arXiv:2601.21917v1 [math.OC] 29 Jan 2026

关于临界点近似计算

Amir Ali Ahmadi* Georgina Hall†

摘要

我们证明，对于一些简单的非凸函数类，即使是临界点的非常粗略的近似计算也难以处理。更具体地说，我们证明：如果存在一个多项式时间算法，其输入为一个关于 n 个变量、次数为常数（最低可至三次）的多项式，并在该多项式存在临界点时输出一个梯度的欧几里得范数至多为2^n 的点，那么 P=NP。若不存在临界点，该算法可以返回任意点。我们还证明了在额外结构假设下近似计算临界点的困难性结果，其中包括保证临界点存在且唯一、函数有下界，以及以到临界点的距离来衡量近似程度的情形。总的来说，我们的结果与一种普遍看法形成对照：在非凸优化中，临界点的近似计算是一项易于处理的任务。

关键词： 临界点、非凸优化、近似困难性

1 引言

可微函数的临界点（或驻点）是指其梯度为零的点。计算临界点是数学优化中最基本的算法任务之一。一个关键原因是，可微函数的任何全局（甚至局部）极小化点都必然是临界点。因此，寻找临界点通常被视为寻找极小化点的前置步骤。毫不意外的是，函数的某些结构性质——例如凸性或 Polyak-Łojasiewicz 性质——若能使寻找临界点这一任务足以找到全局极小点，往往会在优化理论中备受关注。

多年来，优化文献提出了许多算法，在各种假设下搜索临界点；例如参见（Bertsekas, 1995； Boyd 和 Van- denberghe，2004年；Nocedal 和 Wright，2006年；Cartis 等人，2022年）及其中的参考文献。其中最广为人知的两种算法或许是梯度下降法，它构成了现代机器学习中许多方法的计算基础；以及牛顿法，它是内点法的重要组成部分。自然地，临界点是这两种算法的不动点。对于牛顿法而言，每次迭代的目标实际上也是寻找一个临界点，即当前迭代点处原函数的二阶泰勒近似的临界点。

临界点——以及更一般而言的局部最优性概念——近年来在优化领域中愈发受到重视，这是由于在机器学习的大力推动下，研究范式从凸优化转向了非凸优化。事实上，在非凸情形中，人们已充分认识到，全局最小化往往是一个NP难问题。一个常见

*Amir Ali Ahmadi 就职于普林斯顿大学运筹学与金融工程系。电子邮箱：aaa@princeton.eduAmir Ali Ahmadi 的部分研究得到斯隆研究奖、普林斯顿人工智能实验室种子基金、普林斯顿工程与应用科学学院创新基金以及数学优化领域研究赠款的支持。

†Georgina Hall 就职于 INSEAD 决策科学领域。电子邮箱：georgina.hall@insead.edu

误解则是，这种困难主要源于伪局部极小值的存在，即那些并非全局极小值、并可能使优化算法陷入其中的局部极小值。这一观点忽视了如下事实：即便对于低次多项式，找到一个局部极小值也已是NP难任务（Ahmadi and Zhang,2022b）。同样的复杂度结论对于判定一个点是否为局部极小值也成立（Murty and Kabadi, 1987）。在了解局部最优性困难性结果的研究者中，一种常听到的解释是，所遇到的困难源于鞍点或其他阻碍算法进展的临界点。这一点可由例如（Ge等人, 2015年）：

> “优化非凸函数的难点 [来自两个方面] 。首先，非凸函数可能有许多局部极小值，而从中找出最优的那个（全局最小值）可能很困难。其次，即使是寻找局部极小值也可能很困难，因为可能存在许多梯度为 0、但并非局部极小值的鞍点。”

这段话所反映的一种广泛看法是，寻找临界点应该相对容易——或许是因为临界点很容易识别：按照定义，只需检查单个点处的方程，而不必像寻找局部极小值那样检查某个邻域内的不等式；又或许是因为人们预期上述算法能在合理的迭代次数内收敛。

与这种普遍看法相反，最近的研究通过从组合优化中的最大割问题进行相对简单的多项式时间归约，证明了判定一个次数为 3 或更高的多项式是否存在临界点是 NP难的（Ahmadi and Zhang，2022a，定理 2.1）。[1] 为调和这一结果与先前的看法，研究人员通常作出如下假设：（Ahmadiand Zhang，2022a，定理 2.1）排除了（除非 P=NP）存在用于寻找精确的临界点的多项式时间算法，而困难很可能源于对精确计算的需求。然而在实践中，人们往往满足于对临界点进行近似计算，并认为这是一项简单得多的任务。本文旨在考察这一假设。结果表明，我们的发现与这种直觉相悖。

1.1 概述、贡献及相关文献

本文分为两节，分别讨论与计算 n 元可微函数f的临界点的不同近似计算概念相关的困难性结果。我们的所有结果都基于标准图灵计算模型，其中函数 f属于某个特定的参数化函数类（例如低次多项式），我们会在定理的正式陈述中具体说明该函数类。第 2 节中，我们考虑 ϵ-近似临界点的概念，即梯度 f 的 2-范数不超过 ϵ. 的点。我们在定理 3 中证明，除非 P=NP，否则不存在多项式时间算法能够计算出一个近似质量极差的近似临界点（更准确地说，具有 $\epsilon = 2^n$）。证明这一近似困难性结果，远比证明临界点精确计算的困难性（如 (Ahmadi and Zhang, 2022a, Theorem 2.1)）复杂。在定理 3 的证明之后，我们通过讨论论证，即使对于已知具有唯一临界点的函数，计算2^n-近似临界点的问题仍然难以处理。在命题10中，我们证明，如果存在多项式时间算法，能够为没有伪临界点（因而也没有伪局部极小值）的下有界函数寻找临界点，那么 PPAD=P。最后，在定理 11 中，我们证明，即使对于在单位超立方体中有临界点的下有界函数，除非 P=NP，否则不存在多项式时间算法能够在半径为 2^n的球内找到一个 2^n-近似临界点。

[1]对于二次多项式，同一判定问题属于 P，因为只需检验一个线性方程组是否可行。

第 3 节中，我们考虑函数 f 的 ϵ-近临界点这一概念，即与 ϵ 的某个临界点距离不超过 f 的点。我们的困难性结果定理 13排除了存在此类算法的可能性：它们无法在关于n 和$\frac{1}{\epsilon}$ 的多项式时间内计算出 ϵ-近临界点。我们的最终结果定理 15表明，即使函数 f 下有界，同样的结论仍然成立。第 4 节给出了一些结论性讨论。

我们在引言的最后回顾有关计算近似临界点（即第2节所讨论的概念）困难性的相关文献。据我们所知，文献中尚无关于计算近临界点困难性的结果；第3节将对此进行讨论。在（Carmon等，2020）中，作者研究了一种基于预言机的复杂性模型：通过在查询点处评估函数及其截至某一阶的导数来获取函数信息。他们建立了该模型中任何算法为计算 ϵ近似临界点所需查询次数的下界。在（Carmon等，2021）中，作者在额外假设凸性的条件下推导出了类似的下界。在这两篇论文中，作者构造的困难实例的定义域维数均随$\frac{1}{\epsilon}$ 增长。在（Hollender和Zampetakis，2023；Chewi等，2023）中，作者同样给出了预言机模型中恢复 ϵ近似临界点所需查询次数的下界，但他们是在固定且较低的维数下完成的。Hollender和Zampetakis（2023）的工作进一步在他们称为“白盒”模型的模型中，给出了计算近似临界点的困难性结果。在该模型中，给定图灵机；对于输入点x，这些图灵机能够在关于 x大小的多项式时间内计算函数 f 或其在 x处的梯度。他们证明，在这一模型中，对于任意固定维数 $n \geq 2$，寻找近似临界点的问题对于一个称为多项式局部搜索（PLS）的复杂度类是完全的；该复杂度类介于P和NP（的函数版本）之间。相比之下，我们既不采用预言机模型，也不采用白盒模型，而是采用上文所述的图灵计算模型，其中输入是给定（参数化）函数类中的函数实例的参数。我们的困难性结果表明，除非P=NP，否则不仅基于函数或导数评估的方法（例如梯度下降）不存在多项式时间算法，任何计算方法（例如先进行凸松弛再执行舍入过程的方法）也不存在多项式时间算法。我们注意到，在我们的设定下，无法获得Hollender和Zampetakis（2023）所得到的固定维数困难性结果，因为我们关注的函数类是多项式。事实上，如果维数固定，即使是更一般的求解多项式方程组与不等式组问题，也可以利用量词消去理论及相关算法（参见例如Grigor’ev和Vorobjov Jr，1988）在关于其次数和系数大小的多项式时间内完成。最后，我们指出，关于计算Karush-Kuhn-Tucker点也存在困难性结果；Karush-Kuhn-Tucker点是约束优化问题中临界点的类似物（Vavasis，1993；Fearnley等，2022年，2025年）。

2 计算 ϵ-近似临界点的复杂度

在整篇论文中，我们用 ∇f表示标量值函数 f 的梯度向量。

定义 1 （临界点）。 设 $f : \mathbb{R}^n \to \mathbb{R}$ 是一个可微函数。f 的临界点 是满足 $\nabla f(x) = 0$.的点 $x \in \mathbb{R}^n$

如引言所述，论文考虑了两种不同的临界点近似计算概念。本节聚焦于 ϵ-近似 临界点这一概念，现给出其定义。在整篇论文中，记号 $||\cdot||$ 表示 2-范数。

定义 2 （ϵ-近似临界点）。 设 $f : \mathbb{R}^n \to \mathbb{R}$ 是一个可微函数，且$\epsilon > 0$。 f 的 ϵ-近似临界点 是满足 $||\nabla f(x)|| \leq \epsilon$.的点 $x \in \mathbb{R}^n$

本节将介绍关于计算 ϵ-近似临界点复杂度的两个结果。正如引言中所述，所有这些结果以及后续结果均

在标准图灵计算模型中表述。在该模型中，问题的每个实例都有一个可用有限比特数表示的输入（例如，参见（Sipser，1996年））。因此，我们考虑有限参数化的可微函数。我们的大多数主要结果都涉及固定次数多项式（定理 3、11、13）。其他一些结果（定理 15和命题 9、10）则涉及固定次数多项式与固定次数多项式的指数函数的混合。于是，问题实例的输入由构成这些函数的单项式系数组成，并且这些系数均取有理数。我们再次说明，多项式时间算法是指运行时间受输入规模的某个多项式所界定的算法；也就是说，输入规模是对所有这些有理数进行二进制编码所需的比特数。我们还再次说明，伪多项式时间算法是指运行时间受输入数值大小的某个多项式所界定的算法。

2.1 计算一般函数 ϵ 近似临界点的复杂度

我们首先陈述本节的主要结果。其部分推论将在证明之后讨论。

定理 3. 如果存在一个算法 $\mathcal{A}$，它以含有 n 个变量的三次多项式 p 为输入，在多项式时间内运行，并且满足

(i) 当 p 没有临界点时，$\mathcal{A}$ 可以自由返回任意一点，

(ii) 当 p 有临界点时，$\mathcal{A}$ 返回一个 2^n 近似临界点，

那么 $P = NP$.

