arXiv:2603.24890v1 [math.PR] 26 Mar 2026

二元域上稀疏方程的不一致概率

Peter Horak[1] |Igor Semaev[2]1美国华盛顿大学 | 2挪威卑尔根大学 通讯作者：Peter Horak（horak@uw.edu）
关键词：稀疏多项式系统；二元域上的随机方程，随机可满足性问题。

摘要

设 n 表示稀疏多项式系统中的变量数， m 表示方程数，且该系统定义在二元域上。我们研究在二元域上随机生成的稀疏多项式系统的不一致概率，其中每个方程至多依赖于 k 个变量，且变量数不断增长。将该系统与一个超图关联起来后，我们证明，不一致概率强烈依赖于该超图的结构性质，而不仅仅取决于 n, m, 和 k. 。利用容斥原理，我们推导出一般界，并获得完全 k-均匀超图的紧渐近结果。在 2-稀疏情形下，我们给出路径和星形图的显式公式，刻画极值树和森林，并猜想一个关于环的公式。这些结果对 SAT 求解和密码分析具有启示意义。

1 引言

设 n, m, k 为自然数， X 表示一个由 n 个变量组成的集合，这些变量的取值属于二元域 $\mathbb{F}_2$。此外，设 $X_i, 1 \le i \le m,$ 为 X 的非空子集，并设 $f_i(X_i)$ $i = 1, \ldots, m,$ 为 $\mathbb{F}_2$ 上的多项式，它们仅依赖于 X_i中的变量。

1.1 问题

在本文中，我们旨在分析 $\mathbb{F}_2^n$ 中方程组的解

$$f_1(X_1) = 0,\ f_2(X_2) = 0, \ldots, f_m(X_m) = 0. \tag{1}$$

若 $|X_i| \le k$，则称该系统为 k-稀疏系统，其中 k 是一个固定的小数（例如 3 或 4），而 n 很大。照常，若方程组 (1) 在 $\mathbb{F}_2^n$ 中有解，则称其为相容（可解）；否则称其为不相容。

若变量子集 X_i 固定（允许重复），且多项式 f_i 均匀且相互独立地随机选取，则称方程组 (1) 为随机生成。可将一个超图 $\mathcal{X} = (X, \{X_1, X_2, \ldots, X_m\})$，与之关联，其中顶点集为 X，边为子集 X_i。本文研究随机生成的方程组 (1) 的不一致概率 $p(\mathcal{X})$。特别地，我们证明 $p(\mathcal{X})$ 显著取决于底层超图 $\mathcal{X}$ 的结构性质，而不仅仅取决于参数 n, m, 和 k。

1.2 动机

据 [10]观察，随着变量数量增加，现代 SAT 求解器（如 MiniSat） [5]在处理随机生成的稀疏方程组 (1) 时，表现出次指数级而非指数级的运行特性。鉴于求解此类方程组（即确定其可满足性）属于 NP 完全问题，且该问题可归约为 SAT 问题，这一结果令人惊讶。

在这项工作中，我们初步尝试理解并解释这一现象。在第 2 节中，我们证明，稀疏情形下的不一致概率 $p(\mathcal{X})$ 显著更高。这很可能解释了 SAT 求解器在此类实例中为何效率惊人。将少量变量固定为常数后，求解器往往会在由此产生的不一致子系统中检测到矛盾。一旦发现矛盾，当前赋值就会被舍弃，并尝试新的赋值——实际上，这相当于沿搜索树遍历以求解该系统。

基于这一思路，我们证明，对于某些随机生成的系统，不一致概率的下界可以导出求解（1）的期望复杂度上界。假设系统（1）分为两个子系统:

$$f_1(X_1) = 0, \ldots, f_k(X_k) = 0, \tag{2}$$

and

$$f_{k+1}(X_{k+1}) = 0, \ldots, f_m(X_m) = 0, \tag{3}$$

其中 $Y_1 = X_1 \cup \ldots \cup X_k$ 和 $Y_2 = X_{k+1} \cup \ldots \cup X_m$，以及 $Y \subseteq Y_1 \cap Y_2$。

