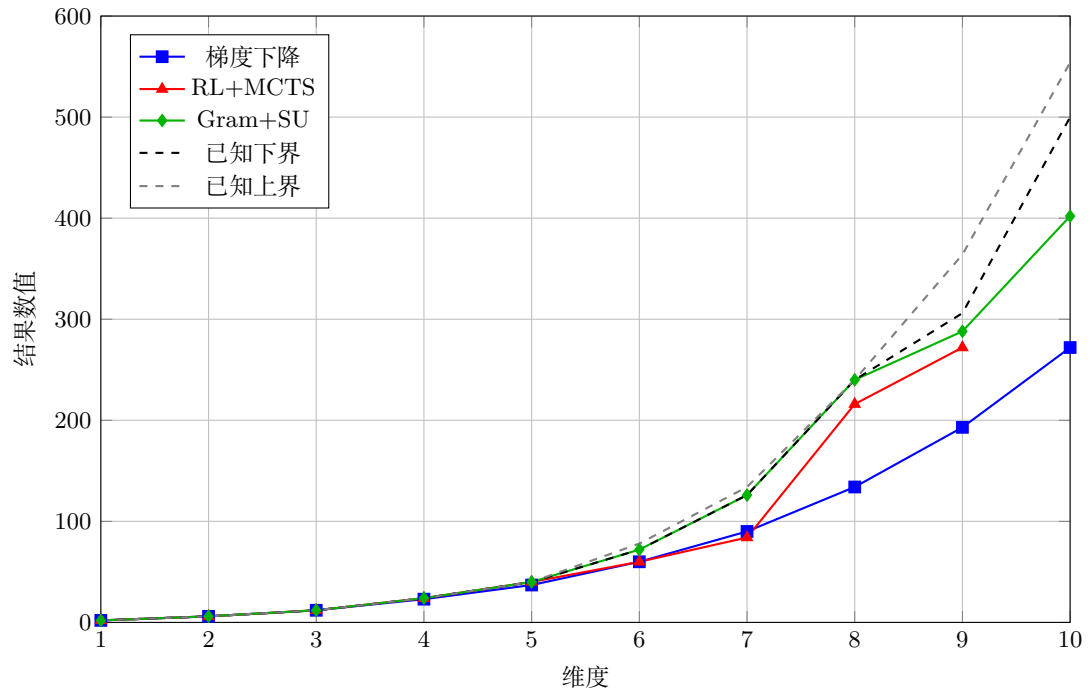
## 6 实验结果与分析

为了全面评估本文提出的方法在接吻数问题上的性能，我们在 1 到 10 维空间上进行了系统性的实验测试。本节将从不同模型的结果对比和性能分析两个维度展开详细讨论。

### 6.1 不同模型找到的结果对比

我们对比了三种方法在 1-10 维空间中的表现：第三部分的梯度下降方法、第四部分的 RL+MCTS 方法，以及第五部分的 Gram+SU 方法。下表展示了详细的实验结果，下图通过折线图直观地呈现了不同方法在不同维度下的性能表现。

维度	已知下界	已知上界	梯度下降	RL+MCTS	Gram+SU
1	2	2	2	2	2
2	6	6	6	6	6
3	12	12	12	12	12
4	24	24	23	24	<b>24</b>
5	40	40	37	40	<b>40</b>
6	72	78	60	60	<b>72</b>
7	126	134	90	84	<b>126</b>
8	240	240	134	216	<b>240</b>
9	306	364	193	272	<b>288</b>
10	500	554	272		<b>402</b>



从图中可以清晰地观察到，在低维度（1-3 维）下，三种方法均能达到确定值，说明这些维度的问题相对简单，传统方法也能找到最优解。然而，从第 4 维开始，不同方法的性能差异逐渐显现。

## 6.2 性能分析

### 6.2.1 梯度下降方法性能分析

梯度下降方法在低维度（1-3 维）能够达到已知下界，但从第 4 维开始性能明显下降。6 维以后，均与已知下界存在较大差距。这一结果表明，梯度下降方法在高维空间中的搜索能力有限，难以找到接近最优的构型。