为证明该定理，我们对以下称为 CLIQUE的问题进行多项式时间归约：给定一个图 G，其具有 m 个顶点和一个正整数 $k \leq m$，判定 G 是否具有大小为 k的团，即一组由 k 两两相邻的顶点组成的集合。众所周知，CLIQUE 是 NP 完全的（Karp，1975），这尤其意味着，除非 P=NP，否则该问题不存在多项式时间算法。

在介绍这一归约之前，我们先证明几个技术引理。其中除一个之外，其余引理都涉及四次多项式 g 的性质：$\mathbb{R}^m \to \mathbb{R}$，该多项式的根与图中的团之间存在一一对应关系（如引理 4中形式化所述）。该多项式在定理 3 的证明以及第 2.2 节和第 3节的结果中都发挥着关键作用。我们先给出其定义。设 G 是一个具有 m 个顶点和邻接矩阵 A 的图，并设$x = (x_1, \ldots, x_m)$。对于常数 $k \in \{1, \ldots, m\}$、$C \in \mathbb{R}$、和 $N = 160m^2 \cdot (m^2 + 1)$，定义

$$g_{G,C,k}(x) := C^2 \left(\left(\sum_{i=1}^{m} x_i - 2k + m \right)^2 + N^2 \sum_{i=1}^{m} (x_i^2 - 1)^2 + \sum_i \sum_{j>i} (1 - A_{ij})^2 (x_i + 1)^2 (x_j + 1)^2 \right). \tag{1}$$

为简化记号，我们在证明中省略对 G, C, k 的依赖，简写为 g.

引理 4. 一个具有 m 个顶点和边集 E 的图 G 存在一个大小为 k 的团，当且仅当方程组

$$\begin{cases} \sum_{i=1}^{m} x_i = 2k - m \\ x_i^2 - 1 = 0, \ i = 1, \ldots, m \\ (x_i + 1)(x_j + 1) = 0, (i, j) \notin E \end{cases} \tag{2}$$

是可行的，或者等价地，当且仅当多项式 $g = g_{G,C,k}$ 在 *(1)* 的某个点处取值为零。

该引理的证明留给读者，读者可以验证，方程 (2) 的唯一可能解是 k 中大小为 G 的团的指示向量（即，$\{\pm 1\}^m$ 中恰有 k 个分量等于 $+1$,其下标对应于 G 中构成大小为 k 的团的顶点）。

在本文余下部分中，对于 $x \in \mathbb{R}^m$，我们令 $sgn(x)$ 表示 $\mathbb{R}^m$ 中的向量，其第 i 个分量在 $x_i \geq 0$ 时等于 1，在 $x_i < 0$ 时等于 -1。接下来的引理将给出单位超立方体顶点附近多项式 g 的值的界。

引理 5. 设 $\epsilon > 0$，G 是一个具有 m 个顶点的图，$k \in \{1, \ldots, m\}$，且 $C \in \mathbb{R}$. 回忆定义 $g = g_{G,\ C,\ k}$ 见 (1) 并假设 $\bar{x} \in \mathbb{R}^m$ 满足 $|\bar{x}_i^2 - 1| \leq \epsilon$ 对 $i = 1,\ \ldots, m$。那么，

$$g(\bar{x}) \geq g(sgn(\bar{x})) - 48C^2m^2 \cdot \epsilon.$$

证明。令 $\bar{x} \in \mathbb{R}^m$ 满足 $|\bar{x}_i^2 - 1| \leq \epsilon$，其中 $i = 1,\ \ldots\ldots,\ m$. 固定 $i \in \{1, \ldots, m\}$。那么，$1 - \epsilon \leq (\bar{x}_i)^2 \leq 1 + \epsilon$. 由于 $0 \leq 1 - \epsilon \leq 1$ 和 $1 \leq 1 + \epsilon \leq 2$，我们有

$$1 - \epsilon \leq \sqrt{1 - \epsilon} \leq |\bar{x}_i| \leq \sqrt{1 + \epsilon} \leq 1 + \epsilon. \tag{3}$$

如果 $sgn(\bar{x}_i) > 0$，则由 $|\bar{x}_i| = \bar{x}_i$ 和 (3) 可得 $sgn(\bar{x}_i) - \epsilon \leq \bar{x}_i \leq sgn(\bar{x}_i) + \epsilon$. 。同理，如果 $sgn(\bar{x}_i) < 0$，则由 $|\bar{x}_i| = -\bar{x}_i$ 和 (3) 可得 $sgn(\bar{x}_i) + \epsilon \geq \bar{x}_i \geq sgn(\bar{x}_i) - \epsilon$。因此，$|\bar{x}_i - sgn(\bar{x}_i)| \leq \epsilon$。令 $\epsilon_i = \bar{x}_i - sgn(\bar{x}_i)$，可将 $\bar{x}_i = sgn(\bar{x}_i) + \epsilon_i$ 写作，其中 $|\epsilon_i| \leq \epsilon$. 。利用对所有 $i = 1,\ \ldots\ldots,\ m$ 均有 $N^2(\bar{x}_i^2 - 1)^2 \geq 0$ 这一事实，我们得到

$$g(\bar{x}) \geq C^2 \left(\left(\sum_{i=1}^{m} \bar{x}_i - 2k + m \right)^2 + \sum_i \sum_{j>i} (1 - A_{ij})^2 (\bar{x}_i + 1)^2 (\bar{x}_j + 1)^2 \right). \tag{4}$$

现在，我们将上述求和中的每一项分别替换为 $\bar{x}_i = sgn(\bar{x}_i) + \epsilon_i$。我们可以写成

$$\begin{aligned}
&\left(\sum_{i=1}^{m} \bar{x}_i - 2k + m \right)^2 = \left(\sum_{i=1}^{m} sgn(\bar{x}_i) + \sum_{i=1}^{m} \epsilon_i - 2k + m \right)^2 \\
&\geq \left(\sum_{i=1}^{m} sgn(\bar{x}_i) - 2k + m \right)^2 - 2 \left| \sum_{i=1}^{m} \epsilon_i \right| \cdot \left| \sum_{i=1}^{m} sgn(\bar{x}_i) - 2k + m \right| \\
&\geq \left(\sum_{i=1}^{m} sgn(\bar{x}_i) - 2k + m \right)^2 - 8m^2\epsilon,
\end{aligned} \tag{5}$$

其中，第一个不等式用到了 $(a+b)^2 \geq a^2 - 2|a| \cdot |b|$ 这一事实；第二个不等式则结合三角不等式以及 $|\epsilon_i| \leq \epsilon$ 和 $k \leq m$ 得到。同样，我们有

$$\begin{aligned}
&(\bar{x}_i + 1)^2 (\bar{x}_j + 1)^2 = ((sgn(\bar{x}_i) + \epsilon_i + 1)(sgn(\bar{x}_j) + \epsilon_j + 1))^2 \\
&\geq (sgn(\bar{x}_i) + 1)^2 (sgn(\bar{x}_j) + 1)^2 \\
&- 2|sgn(\bar{x}_i) + 1| \cdot |sgn(\bar{x}_j) + 1| \cdot |\epsilon_i(sgn(\bar{x}_j) + 1) + \epsilon_j(sgn(\bar{x}_i) + 1) + \epsilon_i\epsilon_j| \\
&\geq (sgn(\bar{x}_i) + 1)^2 (sgn(\bar{x}_j) + 1)^2 - 40\epsilon,
\end{aligned}$$

其中，第一个不等式用到了以下事实：$((a+b)(c+d))^2 = (ac+ad+bc+bd)^2 \geq a^2c^2 - 2|a| \cdot |c| \cdot |ad+bc+bd|$；第二个不等式则用到了三角不等式，以及以下事实：$|\epsilon_i| \leq \epsilon \leq 1$、$|\epsilon_i\epsilon_j| \leq \epsilon^2 \leq \epsilon$，并且对于任意 $i,\ j$，都有 $|sgn(x_i^*) + 1| \leq 2$。这意味着

$$\sum_i \sum_{j>i} (1 - A_{ij})^2 (\bar{x}_i + 1)^2 (\bar{x}_j + 1)^2 \geq \sum_i \sum_{j>i} (1 - A_{ij})^2 (sgn(\bar{x}_i) + 1)^2 (sgn(\bar{x}_j) + 1)^2 - 40m^2 \cdot \epsilon.$$