令 q_1 和 q_2 分别表示对于给定赋值 $Y = a$，(2) 和 (3) 一致的概率。由于方程是随机生成的，这些概率不依赖于 a。令 Q_1 和 Q_2分别表示固定之后求解 (2) 和 (3) 的平均复杂度。那么，求解 (1) 的平均复杂度为:

$$2^{|Y|} \left(Q_1 + q_1 Q_2 + q_1 q_2 Q \right), \tag{4}$$

其中， Q 表示将 (2) 和 (3) 的解组合成 (1) 的解所需的复杂度，这些解在 Y 中的变量取值一致。

为说明这一点，请注意：

- $2^{|Y|}Q_1$ 是对 Y 的所有固定赋值求解 (2) 的期望复杂度。
- (2) 相容的固定赋值的期望数量为$2^{|Y|}q_1$.
- 对于每个这样的固定赋值，都必须求解 (3)，这为总复杂度贡献了 $2^{|Y|}q_1Q_2$。
- 两个部分都相容的固定赋值数量为 $Y = a\ 2^{|Y|}q_1q_2Q$，合并它们的解会产生最后一项。

由此得到 (4) 中给出的总复杂度界。

显然，该系统可以拆分为两个以上的子系统，从而可能进一步降低复杂度。在实践中，可以通过针对每个变量固定值，尝试识别不一致子系统来动态完成这一过程。因此，一致性概率的上界q_1， q_2，或相应不一致概率的下界，都意味着求解（1）的平均复杂度存在上界。

1.3 密码分析中的稀疏方程

稀疏方程组（1）的定义源于密码分析中的应用，其目标是求离散函数的逆，以破解密码。密码所实现的映射通常由多个函数复合而成，每个函数仅涉及少量变量，例如现代分组密码中的 S 盒、置换和线性层。这些映射可以用少量简单门表示，因为它们的设计目标是能够高效计算。为了简化描述密码的方程并得到（1）型稀疏系统，可以引入中间变量。尽管所得系统可能涉及大量变量，但在内存使用和计算操作方面通常仍然易于处理。感兴趣的读者可参阅[2]和 [8]。

1.4 稀疏方程与 k-SAT问题

关于所谓的相变现象， k-SAT 问题的一致性（可满足性）概率已得到广泛研究。布尔 k-可满足性（k-SAT）问题涉及判定一个由子句合取而成的合取范式（CNF）公式是否可满足——该公式由

长度为 k -的子句合取而成。每个子句都是由 k 个文字构成的析取式，其中一个文字要么是变量 $x \in X$ ，要么是它的否定 $\bar{x}$。我们假设没有子句同时包含某个变量及其否定。求解 k-稀疏方程系统（1）的问题可表示为一个 k-SAT 问题实例，反之亦然。

对于变量集 $|X| = n$，大小为 k 的可能子句数量为 $\binom{n}{k}2^k$。设 m 为自然数。随机 k-SAT 实例是指其合取范式由 m 个子句组成，这些子句从所有可能的 k-子句中有放回地独立、均匀随机选取。（其他随机性定义也已有研究。）

人们普遍认为，随机 k-SAT 实例会呈现相变。具体而言， [3] 年有人猜想，对于每个 $k \geq 2$，都存在一个阈值 r_k ，使得对于任意 $\varepsilon > 0$ ，当 $n \to \infty$ 时有如下结论：若 $m/n < r_k - \varepsilon$，则合取范式以高概率可满足；若 $m/n > r_k + \varepsilon$ ，则合取范式以高概率不可满足。

对于 $k = 2$ ，这一猜想已在阈值为 $r_2 = 1$ [6] 时得到证明。对于 $k = 3$，已证明，若 $m/n < 3.145$，随机合取范式几乎总是可满足；若 $m/n > 4.596$，则几乎总是不可满足， [1] 且 [7]。人们认为，3-SAT 的阈值约为 4。2， [11]。对于足够大的 k，这一猜想已于 [4] 年得到证明。