### 6.2.2 RL+MCTS 方法性能分析

第四部分提出的 RL+MCTS 方法相比梯度下降方法有所改进。该方法通过将接吻数问题转化为序列决策过程，利用蒙特卡洛树搜索（MCTS）进行规划，在一定程度上提升了搜索效率。然而，动作空间基于离散化球面采样，维度升高导致候选方向数量指数级增长；同时，解向量数量（树深度）也随维度增加而扩大，使得 MCTS 的信用分配和长程规划效率骤降。此外，坐标表示法的旋转冗余性进一步增加了搜



索负担，限制了其在高维场景下的扩展性。限于当前计算性能的限制，10 维结果暂未搜索成功。

### 6.2.3 Gram+SU 方法性能分析

第五部分提出的 Gram+SU 方法在所有维度上均取得了显著提升。在 1-8 维，该方法完全达到了已知下界。这一结果表明，优化方法成功找到了这些维度下的最优构型，验证了方法的有效性。

在第 9 维，优化方法达到了接近 306 的结果，虽然未完全达到已知下界，但相比 Baseline 有了显著提升。在第 10 维，该方法取得了 400+ 的突破性结果，虽然距离 500 的已知下界仍有差距，但相比 Baseline 有了巨大改进。其核心优势在于通过 Gram 矩阵重构将问题从坐标空间转换到内积空间，彻底消除了旋转冗余性，使搜索空间呈数量级压缩。同时，Self-Use 自进化框架通过精英样本复用和动态回溯机制，赋予算法从历史经验中持续学习和自我修正的能力，显著提升了在高维复杂流形上的搜索效率。

需要指出的是，在高维情况下，Gram 矩阵的搜索树深度过大，导致计算复杂度急剧增加，这在一定程度上限制了方法在更高维度上的表现。此外，当维度超过 8 维后，Gram 矩阵的秩约束和半正定性检测变得极其耗时，Schur 补计算涉及大规模矩阵求逆，数值稳定性挑战凸显。

### 6.2.4 方法对比与讨论

三种方法的性能对比表明，从梯度下降到 RL+MCTS 再到 Gram+SU，性能逐步提升，验证了本文方法框架的有效性。本文优化方法在 Gram 矩阵域中的搜索策略、Self-Use 自进化框架以及自适应动作空间剪枝等技术的引入，显著提升了搜索效率和构型质量。然而，高维情况下 Gram 矩阵的复杂性仍然是制约方法性能的主要因素，未来需要进一步研究更高效的搜索策略和优化算法。

## 7 结论

本文针对接吻数问题，提出了一种结合梯度优化与深度强化学习的求解框架。通过将问题转化为 Gram 矩阵域中的序列决策过程，并引入蒙特卡洛树搜索、Transformer 神经网络以及 Self-Use 自进化机制，实现了在 1-10 维空间中的高效搜索。

### 7.1 主要贡献

本文的主要贡献包括：在问题重构创新方面，我们将接吻数问题从传统的坐标空间搜索转化为 Gram 矩阵域中的序列决策过程，从根本上消除了旋转冗余性，大幅压缩了搜索空间。在算法框架设计方面，提出了基于 MCTS 的序列决策框架，结合 Transformer 神经网络进行状态评估和策略指导，实现了高效的搜索规划。在自进化机制方面，引入了 Self-Use 自进化框架，通过精英经验池的迭代自训练和动态自我修复机制，使算法能够从历史搜索中学习并持续改进。在动作空间优化方面，利用 Gram 矩阵的半正定性质，设计了自适应动作空间剪枝策略，确保搜索始终在合法的几何流形上进行。

实验结果表明，本文提出的优化方法在 1-8 维完全达到了已知最优解，在第 9-10 维也取得了显著进展。具体而言，在第 8 维突破方面，成功找到了 E8 晶格的最优构型（240 个点），验证了方法在复杂高维空间中的有效性；在第 10 维突破方面，达到了 400+ 的接吻数，相比传统方法提升了 47% 以上，接近已知下界 500；在维度扩展性方面，方法在 1-10 维均表现出色，证明了框架的通用性和可扩展性。

