arXiv:2601.03433v1 [math.CO] 6 Jan 2026

一些小图的半诱导性

J´ozsef Balogh* Bernard Lidick´y† Dhruv Mubayi‡ Florian Pfender§ Jan Volec¶

2026年1月8日

摘要

设 H 是一个固定的图，其边被染成红色和蓝色，并设 $\beta \in [0,1]$。令 $I(H,\ \beta)$ 表示在大型红蓝边着色的完全图 G 中，H 的副本数的（渐近归一化）最大值，其中 G 中红边的密度为 β。这细化了确定 H半诱导性的問題，而后者本身又是经典问题——确定 H的可诱导性——的推广。对于 $\beta \in [0,1]$，函数 $I(H,\ \beta)$ 在任何具有三个以上顶点的图 H 上都尚未知晓，除非 H 是单色团（Kruskal-Katona）或单色星（Reiher-Wagner）。我们对一些四顶点和五顶点图取得了精确结果，回答了不同作者近期提出的若干问题。我们还获得了关于树和星的一些一般性结果。仍有许多开放问题有待解决。

1 引言

在宿主图中计数子图是极值组合学中的一个基本问题。一个经典问题如下：给定一个固定图 F，在一个有 n 个顶点的宿主图中，诱导子图 F 的副本数量最大值是多少？这是 Pippenger 和 Golumbic [9] 提出的可诱导性问题。本文关注这一经典问题的红蓝（边二染色）推广。设 H 是一个边被染成红色或蓝色的图，G 是一个有 n 个顶点的红蓝染色完全图。

*伊利诺伊大学厄巴纳-香槟分校数学系，美国伊利诺伊州厄巴纳；以及基础科学研究院（IBS）极值组合学与概率研究组（ECOPRO），韩国大田。电子邮箱：jobal@illinois.edu。研究得到美国国家科学基金会资助，项目编号为 RTG DMS-1937241、FRG DMS-2152488；并获阿诺德·O·贝克曼研究奖（UIUC 校园研究委员会 RB 24012）、西蒙斯学者奖以及基础科学研究院（IBS-R029-C4）的支持。

†爱荷华州立大学数学系，美国爱荷华州埃姆斯。电子邮箱：lidicky@iastate.edu。该作者的研究部分由美国国家科学基金会 FRG DMS-2152490、 西蒙斯基金会 TSM-00013439 以及斯科特·汉纳教授席位资助。

‡伊利诺伊大学芝加哥分校数学、统计与计算机科学系，美国伊利诺伊州芝加哥，邮编 60607。电子邮箱：mubayi@uic.edu。研究部分得到美国国家科学基金会 DMS-2153576 项目资助。§科罗拉多大学丹佛分校数学与统计科学系，美国科罗拉多州丹佛。电子邮箱：Florian.Pfender@ucdenver.edu。该作者的研究部分由美国国家科学基金会 DMS-2152498 项目资助。

¶布拉格捷克理工大学信息技术学院理论计算机科学系，Th´akurova 9，布拉格，160 00，捷克共和国。电子邮件：jan@ucw.cz。研究部分得到了捷克共和国资助机构 23-06815M 号资助。

定义 1。 记 $(\#H, G)$ 为满足以下条件的单射 $f\colon\ V(H) \to V(G)$ 当且仅当 $f(u)f(v) \in E(G)$ 为红色时， $uv \in E(H)$为红色的数量。此外，若 H 有 h 个顶点，

$$\max(H,n) := \max_{G \text{ a red-blue } K_n} (\#H, G) \qquad and \qquad \max(H) := \lim_{n\to\infty} \frac{\max(H,n)}{n^h}.$$

为澄清定义 1 中的归一化因子，若 H 是一个三条边均为红色的 K_3，则 $(\#H, G)$ 等于 G 中红色 K_3 数量的六倍。对于 $\max(H)$，我们选择使用归一化因子 n^h，而非 $(n)_h$，以便与此前在 [1] 中定义的 $\max(H)$ 保持一致。由于我们的结果和问题都是渐近性质的，这不会造成影响。

确定 $\max(H, n)$ 和 $\max(H)$ 的问题称为半诱导性 问题。在经典诱导性设定中，要求在 G 的所选 $|V(H)|$ 个顶点上，指定的边必须存在，而所有其他边都不存在。相比之下，在半诱导性设定中， H 的某些边必须出现在副本中（例如红色边），而某些边必须不存在（蓝色边）；但所有其他顶点对都可以有边，也可以没有边，因此允许出现“半诱导”结构。

半诱导性问题由 Basit、Granet、Horsley、K¨undgen 和 Staden [1],提出；他们还确定了交替着色路径和游走的 $\max(H)$，以及若干个四个顶点的图 H 的 $\max(H)$。随后，Chen 和 Noel [5] 确定了交替着色六圈的 $\max(H)$，Bodn´ar 和 Pikhurko [2]则确定了另外若干个四顶点图的 $\max(H)$。两项研究 [5] 和 [2] 几乎完全采用了旗代数。Chen、Clemen 和 Noel [4] 也使用熵研究了交替路径这一具体情形。

我们的目标是着手系统研究这一问题的一个更精细版本：对于各种小型红蓝图 H，当底层图G 具有给定密度 β时，我们确定或界定 $\max(H, n)$（从而也确定或界定 $\max(H)$）。更确切地说，给定 $\beta \in [0, 1]$ 以及一个有 h 个顶点的红蓝图 H，我们考虑参数 $I(H, \beta)$，其形式定义如下。

定义 2. 红色/蓝色团序列 $(G_n) = (G_n)_{n=1}^{\infty}$ 的密度为 β，若 $|V(G_n)| = n$，且 G_n 中边的红边密度趋于 β。序列 (G_n) 是 H-优的，若 $\rho(H,(G_n)) :=\lim_{n\to\infty}(\#H, G_n)/n^h$ 存在。于是

$$I(H, \beta) = \sup \rho(H, (G_n)),$$

其中，上确界取遍所有密度为 β 的 H-优序列。

显然，$\max(H) = \sup_\beta I(H, \beta)$ 关于 $\beta \in [0, 1]$ 取得，但确定对所有 β 而言的函数 $I(H, \beta)$ 更具挑战性，并且能够更完整地展现大型图 G 中可能出现的 H 的副本数量。为统一我们的记号，如果一对顶点被染成红色，我们就称其为一条边；如果被染成蓝色，则称其为一条非边。这样，我们无需再使用红/蓝着色，而只需考虑指定了某些非边的图。在图示中，我们使用红色和点状蓝色[1] 来表示红边和蓝边是固定的，而其余非边的颜色

[1]我们也将蓝边称为非边。

尚未指定；或者使用黑色和非边来描述所有顶点对均已指定的诱导子图。

函数 $I(H,\ \beta)$ 对所有单色团和单色星都已知。前者由克鲁斯卡尔-卡托纳定理得出，后者是 Reiher 和 Wagner 的结果 [11]。此外，Liu、Mubayi 和 Reiher 最近的一项观察 [7]，确定了剩余的非单色三顶点图 H对应的 $I(H,\ \beta)$；当 H 是非单色三角形时，$I(H,\ \beta)$ 与在给定密度下使图中的三角形数量最少这一问题密切相关，而该问题已由 Razborov 的经典结果解决 [10]。

因此，使 $I(H,\ \beta)$ 尚未可知的最小图 H 是四个顶点上的图。第一个有趣的情形或许是由四个顶点组成的 3 边交替路径 AP_4，其顶点为 $u,\ v,\ w,\ x$，其中 uv和 wx 为红色，vw 为蓝色（见图 1）。更一般地，令AP_k 表示一个红蓝染色的、具有 $k+1$ 个顶点和 k 条边的路径，其中任意两条关联边颜色不同；若 k 为偶数，则两端的边颜色不同；若 k 为奇数，我们还规定第一条边和最后一条边均为红色。



图1：交替路径 AP_4 和交替 4-环 AC_4。

Basit、Granet、Horsley、K¨undgen 和 Staden [1, 推论 1.4] 确定了AP_k为偶数时的 max(AP_k)。Chen、Clemen 和 Noel [4, 定理 1.2] 确定了所有 $k \geq 1$ 下的 max(AP_k)，而 Bodn´ar 和 Pikhurko [3, 定理 6.2] 也确定了 max(AP_4)。尽管关于 max(AP_k) 已有这些结果，但对于任何 $k \geq 4$，函数 $I(AP_k,\ \beta)$ 都尚不为人所知。

我们的第一个结果完全确定了 $I(AP_4,\ \beta)$，回答了 Basit、Granet、Horsley、K¨undgen 和 Staden 提出的一个问题 [1, 问题 9.2]。我们在第2节给出的简短证明只使用了对图中与顶点度有关的各种量进行双重计数。

定理 1.1. 对于所有 $\beta \in [0,1]$，我们有 $I(AP_4,\ \beta) = \beta^2(1-\beta)$。只有（渐近）正则图才能达到等号。

