arXiv:2611.28664v1 [math.CO] 21 Nov 2026

遗传超图族多色着色札记 III

D¨om¨ot¨or P´alv¨olgyi *

免责声明。 本文由 ChatGPT 撰写，我仅提供了非常有限的反馈。

摘要

对于每个整数 $h \geq 4$，我们构造了一个具有 $m_2 = h$ 且.....的超图遗传族

$$m_3 \geq 2(h-1) + 2\left\lfloor \frac{h-1}{3} \right\rfloor + 1 \approx \frac{8}{3} m_2.$$

这改进了前述笔记中针对每个 $m_2 \geq 4$ 得到的下界 $m_3 \geq 2m_2$。我们还证明，在满足 $m_2 \leq h$ 的遗传族上，加权横截问题的最坏情况整性间隙介于 $h-1$ 和 $5000h$ 之间。因此，每个这样的族都存在大小至多为 $5000h/\varepsilon$ 的强 ε-网。

关键词： 超图，多色着色，性质 B，横截集，整性间隙**数学主题分类（MSC）：** 05C15, 05C65

关于定义和记号：本文中所有超图均为有限超图。对于 $X \subseteq V(\mathcal{H})$，其迹$\mathcal{H}[X]$ 的边为 $E \cap X$，其中 $E \in E(\mathcal{H})$。迹的子超图称为受限子超图。若超图的每条边大小至少为 h，则称其为 h-重；若每条边大小均为 r，则称其为 r-均匀。若每条边都包含所有颜色，则称一种着色为多色着色。特别地，多色二着色就是恰当二染色。

超图族 $\mathcal{F}$ 若包含其每个成员的所有受限子超图，则称其为遗传的。按照 [1], 令 $m_k(\mathcal{F})$为最小正整数 m，使得每个 m-重成员 $\mathcal{F}$ 都具有一个多色 k-着色，并在不存在这样的 $m_k(\mathcal{F})$时 $= \infty$ 令 m。因此，$m_2(\mathcal{F})$是可适当进行二着色的阈值。对于单个超图 $\mathcal{H}$，我们用 $m_k(\mathcal{H})$表示由其所有受限子超图组成的遗传族的阈值。

第一篇短文 [1] 给出了一个具有 $m_2 = 3$ 和 $m_3 = 6$的5-均匀超图。第二篇短文 [2]对每个 $h \geq 4$都给出了一个具有 $m_2 \leq h$ 和 $m_3 = 2h$的$(2h-1)$-均匀超图。对于每个 $h \geq 4$，我们改进了后者的界。

定理 1。 对于每个整数 $h \geq 4$，令

$$r(h) = 2(h-1) + 2\left\lfloor \frac{h-1}{3} \right\rfloor.$$

存在一个 $r(h)$-均匀超图 $\mathcal{H}$，它没有多彩3染色，但其所有h-重受限子超图都可以2着色。特别地，$m_2(\mathcal{H}) \leq h$ 以及 $m_3(\mathcal{H}) = r(h)+1$。

*匈牙利布达佩斯的厄特沃什·罗兰大学和阿尔弗雷德·雷尼数学研究所。电子邮箱：domotor.palvolgyi@ttk.elte.hu。研究获NRDI EXCELLENCE-24项目（编号151504）“组合学与几何”资助，以及欧洲研究委员会高级资助项目（编号101054936）ERMiD资助。

要得到抽象图中的遗传族，在 $\mathcal{H}$ 的顶点集上添加所有$(h-1)$-边，并取其生成的遗传族。这些添加的边没有 h-重迹，而顶点数为 $2h-3$ 的完全$(h-1)$-均匀子超图没有恰当二染色。顶点数足够，因为$r(h) \geq 2h$。原超图仍然是一个子超图，因此 $m_2 = h$ 且 $m_3 \geq r(h) + 1$。