2 不一致概率的界

本节的主要结果给出了针对一般超图 $\mathcal{X}$ 和完全 k-均匀超图 $\mathcal{X}_k = (X, \{X_1, X_2, \ldots, X_m\})$ 的不一致概率 $p(\mathcal{X})$ 的界，其中 X_i 是 X 的所有不同的 k-子集；即 $m = \binom{n}{k}$。我们先介绍两个命题，它们将在本节后续证明中频繁用到。

2.1 辅助结果

命题 1 设 *(1)* 是一个含有 m 个方程且满足 $|X| = n$ 的随机生成的方程组。那么

1. 每个方程 $f_i(X_i) = 0$单独来看是不一致的，概率为 $\frac{1}{2^{2^{|X_i|}}}$ ；*2.* 任意 $a \in \mathbb{F}_2^n$是方程 $f_i(X_i) = 0$ 的解，概率为$1/2$；*3.* 任意 $a \in \mathbb{F}_2^n$ 是整个系统 *(1)* 的解，概率为 $\frac{1}{2^m}$.

证明。 设 $|X_i| = k$. 所有 2^{2^k} 变量上的 $\mathbb{F}_2$ 布尔多项式集合，与 k 的所有子集集合之间存在自然的一一对应关系 $\mathbb{F}_2^k$，其中每个多项式都对应于其根集。恰有一个多项式，即 $f(X_i) = 1$，对应于空集。因此，随机选取一个多项式$f_i(X_i)$后，得到不一致方程 $f_i(X_i) = 0$ 的概率为 $\frac{1}{2^{2^k}}$. 由于任意$a \in \mathbb{F}_2^k$ 恰好属于 $\mathbb{F}_2^k$的所有子集中的一半，因此~~a~~是随机选取的多项式的根的概率为 $_1{}_2$。 由于系统 (1) 中的方程是独立生成的， a 是全部 m 个方程的解的概率为 $\frac{1}{2^m}$.

— ■

以下命题几乎显而易见，且已用于估算求解 [9, 10] 中 (1) 的复杂度。

命题 2 设 $\mathcal{X} = (X, \{X_1, X_2, \ldots, X_m\})$，并设 *(1)* 是随机生成的方程组。那么

1. $p(\mathcal{X}) \geq 1 - \prod_{i=1}^{m}(1 - 2^{-2^{|X_i|}})$。
2. 如果 $X_1, \ldots, X_m$ 两两不相交，那么 $p(\mathcal{X}) = 1 - \prod_{i=1}^{m}(1 - 2^{-2^{|X_i|}})$。

证明。 如果系统 (1) 相容，那么每个方程 $f_i(X_i)$ =0 都相容。根据命题 1， $f_i(X_i) = 0$ 有解（即相容）的概率为 $1-2^{-2^{|X_i|}}$。因此， $1-p(\mathcal{X}) \leq \prod_{i=1}^{m}(1-2^{-2^{|X_i|}})$。这就推出了第一个结论。

如果变量集 X_1、.....、 X_m 两两不相交，那么这些事件相互独立，这意味着第二个结论。显然，在这种情况下，系统 (1) 的一致概率最大。

■

我们说明变量子集的大小如何影响不一致概率 $p(\mathcal{X})$。的收敛。

令 $\mathcal{X}' = (X, \{X_1, X_2, \ldots, X_m\})$，其中 $|X_i| = k$，并令 $\mathcal{X}'' = (X, \{X_1, X_2, \ldots, X_m\})$，其中 $X_i = X$，对所有 i. 均成立。命题 2 意味着 $p(\mathcal{X}') \geq 1 - \left(1 - 2^{-2^k}\right)^m$，因此，随着 m 增大且 k 固定， $p(\mathcal{X}') \to 1$。