我们将定理 1.1 的一部分推广到如下定义的双星族。对于一个整数s，我们将交替双星 $D(s)$ 定义为如下的带标号（且有序）图，它有 $2s+2$ 个顶点：$V(D(s)) = \{v_1, \ldots, v_s, v, u, u_1, \ldots, u_s\}$，以及 $E(D(s)) = \{v_1v, \ldots, v_sv, uu_1, \ldots, uu_s\}$ 和$uv \notin E(D(s))$（见图 2）。

显然 $D(1) = AP_4$。以下结果确定了一系列 β 的 $I(D(s)\ \beta)$，并且还确定了 max($D(s)$)。

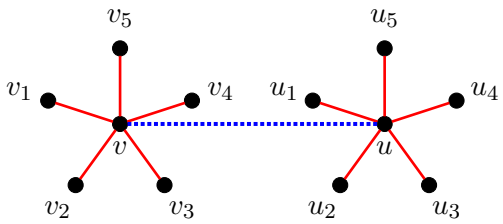


图 2：图 $D(5)$。

定理 1.2. 固定 $s \geq 1$ 和 $1 - 1/2s \leq \beta \leq 1$。则 $I(D(s)\ \beta) = \beta^{2s}(1-\beta)$。等号对（渐近）正则图成立。此外，$\max(D(s)) = (2s)^{2s} \cdot (2s+1)^{-2s-1}$。

我们的下一个结果涉及交替 4-环，它是从 AP_4 出发，在其两个端点之间添加一条非边得到的图。我们将这个有色图称为 AC_4，其顶点为 u、v、w、x，红边为 uv 和 wx，蓝边为 vw 和 xu（见图 1）。图 G中的 AC_4 数量与补图 G中的 AC_4 数量相同。因此，$I(AC_4、\ \beta) = I(AC_4、\ 1-\beta)$。

在 [1], 中，已确定了 max(AC_4)，并在 [1] 的问题 9.3 中再次提出了确定所有 β 的 $I(AC_4、\ \beta)$这一问题（更早之前，第三作者也曾提出过这个问题）。下面的结果给出了这一问题的答案，适用于每个正整数 k都满足 $\beta \in \{1/k, 1-1/k\}$ 的情形。

定理 1.3。 我们有 $I(AC_4,\ \beta) \leq \beta^2(1-\beta)$，当 $\beta \in \bigcup_{k\in\mathbb{N}}\{1/k, 1-1/k\}$时等号成立。

假设 $\beta \leq 1/2$。定理 1.3 的证明表明，只有当底层图几乎是正则的，且只有少数端点不相邻的三边路径时，等号才成立。实现这种构造的唯一方式是取大小为 $(\beta + o(1))n$ 的顶点不交团。这就要求 $1/\beta$ 是整数。我们猜想，当 $1/\beta$ 不是整数时，大致相同的构造是最优的。下面的问题将这一点精确化：它提出，当 β 不是 $1/k$ 或 $1-1/k$ 的形式时，$I(AC_4,\ \beta)$ 的行为与 Razborov [10] 所解决的三角形密度问题类似。

问题 1. 对于所有 $\beta < 1/2$ 以及密度为 β 的 AC_4-良序列(G_n)，$\rho(AC_4,(G_n))$ 的上确界是否在 G_n 由 $k = \lceil 1/\beta \rceil$ 个两两不交团构成时取得，其中 $k-1$ 个团的大小（渐近地）相同，而剩余团的大小至多为 n/k?

Liu、Mubayi 和 Reiher [7, 问题 7.1。] 询问是否存在一个红色/蓝色着色的完全图 H，使得 $I(H$ <style id="8">, $\beta)$ 有两个全局最大值（即在可诱导性设定中）。我们证明，如果不要求 H 为完全图（即在半诱导性设定中），他们问题的答案是肯定的。

用 P_{EENN} 表示一个由五个顶点组成、模式为边-边-非边-非边的路径（见图 3）。定义 $K(a,\ n)$ 为由一个含 a 个顶点的团和 $n-a$ 个孤立顶点组成的 n 阶图。给定定义在同一顶点集上的图 G 和 H，回顾一下，图 G 和 H 之间的编辑距离 是将 G 转变为 H 所需删除或添加的边数的最小值。

以下定理表明，当且仅当 $\beta \in \{9/16, 7/16\}$ 时， $I(P_{EENN}, \beta)$ 取最大值 $27/252$；其渐近极大化图为 $K(a, n)$ 或其补图，其中 $a = 3/4$。



图 3：路径 P_{EENN}。

定理 1.4. 对于所有 $\beta \in [0,1]$, $I(P_{EENN}, \beta) = \max\{\beta^{3/2} - \beta^2, (1-\beta)^{3/2} - (1-\beta)^2\}$。此外，对于固定的 β 和 $\varepsilon > 0$，存在 $\delta > 0$ 和 n_0，使得当 $n > n_0$时，每个 n顶点的红蓝团G，若其红边密度在 $[\beta - \delta, \beta + \delta]$ 中使 $(\#P_{EENN}, G)$ 最大，则其与$K(a,n)$ 或其补图的编辑距离至多为 εn^2，其中某个 a。

将 $P_{E^\ell N^k}$ 记作一条长度为 $\ell + k$ 的路径，其前 ℓ 条边为红色，剩余 k 条边为蓝色。注意， P_{EENN} 是 $P_{E^2N^2}$。旗代数实验表明， $K(a,n)$及其补图也是 $P_{E^3N^3}$的极值构造。我们预计更长的对称路径也会如此。

猜想 1. 对于 $\ell \geq 2$, $I(P_{E^\ell N^\ell}, \beta)$，当 $K(a,n)$ 或其补图时渐近达到该值。

接下来我们考虑单色树的情形。这里无需提及颜色，我们的问题只是：在密度为 β 的图中，如何使给定树 T 的副本数最大。即使在 T 是星这一简单情形下，确定 $I(T,\beta)$ 也相当困难。事实上，视 β 的取值而定，极值图为 $K(a,n)$，或为 $K(a,n)$ 的补图，其中 a 取适当值，使所得图的密度为 β。这是 Reiher 和 Wagner [11] 证明的。只要 β 接近 1，我们就将他们结果的一部分推广到任意树。

定理 1.5. 对于每棵树 T，都存在 $\beta_T \in (0,1)$，使得对于 $\beta > \beta_T$，图序列(G_n)（其中 $G_n = K((1+o(1))\sqrt{\beta}n, n)$）达到 $I(T, \overline{\beta})$。

请注意，在定理 1.5 中，对于某些 β_T，我们需要 $\beta > \beta_T$；因为即使 T 由两条边和 $\beta < 1/2$组成，$K((1+o(1))\sqrt{1-\beta}n, n)$ 的补图也具有 β 的边密度，且其包含的 T副本比 $K((1+o(1))\sqrt{\beta}n, n)$ 更多。

如果考虑双色边树 T，问题似乎要复杂得多，甚至可能一般而言毫无希望。我们针对某些 β，提出关于 $I(T,\beta)$ 的以下问题。

问题 2. 设 T 是一棵红蓝着色树。是否存在满足 $0 \leq \beta_T < \beta'_T \leq 1$ 的 β_T、β'_T，使得下述结论成立？设 $\beta \in [\beta_T, \beta'_T]$。那么， $I(T,\beta)$ 可由以下图序列渐近达到 (G_n)。 G_n 的红色子图 R 具有如下形式：存在一个顶点划分 $A \cup B = V(G_N)$， A 中的所有顶点的度渐近相同， A 中的顶点与 B 中的顶点之间的所有边均存在，并且 B 是一个独立集。

对于给定密度为 β 的图， T 副本的最小数量也有可能渐近地由问题 2 中的图达到。文献 [6] 证明，当 T 为单色时，正则图能够达到这一最小值。

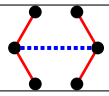
图 H	$I(H,\ \beta)$	构造	参考文献
AP_4 	如果 $\beta \in [0,1]$，则为 $\beta^2(1-\beta)$	正则	定理 1.1
$D(s)$ 	$\beta^{2s}(1-\beta)$ 若 $\beta \in [1-1/2s, 1]$	正则	定理 1.2
AC_4 	$\beta^2(1-\beta)$ 若 $\beta \in \{1/k, 1-1/k\}$	正则	定理 1.3
P_{EENN} 	$(1-\beta)^{3/2} - (1-\beta)^2$ if $\beta \in [0,1/2]$ $\beta^{3/2} - \beta^2$ if $\beta \in [1/2,1]$	团 + 孤立点 团 ∪ 孤立点	定理 1.4
树 T	$\beta \in [\beta_T, 1]$	团 ∪ 孤立点	定理 1.5
$S_{2,1}$ 	$\beta/4$ if $\beta \in [1/4,1/2]$ $\beta^2(1-\beta)$ if $\beta \in [1/2,1]$	正则 ∪ 孤立点 正则	第7节

Table 1：本文结果概述。所示的 $D(s)$实例是 $D(2$).