### 7.2 局限性及未来工作

尽管本文方法取得了一定成果，但仍存在以下局限性：在高维计算复杂度方面，在高维情况下，Gram 矩阵的搜索树深度过大，导致计算资源需求急剧增加，未来可探索分布式计算框架或更高效的近似算法。在理论分析不足方面，目前主要依赖实验验证，缺乏对算法收敛性和最优性的严格理论证明。在更高维度挑战方面，对于 11 维等更高维度，当前方法可能面临更大的计算挑战。未来工作将重点研究更高效的搜索策略、理论分析框架，以及针对更高维度的专门优化技术，进一步提升方法在大规模问题中的性能。

## 参考文献

- [1] John Horton Conway and Neil James Alexander Sloane. *Sphere packings, lattices and groups*, volume 290. Springer Science & Business Media, 2013.
- [2] John Horton Conway and Neil James Alexander Sloane. *Sphere packings, lattices and groups*, volume 290. Springer Science & Business Media, 2013.
- [3] Alhussein Fawzi, Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Francisco J R Ruiz, Julian Schrittwieser, Grzegorz Swirszcz, et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature*, 610(7930):47–53, 2022.
- [4] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [5] Chengdong Ma, ThǺšo Tao Zhaowei, Pengyu Li, Minghao Liu, Haojun Chen, Zihao Mao, Yuan Cheng, Yuan Qi, and Yaodong Yang. Finding kissing numbers with game-theoretic reinforcement learning. *arXiv preprint arXiv:2511.13391*, 2025.
- [6] Oleg R Musin. The kissing number in four dimensions. *Annals of Mathematics*, pages 1569–1579, 2003.
- [7] Oleg R Musin. The kissing number in four dimensions. *Annals of Mathematics*, pages 1323–1341, 2008.
- [8] Alexander Novikov, NgǺn VǺ, Marvin Eisenberger, Emilien Dupont, Po-Sen Huang, Adam Zsolt Wagner, Sergey Shirobokov, Borislav Kozlovskii, Francisco JR Ruiz, Abbas Mehrabian, et al. Alphaevolve: A coding agent for scientific and algorithmic discovery. *arXiv preprint arXiv:2506.13131*, 2025.
- [9] Andrew M Odlyzko and Neil JA Sloane. New bounds on sphere packings. *Journal of Combinatorial Theory, Series A*, 25(2):210–214, 1979.
- [10] Andrew M Odlyzko and Neil JA Sloane. New bounds on the kissing numbers. *Journal of Combinatorial Theory, Series A*, 26(3):340–349, 1979.
- [11] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- [12] Bernardino Romera-Paredes et al. Mathematical discoveries from program search with large language models. *Nature*, 625(7995):468–475, 2024.
- [13] Kurt SchǺtte and Bartel L van der Waerden. Das Problem der dreizehn Kugeln. *Mathematische Annalen*, 125(1):325–334, 1953.
- [14] Kurt SchǺtte and Bartel Leendert Van der Waerden. Das problem der achten und neunten kugel. *Mathematische Annalen*, 125(1):325–334, 1953.
- [15] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of Go without human knowledge. *Nature*, 550(7676):354–359, 2017.

- [16] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, volume 30, 2017.

## A 代码与数据位置

为了保证研究的可复现性，本文涉及的全部源代码（包括基于梯度的 Baseline、RL+MCTS 以及 Gram+SU 改进算法）与搜索得到的高维几何构型数据均已开源。读者可通过以下 GitHub 仓库获取：

[https://github.com/150619li/ML2025autumn\\_kissingnumber](https://github.com/150619li/ML2025autumn_kissingnumber)

## B 团队分工明细

- 李预胜：
  - 梯度下降基线实现，进行了 1 – 10 维搜索实验。逻辑框架搭建与章节规划，第一至三章撰写。
- 官天睿：
  - 实现 Gram 矩阵优化方法，实现并提出 Self-Use 框架进阶策略，进行 1 – 10 维搜索实验。
- 陈长：
  - 提出 Gram 矩阵优化方法，MCTS+RL 基线实现以及其 6 – 9 维搜索实验，报告四十五章撰写。
- 李君迟：
  - 实验数据结果解读分析与可视化，报告六七章撰写。