将此不等式与 (5) 中的 (4) 结合起来，即可得到我们的结果：

$$g(\bar{x}) \geq C^2 \left(\left(\sum_{i=1}^{m} sgn(\bar{x}_i) - 2k + m \right)^2 + \sum_i \sum_{j>i} (1 - A_{ij})^2 (sgn(\bar{x}_i) + 1)^2 (sgn(\bar{x}_j) + 1)^2 - 48m^2 \cdot \epsilon \right)$$
$$= g(sgn(\bar{x})) - 48C^2 m^2 \cdot \epsilon.$$

■

下面的引理给出了多项式 g 在单位超立方体上那些不对应于指定大小的团的顶点处的值的界。

引理 6. 设 G 是一个具有 m 个顶点的图， $k \in \{1, \ldots, m\}$，并且 $C \in \mathbb{R}$. 回忆一下$g = g_{G、C、k}$在(1)中的定义。对于任意 $s \in \{\pm 1\}^m$，若支撑集 $\{i \in \{1, \ldots, m\} \mid s_i = 1\}$不对应于 k 中大小为 G的团，则有

$$g(s) \geq \frac{C^2}{(m^2+1)^2}.$$

证明。 设 $s \in \{\pm 1\}^m$ 满足集合 $\{i \in \{1, \ldots, m\} \mid s_i = 1\}$ 不对应于 G中大小为 k 的团。必有

$$\left| \sum_{i=1}^{m} s_i - 2k + m \right| + \sum_i \sum_{j>i} |(1 - A_{ij})(s_i + 1)(s_j + 1)| \geq 1. \tag{6}$$

事实上，当 $s \in \{-1, 1\}^m$ 时，前一个表达式只取整数值；根据引理 4，它不可能为零。由于等式左边的求和中至多有 $m^2 + 1$ 项，(6)，因此必有 $|\sum_{i=1}^{m} s_i - 2k + m| \geq \frac{1}{m^2+1}$，或者 $(1 - A_{ij})|(s_i + 1)(s_j + 1)| \geq \frac{1}{m^2+1}$，其中 (i, j) 满足 $j > i$。换言之，

$$\left(\sum_{i=1}^{m} s_i - 2k + m \right)^2 \geq \frac{1}{(m^2+1)^2} \text{ or } (1 - A_{ij})^2 (s_i + 1)^2 (s_j + 1)^2 \geq \frac{1}{(m^2+1)^2} \text{ for } (i, j),\ j > i. \tag{7}$$

由（1），我们得到结论。 ■

最后一条不涉及多项式 g的引理，是证明定理3所必需的。它由（Fujiwara，1916年）关于一元多项式根的模的一个界导出。

引理 7. 考虑一元函数 $f(t) = -\alpha_3|t|^3 + \alpha_2|t|^2 + \alpha_1|t| + \alpha_0$，对于 $t \in \mathbb{R}$，其中 α_0、α_1、α_2、$\alpha_3 > 0$.。若 $t^* \in \mathbb{R}$ 使得 $f(t^*) \geq 0$，则

$$|t^*| \leq 2 \max \left\{ \frac{\alpha_2}{\alpha_3}, \sqrt{\frac{\alpha_1}{\alpha_3}}, \sqrt[3]{\frac{\alpha_0}{2\alpha_3}} \right\}.$$

证明。设

$$u := 2 \max \left\{ \frac{\alpha_2}{\alpha_3}, \sqrt{\frac{\alpha_1}{\alpha_3}}, \sqrt[3]{\frac{\alpha_0}{2\alpha_3}} \right\} > 0.$$

假设存在 $t^* \in \mathbb{R}$，满足 $f(t^*) \geq 0$ 且 $|t^*| > u$，即 $t^* > u$ 或 $t^* < -u$。不失一般性，我们取 $t^* > u$；否则，只需将 $-t^*$ 视为 $f(t^*) = f(-t^*)$。令$\bar{f}(t) = -\alpha_3 t^3 + \alpha_2 t^2 + \alpha_1 t + \alpha_0$。根据介值定理，由于 $\bar{f}$ 连续，且 $\bar{f}(t^*) = f(t^*) \geq 0$，并且 $\lim_{t\to\infty} f(t) = -\infty$，因此必定存在 f 的一个根 $\bar{t}$，使得$\bar{t} \geq t^* > u$。这与 Fujiwara 界矛盾（Fujiwara，1916年）。■

现在我们可以证明定理 3 了。除 $x \in \mathbb{R}^m$ 外，还需要引入两个新的变量向量：

$$y := (y_0, \ldots, y_m) \in \mathbb{R}^{m+1} \text{ and } z := (z_{12}, \ldots, z_{1m}, z_{23}, \ldots z_{(m-1)m}) \in \mathbb{R}^{m(m-1)/2}.$$

定理的证明 *3.* 我们证明，这样的算法能够在多项式时间内求解 CLIQUE 。设 G 是一个有 m 个顶点、邻接矩阵为 A 的图，并设 $k \in \{1, \ldots, m\}$. 令 $n := m + (m+1) + {}^{m}_{(m-1)\,2}$，$N = 160m^2 \cdot (\overline{m^2+1})$，并令 $C = 2^{n+1}(m^2+1)$。定义 p 为一个n元三次多项式

$$p(x,y,z) := C\left(y_0\left(\sum_{i=1}^{m} x_i - 2k + m\right) + N \cdot \sum_{i=1}^{m} y_i\left(x_i^2 - 1\right) + \sum_i \sum_{j>i} z_{ij}(1 - A_{ij})(x_i+1)(x_j+1)\right). \tag{8}$$

从 (G,k) 到 p 的归约在长度上是多项式级的。事实上，p 的所有系数都形如 $C \cdot O(n^2)$，因此写出它们需要 $\lceil \log_2(C) \rceil + \lceil \log_2(O(n^2)) \rceil = O(n)$ 比特。此外，p 有 $O(n^2)$个这样的系数。p 的梯度为

$$\nabla p(x,y,z) = \begin{pmatrix} \nabla_x p(x,y,z) \\ \nabla_y p(x,y,z) \\ \nabla_z p(x,y,z) \end{pmatrix},$$

其中

$$\begin{aligned} \nabla_x p(x,y,z) &= C \begin{pmatrix} y_0 + 2Ny_1x_1 + \sum_{j>1} z_{1j}(1 - A_{1j})(x_j+1) \\ \vdots \\ y_0 + 2Ny_mx_m + \sum_{j>m} z_{mj}(1 - A_{mj})(x_j+1) \end{pmatrix}, \\ \nabla_y p(x,y,z) &= C \begin{pmatrix} \sum_{i=1}^{m} x_i - 2k + m \\ N(x_1^2 - 1) \\ \vdots \\ N(x_m^2 - 1) \end{pmatrix}, \end{aligned} \tag{9}$$

最后，$\nabla_z p(x,y,z)$的各项为 $C(1 - A_{ij})(x_i+1)(x_j+1)$，其中 $(i,j) \in \{1, \ldots, m\}^2, i < j$.

我们证明，如果 G中存在一个大小为 k 的团，那么 p 有一个临界点；如果 G中不存在一个大小为 k 的团，那么 p 甚至没有 2^n-近似临界点。证明这一断言后，定理便随之成立。为此，设$(\bar{x},\ \bar{y},\ \bar{z}) \in \mathbb{R}^n$ 是算法 $\mathcal{A}$返回的点。如果 $||\nabla p(\bar{x},\ \bar{y},\ \bar{z})|| \leq 2^n$，那么根据该断言，$G$ 中存在一个大小为 k 的团。如果$||\nabla p(\bar{x},\ \bar{y},\ \bar{z})|| > 2^n$，那么 p 没有临界点（否则算法就会返回一个 2^n-近似临界点）。根据该断言，G中不存在一个大小为 k 的团。因此，如果存在类似 $\mathcal{A}$ 的算法，我们就能在多项式时间内求解 CLIQUE，这将意味着 P=NP。

我们首先注意到，如果 k 中存在大小为 G 的团，那么 p 有一个临界点。事实上，根据引理 4，如果 G 中存在大小为 k 的团，则（2）是可行的。令 $\bar{x}$ 为（2）的一个解，$\bar{y} = 0$，$\bar{z} = 0$，我们有 ∇p（$\bar{x}$，$\bar{y}$，$\bar{z}$）$= 0$。

反之，假设 G 中不存在大小为 k 的团。我们证明，在这种情况下 ，

$$||\nabla p(x,y,z)|| > 2^n, \forall x, y, z. \tag{10}$$

令 x，y，$z \in \mathbb{R}^n$。我们有

$$||\nabla p(x,y,z)||^2 \geq ||\nabla_y p(x,y,z)||^2 + ||\nabla_z p(x,y,z)||^2 = g(x),$$

其中 $g(x)$ 是在 (1) 中定义的多项式。注意，g 是一个最高次项为 $C^2\left(N^2 \sum_{i=1}^{m} x_i^4 + \sum_i \sum_{j>i} (1 - A_{ij})^2 x_i^2 x_j^2\right)$ 的四次多项式，该最高次项是正定的，即取正值

对非零的 $x_\circ$ 这意味着 g 是强制的（Jeyakumar 等人，2014，第 4.2 节），因而达到其下确界（Bertsekas，1995，附录 A.2）。设 x^* 为 g 的一个极小化点。我们给出 $g(x^*)$ 的一个下界，它是 $g(x)$ 的一个下界，而这又是 $||\nabla p(x,y,z)||_2^2.$ 的一个下界。为此，我们通过一系列断言证明：对某个 $\epsilon > 0$ 以及所有 $i = 1, \ldots, m$，有 $|(x_i^*)^2 - 1| \leq \epsilon_\circ$ 这使我们能够利用引理 5 和引理 6 得出结论。