我们以陈述一个与问题 2 相关的猜想作为引言的结尾，该猜想针对的是看似简单的星这一情形。令 $\mathcal{S}$ 为这样一类图：从一个正则图出发，先添加全邻顶点，再添加孤立顶点；或者先添加孤立顶点，再添加全邻顶点。注意，初始正则图可以是一个完全子图，也可以是一组孤立顶点。令 $S_{a,\ b}$ 为一颗含有 a 条边（红色）和 b 条非边（蓝色）的星。

问题3。 对于每个 a、$b \geq 1$ 和 $\beta \in [0,1]$，是否都存在一个$S_{a,b}$-良序列(G_n)，使得 $G_n \in \mathcal{S}$

$$I(S_{a,b}, \beta) = \rho(S_{a,b}, (G_n))?$$

我们在第7节中对 $a = 2, b = 1$ 和 $\beta \geq 1/4$给出了问题3的肯定回答。更确切地说，我们证明了

$$I(S_{2,1}, \beta) = \begin{cases} \beta/4 & \text{if} \quad 1/4 \leq \beta \leq 1/2, \\ \beta^2(1-\beta) & \text{if} \quad 1/2 \leq \beta \leq \ 1. \end{cases} \tag{1}$$

正如我们已经提到的，确定 $I(H,\beta)$包括确定 $\max(H)$。虽然确定 $I(H,\beta)$仍是一个悬而未决的问题，但对于所有四顶点图 [2] ，$\max(H)$都已确定，唯一显著的例外是具有边-边-非边模式的路径。这似乎是一个颇具挑战性的开放问题；参见 [1]中的问题9.2和 [2]中的猜想1.1。

我们的结果总结于表1。第2节中，我们证明定理 1.1，描述 $I(AP_4,\ \beta)$。第3节包含定理 1.2 的证明。第4节讨论定理 1.3 和问题 1。第5节讨论定理 1.4，第6节包含定理 1.5。最后，第7节讨论星的可能极值构造 $\mathcal{S}$ 。

2 交替三边路径

我们先介绍 Nikiforov 的一项结果，它是计数图中三边路径的一个非常有用的工具。为完整起见，我们给出一个证明。

引理 2　.1 (Nikiforov [8])**.** 设 $G=(V,E)$ 是一个有 n 个顶点和 m 条边的图。则

$$\sum_{uv\in E} d_u d_v \;\geq\; \frac{4m^3}{n^2},$$

其中 d_w 表示顶点 w的度。当且仅当 G是正则图时，等号成立。

证明。 我们可以假设 G 没有孤立顶点。设

$$S := \sum_{uv\in E} d_u d_v, \qquad U := \sum_{uv\in E} \frac{1}{\sqrt{d_u d_v}}.$$

首先注意，对每条边 uv,

$$\frac{1}{\sqrt{d_u d_v}} \leq \frac{1}{2}\left(\frac{1}{d_u}+\frac{1}{d_v}\right).$$

对所有边求和，得到

$$U \leq \frac{1}{2}\sum_{uv\in E}\left(\frac{1}{d_u}+\frac{1}{d_v}\right) = \frac{1}{2}\sum_{u} d_u\cdot\frac{1}{d_u} = \frac{n}{2}.$$

对于每条 m 边 uv，令 $x_{uv}=(d_u d_v)^{1/4}$ 和 $y_{uv}=(d_u d_v)^{-1/4}$。令 $x=(x_{uv})$和 $y=(y_{uv})$为相应的向量。将柯西-施瓦茨不等式 $|x\cdot y|\leq ||x||\cdot||y||$ 应用于向量 x 和 y，得到

$$\begin{aligned} m^2 \;&=\; \Big(\sum_{uv\in E} x_{uv}\cdot y_{uv}\Big)^2 \;=\; |x\cdot y|^2 \;\leq\; ||x||^2\cdot||y||^2 = \Big(\sum_{uv\in E} x_{uv}^2\Big)\cdot\Big(\sum_{uv\in E} y_{uv}^2\Big) \\ &=\; \Big(\sum_{uv\in E}\sqrt{d_u d_v}\Big)\cdot\Big(\sum_{uv\in E}\frac{1}{\sqrt{d_u d_v}}\Big). \end{aligned}$$

利用 $U\leq n/2$，得到

$$\sum_{uv\in E}\sqrt{d_u d_v} \geq \frac{m^2}{U} \geq \frac{2m^2}{n}.$$

将此与　　应用平方平均数与算术平均数之间的不等式相结合，得到　　s

$$S=\sum_{uv\in E} d_u d_v \geq \frac{1}{m}\Big(\sum_{uv\in E}\sqrt{d_u d_v}\Big)^2 \geq \frac{1}{m}\left(\frac{2m^2}{n}\right)^2 = \frac{4m^3}{n^2},$$

这就完成了证明。由于我们的应用，只有当所有 d_u 都相同时等号才成立

算术平均数/二次平均数不等式。 □

回顾一下， AP_4 是长度为三的交替路径。在 G 中， G 的一个 AP_4 副本是与 AP_4同构的（带颜色）子图，因此，(#AP_4， G)等于 G中 AP_4 副本数量的两倍。利用引理 2.1，我们得到具有给定边密度的图中 AP_4 副本数量的一个上界；对于正则图，该上界是紧的。这完全确定了图中AP_4 数量的分布。因此，定理 1.1（即 $I(AP_4, \beta) = \beta^2(1-\beta)$）是下面定理 2.2 的直接推论。

定理 2.2。 设 $G = (V, E)$ 是一个有 m 条边的图。那么

$$\sum_{uv \notin E} d_u d_v \leq 2m^2 - \frac{2m^2}{n} - \frac{4m^3}{n^2}.$$

等号成立 当且仅当 G 是正则的。此外， G 中 AP_4 的副本数至多为 st

$$2m^2 - \frac{2m^2}{n} - \frac{4m^3}{n^2} - t,$$

其中 t 是 G 中恰好包含两条边的 3 元顶点集的数量。

证明。

$$2m^2 - \frac{1}{2}\sum_v d_v^2 = \frac{1}{2}\left(\sum_v d_v\right)^2 - \frac{1}{2}\sum_v d_v^2 = \sum_{uv \in \binom{V}{2}} d_u d_v = \sum_{uv \notin E} d_u d_v + \sum_{uv \in E} d_u d_v.$$

重新整理可得

$$\sum_{uv \notin E} d_u d_v = 2m^2 - \left(\frac{1}{2}\sum_v d_v^2 + \sum_{uv \in E} d_u d_v\right).$$

现在，若固定 n 和 m，则当 $(\sum_v d_v^2)/2 + \sum_{uv \in E} d_u d_v$ 最小时， $\sum_{uv \notin E} d_u d_v$ 取最大值。根据凸性，第一个量仅在正则图的情况下取最小值；根据引理 2.1，第二个量也仅在正则图的情况下取最小值。因此， $\sum_{uv \notin E} d_u d_v$ 仅在正则图的情况下取最大值。关于 AP_4 的结论可通过观察得出: $\sum_{uv \notin E} d_u d_v$ 等于 AP_4 的副本数，加上恰有两条边的三元素集合数（若 u 和 v 的邻域相交）。

□

2.1 利用旗代数求长度为3 的交替路径的密度

Chen 和 Noel [5]以及 Bodn´ar 和 Pikhurko [2]的研究表明，旗代数可用于解决半诱导性问题。在找到上文给出的定理 1.1 的证明之前，我们已利用至多含 4个顶点的旗确定了 $I(H, \beta)$（ $1/2 \leq \beta \leq 1$ ）。针对若干固定的 $\beta < 1/2$ 值所做的数值实验表明，旗代数或许也可用于 $\beta < 1/2$。不过，这样的证明需要用到至少含 6个顶点的旗，而非 4个顶点。我们预计，这种潜在的旗代数

证明若以 β 为参数并使用含 6个顶点的旗，将会复杂得多。在发现了无需使用旗代数即可覆盖整个范围的证明后，我们便没有继续尝试寻找这样的旗代数证明。我们仍在此给出关于 $\beta \geq 1/2$ 的简明旗代数证明，供熟悉该方法的读者参考；但为节省篇幅，省略了标准记号的介绍。

定理 2.3. $I(AP_4, \beta) = \beta^2(1-\beta)$ 当 $\beta \in [1/2, 1]$。

证明。回顾一下，在 $I(AP_4, \beta)$ 的定义中，子图计数按 n^4 缩放。然而，4顶点子图的旗代数密度按 $\binom{n}{4}$ 缩放。因此，我们的目标是证明上界 $24\beta^2(1-\beta)$。在旗代数中， AP_4 的数量可以写成

$$O := 4 + 4 + 8 + 6 + 8 .$$

我们的目标是证明 $O \leq 24\beta^2(1-\beta)$。设 $\alpha = 1-\beta$ 为非边的密度。我们将把 O 与以下四个平方项的线性组合以及一个用于固定非边密度的表达式结合起来。

$$C_1 := \left[\!\!\left[\left(\; - \; \right)^2 \right]\!\!\right]$$

$$C_2 := \left[\!\!\left[\left(\; - \; \right)^2 \right]\!\!\right]$$

$$C_3 := 12 \left[\!\!\left[\left(- \alpha \; + (1-2\alpha) \; + (1-\alpha) \; \right)^2 \right]\!\!\right]$$

$$C_4 := 12 \left[\!\!\left[\left(- \alpha \; + (1-2\alpha) \; + (1-\alpha) \; \right)^2 \right]\!\!\right]$$

$$E := 6 \left(\; - \alpha \right)$$

注意，所有五项都是非负的。它们可以展开为 4顶点图的线性组合。这些线性组合在 4顶点旗上展开后的系数见表 2。

以下线性组合表明，每个含 4个顶点 的图的系数都等于$24\alpha(1-\alpha)^2$。

$$O + (48\alpha^3 - 96\alpha^2 + 48\alpha)C_1 + (48\alpha^3 - 72\alpha^2 + 24\alpha + 12)C_2 + (-4\alpha + 4)C_3 \\ +(-4\alpha + 2)C_4 + (-12\alpha^2 + 16\alpha - 4)E = 24\alpha(1-\alpha)^2.$$