对于 $\mathcal{X}''$，根据命题 1，固定赋值 $X = a$ 是 (1) 的解的概率为 $1/2^m$。因此， a 不是解的概率为 $1 - 1/2^m$。总之，由于 a 可以取 2^n 个可能的值，我们有 $p(\mathcal{X}'') = \left(1 - \frac{1}{2^m}\right)^{2^n}$。

因此，只有当 m 显著大于 n 时，概率 $p(\mathcal{X}'')$ 才接近 1。相反，随着 m 增大， $p(\mathcal{X}') \to 1$ 。我们注意到，在估计 $p(\mathcal{X}')$ 时，并未考虑超图 $\mathcal{X}'$ 的结构性质。

2.2 含2 个和3 个方程的系统的不一致概率

命题 3 给出了在 $m = 2$ 时计算 $p(\mathcal{X})$ 的一个便捷公式。在命题 4 中，我们证明， $m = 3$ 的情形可归结为计算随机三部图中不含 3-环的概率。

命题 3 设 $\mathcal{X} = (X, \{X_1, X_2\})$，且设 (1) 是一个随机生成的方程组。则

$$p(\mathcal{X}) = \left(1 - (1 - \frac{1}{2^{2^{|X_1 \setminus X_2|}}})(1 - \frac{1}{2^{2^{|X_2 \setminus X_1|}}})\right)^{2^{|X_1 \cap X_2|}}.$$

证明。对于 $X_1 \cap X_2$ 中变量的任意固定赋值，方程 $f_1(X_1) = 0$ 和 $f_2(X_2) = 0$中涉及的其余变量彼此不相交。在这样的固定赋值下， f_1 和 f_2 成立的事件相互独立。因此，该系统不一致（即至少有一个方程不成立）的概率为 $1 - (1 - \frac{1}{2^{2^{|X_1 \setminus X_2|}}})(1 - \frac{1}{2^{2^{|X_2 \setminus X_1|}}})$。由于对 $X_1 \cap X_2$ 的赋值相互独立，这就证明了所述结论。■

令 $\mathcal{X} = (X, \{X_1, X_2, X_3\})$，且设

$$f_1(X_1) = 0, f_2(X_2) = 0, f_3(X_3) = 0 \tag{5}$$

是一个由 $m = 3$ 个方程组成的系统。记 $\{i, j, k\} = \{1, 2, 3\}$。对于变量 $X_i \cap X_j \cap X_k$ 的每一种固定方式（由常数给定），我们将定义一个三部图 G，其顶点集为 Y_{ij}、 $\{i, j\} \subseteq \{1, 2, 3\}$，定义如下。令

$$X_{ij} = (X_i \cap X_j) \setminus (X_i \cap X_j \cap X_k).$$

顶点 Y_{ij}、都是二进制 $|X_{ij}|$串，因此其取值为 X_{ij}。如果$X_{ij} = \emptyset$，那么 Y_{ij} 只包含一个顶点。因此， $|Y_{ij}| = 2^{|X_{ij}|}$ 无论如何都成立。假设方程组 (5) 有一个解。将该解投影到变量 $X_{12} \cup X_{13}$ 上，记为 (a、 b)。类似地，将该解投影到变量 $X_{13} \cup X_{23}$ 上，记为 (b、 c)；投影到变量 $X_{23} \cup X_{12}$ 上，记为 (c 、 a)。因此，(a 、 b) 是图 G 中连接 Y_{12} 和 Y_{13} 的一条边，(b、 c) 是连接 Y_{12} 和 Y_{23} 的一条边，而 (c、 a) 是连接Y_{13} 和 Y_{23} 的一条边。这些边在该图中构成一个长度为 3 的环。显然，系统 (5) 有解，当且仅当图 G 中存在一个长度为 3 的环。

当系统 (5) 随机生成时， G是一个随机生成的三部图，其中 Y_{ik} 与 Y_{kj} 之间存在一条边的概率为 $q_k = 1 - 2^{-2^{|X_k \setminus (X_i \cup X_j)|}}$。令 q 表示以这种方式生成的图 G 不含长度为 3 的环的概率。因此，

q 是系统 (5) 在变量 $X_i \cap X_j \cap X_k$固定后不一致的概率。概率 q 不取决于固定本身。由于对 $X_1 \cap X_2 \cap X_3$ 的赋值彼此独立，我们有