我们的断言证明使用了无约束极小化的一阶必要条件（一阶必要条件）和二阶必要条件（SONC）（Bertsekas，1995，命题 1.1.1）。因此，我们计算 g的一阶和二阶导数。对于$i = 1, \ldots\ldots, m$，我们有

$$\frac{\partial g(x)}{\partial x_i} = C^2 \left(2 \left(\sum_{j=1}^{m} x_j - 2k + m \right) + 4N^2 x_i (x_i^2 - 1) + 2 \sum_{j>i} (1 - A_{ij})^2 (x_i + 1)(x_j + 1)^2 \right). \tag{11}$$

此外，同样对于 $i = 1, \ldots\ldots, m$，我们有

$$\frac{\partial^2 g(x)}{\partial x_i^2} = C^2 \left(2 + 4N^2 (3x_i^2 - 1) + 2 \sum_{j>i} (1 - A_{ij})^2 (x_j + 1)^2 \right). \tag{12}$$

断言 1： $0.3 \leq |x_i^*| \leq 2.1$ **对于** $i = 1, \ldots\ldots, m_\circ$

我们首先证明上界。由于 x^* 是 g的极小化点，将一阶必要条件应用于 x^* 处的 g，可得 $\nabla g(x^*) = 0$，也就是说，对于任意 $i = 1, \ldots\ldots, m$，

$$-4C^2 N^2 x_i^* ((x_i^*)^2 - 1) = C^2 \left(2 \left(\sum_{j=1}^{m} x_j^* - 2k + m \right) + 2 \sum_{j>i} (1 - A_{ij})^2 (x_i^* + 1)(x_j^* + 1)^2 \right).$$

对等式两边取绝对值，并通过三角不等式对等式右侧进行上界估计，我们得到：

$$4C^2 N^2 |x_i^*| \cdot |(x_i^*)^2 - 1| \leq C^2 \left(2 \sum_{j=1}^{m} |x_j^*| + 4k + 2m + 2 \sum_{j>i} (1 - A_{ij})^2 (|x_i^*| + 1)(|x_j^*| + 1)^2 \right). \tag{13}$$

随后，我们可以利用反三角不等式对不等式左边给出下界。对于任意 $i = 1, \ldots, m$，可得

$$4C^2 N^2 |x_i^*| \cdot (|x_i^*|^2 - 1) \leq C^2 \left(2 \sum_{j=1}^{m} |x_j^*| + 4k + 2m + 2 \sum_{j>i} (1 - A_{ij})^2 (|x_i^*| + 1)(|x_j^*| + 1)^2 \right).$$

令 $i_+ \in \arg\max_{i=1,\ldots,m} |x_i^*|$. 将 $i = i_+$ 代入前一个不等式，并注意到对于任意$j = 1, \ldots, m$，$|x_j^*| \leq |x_{i_+}^*|$，可得:

$$4C^2 N^2 |x_{i_+}^*| \cdot (|x_{i_+}^*|^2 - 1) \leq C^2 \left(2m |x_{i_+}^*| + 6m + 2m (|x_{i_+}^*| + 1)^3 \right),$$

其中我们还使用了 $k \leq m$ 和 $(1 - A_{ij})^2 \leq 1$. 展开并重新整理可得

$$C^2 \left((2m - 4N^2) |x_{i_+}^*|^3 + 6m |x_{i_+}^*|^2 + (8m + 4N^2) |x_{i_+}^*| + 8m \right) \geq 0.$$

回顾 $N = 80m^2 \cdot 2(m^2+1)$，因此 $4N^2 = 25600m^4 \cdot 4(m^2+1)^2$。对于任意 $m \geq 1$，可得 $C^2(2m - 4N^2) < 0$，$6mC^2 > 0$，$C^2(8m + 4N^2) > 0$，且 $8mC^2 > 0$，引理 7适用。容易证明，对于任意 $m \geq 1$，

$$\max\left\{\frac{6m}{4N^2-2m}, \sqrt{\frac{8m+4N^2}{4N^2-2m}}, \sqrt[3]{\frac{8m}{4N^2-2m}}\right\} = \sqrt{\frac{8m+4N^2}{4N^2-2m}}.$$

引理 7 则意味着

$$|x^*_{i_+}| \leq 2\sqrt{\frac{8m+4N^2}{4N^2-2m}} = 2\sqrt{1+\frac{10m}{4N^2-2m}} = 2\sqrt{1+\frac{10}{25600m^3\cdot 4(m^2+1)^2-2}} \text{ for } i = 1,\ldots,m.$$

我们有 $25600m^3 \cdot 4(m^2+1)^2 - 2 \geq 25600$ 对任意 $m \geq 1$，因此，

$$|x_i| \leq |x^*_{i_+}| \leq 2 \cdot \sqrt{1+\frac{10}{25600}} \leq 2.1 \text{ for } i = 1,\ldots,m. \tag{14}$$

为了求得 $|x^*_i|$的下界，我们将 SONC 应用于 g 在 x^*处。这表明$\nabla^2 g(x^*)$ 是半正定的，因此其对角线元素非负。回顾 (12)，对任意$i = 1,\ldots,m$，我们有

$$C^2\left(2 + 4N^2(3(x^*_i)^2 - 1) + 2\sum_{j>i}(1-A_{ij})^2(x^*_j+1)^2\right) \geq 0,$$

根据三角不等式和 (14)，可知

$$C^2\left(2 + 4N^2(3(x^*_i)^2 - 1) + 2m(2.1+1)^2\right) \geq 0.$$

整理后可得

$$(x^*_i)^2 \geq \frac{1}{3} - \frac{2+2m(2.1+1)^2}{12N^2} \geq \frac{1}{3} - \frac{22m}{3\cdot 25600m^4\cdot 4(m^2+1)^2} \geq \frac{1}{3} - \frac{22}{3\cdot 25600} \geq 0.3,$$

其中，第二个不等式用到了对任意 $m \geq 1$，$2+2m(2.1+1)^2 \leq 22m$ 这一事实；第三个不等式用到了对所有 $m \geq 1$，$m^3 \cdot 4(m^2+1)^2 \geq 1$ 这一事实。对两边开平方，便得到断言 1 中的下界。

断言 2：设 $\epsilon := \frac{0.005}{m^2(m^2+1)^2}$。**那么** $|(x^*_i)^2 - 1| \leq \epsilon$，**对** $i = 1,\ldots,m$.

我们首先利用断言 1，以更精细的方式对 $|(x^*_i)^2 - 1|$ 进行界定。设 $i \in \{1,\ldots,m\}$.。我们有以下不等式组：

$$\begin{aligned}
&4C^2N^2 \cdot 0.3 \cdot |(x^*_i)^2 - 1| \leq 4C^2N^2 \cdot |x^*_i| \cdot |(x^*_i)^2 - 1| \\
&\leq C^2\left(2\sum_{j=1}^{m}|x^*_j| + 4k + 2m + 2\sum_{j>i}(1-A_{ij})^2(|x^*_i|+1)(|x^*_j|+1)^2\right) \\
&\leq C^2\left(2m\cdot 2.1 + 6m + 2m(2.1+1)^3\right),
\end{aligned}$$

其中，第一个不等式用到了断言 1，第二个不等式恰好是（13），第三个不等式则利用了（14）以及以下事实：$k \leq m$和 $1 - A_{ij} \leq 1$.。然后，我们将不等式两边同时除以 $4C^2N^2 \cdot 0.3$，即可得到断言 2：

$$|(x^*_i)^2 - 1| \leq \frac{2m\cdot 2.1 + 6m + 2m(2.1+1)^3}{4\cdot 0.3N^2} \leq \frac{235m}{25600m^4\cdot 4(m^2+1)^2} \leq \frac{0.005}{m^3(m^2+1)^2} \leq \epsilon. \tag{15}$$

由断言 2 可知，引理5 推出 $g(x^*) \geq g(sgn(x^*)) - 48C^2m^2\epsilon$。注意，$sgn(x^*) \in \{\pm 1\}^m$，但集合 $\{i \in \{1,\ldots,m\} \mid sgn(x^*)_i = 1\}$ 不可能是 G 中大小为 k 的团（根据假设，k 中不存在大小为 G 的团）。因此，由引理6可得:

$$g(sgn(x^*)) \geq \frac{C^2}{(m^2+1)^2}.$$

结合这两个不等式，我们得到 $g(x^*) \geq \frac{C^2}{(m^2+1)^2} - 48C^2m^2\epsilon$. 回想

$$||\nabla p(x,y,z)||_2^2 \geq g(x)$$

且 x^* 是 g 的极小化点，我们得到:

$$||\nabla p(x,y,z)||_2^2 \geq g(x) \geq g(x^*) \geq \frac{C^2}{(m^2+1)^2} - 48C^2m^2\epsilon.$$

由于 $\epsilon = \frac{0.005}{m^2(m^2+1)^2}$, $C = 2^{n+1}(m^2+1)$ 和 $48 \cdot 0.005 \leq 1/4$，因此

$$||\nabla p(x,y,z)||_2^2 \geq \frac{2^{2n+2}(m^2+1)^2}{(m^2+1)^2} - \frac{2^{2n+2}(m^2+1)^2 \cdot 48m^2 \cdot 0.005}{m^2(m^2+1)^2} = 2^{2n+2}(1-\frac{1}{4}) \geq 2^{2n}.$$

然后对两边开平方，便得到我们所需的不等式（10）。 ■

我们指出，如果在证明中将 $C = 2^{\text{poly}(n)}(m^2+1)$ 设为 $C = 2^{n+1}(m^2+1)$，那么除非 P = NP，否则不可能在多项式时间内计算出一个 $2^{\text{poly}(n)}$-近似临界点。我们对常数 C作了更具体的选择，因为我们认为定理 3 的结论已经足够负面。更有意思的是，在定理 3的证明中，对某个 $d \in \mathbb{N}$，将$C = n^d(m^2+1)$ 设为 n^d-近似临界点时，若目标是得到一个不那么粗糙的 n^d-近似临界点，那么我们甚至可以排除伪多项式时间算法（除非 P =NP）。

推论 8. 设 d为任意正整数。如果存在一个伪多项式时间算法 $\mathcal{A}$，它以含 n 个变量的三次多项式 p 为输入，并且满足：(i) 当 p没有临界点时，$\mathcal{A}$ 可以返回任意一点，(ii)当 p有临界点时，$\mathcal{A}$返回一个 n^d-近似临界点，则 $P = NP$.