请注意，证明仅对 $0 \leq \alpha \leq 1/2$ 成立，因为线性组合中 C_1， C_2， C_3 和 C_4 对应的系数必须非负，而 C_4 对应的系数在 $\alpha > 1/2$ 中为负。由于 $\alpha = 1-\beta$，结论得证。

□

	O	C_1	C_2	C_3	C_4	E
	0	0	0	$12\alpha^2$	0	$0-6\alpha$
	4	0	0	$10\alpha^2-4\alpha$	$2\alpha^2$	$1-6\alpha$
	4	1/6	0	$10\alpha^2-8\alpha+1$	$4\alpha^2-2\alpha$	$2-6\alpha$
	0	0	1/2	$6\alpha^2-6\alpha$	$12\alpha^2-12\alpha+3$	$3-6\alpha$
	0	1/2	0	$12\alpha^2-12\alpha+3$	$6\alpha^2-6\alpha$	$3-6\alpha$
	8	−2/3	0	0	$4\alpha^2$	$2-6\alpha$
	6	−1/6	−1/6	$4\alpha^2-6\alpha+2$	$4\alpha^2-2\alpha$	$3-6\alpha$
	0	0	1/6	$4\alpha^2-6\alpha+2$	$10\alpha^2-12\alpha+3$	$4-6\alpha$
	8	0	−2/3	$4\alpha^2-8\alpha+4$	0	$4-6\alpha$
	0	0	0	$2\alpha^2-4\alpha+2$	$10\alpha^2-16\alpha+6$	$5-6\alpha$
	0	0	0	0	$12\alpha^2-24\alpha+12$	$6-6\alpha$

表 2：展开为含4个顶点的无标号图时的系数。

3 交替双星的密度

本节证明定理 1.2。该结论可由以下定理推出。回顾一下，对于整数 s，交替双星 $D(s)$ 是如下的一个带标号（且有序）的图，它有 $2s+2$ 个顶点： $V(D(s)) = \{v_1, \ldots, v_s, v, u, u_1, \ldots, u_s\}$，以及 $E(D(s)) = \{v_1 v, \ldots, v_s v, u u_1, \ldots, u u_s\}$ 和$uv \notin E(D(s))$（见图 2）。

定理3.1。固定 $s \geq 1$。则

$$\max(D(s)) \leq \frac{(2s)^{2s}}{(2s+1)^{2s+1}}. \tag{2}$$

若 G 是一个具有 n个顶点的 图，其平均度为 $d \geq (1-1/(2s))n$，则 $(\#D(s)\ G) \leq (n-d)\cdot n \cdot d^{2s}$.

证明。 设 G的平均度为 d。我们有

$$(\#D(s), G) \leq 2 \sum_{uv \notin E} d_u^s d_v^s \leq \sum_{uv \notin E} \left(d_u^{2s} + d_v^{2s}\right) = \sum_u d_u^{2s} \cdot (n-1-d_u) < \sum_u d_u^{2s} \cdot (n-d_u). \tag{3}$$

容易给出一个上界：

$$\sum_u d_u^{2s} \cdot (n - d_u) \le \frac{1}{2s} \sum_u d_u^{2s} \cdot (2sn - 2sd_u) \le \frac{n}{2s} \Big(\frac{2sn}{2s+1} \Big)^{2s+1} = \frac{(2s)^{2s}}{(2s+1)^{2s+1}} n^{2s+2}, \qquad (4)$$

从而得到 $\max(D(s)) = (2s)^{2s} \cdot (2s+1)^{-2s-1}$。

为了将式 (3) 中的和表示为图的密度的函数并给出其上界，我们使用一个广为人知的一般优化技巧（并给出其证明）。

引理 3.2. 对于一个 $\alpha > 0$，定义 X_α 为向量 $(x_1, \dots, x_n)$ 的集合，其中对所有 i，除至多一个坐标外，均有 $x_i \in \{0, \alpha\}$。假设 $0 < \gamma < 1$ 且 $f\colon [0,1] \to \mathbb{R}$ 在 $[0, \gamma]$ 上是凸的，在 $[\gamma, 1]$ 上是凹的。那么，在限制条件 $x_i \in [0,1]$ 且对某个 $0 \le D \le n$ 有 $\sum x_i = D$ 的情况下，$\sum_{i=1}^n f(x_i)$ 在 X_α 的某个元素处取得最大值，其中 $\alpha \ge \gamma$ 取某个值。

证明。 假设 $x = (x_1, \dots, x_n)$ 在满足 $\sum x_i = D$ 的条件下使该和最大。若存在 $0 < x_i \le x_j < \gamma$，那么我们可以将 x_i 减小，并将 x_j 增加相同的量，从而增大该和 $\sum_{i=1}^n f(x_i)$（由凸性可知），因此不存在两个这样的 i, j。类似地，若对某个 i, j 有 $\gamma \le x_i < x_j$，那么我们可以增加 x_i 并减小 x_j，从而增大该和（由凹性可知），因此不存在两个这样的 i, j；也就是说，所有 $x_i \ge \gamma$ 彼此相等（我们将 $\alpha \ge \gamma$ 定义为它们的值；若不存在 $x_i \ge \gamma$，则令 $\alpha = \gamma$），而其余变量中至多有一个不为 0。由此得到 $x \in X_\alpha$，如所述。

□

我们对 $f(x) = (1-x)x^{2s} = x^{2s} - x^{2s+1}$ 应用引理 3.2。利用

$$f'(x) = 2s \cdot x^{2s-1} - (2s+1) \cdot x^{2s} \qquad \text{and} \qquad f''(x) = 2s(2s-1) \cdot x^{2s-2} - 2s(2s+1) \cdot x^{2s-1},$$

我们有 $f''(x) > 0$ 当且仅当 $\frac{2s-1}{2s+1} > x$，也就是说，我们可以选择 $\gamma = \frac{2s-1}{2s+1}$ 和 $D = d$。注意，我们定义 $x_u = d_u/n$。利用引理 3.2，存在一个 $\alpha > \gamma$，使得 f 的最大值在 m 个变量为 α 时取得（对于某个 m），其中至多一个变量介于 0 和 γ 之间，其余变量均为 0。注意，$\alpha \ge d/n$ 也成立。忽略这个离群变量（当 n 较大时，它的作用可以忽略不计），我们有 $\alpha m = d = D$，并由引理 3.2 得到以下上界：

$$\sum_{i=1}^n f(x_i) \le m \cdot f(\alpha) = \frac{d}{\alpha} (\alpha^{2s} - \alpha^{2s+1}).$$

这意味着 (3) 有以下上界，而这实质上证明了 (2)：

$$n^{2s+1} \frac{d}{\alpha} \left(\alpha^{2s} - \alpha^{2s+1} \right) = n^{2s+1} \cdot d \cdot \alpha^{2s-1} \cdot (1 - \alpha).$$

现在假设 $d \ge (1 - 1/(2s))n$。那么 $\alpha \ge d/n \ge (2s-1)/(2s)$，并利用 $\alpha^{2s-1} \cdot (1-\alpha)$ 在 $\alpha > (2s-1)/2s$ 时递减这一点，我们得到

$$n^{2s+1} \cdot d \cdot \alpha^{2s-1} \cdot (1 - \alpha) \le n \cdot (n-d) \cdot d^{2s} \le n^{2s+2} \cdot \frac{(2s)^{2s}}{(2s+1)^{2s+1}}. \qquad (5)$$

□

当 $s = 1$ 时，定理 1.1 表明，对于每个 d，更强的界 $d^2(n-d)n$ 都成立；但对于 $s \geq 2$，类似的结论并不成立，因为与正则图相比，顶点数较少且含有额外孤立顶点的正则图包含更多个 $D(s)$。