现在我们转而关注如何界定 ε-网的大小。

对于边非空的超图 $\mathcal{H}$ 以及非负顶点代价 c，令 $\tau_c(\mathcal{H})$ 表示横截集的最小代价，即一个与每条边相交的集合。其分数松弛为

$$\tau_c^*(\mathcal{H}) = \min \left\{ \sum_{v \in V(\mathcal{H})} c(v)x(v) : x(v) \geq 0,\ \sum_{v \in E} x(v) \geq 1 \text{ for every } E \in E(\mathcal{H}) \right\}.$$

令 $G(h)$ 表示在满足 $\tau_c^*(\mathcal{H}) > 0$ 且 $m_2(\mathcal{H}) \leq h$ 的代价和超图上，$\tau_c(\mathcal{H})/\tau_c^*(\mathcal{H})$ 的上确界。

定理 2。 对于每个整数 $h \geq 2$,

$$h - 1 \leq G(h) \leq 5000h.$$

特别地，对于每个满足 $2 \leq m_2(\mathcal{F}) < \infty$ 的遗传族 $\mathcal{F}$、每个边非空的 $\mathcal{H} \in \mathcal{F}$ 以及每个非负顶点代价 c，都有

$$\tau_c(\mathcal{H}) \leq 5000\, m_2(\mathcal{F})\, \tau_c^*(\mathcal{H}).$$

如果 $m_2(\mathcal{F}) \leq h$，则 $\mathcal{F}$ 的每个成员的 VC 维数至多为 $2(h-1)$。一个由$2h-1$ 个顶点组成的打散集将给出一个受限子超图，该子超图包含该集合的所有 h-子集，而任意 2-着色都会使其中某个 h-子集成为单色集。Haussler–Welzl 定理 [3]因而给出大小为 $\varepsilon\log(1/\varepsilon))$ 的强 $\varepsilon \to 0$-网。下面的推论消除了对数因子。

推论 3. 设 $\mathcal{F}$ 是一个遗传族，满足 $m_2(\mathcal{F}) \leq h < \infty$，其中 $h \geq 2$。对于每个 $\mathcal{H} \in \mathcal{F}$、其顶点上的每个概率测度 μ 以及每个 $0 < \varepsilon \leq 1$，都存在一个大小至多为 $5000h/\varepsilon$的强 ε-网。也就是说，存在一个至多包含 $5000h/\varepsilon$个顶点的集合，它与每条满足$\mu(E) \geq \varepsilon$的边 E 都相交。

证明。 只保留满足 E $\mu(E) \geq \varepsilon$的边。所得子超图根据遗传性属于 $\mathcal{F}$。权重 $x(v) = \mu(v)/\varepsilon$ 构成一个总权重为 $1/\varepsilon$的分数横截。应用定理 2 并取单位代价。

□

$m_k(\mathcal{F})$的线性增长还将意味着均匀测度下的线性大小 ε-网。事实上，假设对所有 $k \geq 2$都有 $m_k(\mathcal{F}) \leq Ck$。在 $\mathcal{F}$的一个具有 n个顶点的成员中，保留大小至少为εn的边。如果 $\varepsilon n \geq 2C$，则令 $k = \lfloor \varepsilon n/C \rfloor$。保留下来的子超图是 $m_k(\mathcal{F})$-重的，而多色 k-着色中最小的颜色类会与每条保留的边相交，其大小至多为 $n/k \leq 2C/\varepsilon$。如果$\varepsilon n < 2C$，则整个顶点集都满足该大小界。因此，若均匀测度不存在具有线性大小的 ε-网，即使在几何实现中不存在弱 ε-网，也将排除 $m_k(\mathcal{F}) = O(k)$。

1 构造

我们描述一个辅助超图 $\mathcal{G}$，它将用于定义 $\mathcal{H}$。通过计算机搜索发现，它具有引理 4中所述的性质。它有14个顶点、五条2-边和140条8-边。其顶点集 V 由 $\mathbb{Z}_7$的两个副本组成。第一个副本的标签为 i，第二个副本的标签为i'，其中 $i \in \mathbb{Z}_7$。记作 $P_i = \{i, i'\}$。短边为 P_2、P_3、P_4、P_5、P_6。要定义长边，为每个四元子集 $Q \subset \mathbb{Z}_7$分配一对 $D(Q) \subset Q$。以下五对（Q、$D(Q)$）通过循环平移规定了该分配。