命题 4 设 $\mathcal{X} = (X, \{X_1, X_2, X_3\})$ 和 (1) 是随机生成的方程组。那么

$$p(\mathcal{X}) = q^{2^{|X_1 \cap X_2 \cap X_3|}}$$

估计 q 似乎是一个困难的问题。

2.3 主要结果

根据命题 2，对于每个超图 $\mathcal{X} = (X, \{X_1, X_2, \ldots, X_m\})$，都有 $p(\mathcal{X}) \geq 1 - \prod_{i=1}^{m}(1 - 2^{-2^{|X_i|}})$。因此，随着 m 增大，$p(\mathcal{X}) \to 1$。该界将各集合独立处理，没有体现它们如何相交或重叠；当交结构非平凡时，这可能导致次优估计。一般而言，$p(\mathcal{X})$ 趋于 1 的速度要快得多。在本小节中，我们将针对一般超图 $\mathcal{X} = (X, \{X_1, X_2, \ldots, X_m\})$, 以及完全 k 均匀超图 $\mathcal{X}_k$ 证明这一事实。

这两个界的证明都基于容斥原理。令 S 表示方程 (1) 的所有解的集合，并令 $q(\mathcal{X})$ 表示至少存在一个解的概率。因此，$q(\mathcal{X}) = 1 - p(\mathcal{X}) = \mathbf{Pr}(|S| > 0)$。在我们的设定中，容斥原理可表述如下：

定理 5

$$q(\mathcal{X}) = \sum_{a} \mathbf{Pr}(a \in S) - \sum_{a,b} \mathbf{Pr}(a, b \in S) + \sum_{a,b,c} \mathbf{Pr}(a, b, c \in S) - \ldots,$$

其中，第一项求和遍历所有 n 位字符串 a，第二项求和遍历所有不同的 n 位字符串对 a, b，第三项求和遍历所有不同的 n 位字符串三元组 a, b, c，依此类推。特别地，

$$\sum_{a} \mathbf{Pr}(a \in S) - \sum_{a,b} \mathbf{Pr}(a, b \in S) \leq q(\mathcal{X}) \leq \sum_{a} \mathbf{Pr}(a \in S). \quad (6)$$

定理 6 对于每个超图 $\mathcal{X} = (X, \{X_1, X_2, \ldots, X_m\})$，均有

$$\frac{3 \cdot 2^n}{2^{m+1}} - \frac{2^{2n}}{2^{2m+1}} \sum_{\{i_1, .., i_r\}} 2^{-|X_{i_1} \cup .. \cup X_{i_r}|} \leq q(\mathcal{X}) \leq \frac{2^n}{2^m}.$$

其中求和遍历所有子集 $\{i_1, .., i_r\} \subseteq \{1, .., m\}$。

注记 1. 该定理意味着 $p(\mathcal{X}) \geq 1-\frac{2^n}{2^m}$，而命题 2 给出了界 $p(\mathcal{X}) \geq 1-\prod_{i=1}^{m}(1-2^{-2^{|X_i|}})$。对于 $m > {}^{n}\log_2(3/2)$，我们有

$$2^n < \prod_{i=1}^{m} 2(1-2^{-2}) < \prod_{i=1}^{m} 2(1-2^{-2^{|X_i|}}).$$

因此，$1-\frac{2^n}{2^m} > 1-\prod_{i=1}^{m}(1-2^{-2^{|X_i|}})$ 和 $p(\mathcal{X}) \to 1$ 的速率快于根据命题 2 可能推导出的速率。

注记 2。 在一般情况下，很难用 n 和 m 来估计 $\sum_{\{i_1,..,i_r\}} 2^{-|X_{i_1}\cup..\cup X_{i_r}|}$，因此也难以评估 $q(\mathcal{X})$ 的下界与上界有多接近。