围绕定理3的讨论。

为具体说明定理3的含义，考虑梯度下降算法。定理3 表明，除非 P =NP，否则存在 n元三次多项式；对于这些多项式，无论使用何种可高效计算的步长，梯度下降都需要指数级的迭代次数，才能返回哪怕一个 2^n-近似临界点。事实上，由复杂性理论中更为精细的结果可知，定理 3 还进一步表明，这种行为必然出现在三次 n元多项式的“相当大一部分”上；其精确的复杂性理论表述见（Hemaspaandra 和 Williams，2012年，推论 2.2）。

更一般地，除非 P=NP，否则定理 3 排除了存在这样一种算法的可能性：该算法能够在关于（ϵ）次变量的函数中，找到一个 n-近似临界点，且运行时间关于（$n, 2^{\frac{1}{\epsilon}}$）为多项式（同时也关于函数实例的自然描述大小为多项式）。事实上，若这样的算法存在，那么令 $\epsilon = 2^n$，所得运行时间将关于 n 和实例的描述大小为多项式，这与定理 3矛盾。

我们还声称，即使在保证存在 p 的临界点的情况下，计算 2^n-近似临界点仍然是困难的。事实上，我们可以证明，如果存在一种多项式时间算法，其输入为一个 n元三次多项式，并承诺该多项式存在临界点，且输出一个 2^n-近似临界点，

那么 PLANTED CLIQUE猜想（Jerrum，1992；Kuˇcera，1995）就会被推翻。这个著名猜想的输入是两个正整数 m 和 $k \in \{1, \ldots, m\}$，并通过以概率 $1/2$独立连接每一对顶点，在 m 个顶点上构造一个 Erd¨os-R´enyi 图。然后，从原有的 m 个顶点中随机选取一个包含 k 个顶点的子集，并添加这些顶点之间的所有边，从而在图中植入一个大小为 k 的团。该猜想被广泛认为成立，其内容是：除非 P=NP，否则只要 $k = o(\sqrt{m})$（且对任意常数$\epsilon > 0$，$k \geq (2+\epsilon)\log(m)$），就不可能在多项式时间内找回这个团（Chen 等人，2025）。我们省略上述论断的证明：该证明采用与定理 3 的证明相同的归约，随后证明，如果算法返回的、式 (8) 中多项式 p 的一个 2^n-近似临界点 $(\bar{x}, \bar{y}, \bar{z})$ 满足条件，那么 $sgn(\bar{x}) \in \{\pm 1\}^m$ 就是 G中大小为 k 的团的指示向量。

需要注意的是，在植入团设定下，该图以高概率存在一个大小为 k (Bollob´as 和 Erd¨os，1976年）的唯一团，这意味着在上述归约中构造的多项式 p 有唯一临界点。因此，在存在临界点的基础上，再额外保证临界点唯一，并不会让寻找近似临界点这一任务变得更容易。事实上，将我们的归约与 Valiant 和 Vazirani 的开创性结果（Valiant 和 Vazirani，1985）结合起来，我们可以证明：若存在一种多项式时间算法，能够为一个保证具有唯一临界点的 2^n-元三次多项式找到 n-近似临界点，那么任何 NP 问题都可以在随机多项式时间内求解（即 NP= RP）。

作为关于定理3的最后一点说明，我们简要讨论其对高阶牛顿法设计的影响（Nesterov，2021年；Silina 和 Zhang，2022年；Ahmadi 等人，2024年；Cartis 等人，2022年）。回想一下，在经典牛顿法的每次迭代中，我们都会计算目标函数在当前迭代点处的二阶泰勒展开的一个临界点。这种计算是可行的，因为它可归结为求解一个线性系统，且所需时间关于维数为多项式级。高阶牛顿法旨在通过改为使用三阶或更高阶的泰勒展开，提升牛顿法的收敛速率和预言机复杂度。尽管近期研究提出了每次迭代复杂度可行的高阶方法（Nesterov，2021年；Ahmadi 等人，2024年），但这些算法比牛顿法的直接推广更为复杂；所谓直接推广，是指计算高阶泰勒展开的一个（近似）临界点。定理 3 表明，除非 P=NP，否则这种最自然的牛顿法高阶推广，在泰勒展开为三阶时就已经不具备可行的每次迭代复杂度。

2.2 计算下有界函数的 ϵ-近似临界点的复杂度

在本节中，我们探讨：如果已知函数下有界，计算其近似临界点是否仍然难以处理。首先注意，判定一个下有界函数是否存在临界点是 NP 难的。

命题9。 如果存在一个多项式时间算法，能够判定三次多项式的指数函数（它总是下有界的）是否存在临界点，那么 *P=NP*。

证明。 函数 $f : \mathbb{R}^n \to \mathbb{R}$ 有临界点，当且仅当 e^f（它是下有界的）有临界点。事实上，$\nabla e^{f(x)} = e^{f(x)} \nabla f(x)$。已知判定一个三次多项式是否存在临界点是 NP 难的（Ahmadi and Zhang，2022a），由此证明得证。 ■

下一个命题表明，即使增加额外假设，即保证临界点存在且不存在任何伪临界点[2]，，计算问题仍然

[2]我们称 如果函数的任意临界点都是全局最小值，则称该函数不存在伪临界点。注意， at 不存在 不存在伪临界点意味着不存在伪局部极小值。

下有界函数的临界点依然很难。事实上，对于 $x \in \mathbb{R}^n$，令$\{\tilde{q}_i(x) = 0,\ i = 1, \ldots\ldots, r\}$为一组已知存在（有理数）解、但求解困难的二次方程。例如，这样一组二次方程可以编码 Planted Clique问题（见第2.1节），也可以编码在双矩阵博弈中寻找纳什均衡的问题[3]（参见（Daskalakis等人，2009年；Chen与Deng，2006年），其中给出了计算纳什均衡的困难性结果以及复杂性类PPAD的定义）。

命题10。 如果存在一种多项式时间算法，其输入为一个下有界函数（一个 n 元四次多项式与一个单变量指数函数的乘积），且保证该函数具有临界点但没有伪临界点，并返回一个临界点，那么 Planted Clique 猜想将被推翻，且 *PPAD=P*。

证明。如上文所述，令 $\{\tilde{q}_i(x) = 0,\ i = 1, \ldots\ldots, r\}$ 为 n 个变量上的一组二次方程，其解对应于 PlantedClique 问题中的植入团，或对应于双矩阵博弈的纳什均衡。定义函数 $f: \mathbb{R}^{n+1} \to \mathbb{R}$ 为

$$f(x, w) = e^w \cdot \left(\sum_{i=1}^{r} \tilde{q}_i^2(x) \right).$$

显然，f 的下界为零。此外，我们有

$$\nabla f(x, w) = \begin{pmatrix} 2e^w \cdot \sum_{i=1}^{r} \tilde{q}_i(x) \nabla q_i(x) \\ e^w (\sum_{i=1}^{r} \tilde{q}_i^2(x)) \end{pmatrix}.$$

因此，f 的任意临界点（$\bar{x}$，$\bar{w}$）都必须满足：对于 $i = 1, \ldots\ldots, r$，$\tilde{q}_i(\bar{x}) = 0$。这意味着，在任意临界点（$\bar{x}$，$\bar{w}$）处，我们有 $f(\bar{x}, \bar{w}) = 0$，因此 f 没有伪临界点。这还意味着，任何返回 f 的临界点（$\bar{x}$，$\bar{w}$）的多项式时间算法，都可以通过向量 $\bar{x}$，在多项式时间内为我们提供所关注的植入团或该博弈的纳什均衡。由此得证两个结论。■

我们指出，命题 10 与一种相当普遍的看法相悖：最小化一个没有伪局部极小值的函数是易于处理的。现在我们研究寻找下有界函数的一个 近似临界点的复杂度。已知下有界函数总是存在一个 ϵ-近似临界点，任意 $\epsilon > 0$ (Ekeland，1974年）。因此，问题在于寻找这样的点是否易于处理。我们现在证明，在一个大小合理的集合内寻找近似临界点是难以处理的。

定理 11。 如果存在一个算法 $\mathcal{A}$，其输入为一个含有 n 个变量的下有界四次多项式p，运行时间为多项式时间，并满足

(i) 当 p 在超立方体 $[-1, 1]^n$中没有临界点时，$\mathcal{A}$ 可以返回任意点；*(ii)* 当 p 在超立方体 $[-1, 1]^n$中有临界点时，$\mathcal{A}$ 会在以原点为中心、半径为 2^n 的球内返回一个 2^n-近似临界点，

那么 *P=NP*。

请注意，对 $\mathcal{A}$ 的保证要求非常宽松：除非超立方体 $[-1, 1]^n$中存在临界点，否则算法可以任意运行；即使在这种情况下，也只要求它返回一个非常粗略的近似临界点，且该点位于一个

[3]纳什均衡可以用一组二次不等式来定义，这些不等式确保转而采用纯策略时不会出现单边收益改进。随后，可以通过引入松弛变量将其转化为一组二次方程。

非常宽松的球内。我们的看法是，在大多数应用中，人们关心的是找到 ϵ-近似临界点，且其大小不至于大得不合理。这类情形已由定理11涵盖。若只要求算法输出的范数在维数上呈指数级增长，我们尚未解决下有界函数的近似临界点搜索复杂度。