稍后我们将用到以下命题，它是式 (3) 的 $s = 1$ 情形以及式 (5) 中的第一个不等式。

命题 3.3。设 G 是一个图，其度序列为 $d_1, \ldots, d_n$，平均度为 $d \geq n/2$。那么

$$\sum_i d_i^2(n - d_i) \leq nd^2(n - d).$$

4 交替 C_4

本节将给出定理 1.3 的一个相对简短的证明，然后讨论密度 $\beta \neq 1/k$ 的情形。

定理 1.3 的证明。回顾一下，我们有一个 n 阶图 G，其密度为 $\beta + o(1)$，并计划证明 $(\#AC_4, G) \leq (1 + o(1))\beta^2(1 - \beta)n^4$。我们的主要（简单）不等式是

$$(\#AC_4, G) \leq 2 \sum_{uv \notin E} d_u d_v.$$

通过选取一条非边 uv、u 的邻点 u' 以及 v 的邻点 v'，可以得到一个 AC_4。若 $u' \neq v'$ 且 $u'v' \notin E$，则 $uu'v'v$ 给出一个 AC_4 的副本。由于每个 AC_4 都有两条非边，因此每个副本都被计数两次。此外，$(\#AC_4, G)$ 是 AC_4 的副本数的四倍，从而得到该不等式。更准确地说，若记 t 为恰有两条边的三元顶点集数量，并记 s 为端点构成一条非边的三边路径数量，则

$$\sum_{uv \notin E} d_u d_v = t + s + \frac{1}{2} \cdot (\#AC_4, G).$$

设 $m = |E(G)| = (1 + o(1))\beta\binom{n}{2}$。应用定理 2.2

$$(\#AC_4, G) \leq 2 \sum_{uv \notin E} d_u d_v \leq 4m^2 - \frac{4m^2}{n} - \frac{8m^3}{n^2} \leq 4m^2 - \frac{8m^3}{n^2} = (1 + o(1))\,\beta^2(1 - \beta)n^4.$$

当 $s = o(n^4)$ 时，上述第一个不等式是渐近紧的；当 G 接近正则时，第二个不等式是渐近紧的。因此，若 G 是由大小为 $(\beta + o(1))n$ 的团构成的不交并，那么当 $1 - \beta$ 接近正则且 $s = 0$ 时，$(\#AC_4, G) = (1 + o(1))\beta^2 n^4$。这样的 G 仅当 $k = 1/\beta$ 是整数时才存在，从而完成了定理的证明。□

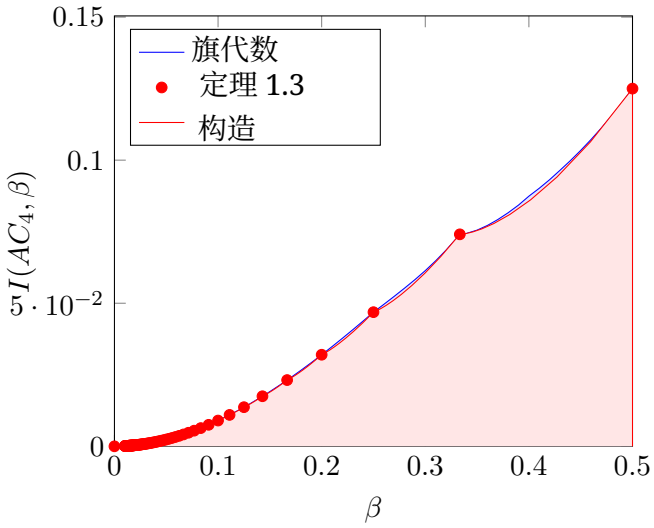


图 4：上界是通过对 7 个顶点进行旗代数计算得到的，下界则由一个构造得到。

如果 $1/\beta$ 不是整数，那么就不可能构造出具有定理 1.3 的证明中所讨论性质的正则图。这里我们提出，极值构造或许可以通过再次取团的不交并来得到，其中除一个团外，其余团的大小都相同，而剩下的那个团更小，见问题 1。

我们尝试针对许多不同的 β 值，将 7 个顶点上的旗代数计算结果与这一构造进行比较，发现数值很接近，但并不相等（见图 4）。例如，即使在 8 个顶点上进行计算，我们也得到了数值上界 $I(AC_4, 2\,5) \leq 0.0859499752$（优于通过 7 个顶点的计算得到的上界），而该构造给出的数值为 $I(AC_4, 2\,5) \geq 0.08566600788$。

此处的构造（三个不交团）足够简单，因此人们可能会认为，如果它确实是唯一的极值构造，那么对适量顶点进行旗代数计算就能“发现”它。也许还存在我们尚未意识到的其他构造，其结果与我们的构造相当或优于它。另一方面，这类构造肯定不可能是正则的，因为在我们的旗代数计算中加入正则性约束后，得到的数值上界为 0.08409772。

5 路径 P_{EENN}。

定理 1.4 是以下两个关于 $\beta \in [1/2, 1]$ 的定理的推论。结果的界限示意见图 5。情形 $\beta \in [0, 1/2]$ 可由对称性得出。回想一下，$K(a,\ n)$ 是一个有 a 个顶点的图，其中包含一个团和 $n - a$ 个度为 0 的顶点。

定理 5.1。 $I(P_{EENN},\ \beta) = \beta^{3/2} - \beta^2$ 对于 $\beta \in [1/2, 1]$。

证明。 为简化证明中的表达式，我们令 $a = \sqrt{\beta}$。我们的目标是证明 $I(P_{EENN},\ a^2) =$

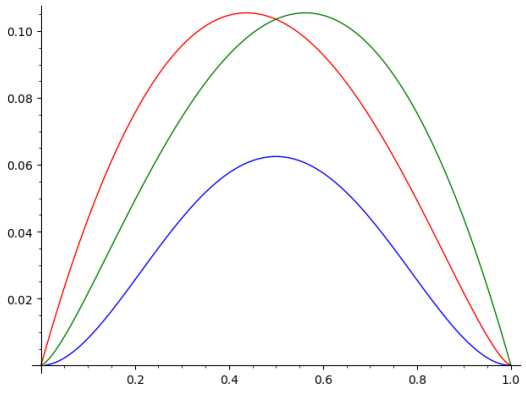


图 5：红色和绿色分别表示由团、孤立顶点及补图给出的界。它们的最大值为 $I(P_{EENN}, \beta)$。蓝色表示一个正则图，它或许是最小值。

$a^3 - a^4$ 对于 $a \in [1/\sqrt{2}, 1]$。下界由 $K(a,\ n)$ 得出。对于一个收敛序列 (G_n) 及其对应的同态 ϕ，旗代数的缩放为

$$\rho(P_{EENN}, G_n) = \frac{1}{120}\phi(P_{EENN}).$$

在旗代数计算中，我们省略了 ϕ，因为它们对所有 ϕ 都成立。我们可以将P_{EENN} 改写为五顶点子图密度的线性组合，如下所示。

$P_{EENN} =$  $= 4 \cdot$  $+ 12 \cdot$  $+ 24 \cdot$  $+ 6 \cdot$  $+ 8 \cdot$  $+ 16 \cdot$  $+ 20 \cdot$

 $+ 12 \cdot$  $+ 12 \cdot$  $+ 20 \cdot$  $+ 16 \cdot$  $+ 2 \cdot$  $+ 4 \cdot$  $+ 8 \cdot$  $+ 8 \cdot$

 $+ 6 \cdot$  $+ 8 \cdot$  $+ 4 \cdot$  $+ 24 \cdot$  $+ 12 \cdot$  $+ 2 \cdot$  $+ 12 \cdot$  $+ 4 \cdot$ 

用旗代数语言来说，我们的主要目标是证明

$$\frac{1}{120}\ \text{  }\ \leq a^3 - a^4. \tag{6}$$

由于 $\beta = a^2$ 是边密度，我们可以使用 $-a^2 = 0$。不等式 (6) 可由下式推出

$$0 \geq a^2(1-a^2)\cdot\left(\quad - 120(a^3-a^4)\right) + \left(120a^2(a^3-a^4)\cdot\quad - 120(1-a^2)(a^3-a^4)\cdot\quad\right.$$
$$\left. + (-12a^6+120a^5+80a^4-80a^3+20a^2-20a)\cdot\quad + 15\cdot B\cdot\quad\right)\times\left(\quad - a^2\right) \tag{7}$$
$$+ 60(a-a^2)\cdot\left[\!\left[\left(a\quad - (1-a)\quad\right)^2\right]\!\right] + 30\cdot C\cdot\left[\!\left[\left(a\quad - (1-a)\quad\right)^2\right]\!\right],$$

其中 $B = C = \sqrt{2}-1$ 若 $\beta \in [1/2, 0.8^2]$ 且 $B = 0.361$，$C = 0$ 若 $\beta \in (0.8^2, 1])$。我们注意到，当 $\beta > 1/2$ 时，B 和 C 的选择具有一定自由度，但对于 $\beta = 1/2$，它们是固定的。不等式 (7) 有效性的一个关键点在于，$(a-a^2)$ 和 C（即平方和的乘法因子）均为非负数。

对 (7) 的求值类似于定理 2.3 的证明。不过，该过程是在 5 顶点图上进行的。令 $\mathcal{F}_5$ 表示所有不计同构的 5 顶点双色边图的集合。繁琐但直接的枚举表明，$|\mathcal{F}_5| = 34$。此外，相应的系数更加复杂。另一方面，它们只需在极值构造的子图上取紧。更确切地说，(7) 的右侧可以表示为 $\mathcal{F}_5$ 中各元素的线性组合

$$\sum_{F\in\mathcal{F}_5} c_{a,F}\cdot F \leq \max_{F\in\mathcal{F}_5} c_{a,F} = 0. \tag{8}$$