定理 6 的证明。 根据命题 1，$\mathbf{Pr}(a \in S) = \frac{1}{2^m}$，因此 $\sum_a \mathbf{Pr}(a \in S) = \frac{2^n}{2^m}$，其中求和遍历所有 n 位字符串 a。

现在我们求出 $\sum_{[a,b]} \mathbf{Pr}(a,b \in S)$ 的表达式，其中 $[a,b]$ 表示有序对 a, b，包括 $a = b$。根据恒等式 $\sum_{[a,b]} \mathbf{Pr}(a,b \in S) = 2\sum_{a,b} \mathbf{Pr}(a,b \in S) + \sum_a \mathbf{Pr}(a \in S)$，可得 $\sum_{a,b} \mathbf{Pr}(a,b \in S)$ 的值，因此定理由下面的引理得证。

引理 7

$$\sum_{[a,b]} \mathbf{Pr}(a,b \in S) = \frac{2^{2n}}{2^{2m}} \sum_{\{i_1,..,i_r\}} 2^{-|X_{i_1}\cup..\cup X_{i_r}|},$$

其中求和遍历所有子集 $\{i_1,..,i_r\} \subseteq \{1,..,m\}$。

设 T 是一个变量集合，是 X 的子集，并令 m_T 表示 $X_i \subseteq T$ 的数量。那么

$$\begin{aligned}\sum_{[a,b]} \mathbf{Pr}(a,b \in S) &= \sum_T \sum_{a=b(T)} \mathbf{Pr}(a,b \in S) \\ &= \sum_T \frac{2^n}{2^{m_T}\, 4^{m-m_T}} = \frac{2^n}{4^m} \sum_T 2^{m_T},\end{aligned}$$

其中 $a = b(T)$ 表示那些在 T 上完全一致的 a, b。第二个等式成立，因为 $\mathbf{Pr}(a,\ b \in S) = \frac{1}{2^{m_T}\, 4^{m-m_T}}$，如果 $a = b(T)$。现在

$$\sum_T 2^{m_T} = \sum_{s=0}^{m} 2^s \sum_{\{j_1,..,j_s\}} C_{j_1,..,j_s},$$

其中$C_{j_1,..,j_s}$ 是恰好包含 s个变量集 $X_{j_1},..,X_{j_s}$的子集 T的数量。根据容斥原理，

$$C_{j_1,..,j_s} = \sum_{\{i_1,..,i_t\}} (-1)^t 2^{n-|X_{j_1}\cup..\cup X_{j_s}\cup X_{i_1}\cup..\cup X_{i_t}|},$$

其中求和遍历所有子集 $\{i_1,..,i_t\} \subseteq \{1,..,m\} \setminus \{j_1,..,j_s\}$。记 $\{i_1,..,i_t,j_1,..,j_s\} = \{i_1,..,i_r\}$，得到

$$\sum_T 2^{m_T} = \sum_{\{i_1,..,i_r\}} 2^{n-|X_{i_1}\cup..\cup X_{i_r}|} \sum_{s=0}^{r} 2^s \binom{r}{s} (-1)^{r-s} = \sum_{\{i_1,..,i_r\}} 2^{n-|X_{i_1}\cup..\cup X_{i_r}|}.$$

因此，

$$\sum_{[a,b]} \mathbf{Pr}(a,b \in S) = \frac{2^{2n}}{2^{2m}} \sum_{\{i_1,..,i_r\}} 2^{-|X_{i_1}\cup..\cup X_{i_r}|}.$$

其中，求和遍及所有子集 $\{i_1,..,i_r\} \subseteq \{1,..,m\}$。引理得证。

我们有

$$\begin{aligned}
&\sum_a \mathbf{Pr}(a \in S) - \sum_{a,b} \mathbf{Pr}(a,b \in S) \\
&= \sum_a \mathbf{Pr}(a \in S) - \frac{1}{2}(\sum_{[a,b]} \mathbf{Pr}(a,b \in S) - \sum_a \mathbf{Pr}(a \in S)) \\
&= \frac{3}{2}\sum_a \mathbf{Pr}(a \in S) - \frac{1}{2}\sum_{[a,b]} \mathbf{Pr}(a,b \in S).
\end{aligned}$$