Q	0123	0124	0134	0234	0135
$D(Q)$	13	01	34	02	03

这里，像 0123 这样的字符串表示集合 $\{0,1,2,3\}$。首先，我们通过将五个表格条目模 7 平移，为 D 赋予全部 35 个值。例如，将 $Q = 0123$ 和 $D(Q) = 13$ 平移 6，得到 $Q = 0126$，其暂定值为 $D(0126) = 02$。赋予全部 35 个值后，我们将 $D(0126)$从 02 改为 06，并将 $D(0124)$从 01 改为 24，其余 33 个值保持不变。

对于每个 Q，从每个由 $D(Q)$索引的点对中省去两个点，并从每个由 $Q \setminus D(Q)$索引的点对中省去一个点。这样省去了 $2+2+1+1 = 6$ 个点。对这两个单点的四种选择，各自都通过在 V 中取补集得到一条 8 边。例如，当 $Q = 0126$ 时，我们省去所有 P_0 和 P_6。如果再省去 1 和 2，得到的边是

$$P_3 \cup P_4 \cup P_5 \cup \{1', 2'\}.$$

每条这样的边都确定 Q，它恰好由未完全包含在该边中的四对的指标组成。对于固定的 Q，四种选择会给出不同的边。因此，共有$35 \cdot 4 = 140$ 条长边。

引理 4. 超图 $\mathcal{G}$ 具有以下性质。

1. 每个 至多含四个点的子集都与某条长边及某条短边不相交 e.
2. 对于每个 $X \subseteq V$，都存在一种 X 的红蓝着色，使得每条长边的单色迹大小至多为三，且每条短边的单色迹大小至多为一。

证明。 五条短边两两不相交，因此没有四个点能与它们全部相交。对于长边，只需检验四元子集，因为较小的集合可以扩充为四元子集。附录中的确定性程序A 检验所有$\binom{14}{4} = 1001$ 个四元子集与 140 条长边的关系。对于第二个断言，它检验所有 $2^{14} = 16384$ 种 X。对于每个 X，它都会寻找一种着色，使每条大小至少为四的长边迹都被分开，并且每条包含于 X中的短边也都被分开。 □

定理 1的证明。 设 $h - 1 = 3q + t$，其中 $q = \lfloor (h-1)/3 \rfloor$ 且 $t = (h-1)$ 模 3，因此$0 \le t \le 2$。取 $3(q + t - 1) + 1$ 个互不相交的 $\mathcal{G}$副本。 $\mathcal{H}$ 中的一条边是分别从 q 个副本中各取一条长边，并从 t 其他副本中各取一条短边所得的并集。每条这样的边的大小为$8q + 2t = 2(h-1) + 2q = r(h)$。

假设存在一种多色 3-着色。在 $\mathcal{G}$的每个副本中，选择一个最小的颜色类，该颜色类至多包含四个顶点。根据鸽巢原理，至少有 $q + t$ 个副本选择了同一种颜色。根据引理 4，我们可以在其中每个 q副本中选择一条避开该颜色的长边

个副本，并在其余 t个副本中各选择一条避开该颜色的短边。它们的并集是 $\mathcal{H}$ 的一条边，其中不包含所选颜色，矛盾。

现在令 X 为顶点集的任意子集。按照引理4，对它与每个副本的交进行着色。边 $\mathcal{H}$ 的单色迹在每条长边中至多包含三个点，在每条短边中至多包含一个点。因此，其大小至多为 $3q + t = h - 1$。于是 $m_2(\mathcal{H}) \leq h$。由于不存在多色 3-着色，得到 $m_3(\mathcal{H}) > r(h)$，且没有受限子超图包含大小超过 $r(h$的边，因此 $m_3(\mathcal{H}) = r(h) + 1$。 □