感兴趣的读者可以在本文 arXiv 预印本的附录中找到所有系数 $c_{a,\,F}$。

□

定理 5.1 的证明还蕴含了极值构造的（渐近）稳定性。

定理 5.2。 固定 $\beta \in [1/2, 1)$。如果 G 是一个充分大的图，其边密度为 β，并且使 P_{EENN} 的副本数最大，那么 G 与某个图 $K(\sqrt{\beta}n,\ n)$ 的编辑距离至多为 εn^2；当 $\beta = 1/2$ 时，也可能与该图的补图满足此条件。

证明。如果 (G_n) 是一个边密度为 β 且达到 $I(P_{EENN},\ \beta)$ 的收敛图序列，那么所有具有非零密度的 5 个顶点子图都必须在 (8) 中具有 $c_{a,\,F} = 0$。检查 (8) 中的 $c_{a,\,F}$ 可知，对于所有 $a \in [\sqrt{\beta}, 1)$，仅当 F 时，$c_{a,\,F} = 0$。

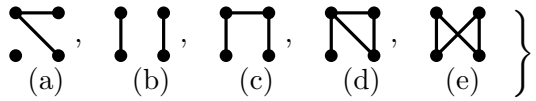
$$\mathcal{A} := \left\{\quad,\quad,\quad,\quad,\quad\right\}.$$

其他图 F 至少有一个 $a \in [\sqrt{\beta}, 1)$，此类图 $c_{a,\,F} = 0$ 仅出现在 $a = 1/\sqrt{2}$，并且它们

如下四个

$\mathcal{A}_{1/2} :=$  .

然而，$c_{1/\sqrt{2},\ C_5} = 0$ 是证明中的一个产物。针对$a = 1/\sqrt{2}$的单独旗代数计算得出了$c_{a,\ C_5} < 0$。由此可知，¯若 (G_n) 是一个收敛序列，且对于 $\beta \in [1/2, 1)$，它达到 $I(P_{EENN},\ \beta)$，那么下列图的密度

$\mathcal{F} =$ 

t趋于零，因为它们不是 $\mathcal{A} \cup \mathcal{A}_{1/2} \setminus \{C_5\}$的诱导子图 .

令 G 为一个足够大的极值例子。对 G应用诱导删除引理后，我们得到一个图G'，其与 G的编辑距离为 εn^2，且不含 $\mathcal{F}$ 中任何图作为诱导子图。

令 K 为 G'中一个最大团的顶点集，令 I 为 G'中的孤立顶点集，并令 $R = V(G') - K - I$ 为其余顶点。

命题 5.3. *K 中至多有一个顶点 k，使得 R 中的所有顶点 r 都与 $K \setminus \{k\}$ 中的所有顶点相邻，且与 k不相邻。*

证明。根据 K 的极大性，R 中的每个顶点都必须在 K 中有一个非邻点。设 R 中有一个顶点 r。反设 r 与两个顶点 $u,\ v \in K$ 均不相邻。由于 r 不是孤立点，它有一个邻点 w。这四个顶点 $u,\ v,\ w,\ r$ 诱导出 (b)、(c) 或 (d) 中的一种情形，矛盾。因此，对于 r 中的每个顶点，K 中至多有一个顶点 u 使得 ur 不是一条边。若 $|R| = 1$，则命题得证。

设 R 中有两个不同的顶点 r_1 和 r_2。反设 K 中有两个不同的顶点 k_1 和 k_2，使得 r_1k_1 和 r_2k_2 不是边。若 r_1r_2 是一条边，则 $r_1r_2k_1k_2$ 诱导出 (e)；若 r_1r_2 不是一条边，则 $r_1r_2k_1k_2$ 诱导出 (c)。两种情况下都会得到矛盾。 □

断言 5.4。 *由 R 诱导的图是一个独立集。*

证明。如果 r_1r_2 是 $G[R]$中的一条边，那么 $K \setminus k \cup \{r_1, r_2\}$ 是一个团，这与K的极大性矛盾。

□

断言 5.5。 *要么 $R = \emptyset$，要么 $I = \emptyset$。*

证明。反设 $r \in R$ 和 $i \in I$ 都存在。令 k 为 R中 r 的一个非邻点，令 u为 K中 r的一个邻点。由于 $\{r, i, k, u\}$ 诱导出(a)，我们得到矛盾。 □

这些断言蕴含了定理的结论。 □

6 一般树

本节我们证明定理 1.5。首先陈述 Reiher 和 Wagner [11] 的以下结果。

定理 6.1 (Reiher-Wagner [11])。设 S_k 为具有 k 条边的星，$\beta \in [0,1]$ 且 $\eta = 1-\sqrt{1-\beta}$。于是

$$I(S_k,\beta) \le \max\{\beta^{(k+1)/2}, \eta + (1-\eta)\eta^k\}.$$

注意，最大值中的两个量分别在 G_n 为 $K(a,\ n)$ 或其补图时取得，其中 a 取适当值；此外，当 β 充分接近 1 时，最大值中的第一项（对应于 $\rho(S_k,\ (K(a,\ n)))$）较大。

定理 1.5 的证明。回顾一下，给定一棵树 T，我们的目标是证明存在 $\beta_T \in (0,1)$，使得对于 $\beta > \beta_T$，图序列 (G_n)（其中 $G_n = K((1+o(1))\sqrt{\beta}\, n, n)$）达到 $I(T,\beta)$。我们对 T 的边数进行归纳。对于基础情形，使用定理 6.1，它适用于所有星。对于归纳步骤，假设 T 不是星。那么 T 中存在一条边 e，使得 $T-e$ 由两棵树 T_1 和 T_2 组成，且二者都至少有一条边。由归纳法，存在 β_{T_1} 和 β_{T_2}，于是令 $\beta_T =$ 为 $\max\{\beta_{T_1}, \beta_{T_2}\}$。考虑密度为 $\beta > \beta_T$ 的 G。显然，$(\#T,G) \le (\#T_1,G)\cdot(\#T_2,G)$。根据归纳法，后两个量在 G 为一个团加上若干孤立顶点时渐近地达到最大值。此外，对于这样选取的 G，忽略低阶项，有 $(\#T,G) \ge (\#T_1,G)\cdot(\#T_2,G)$。事实上，对于 G 中的任意两个互不相交的 T_1 副本和 T_2 副本，以及分别属于 T_1 和 T_2 的任意两个顶点 v_1 和 v_2，顶点对 v_1v_2 在 G 中相邻。适当选择 v_1 和 v_2 即可得到 T 的一个副本。

□

7 星

本节中，我们探讨问题3。这里，边为红色，非边为蓝色。设 $S_{a,\ b}$ 为一颗含有 a 条边（红色）和 b 条非边（蓝色）的星，其中 $a \ge b$ 且 $a \ge 2$。我们将考虑

$S_{2,1} =$ 

作为主要示例。

回顾引言中的内容，$\mathcal{S}$ 是这样一类图：先向正则图添加全邻顶点，再添加孤立顶点；或者先添加孤立顶点，再添加全邻顶点，由此得到的图类。图类 $\mathcal{S}$ 结合了若干种不同的构造；对于某些 β，这些构造似乎能使 n 顶点图的 $\#S_{a,\ b}$ 达到最大（或最小）。这里我们列出 $\mathcal{S}$ 的这些特殊情况及其对应的密度。

(i) n 个孤立顶点与一个具有 sn 个顶点、R 为正则图的图的不交并，其中 $s \in (0,1)$。图 R 的密度为 x，因此其中每个顶点的度均为 snx。这意味着

$\beta = s^2 x$，解得 x 为 $x = \beta/s^2$。标记图 $S_{a,b}$ 的密度作为 s 的函数为

$$f(s) = s(sx)^a(1-sx)^b = s\left(\frac{\beta}{s}\right)^a\left(1-\frac{\beta}{s}\right)^b.$$

通过求解 $f'(s) = 0$ 并使用 Wolfram Alpha[2]，我们得到

$$s = \frac{\beta(a+b-1)}{a-1}, \quad x = \frac{(a-1)^2}{\beta(a+b-1)^2}. \tag{9}$$

由于 $s < 1$，(9) 给出了 β 的上界。同样，由于 $x \le 1$，(9) 给出了β的下界。这些界为

$$\beta \in \left[\frac{(a-1)^2}{(a+b-1)^2}, \frac{a-1}{(a+b-1)}\right].$$

将 (9) 中的 s 代入 f 后，我们得到标记$S_{a,b}$ 的密度关于 β 的线性估计

$$\ell(\beta) := \beta\frac{(a-1)^{a-1}b^b}{(a-1+b)^{a-1+b}}.$$

(ii) 一个大小为 $(1-s)n$ 的团，与一个大小为 sn 的正则图 R 完全连接。这是 (i) 的补图，因此前面的公式给出了边密度为 $1-\beta$ 时 $S_{b,a}$ 的密度。在这种情况下，$S_{a,b}$ 的密度关于 β 的函数为

$$\ell c(\beta) := (1-\beta)\frac{a^a(b-1)^{b-1}}{(a+b-1)^{a+b-1}} \qquad \text{for} \quad \beta \in \left[1-\frac{b-1}{(a+b-1)}, 1-\frac{(b-1)^2}{(a+b-1)^2}\right].$$