结合式 (6) 和引理 7，得到

$$\frac{3\cdot 2^n}{2^{m+1}} - \frac{2^{2n}}{2^{2m+1}} \sum_{i_1,..,i_r} 2^{-|X_{i_1}\cup..\cup X_{i_r}|} \le q(\mathcal{X}) \le \frac{2^n}{2^m}.$$

定理 6 的证明至此完成。 ■

3 完全 k-超图的不一致概率

设 $\mathcal{X}_k = (X, \{X_1, X_2, \ldots, X_m\})$ 为完全 k-均匀超图，其中 $m = \binom{n}{k}$ 和 X_i 是 X 的所有不同的 k-子集。在这种情况下，紧渐近界成立。

Theorem 8 *Let $k \ge 2$. Then $p(\mathcal{X}_k) = 1 - \frac{2^n}{2^{\binom{n}{k}}} + O\left(\frac{n2^n}{2^{\binom{n}{k}+\binom{n-1}{k-1}}}\right)$.*

证明。该定理的证明同样基于不等式 (6)。$\mathcal{X}_k$ 中的边（即方程）总数为 $\binom{n}{k}$. 根据命题 1，我们有 $\mathbf{Pr}(a \in S) = \frac{1}{2^{\binom{n}{k}}}$，因此

$$\sum_a \mathbf{Pr}(a \in S) = \frac{2^n}{2^{\binom{n}{k}}}. \tag{7}$$

为了计算 $\sum_{a,\ b}\mathbf{Pr}(a,\ \ b\in S)$，我们将所有不同的 n 位字符串的无序对集合 $a,\ b$ 划分为 n 个类。若一对 $\{a,b\}$ 满足 a 和 b 恰好在 $0\le t\le n-1$ 个位置上相同，则它属于类 C_t。每个类的大小为 $|C_t|=\binom{n}{t}2^{n-1}$.

令 $\{a,b\}\in C_t$。那么，数对 $\{a,b\}$ 在变量集 X_i上的大小为 1 的投影有 $\binom{t}{k}$ 个，大小为 2 的投影有 $\binom{n}{k}-\binom{t}{k}$ 个。因此，

$$\mathbf{Pr}(a,b\in S)=\frac{1}{2^{\binom{t}{k}+2\left(\binom{n}{k}-\binom{t}{k}\right)}}=\frac{1}{2^{-\binom{t}{k}+2\binom{n}{k}}}.$$

Then

$$\begin{aligned}\sum_{a,b}\mathbf{Pr}(a,b\in S)&=\sum_{t=0}^{n-1}\sum_{\{a,b\}\in C_t}\frac{1}{2^{-\binom{t}{k}+2\binom{n}{k}}}\\&=\sum_{t=0}^{n-1}\frac{\binom{n}{t}2^{n-1}}{2^{-\binom{t}{k}+2\binom{n}{k}}}=\frac{2^n}{2^{1+2\binom{n}{k}}}\sum_{t=0}^{n-1}\binom{n}{t}2^{\binom{t}{k}}\\&=O\left(\frac{n2^n}{2^{\binom{n}{k}+\binom{n-1}{k-1}}}\right).\end{aligned}$$

最后一个估计由 $\sum_{t=0}^{n-1}\binom{n}{t}2^{\binom{t}{k}}=O(n2^{\binom{n-1}{k}})$ 得出，因为 $k\ge 2$ 固定且 n趋于无穷大。为完成定理的证明，现在只需将 (7) 中给出的 $\sum_{a,b}\mathbf{Pr}(a,b\in S)$ 的上界以及 $\sum_a\mathbf{Pr}(a\in S)$ 的值代入 (6) 中的表达式即可。