2 加权横截间隙

为定理2中的下界，取顶点集 n 上所有的 $(h-1)$-子集，并赋予单位代价。关于 h-重限制的条件是空泛成立的。横截数为 $n-h+2$，分数横截数为 $n/(h-1)$，由对称性或对边约束取平均可得。它们的比值在 n 趋于无穷时趋于 $h-1$ 。

对于上界，我们先处理取二的整数次幂的代价。在抽样论证中，分别对每个成本类别进行抽样。我们同时对所有红蓝标记给出抽样误差代价的上界，这样就可以将该上界应用于遗传假设所给出的着色。

这一论证与 Haussler 和 Welzl 的 ε-网定理的双重抽样证明有关 [3]，；Matou˘sek [4,第 10] 章中也介绍了该证明。Agrawal、Wang 和 Ye [5, 引理 2]中也采用了类似的方法：利用集中不等式，并对由线性规划对偶解诱导的模式应用并合界。我们证明中特有的一步，是在定义 F_i 的更大可行集上最小化定义 F_{i+1} 的目标函数。这会在揭示成本类别 i 中哪些位置被标记为蓝色之前，先固定一个对偶打包，并给出一个可在各成本类别间望远镜式累加的比较。

定理 *2* 上界的证明。我们可以删除零代价顶点以及所有与这些顶点相交的边，因为可以将这些顶点免费纳入横截集。这不会改变整数最优值或分数最优值，也不会影响遗传假设。将其余每个顶点的代价向上取整为二的整数次幂。整数最优值只会增大，而分数最优值的增幅至多为两倍。因此，只需证明以下针对代价为二的整数次幂的界 $2500h$即可。

对于给定的代价，考虑所有非空子族中使两个最优值之比最大的边子族。将该比值记为 R。固定一个有理数最优分数横截 x ，并记 $C = \tau_c^*(\mathcal{H}) = \sum_{v \in V(\mathcal{H})} c(v)x(v)$。我们可以取该实例及 x 的足够多个不交副本。对于并集上的分数横截及其代价，仍分别使用记号 x 和 C。由于两个最优值在这些副本上都是可加的，比值仍为 R ，且同一上界 R 对每个非空子族都适用。因此可以假设，对于每个具有正 x-质量的代价 a，$10h \sum_{v:c(v)=a} x(v)$ 是至少为 1000 的整数。遗传假设在不交并下也仍然成立。

将这些成本类别按递增顺序列为 $c_1, \ldots, c_k$。对于每个 i，画出

$$n_i = 10h \sum_{v:c(v)=c_i} x(v)$$

从该类中放回抽样位置，顶点概率与 $x(v)$ 成正比。所有抽样结果彼此独立。计入重复项后，这些位置的总成本为 $\sum_i c_i n_i = 10hC$。对于任一结果，其位置都可以进行红蓝标记，使得每条至少包含 h 个抽样

出现次数的边都同时包含两种颜色。要证明这一点，先将每个至少抽到两次的顶点的出现次数分配到两种颜色中。再对恰好抽到一次的顶点集合应用遗传假设，为这些位置着色。

可行　对偶装箱 y 为每条边分配一个非负权重 y_E，其负载为

$$\ell_y(v) = \sum_{E \ni v} y_E \le c(v).$$

其总权重至　　多为 C，根据线性规划对偶性。对于抽样位置的一种标记，定义松弛量　　s

$$d_i(y) = \sum_{v : c(v) = c_i} x(v)\big(c_i - \ell_y(v)\big), \qquad e_i(y) = \sum_{\text{blue positions } v \text{ in class } i} \big(c_i - \ell_y(v)\big).$$

令 Y_i 为由可行对偶装填构成的集合，这些装填所支持的边避开类别 $j < i$ 中的所有抽样位置，以及类别 $j \ge i$ 中的所有红色位置。特别地，Y_{k+1} 避开每个抽样位置。设

$$N_i = \sum_{j \ge i} n_j, \qquad F_i = \min_{y \in Y_i} \left(30 \cdot 10h \sum_{j < i} d_j(y) + \sum_{j \ge i} e_j(y) \right).$$