(iii) 对于边密度为 β的正则图，标记$S_{a,b}$ 的密度为

$$r(\beta) := \beta^a(1-\beta)^b.$$

(iv) 一个边密度为 β 的图 G，其顶点集划分为若干部分 X, Y, Z，其中 $G[X,Y]$ 是完全二部图，$G[Y]$ 是一个团，且没有其他边。X、Y 和 Z 的最优大小，以及标记 $S_{a,b}$ 的相应密度，由以下多项式规划给出。

$$s(\beta) := \begin{cases} \text{maximum} & xy^a(1-y)^b + y(x+y)^a(1-x-y)^b \\ \text{subject to} & x+y \le 1 \\ & 2xy + y^2 = \beta \\ & x, y \ge 0. \end{cases}$$

对于固定的 a、b 和 β，该程序只有一个变量，可以通过微积分（更准确地说，是借助计算机代数程序）求解。不幸的是，以 β 为自变量的解是一个多项式表达式的根，而该多项式取决于 a 和 b。这种构造有两种特殊情况

[2]命令：对 s(d/s)^a(1-d/s)^b 关于 s 求导，并确定其根

当 $X=\emptyset$ 或 $Z=\emptyset$ 时，这两种情况更容易处理，而且似乎经常出现。

(ivX) 若 $X=\emptyset$，则团 Y 有 $\beta^{1/2}n$ 个顶点，其余顶点在 Z 中都是孤立顶点。标记的 $S_{a,b}$ 的密度为

$$c(\beta) := \beta^{1/2}\beta^{a/2}(1-\beta^{1/2})^b.$$

(ivZ) 若 $Z=\emptyset$，则 Y 是由 $(1-(1-\beta)^{1/2})n$ 个全邻顶点组成的集合，而 $G[X]$ 是一个独立集。注意，这是 **(ivX)** 的补图。标记的 $S_{a,b}$ 的密度为

$$cc(\beta) := (1-\beta)^{1/2}(1-(1-\beta)^{1/2})^a(1-\beta)^{b/2}.$$

(v) 构造 **(iv)** 的补图。也就是说，图 G 的边密度为 β，其顶点集划分为集合 X,Y,Z，其中 Z 是全邻顶点集，$G[X]$ 诱导出一个团，且没有其他边。

$$cs(\beta) := \begin{cases} \text{maximum} & xy^b(1-y)^a + y(x+y)^b(1-x-y)^a \\ \text{subject to} & x+y \le 1 \\ & 2xy+y^2 = 1-\beta \\ & x,y \ge 0. \end{cases}$$

利用这些构造，我们给出关于 $S_{2,1}$ 的分布猜想。

猜想 2. 对于每个$S_{2,1}$-$good(G_n)$

$$\left.\begin{matrix} cc(\beta) & \text{if } \beta \in [0,x] \\ c(\beta) & \text{if } \beta \in [x,1] \end{matrix}\right\} \le \rho(S_{2,1},(G_n)) \le \begin{cases} s(\beta) & \text{if } \beta \in [0,y] \\ c(\beta) & \text{if } \beta \in [y,1/4] \\ \ell(\beta) & \text{if } \beta \in [1/4,1/2] \\ r(\beta) & \text{if } \beta \in [1/2,1], \end{cases}$$

其中 $x \in (0,1)$ 是 $cc(x)=c(x)$ 的一个解，即 $16x^3-40x^2+41x-16=0$，并且 $y \in (0,1/4)$ 是一个六次多项式的根，该多项式由确定最优解在 $x=0$ 处取得时的阈值而来。

该猜想的可视化见图 6。请注意，猜想中的上界描述了 $I(S_{2,1},\beta)$。由于在一个有 n 个顶点的图 $G=(V,E)$ 中，$S_{2,1}$的带标记副本数量以上界 $\sum_{v\in V} d_v^2\cdot(n-d_v)$ 为界，命题 3.3 证明了该猜想对 $\beta \ge 1/2$ 给出的上界。

命题 3.3 的证明方法证明了 $\beta \ge 1/4.$ 的猜想。事实上，应用引理 3.2 得到的极值构造是一个正则图加上孤立顶点，这正是 $I(S_{2,1},\beta)$ 在 $\beta \ge 1/4$时的构造。更准确地说，应用引理 3.2 表明，函数 $f(x)=(1-x)x^2$之和的最大值在存在 m 个值且每个

个值都等于 α时取得。由于 $m\alpha = \beta n$，该和为 $m\alpha^2(1-\alpha) = \beta n\alpha(1-\alpha)$。[3] 由于 $\alpha(1-\alpha)$ 是凹函数，在 $\alpha = 1/2$ 和 $\beta \le \alpha$处取得最大值，因此该和的最大值在 $\alpha = \max\{1/2, \beta\}$处取得。为了使由 m 个值均为 αn 的项组成的度序列得以实现，需要满足 $m \ge \alpha n$，这等价于条件 $\beta n \ge \alpha^2 n$。由于 α 是 $\max\{1/2, \beta\}$，因此该条件对 $\beta \ge 1/4$成立。

我们使用旗代数计算了 β的多个固定取值下的上界和下界，所得数值结果与猜想一致。虽然猜想 2 中没有用到 $\ell c(\beta)$ 和 $cs(\beta)$，但我们仍将它们列出以求完整，因为它们有助于描述其他 a 和 b取值下的 $I(S_{a,b}, \beta)$。

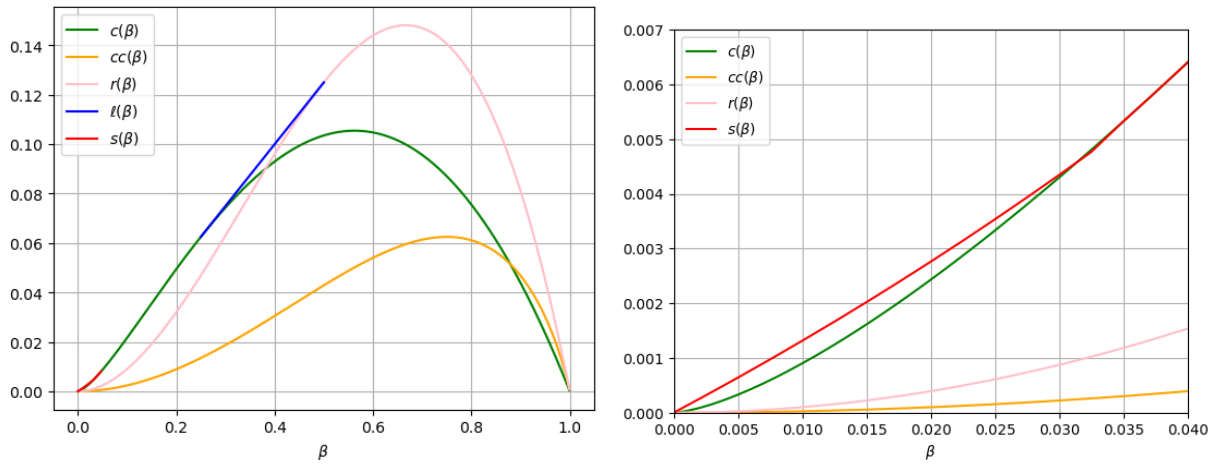


图 6：猜想 2 中使用的函数。

8 致谢

本材料基于美国国家科学基金会 DMS-1928930 号资助所支持的研究成果；2025 年春季，作者驻留于加利福尼亚州伯克利的西蒙斯·劳弗数学科学研究所期间开展了相关工作。计算机计算主要在 Alderaan 集群上完成。Alderaan 集群由美国国家科学基金会校园计算计划 OAC-2019089 号资助以及西蒙斯基金会的一项捐助共同资助。该集群由科罗拉多大学丹佛分校文理学院数学与统计科学系的计算数学中心负责运行。

[3] 此处为 $\beta \in [0,1]$，而参考文献中的证明为 $d \in [0,n]$。

参考文献

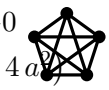
[1] Abdul Basit、Bertille Granet、Daniel Horsley、Andr´e K¨undgen 和 Katherine Staden。《半诱导性问题》，2025 年。arXiv:2501.09842。[2] Levente Bodn´ar 和 Oleg Pikhurko。《四顶点图的半诱导性》，2025 年。arXiv:2510. 24336。[3] Levente Bodn´ar 和 Oleg Pikhurko。《通过旗代数得到的一些图的精确诱导性类型结果》，2025 年。arXiv:2507.01596。[4] Hao Chen、Felix Christian Clemen 和 Jonathan A. Noel。《通过熵最大化交替路径》，2025 年。arXiv:2505.03903。[5] Hao Chen 和 Jonathan A. Noel。《关于边染色图中的交替六边形》，2025 年。arXiv:2505.09809。[6] D. Dellamonica, Jr.、P. Haxell、T. ´Luczak、D. Mubayi、B. Nagle、Y. Person、V. R¨odl 和 M. Schacht。树极小图几乎都是正则图。*J. Comb.*，3(1)：49–62，2012 年。doi: 10.4310/JOC.2012.v3.n1.a2。[7] Xizhi Liu、Dhruv Mubayi 和 Christian Reiher。诱导子图的可行域。*J.Combin. Theory Ser. B*，158：105–135，2023 年。doi:10.1016/j.jctb.2022.09.003。[8] Vladimir Nikiforov。补图无三角形的图中四团的最小数量，2005 年。arXiv:math/0501211。[9] Nicholas Pippenger 和 Martin Charles Golumbic。图的可诱导性。*Journal of Com-binatorial Theory, Series B*，19(3)：189–203，1975 年 12 月。doi:10.1016/0095-8956(75) 90084-2。[10] Alexander A. Razborov。旗代数。*J. Symbolic Logic*，72(4)：1239–1282，2007 年。doi: 10.2178/jsl/1203350785。[11] Christian Reiher 和 Stephan Wagner。最大星密度。*Studia Sci. Math. Hungar.*，55(2)：238–259，2018 年。doi:10.1556/012.2018.55.2.1395。