定理11的证明我们证明，这样的算法能够在多项式时间内求解 CLIQUE 。设 G 是一个有 m 个顶点的图，其邻接矩阵为 A，并令 $k \in \{1,\dots,m\}$. 令$n = m + (m+1) + {}^{m}{}_{(m-1)\,2}$， $N = 80m^2 \cdot 2(m^2+1)$，并设 $C = 2^{n+2}(m^2+1)$。定义 $q(x)$ 为一个大小为 $m+1+{}^{m}{}_{(m-1)\,2}$ 的向量，其分量 $q_i(x)$ 定义如下

$$q_1(x) = \sum_{i=1}^{m} x_i - 2k + m, \quad q_2(x) = N(x_1^2 - 1), \quad \dots, \quad q_{m+1}(x) = N(x_m^2 - 1),$$

$$q_{m+2}(x) = (1 - A_{12})(x_1 + 1)(x_2 + 1)\dots\dots, \qquad q_{m+1+m(m-1)/2} = (1 - A_{(m-1)m})(x_{m-1} + 1)(x_m + 1)。$$

令 $w := (w_1, \dots\dots, w_{m+1+m(m-1)/2})$，然后定义 p 为一个 n 元四次多项式

$$p(x,w) = \frac{1}{2}||w + Cq(x)||^2.$$

显然， $p(x, w)$ 的下界为零。此外，从 (G, k) 到 p 的归约在长度上是多项式的，因为写出 p 的系数所需的比特数关于 (G, k) 的规模是多项式的。

令 B 为以原点为中心、半径为 2^n 的球，并令 $\mathcal{H} = [-1,1]^n$。我们证明：（i'）如果 G 中存在大小为 k 的团，则 p 在 $\mathcal{H}$ 中有一个临界点；以及（ii'）如果 G 中不存在大小为 k 的团，则 p 在 B 内甚至没有 2^n-近似临界点。一旦我们证明了断言（i'）和（ii'），定理即得证。为此，令（$\bar{x}$, $\bar{w}$）为算法 $\mathcal{A}$. 返回的点。如果 $||(\bar{x}, \bar{w})|| \le 2^n$ 且 $||\nabla p(\bar{x}, \bar{w})|| \le 2^n$，则根据断言（$ii'$），$G$ 有一个大小为 k 的团。如果 $||(\bar{x}, \bar{w})|| > 2^n$ 或 $||\nabla p(\bar{x}, \bar{w})|| > 2^n$，则 p 在 $\mathcal{H}$ 中没有临界点（否则算法会在 $\mathcal{B}$ 中返回一个 2^n-近似临界点），并且根据断言（i'），G 中不存在大小为 k 的团。因此，如果存在类似 $\mathcal{A}$ 的算法，我们就能在多项式时间内求解 CLIQUE，这将意味着 P=NP。

我们首先证明命题（i'）；也就是说，如果 G 中存在一个大小为 k 的团，那么 p 在 $\mathcal{H}$ 内有一个临界点。为此，我们再次使用引理 4，得出（2）是可行的。注意到

$$\nabla p(x,w) = \begin{pmatrix} CJ_q(x)^T(w + Cq(x)) \\ w + Cq(x) \end{pmatrix}, \tag{16}$$

其中 $J_q(x)$ 是 q 的雅可比矩阵，并令 $\bar{w} = 0$ 和 $\bar{x}$ 为（2）的一个解，则有$\nabla p(\bar{x}, \bar{w}) = 0$。此外，$\bar{x} \in \{\pm 1\}^m$ 和 $\bar{w} = 0$，这意味着 $(\bar{x}, \bar{w}) \in \mathcal{H}$.

现在我们证明命题（ii'）。我们证明，如果 G 中不存在大小为 k 的团，那么

$$||\nabla p(x,w)|| > 2^n, \forall (x,w) \in B.$$

设（x, w）$\in B$。注意， $||\nabla p(x, w)||^2 \ge ||w + Cq(x)||^2$。利用反三角不等式，我们得到

$$||\nabla p(x,w)|| \ge ||w + Cq(x)|| \ge |C||q(x)|| - ||w||\ |. \tag{17}$$

注意， $C^2||q(x)||^2 = g(x)$如（1）中所定义。与定理 3 的证明类似，由于该证明中的断言 1 和断言 2 对 g 的极小化点 x^* 以及 $\epsilon = {}_{m^2}{}^{0.005}{}_{m^2+1}{}^{()^2}$，均成立，我们可以使用引理5，并得出 $g(x^*) \ge g(sgn(x^*)) - 48C^2m^2\epsilon$. 。根据假设， G 中不存在大小为 k 的团，因此我们可以再次应用引理 6，得到 $g(sgn(x^*)) \ge {}^{C^2}{}_{(m^2+1)^2} \cdot$。结合这两个不等式，可得

$$g(x) \ge g(x^*) \ge \frac{C^2}{(m^2+1)^2} - 48C^2m^2\epsilon.$$

将 ϵ 替换为其值并取平方根，可得

$$C||q(x)|| = \sqrt{g(x)} \geq \frac{\sqrt{3}C}{2(m^2+1)}.$$

将其代入（17），并注意到当（x， w）$\in B$时，我们有 $||w|| \leq 2^n$，从而得到

$$||\nabla p(x,w)|| \geq \frac{\sqrt{3}C}{2(m^2+1)} - 2^n.$$

回想一下 C 的选取，我们可知 $||\nabla p(x,\ w)|| > 2^n$。证明至此完成。 ■

3 计算 ϵ-近临界点的复杂度

I本节聚焦于 ϵ-近临界点这一概念，下面给出其定义 .

定义 12 （ϵ-近临界点）。设 $f\colon \mathbb{R}^n \to \mathbb{R}$ 是一个可微函数，且 $\epsilon > 0$。若存在 f 的某个临界点 x^*，使得 $||\bar{x} - x^*|| \leq \epsilon$，则点 $\bar{x} \in \mathbb{R}^n$ 称为 f 的 ϵ-近临界点。

近临界点的概念是近似临界点概念的一种非常自然的替代。事实上，临界点 x^* 是优化问题 $\min_x ||\nabla f(x)||.$ 的最优解。因此，近似临界点是近似该问题最优值的点，而近临界点则是该问题最优解的近似；一维情形下这两个概念的区别见图 1。与近似临界点相比，近临界点这一概念在文献中或许并没有那么广泛使用。我们猜想这是因为，对于给定的 $\epsilon > 0$，判断某个给定点是否为函数的 ϵ-近临界点是 NP难的[4]，如下述简单归约所示。设 G 是一个具有 m 个顶点的图，且 $k \in \{1, \ldots, m\}$。令 p 为式（8）中定义的 n 元三次多项式， $u = 0$（$\mathbb{R}^n$ 中的原点），且 $\epsilon = \lceil\sqrt{m}\rceil$。我们在定理 3 的证明中已经看到，若 G 有一个大小为 k 的团，则 p 存在一个临界点，满足 $x^* \in \{\pm 1\}^m$、$y^* = 0$、$z^* = 0$（其中 x^* 是该团的指示向量）。因此， u 是 p 的一个 $\sqrt{m}$-近临界点。若 G 没有大小为 k 的团，则 p 没有临界点，因此对于任何 $\epsilon' > 0$ （特别是 $\epsilon' = \lceil\sqrt{m}\rceil$），$p$ 都没有 ϵ'-近临界点。

在接下来的小节中，我们研究计算近临界点的复杂度。

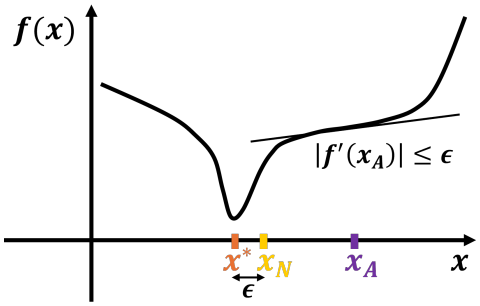

图 1：一维示例展示了近临界点与近似临界点之间的区别： x^* 是 f的临界点；对于给定的 $\epsilon > 0$， x_N 是一个 ϵ-近临界点，而 x_A 是一个 ϵ-近似临界点。

[4]对于判断一个点 $\bar{x}$ 是否为局部极小值（Murty 和 Kabadi，1987）。

3.1 Co计算一般函数的 ϵ-近临界点的复杂度

我们的第一个结果是定理的类比3 ，针对近临界点。

定理 13. 如果存在一个算法 $\mathcal{A}$ ，它以 n 个变量的三次多项式 p 作为输入，在多项式时间内运行，并且满足

（i）当 p 没有临界点时， $\mathcal{A}$ 可以返回任意一点，

（ii）当 p 有临界点时， $\mathcal{A}$ 返回一个 $\frac{1}{12000\cdot n^{3.5}}$-近临界点，

则 $P = NP$.

为证明该定理，我们首先给出一个技术引理。回顾：函数 f 在盒 $B \subset \mathbb{R}^n$ 上是 L-利普希茨的，如果 $|f(x) - f(y)| \leq L||x - y||, \forall x, y \in B$.

引理 14。设 G 是一个有 m 个顶点的图， $k \in \{1, \ldots, m\}$， $C \in \mathbb{R}$，且 $N = 160m^2(m^2 + 1)$。回顾多项式 $g = g_{G,\ C,\ k}$ 在 (1)中的定义。对于 $\epsilon \in (0,\ 1)$， g 是 L-利普希茨的，定义在 $[-1 - \epsilon, 1 + \epsilon]^m$ 上，其中 $L = 64C^2m^2 + 24C^2mN^2\epsilon$.