我们声称，存在一种抽样，使得对每种标记和每个 $i \le k$，以下结论同时成立：

$$F_i \le F_{i+1} + 0.14 c_i N_i. \tag{1}$$

固定 i，并固定类别 $j \ge i$ 中的标签。以类别 i 外的抽样结果以及类别 i 中位置标记为红色的抽样结果为条件。选择 $\widehat{y}$ 以最小化

$$30 \cdot 10h \sum_{j \le i} d_j(y) + \sum_{j > i} e_j(y)$$

关于 Y_i。该目标和 Y_i 现在均已固定，且与类别 i 中位置标记为蓝色的抽样结果无关。由于 $Y_{i+1} \subseteq Y_i$，$\widehat{y}$ 处的值至多为 F_{i+1}。

对于类别 i 中标记为蓝色的位置 v 处的一次抽样，其归一化松弛量 $Z = (c_i - \ell_{\widehat{y}}(v))/c_i$ 属于 $[0,1]$。由凸性，$\exp(5Z) \le 1 + (e^5 - 1)Z$。标记为蓝色的位置上的归一化松弛量之和的条件期望至多为 $10h\, d_i(\widehat{y})/c_i$，因为标记为蓝色的位置数至多为 n_i。独立性和 $1 + z \le e^z$ 给出

$$\mathbb{E}[\exp(5 e_i(\widehat{y})/c_i)] \le \exp\big((e^5 - 1) \cdot 10h\, d_i(\widehat{y})/c_i\big).$$

此处以及下面的概率界中，条件均如上所述。由于 $e^5 - 1 < 5 \cdot 30$，由马尔可夫不等式可得

$$\Pr(e_i(\widehat{y}) > 30 \cdot 10h\, d_i(\widehat{y}) + 0.14 c_i N_i) \le e^{-0.7 N_i}.$$

在此事件之外，我们可以计算 F_i 在 $\widehat{y}$ 处的目标函数，并得到（1）。该界在不进行条件化时也成立。

只有类别 $j \ge i$ 中的标记会进入两个最小值，因此在步骤 i 应用并合界时有 2^{N_i} 种选择。由于 $n_i \ge 1000$，我们有 $N_i \ge 1000(k - i + 1)$。因此，总失败概率至多为

$$\sum_{i=1}^{k} (2e^{-0.7})^{N_i} \le \sum_{j \ge 1} (2e^{-0.7})^{1000 j} < 1.$$

固定一个使得（1）对所有标记都成立的结果。

使用上文所述抽样位置的遗传红蓝标记。如有必要，交换颜色，使蓝色位置的成本至少为 $10hC/2$。支持 Y_1 中某个打包的边不会包含红色位置，因此至多有 $h-1$ 个蓝色位置。因此

$$F_1 \geq \frac{10hC}{2} - (h-1)C > 0.4 \cdot 10hC.$$

对于另一端点，舍弃所有与抽样相交的边。将抽样顶点加入剩余子族的一个横截集中，便得到 $\mathcal{H}$ 的一个横截集，额外成本至多为 $10hC$。因此，根据对 R 的选择，剩余子族的整数最优值至少为 $RC-10hC$，分数最优值至少为 $C-10hC/R$。最后一个不等式对空子族也成立，此时 $RC \leq 10hC$。由对偶性，存在一个权重至少为 $C-10hC/R$ 的打包 $y \in Y_{k+1}$。由于 x 可行，

$$\sum_i d_i(y) = C - \sum_E y_E \sum_{v \in E} x(v) \leq C - \sum_E y_E \leq 10hC/R.$$