A P_{EENN}

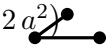
本部分给出（7）的展开。令 $\mathcal{F}_5$ 为5顶点图的集合，按同构分类。我们希望将右侧改写为

$$\sum_{F\in\mathcal{F}_5} c_{a,F}F \le \max_{F\in\mathcal{F}_5} c_{a,F} = 0.$$

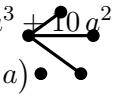
$c_{a,F}$ 的值是关于 a、 B 和 C的以下多项式。

$\sum_{F\in\mathcal{F}_5} c_{a,F}F = 0$  $+0$  $+0$  $+0$  $+0$ 

$+\left(-12\,a^8+12\,a^7+4\,a^6-4\,a^5+12\,a^4-1.5\,Ba^2-16\,a^3+4\,a^2\right)$

$+\left(-36\,a^8+36\,a^7+12\,a^6-12\,a^5+18\,a^4-4.5\,Ba^2-30\,a^3+12\,a^2\right)$ 

$+\left(-72\,a^8+72\,a^7+24\,a^6-24\,a^5+24\,a^4-9\,Ba^2+6\,Ca^2-48\,a^3-12\,Ca+24\,a^2+6\,C\right)$

$+\left(4\,a^2-4\,a\right)+\left(-24\,a^8+24\,a^7+8\,a^6-8\,a^5+12\,a^4-3\,Ba^2-18\,a^3+10\,a^2-4\,a\right)$ 

$+\left(-24\,a^8+24\,a^7+8\,a^6-8\,a^5+12\,a^4-3\,Ba^2-26\,a^3+18\,a^2-4\,a\right)$

$+\left(-12\,a^8+12\,a^7+16\,a^6-16\,a^5+2\,a^4-1.5\,Ba^2-2\,a^3+4\,a^2+1.5\,B-4\,a\right)$ 

$+\left(-48\,a^8+48\,a^7+28\,a^6-28\,a^5+12\,a^4-6\,Ba^2-20\,a^3+12\,a^2+1.5\,B-4\,a\right)$

$+\left(-60\,a^8+60\,a^7+32\,a^6-32\,a^5+16\,a^4-7.5\,Ba^2+Ca^2-36\,a^3-2\,Ca+24\,a^2+1.5\,B+C-4\,a\right)$ 

$)$



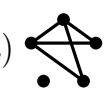




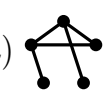


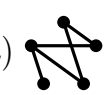


$+\left(-48\,a^8+48\,a^7+16\,a^6-16\,a^5+16\,a^4-6\,Ba^2-24\,a^3+16\,a^2-8\,a\right)$ 

$+\left(-24\,a^8+24\,a^7+8\,a^6-8\,a^5+12\,a^4-3\,Ba^2-28\,a^3+22\,a^2-6\,a\right)$ 

$+\left(-36\,a^8+36\,a^7+36\,a^6-36\,a^5+2\,a^4-4.5\,Ba^2-4\,a^3+6\,a^2+3\,B-4\,a\right)$ 

$+\left(-48\,a^8+48\,a^7+40\,a^6-40\,a^5+8\,a^4-6\,Ba^2-20\,a^3+16\,a^2+3\,B-4\,a\right)$ 

$+\left(-72\,a^8+72\,a^7+60\,a^6-60\,a^5+6\,a^4-9\,Ba^2-10\,a^3+8\,a^2+4.5\,B-4\,a\right)$ 

$+\left(-60\,a^8+60\,a^7+56\,a^6-56\,a^5+8\,a^4-7.5\,Ba^2+Ca^2-24\,a^3-Ca+20\,a^2+4.5\,B-4\,a\right.$ 

$$+\left(-108\,a^8+108\,a^7+108\,a^6-108\,a^5-13.5\,Ba^2+9\,B\right)$$ 

$$+\left(-72\,a^8+72\,a^7+96\,a^6-96\,a^5-9\,Ba^2+6\,Ca^2-12\,a^3-6\,Ca+12\,a^2+9\,B\right)$$ 

$$+\left(6\,a^4-6\,a^3+6\,a^2-6\,a\right)$$ 

$$+\left(-36\,a^8+36\,a^7+24\,a^6-24\,a^5+8\,a^4-4.5\,Ba^2-12\,a^3+10\,a^2+1.5\,B-6\,a\right)$$ 

$$+\left(-48\,a^8+48\,a^7+40\,a^6-40\,a^5+8\,a^4-6\,Ba^2-16\,a^3+14\,a^2+3\,B-6\,a\right)$$ 

$$+\left(-36\,a^8+36\,a^7+48\,a^6-48\,a^5+6\,a^4-4.5\,Ba^2-18\,a^3+18\,a^2+4.5\,B-6\,a\right)$$ 

$$+\left(-48\,a^8+48\,a^7+16\,a^6-16\,a^5+16\,a^4-6\,Ba^2-24\,a^3+16\,a^2-8\,a\right)$$ 

$$+\left(-60\,a^8+60\,a^7+80\,a^6-80\,a^5-10\,a^4-7.5\,Ba^2+10\,a^3+7.5\,B\right)$$ 

$$+\left(-72\,a^8+72\,a^7+72\,a^6-72\,a^5-9\,Ba^2-2\,a^3+4\,a^2+6\,B-2\,a\right)$$ 

$$+\left(-60\,a^8+60\,a^7+56\,a^6-56\,a^5+6\,a^4-7.5\,Ba^2-12\,a^3+10\,a^2+4.5\,B-4\,a\right)$$ 

$$+\left(-84\,a^8+84\,a^7+88\,a^6-88\,a^5-2\,a^4-10.5\,Ba^2+2\,a^3+7.5\,B\right)$$ 

$$+\left(-48\,a^8+48\,a^7+64\,a^6-64\,a^5+2\,a^4-6\,Ba^2-6\,a^3+6\,a^2+6\,B-2\,a\right)$$ 

$$+\left(-72\,a^8+72\,a^7+72\,a^6-72\,a^5-9\,Ba^2+2\,Ca^2+6\,B\right)$$ 

$$+\left(-36\,a^8+36\,a^7+48\,a^6-48\,a^5-4.5\,Ba^2+3\,Ca^2+4.5\,B\right)$$ 

非零多项式的图 $c_{a,\ F}$ 在两种 B 和 C取值方案下的图示于图7。

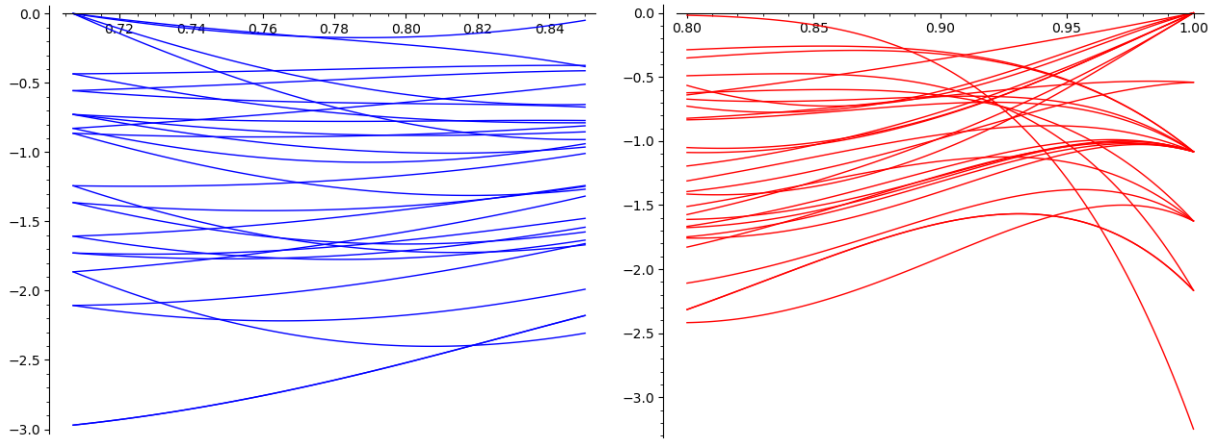


图7：左侧为情况 $B = C = \bar{\sqrt{2}} - 1$，右侧为 $B = 0.361$ 和 $C = 0$。