证明。令 $\epsilon \in (0,\ 1)$ 且 $x \in [-1 - \epsilon, 1 + \epsilon]^m$。只需证明 $||\nabla g(x)|| \leq L$. 我们有

$$||\nabla g(x)|| \leq ||\nabla g(x)||_1 = \sum_{i=1}^{m} \left| \frac{\partial g(x)}{\partial x_i} \right|.$$

固定 $i \in \{1, \ldots, m\}$。由 (11)，可得

$$\begin{aligned}
\left| \frac{\partial g(x)}{\partial x_i} \right| &= \left| C^2 \left(2 \left(\sum_{j=1}^{m} x_j - 2k + m \right) + 4N^2 x_i (x_i^2 - 1) + 2 \sum_{j>i} (1 - A_{ij})^2 (x_i + 1)(x_j + 1)^2 \right) \right| \\
&\leq C^2 \left(2 \left(\sum_{j=1}^{m} |x_j| + 3m \right) + 4N^2 |x_i| \cdot |x_i^2 - 1| + 2 \sum_{j>i} (|x_i| + 1) \cdot (|x_j| + 1)^2 \right) \\
&\leq C^2 \left(2m(1 + \epsilon) + 6m + 4N^2(1 + \epsilon)(3\epsilon) + 2m(2 + \epsilon)^3 \right) \leq C^2(64m + 4N^2 \cdot 6\epsilon),
\end{aligned}$$

这里我们在第一个不等式中使用了三角不等式以及以下事实： $k \leq m$ 且 $(1 - A_{ij}) \leq 1$ ；还使用了以下事实： $|x_j| \leq 1 + \epsilon$ ，因此

$$|x_j^2 - 1| = |\ |x_j|^2 - 1| = |\ |x_j| - 1| \cdot |\ |x_j| + 1| \leq \epsilon(2 + \epsilon) \leq 3\epsilon$$

在第二个不等式中，以及第三个不等式中 $\epsilon < 1$ 这一事实。将此不等式两边对 i求和，我们得到

$$||\nabla g(x)|| \leq \sum_{i=1}^{m} \left| \frac{\partial g(x)}{\partial x_i} \right| \leq mC^2(64m + 4N^2 \cdot 6\epsilon) = L.$$

■

定理 13 的证明。我们证明，这样的算法能在多项式时间内求解 Clique 。设 G 是一个具有 m 个顶点且邻接矩阵为 A 的图，并设 $k \in \{1, \ldots, m\}$。我们令$n := m + (m + 1) + {}^{m}(_{m-1})\ 2$ ， $N = 160m^2(m^2 + 1)$，以及 $C = 1$。定义 $p = p(x, y, z)$为(8)中的n元三次多项式。如前所述，从(G, k)到 p 的归约在长度上是多项式的。

令 $\alpha = \frac{1}{12000\cdot n^{3.5}}$ 我们证明，如果 $(x,\ y,\ z)$ 是 p 的 α-近临界点，那么 $sgn(x)$ 就是一个大小为 k 的团的指示向量。证明这一断言后，定理便得证。事实上，假设算法 $\mathcal{A}$ 在多项式时间内返回的点为 $(\bar{x},\ \bar{y},\ \bar{z})$。

我们只需检查 $sgn(\bar{x})$ 是否是 G 中大小为 k 的团的指示函数。如果是，G 显然有一个大小为 k 的团。如果不是，那么根据该断言，$(\bar{x},\ \bar{y},\ \bar{z})$ 必定不是 α-近临界点。因此，p 没有临界点（否则该算法就会返回一个 α-近临界点）。由定理 3的证明可知，G 中不存在大小为 k 的团。因此，如果存在像 $\mathcal{A}$ 这样的算法，我们就能在多项式时间内求解 CLIQUE，从而推出 P=NP。

现在证明这个结论。设 (x,y,z) 是一个 α-近临界点，即

$$\left|\left|\begin{pmatrix}x\\y\\z\end{pmatrix}-\begin{pmatrix}x^*\\y^*\\z^*\end{pmatrix}\right|\right|\leq\alpha=\frac{1}{12000\cdot n^{3.5}},$$

对于 p 的某个临界点 (x^*,y^*,z^*)。（注意，如果 (x,y,z) 是 p 的 α-近临界点，那么 p 必定存在一个临界点。）我们证明 $s:=sgn(x)$ 是 G 中大小为 k 的团的指示向量。令 $g(x)$ 如 (1) 中所定义。为证明该结论，我们首先证明 $g(s)<\ ^1\!/_{(m+1)^4}$，然后证明对于 G 中任何不是大小为 k 的团的指示向量的 $z\in\{-1,1\}^m$，都有 $g(z)\geq\ ^1\!/_{(m+1)^4}$。

由于 (x,y,z) 是 p 的一个 α-近临界点，因此有 $||x-x^*||\leq\alpha$，从而

$$|x_i-x_i^*|\leq\alpha,\forall i=1,\ldots,m.$$

此外，$x^*\in\{\pm1\}^m$ 由于（x^*，y^*，z^*）是 p 的临界点（见（9））。因此，

$$\begin{aligned}|x_i^2-1|&=|x_i^2-(x_i^*)^2|=|x_i-x_i^*|\cdot|x_i+x_i^*|=|x_i-x_i^*|\cdot|x_i-x_i^*+2x_i^*|\\&\leq|x_i-x_i^*|\cdot(|x_i-x_i^*|+2|x_i^*|)=|x_i-x_i^*|^2+2|x_i-x_i^*|\leq\alpha^2+2\alpha,\ \text{for } i=1,\ldots,m.\end{aligned}$$

应用引理5 并回想起 $C=1$，我们得出

$$g(s)\leq g(x)+48m^2(\alpha^2+2\alpha).\tag{18}$$

现在可以利用引理 g（x）对其作上界估计。由引理14，

$$|x_i-x_i^*|\leq\alpha,\forall i=1,\ldots,m,$$

且 $x^*\in\{\pm1\}^m$，我们有 x，$x^*\in[-1-\alpha,1+\alpha]^m$. 由于 $\alpha<1$，引理14 推出

$$|g(x)-g(x^*)|=|g(x)|\leq(64m^2+24mN^2\alpha)||x-x^*||\leq(64m^2+24mN^2\alpha)\alpha,$$

这里用到了 g（x^*）$=0$ 由 $\nabla_y p$（x^*，y^*，z^*）$=0$. 将其代回（18），并代入 N 和 α的值，得到:

$$\begin{aligned}g(s)&\leq(64m^2+24mN^2\alpha)\cdot\alpha+48m^2(\alpha^2+2\alpha)\\&=\frac{64m^2}{12000n^{3.5}}+\frac{24m\cdot160^2m^4(m^2+1)^2}{12000^2n^7}+48m^2\left(\frac{1}{12000^2n^7}+\frac{2}{12000n^{3.5}}\right)\\&\leq\frac{64m^22^{3.5}}{12000(m+1)^7}+\frac{24\cdot160^2m^5(m^2+1)^2\cdot2^7}{12000^2(m+1)^{14}}+48m^2\left(\frac{2^7}{12000^2(m+1)^{14}}+\frac{2\cdot2^{3.5}}{12000(m+1)^7}\right)\\&\leq\frac{0.1}{(m+1)^5}+\frac{0.6}{(m+1)^5}+\frac{0.01}{(m+1)^{12}}+\frac{0.1}{(m+1)^5}<\frac{1}{(m+1)^4},\end{aligned}$$

这里在第三行使用了不等式 $n\geq\ ^{(m+1)^2}\!/_2$，在第四行使用了 $(m^2+1)\ \leq\ (m+1)^2$。

现在，设 $z\in\{\pm1\}^m$，并假设 z 不是 G中大小为 k 的团的指示向量。我们有

$$g(z)\geq\frac{C^2}{(m^2+1)^2}=\frac{1}{(m^2+1)^2}\geq\frac{1}{(m+1)^4},$$

根据引理6，其中我们用到了以下事实：$m^2+1\leq(m+1)^2$。证明至此完成。■

定理讨论13。

除非 P=NP，否则定理13 排除了存在这样一种算法的可能性：它能在多项式时间内（关于 ϵ，以及函数实例的自然描述大小）找到一个函数的 n 变量的$\bar{n}$- $\frac{1}{\epsilon}$近临界点。事实上，如果存在这样的算法，那么令 $\epsilon = \frac{1}{12000n^{3.5}}$，所得算法的运行时间将关于 n 和实例的描述大小呈多项式级别，这与定理13相矛盾。注意，由于 ϵ 的大小为 $\log(1/\epsilon)$，上述论断实际上排除了运行时间关于 ϵ.的大小为指数级的算法。

3.2 计算下有界函数的 ϵ-近临界点的复杂度

与第 2.2 节类似，我们证明，计算函数 ϵ 的 f-近临界点仍然是困难的，即使额外假设 f 下有界。

定理15。如果存在一个算法 $\mathcal{A}$，以一个下有界函数 f 作为输入，其中该函数是一个 n 变量四次多项式与一个单变量指数函数的乘积，且算法运行时间为多项式时间，并且满足

(i) 当 f 没有临界点时，$\mathcal{A}$ 可以返回任意点，

(ii) 当 f 有临界点时，$\mathcal{A}$ 返回一个 $\frac{1}{1000 \cdot n^7}$-近临界点，

则 *P=NP*。

值得指出的是，尽管定理13与15 中的因子不同，围绕定理13 的讨论所得结论（见上文）同样适用于定理15，即有下界的情形。

定理*15*的证明。我们证明，这样的算法能够求解 CLIQUE，且可在多项式时间内完成。设 G 是一个具有 m 个顶点、邻接矩阵为 A 的图，并设 $k \in \{1, \ldots, m\}$. 令$n = m + 1$，$N = 160m^2\,(m^2 + 1)$，以及 $C = 1$。定义 n元函数

$$f(x, w) = e^w \cdot g(x)$$

其变量为（x，w）的 $\in \mathbb{R}^m \times \mathbb{R}$，其中 g（x）如（1）中所给。显然，f以零为下界，并且如命题10的证明中所述，从（G，k）到 f 的归约在长度上是多项式的。f 的梯度由下式给出

$$\nabla f(x, w) = \begin{pmatrix} e^w \nabla g(x) \\ e^w g(x) \end{pmatrix}.$$

设$\alpha = \frac{1}{1000n^{7\circ}}$与定理13的证明类似，我们证明以下结论：若（$x$，$w$）是 α的 f近临界点，则 sgn（x）是一个大小为 k的团的指示向量。定理的证明与定理13的证明中所述的方法完全相同。