由此可得 $F_{k+1} \leq 30(10h)^2 C/R$。

最后，不同的代价都是 2 的幂，因此 $\sum_{i \leq j} c_i \leq 2c_j$。对（1）求和并使用 g

$$\sum_i c_i N_i = \sum_j n_j \sum_{i \leq j} c_i \leq 2 \cdot 10hC$$

得到

$$0.4 \cdot 10hC < F_1 \leq 30(10h)^2 C/R + 0.28 \cdot 10hC.$$

因此 $R < 30(10h)/0.12 = 2500h$，如所要求。 □

现在我们陈述定理2的又一个加强版本。证明是对上述论证的直接改编，因此省略细节。对于 $k \geq 2$，令 p_k（$\mathcal{F}$）为满足以下条件的最小正整数：h 每个 h-重成员 $\mathcal{F}$ 都允许进行一个适当的 k-着色，其中没有任何边是单色的。若不存在这样的 p_k（$\mathcal{F}$），则令 $=\infty$。于是 p_2（$\mathcal{F}$） $= m_2$（$\mathcal{F}$）。

定理 5。对于每个 $k \geq 2$，存在一个常数 C_k，使得如果 $\mathcal{F}$ 是遗传的且 p_k（$\mathcal{F}$） $\leq h < \infty$，则

$$\tau_c(\mathcal{H}) \leq C_k h \, \tau_c^*(\mathcal{H})$$

对于每个具有非空边集的 $\mathcal{H}$ 成员以及每个非负顶点代价 c，对于 $\mathcal{F}$ 某个成员的顶点上的每个概率测度以及每个 $0 < \varepsilon \leq 1$，都存在一个大小至多为 $C_k h/\varepsilon$ 的强 ε-网。

每当 $p_k(\mathcal{F}) \leq h$ 时，同样的打散论证给出 VC 维数上界 $k(h-1)$。

参考文献

[1] D. P´alv¨olgyi, 《遗传超图族的多色着色注记》, Graphs and Combina- torics 40（2024）, 文章 131。[2] D. P´alv¨olgyi, 《遗传超图族的多色着色注记 II》, arXiv:2604.24412, 2026。[3] D. Haussler 和 E. Welzl, 《ε-nets 与单纯形范围查询》, Discrete and Computational Geometry 2（1987）, 127–151。[4] J. Matouˇsek, 《离散几何讲义》, Graduate Texts in Mathematics 212, Springer, 2002。[5] S. Agrawal、Z. Wang 和 Y. Ye, 《在线性规划的动态近似最优算法》, Operations Research 62（2014）, 876–890。

验证代码

以下 Python 程序验证引理 4。它使用位掩码表示顶点集，且不依赖外部程序包。

```
from itertools import combinations, productdef mask(indices):
return sum(1 << int(i) for i in indices)def shift(indices, j):
return sum(1 << ((int(i) + j) % 7) for i in indices)D = {}for Q,
 pair in [("0123", "13"), ("0124", "01"), ("0134", "34"),key = shift(Q, j)
          ("0234", "02")  ("0135", "03")]: ,assert key not in D
for j in range(7):
D[key] = shift(pair, j)assert len(D) == 35
D[mask("0126")] = mask("06")D[mask("0124")] = mask("24")
V = (1 << 14) - 1long_edges = set()for Q, pair in D.items():
assert Q.bit_count() == 4 and pair.bit_count() == 2i,
assert Q & pair == pair


 j = [k for k in range(7) if (Q ^ pair) & (1 << k)]
omit = sum(3 << (2*k) for k in range(7) if pair & (1 << k))for a,
 b in product(range(2), repeat=2):
long_edges.add(V ^ (omit | (1 << (2*i+a)) | (1 << (2*j+b))))
short_edges = [3 << (2*k) for k in range(2, 7)]
assert len(long_edges) == 140
assert all(e.bit_count() == 8 for e in long_edges)
for four in combinations(range(14), 4):
assert any(e & mask(four) == 0 for e in long_edges)def admissible(S):
heavy = {e for e in long_edges if (e & S).bit_count() >= 4}
if not heavy:
    return True                                   R = Swhile R:


        if all(e & R and e & B for e in heavy):
B = S ^ R    return True                          R = (R - 1) & S
return Falsefor S in range(1 << 14):
    assert admissible(S)  S
print("140 long edges, 1001 four-sets, 16384 restrictions verified.")
```