现在证明该结论。设（x，w）是 f的 α-近临界点，即：

$$||(x, w) - (x^*, w^*)|| \leq \alpha,$$

其中（x^*，w^*）是 f.的临界点。注意，f的任意临界点（x^*，w^*）都必须满足 g（x^*）$= 0$。为证明该结论，我们首先证明 g（sgn（x））$< \frac{1}{(m+1)^4}$，然后证明$g(z) \geq \frac{1}{(m+1)^4}$，对任意一个不是 G中大小为 k 的团的指示向量的 $z \in \{-1, 1\}^m$ 均成立。

重复定理13证明中的相同论证，可得

$$|x_i^2 - 1| \leq \alpha^2 + 2\alpha, \forall i = 1, \ldots, m.$$

利用引理5和引理14，类似地可得

$$g(sgn(x)) \le g(x) + 48m^2(\alpha^2 + 2\alpha) \le L\alpha + 48m^2(\alpha^2 + 2\alpha),$$

其中，将 α、C 和 n 的值代入，可得

$$\begin{aligned} g(sgn(x)) &\le \frac{64m^2}{1000(m+1)^7} + \frac{24m \cdot 160^2 m^4 (m^2+1)^2}{1000^2(m+1)^{14}} + 48m^2\left(\frac{1}{1000^2(m+1)^{14}} + \frac{2}{1000(m+1)^7}\right) \\ &\le \frac{0.1}{(m+1)^5} + \frac{0.65}{(m+1)^5} + \frac{0.01}{(m+1)^{12}} + \frac{0.1}{(m+1)^5} < \frac{1}{(m+1)^4}, \end{aligned}$$

其中我们用到了以下事实：$m^2 + 1 \le (m+1)^2$.

现在令 $z \in \{-1, 1\}^m$，并假设 z 不是 G 中大小为 k 的团的指示向量。由引理 6 可得 $g(z) \ge \frac{1}{(m^2+1)^2} \ge \frac{1}{(m+1)^4}$。证明完毕。■

4 结论

我们已经证明，对于简单类别的非凸函数，即使是计算其临界点的低质量近似也难以处理；即便保证临界点存在且唯一，或函数有下界，结论依然成立。未来的研究或许可以进一步强化我们的困难性结果，例如利用 Clique(H˚astad, 1996) 的近似困难性结果。从积极的一面来看，我们认为这些结果推动了以下方向的研究：对于哪些参数函数类别及其结构假设，能够满足（i）在关于 $\log(1/\epsilon)$ 和参数规模的多项式时间内，计算该函数类别的 ϵ-近似或 ϵ-近临界点；以及（ii）在关于参数规模的多项式时间内检验函数的结构性质？例如，近期的一项结果（Slot 等人，2025年）表明，对于固定次数的凸多项式，可以在关于 $\log(1/\epsilon)$ 和多项式系数规模的多项式时间内，计算其 ϵ-近似且 ϵ-近临界点，这符合（i）。然而，对于次数为 4 或更高的多项式，凸性这一结构性质不满足（ii）（Ahmadi 等人，2013年），尽管存在满足该条件的凸性充分条件（例如，见（Ahmadi 和 Hall，2018年，定义 5））。还有哪些函数类别和结构性质能使临界点的近似计算具有可处理性，并符合我们上述给出的形式化定义？

参考文献

Amir Ali Ahmadi 和 Georgina Hall。利用代数技巧对非凸多项式进行差凸分解。数学规划，169(1):69–94，2018年。Amir Ali Ahmadi 和 Jeffrey Zhang。局部极小值及相关概念的复杂度方面。数学进展，397:108119，2022a。Amir Ali Ahmadi 和 Jeffrey Zhang。关于在多面体上寻找二次函数局部极小点的复杂度。数学规划，195(1):783–792，2022b。Amir Ali Ahmadi、Alex Olshevsky、Pablo A Parrilo 和 John N Tsitsiklis。判定四次多项式凸性及相关问题的 NP 困难性。数学规划，137(1):453–476，2013年。Amir Ali Ahmadi、Abraar Chaudhry 和 Jeffrey Zhang。每次迭代仅需多项式工作量的高阶牛顿法。数学进展，452:109808，2024年。

Dimitri P Bertsekas。《非线性规划》。Athena Scientific，贝尔蒙特，1995。

B´ela Bollob´as 和 Paul Erd¨os。随机图中的团。载于《剑桥哲学学会数学会刊》，第 80 卷，第 419–427 页。剑桥大学出版社，1976年。Stephen P Boyd 和 Lieven Vandenberghe。《凸优化》。剑桥大学出版社，2004年。Yair Carmon、John C Duchi、Oliver Hinder 和 Aaron Sidford。寻找驻点的下界 I。《数学规划》，184(1):71–120，2020年。Yair Carmon、John C Duchi、Oliver Hinder 和 Aaron Sidford。寻找驻点的下界 II：一阶方法。《数学规划》，185(1):315–355，2021年。Coralia Cartis、Nicholas IM Gould 和 Philippe L Toint。《非凸优化算法的评估复杂度：理论、计算与展望》。工业与应用数学学会，2022年。Xi Chen 和 Xiaotie Deng。确定双人纳什均衡的复杂度。载于《计算机科学基础》，第 6 卷，第 261–272 页，2006年。Zongchen Chen、Elchanan Mossel 和 Ilias Zadik。近线性规模的植入团难以被 Metropolis 过程发现。《随机结构与算法》，66(2):e21274，2025年。Sinho Chewi、S´ebastien Bubeck 和 Adil Salim。关于一维光滑函数驻点求解复杂度。载于《算法学习理论国际会议》，第 358–374 页。机器学习研究会论文集，2023年。Constantinos Daskalakis、Paul W Goldberg 和 Christos H Papadimitriou。计算纳什均衡的复杂度。《美国计算机协会通讯》，52(2):89–97，2009年。Ivar Ekeland。论变分原理。《数学分析与应用杂志》，47(2):324–353，1974年。John Fearnley、Paul Goldberg、Alexandros Hollender 和 Rahul Savani。梯度下降的复杂度：CLS=PPAD ∩ PLS。《美国计算机协会学报》，70(1):1–74，2022年。John Fearnley、Paul Goldberg、Alexandros Hollender 和 Rahul Savani。计算二次规划 KKT 解的复杂度。《美国计算机协会学报》，72(5):1–48，2025年。¨Matsusaburˆo Fujiwara。关于代数方程根的绝对值上界。《东北数学杂志，第一辑》，10:167–171，1916年。Rong Ge、Furong Huang、Chi Jin 和 Yang Yuan。逃离鞍点——用于张量分解的在线随机梯度方法。载于《学习理论会议》，第 797–842 页。机器学习研究会论文集，2015年。D Yu Grigor’ev 和 Nicolai N Vorobjov Jr。亚指数时间内求解多项式不等式系统。《符号计算杂志》，5(1-2):37–64，1988年。Johan H˚astad。团的近似求解是困难的，其近似比为 $n^{1-\varepsilon}$。载于《计算机科学基础》，第 627–636 页，1996年。Lane A Hemaspaandra 和 Ryan Williams。典型情形启发式算法的非典型综述。《ACM SIGACT通讯》，43(4):70–89，2012年。Alexandros Hollender 和 Emmanouil Zampetakis。非凸优化中寻找驻点的计算复杂度。载于《学习理论会议》，第 5571–5572 页。机器学习研究会论文集，2023年。

Mark Jerrum. 大团难以被 Metropolis 过程发现. 随机结构与算法, 3 (4):347–359, 1992. Vaithilingam Jeyakumar、Jean B Lasserre 和 Guoyin Li. 非紧半代数集上的多项式优化. 优化理论与应用杂志, 163: 707–718, 2014. Richard M Karp. 论组合问题的计算复杂性. *Networks*, 5(1): 45–68, 1975. Ludˇek Kuˇcera. 图划分问题的期望复杂度. 离散应用数学, 57(2-3):193–212, 1995. Katta G. Murty 和 Santosh N. Kabadi. 二次与非线性规划中的若干 NP 完全问题. 数学规划, 39(2):117–129, 1987. Yurii Nesterov. 无约束凸优化中可实现的张量方法. 数学规划, 186(1):157–183, 2021年. Jorge Nocedal 和 Stephen J Wright. 数值优化. Springer, 2006年. Olha Silina 和 Jeffrey Zhang. 一种无正则化三阶牛顿法. *arXiv* 预印本 *arXiv:2209.10051*, 2022年. Michael Sipser. 计算理论导论. *ACM SIGACT*通讯, 27(1):27–29, 1996年. Lucas Slot、David Steurer 和 Manuel Wiedmer. Hesse 的救赎：高效凸多项式规划. *arXiv* 预印本 *arXiv:2511.03440*, 2025年. Leslie G Valiant 和 Vijay V Vazirani. NP 与检测唯一解同样容易. 载于 *ACM* 计算理论研讨会, 第 458–463 页, 1985. Stephen A Vavasis. 局部最小化的黑箱复杂性. *SIAM* 优化杂志, 3(1):60–80, 1993.