■

4 2-稀疏方程组

本节考虑变量集 $X=\{x_1,...,x_n\}$ 上随机生成的 2-稀疏方程组 (1)。下文将该系统记为 $\mathcal{S}$。对应的超图是 2-均匀的，因此对应于一个简单图 $G=G(V,E)$，其顶点为 n V，边为 m E。顶点 $V=\{v_1,...,v_n\}$ 对应于变量X，每条边 v_iv_j 表示一个涉及活变量x_i 和 x_j.的方程

在本节中，我们将研究对于某些特定图 G，随机生成的系统 $q(G)$可解的概率。

在 $GF(2)$ 上有 2^4 个二元多项式函数，因此图 G 可以表示 2^{4m} 个不同的方程组。因此，$q(G)=\frac{N}{2^{4m}}$，其中 N 是对应于 $G.$ 的可解方程组数量

我们刻画了具有极端一致性概率的树，以及使一致性概率最大化的森林。此外，我们给出了路径和

星形树的一致性概率的显式公式。最后，我们猜想了环的一致性概率的显式公式，并给出了支持这一猜想的证据。

4.1 预备知识

我们首先引入一些记号并陈述若干预备结果。设 G 是与方程组 $\mathcal{S}$ 相关联的图。对于 G 中的任意顶点 v，我们令 $v^0, v^1,$ 和 v^{01} 分别表示在所有解中变量 x_v 等于 0 的概率、在所有解中变量 x_v 等于 1 的概率，以及存在一个解使得 $x_v = 0$，且存在 $\mathcal{S}$ 的另一个解使得 $x_v = 1$ 的概率。显然， $q(G) = v^0 + v^1 + v^{01}.$

引理 9 设 G 是一个图。那么对于 G 中的每个顶点 v，

$$v^0 = v^1$$

因此

$$q(G) = 2v^0 + v^{01}.$$

证明。为简便起见，以下不失一般性地假设顶点 v 对应于变量 $x_1.$。为了证明 $v^0 = v^1,$，我们考虑映射 h_v：$\mathcal{F} \to \mathcal{F},$，其中 $\mathcal{F}$ 是 n 个变量上的、系数取自 $GF(2)$ 的所有多项式构成的集合，满足 $h_v(f) = f_v,$，其中，$(c_1 + 1, c_2, ..., c_n)$ 是 f_v 的一个根，当且仅当 $(c_1, c_2, ..., c_n)$ 是 f 的一个根（$c_1 + 1$ 在 $GF(2)$ 中进行）。显然， h_v 定义良好。

现在我们考虑由 $h_v.$ 诱导的映射 H_v。对于与 G 对应的可解方程组 $\mathcal{S}$，我们设 $H_v(\mathcal{S}) = \mathcal{S}_v,$ 其中 $\mathcal{S}_v = \{f_v, f \in \mathcal{S}\}$。显然， $\mathcal{S}_v$ 可解当且仅当 $\mathcal{S}$ 可解；并且如果对于 $\mathcal{S}$ 的所有解都有 $x_v = 0,$ 那么对于 $\mathcal{S}_v$ 的所有解，我们有 $x_v = 1.$。由此可得 $v^0 = v^1$。

■

一个众所周知且在下文中反复使用的结果指出：

断言 10 设 $f(x、y) = 0$ 为一个随机选取的二元方程。那么 $v_x^0 = \frac{3}{16}$、$v_x^{01} = \frac{9}{16}.$ — —

设 $G = G(V, E)$ 是一个图，且其边 $vw \notin E$。那么我们用 $G + vw$ 表示由 G 添加这条边后得到的图。

引理 11 设 v 是图 G 的一个顶点，且 w 是一个不属于 $G.$ 的顶点，那么，

$$(a)\ w_{G+vw}^0 = \frac{1}{2}v_G^0 + \frac{3}{16}v_G^{01},\ and\ w_{G+vw}^{01} = \frac{1}{2}v_G^0 + \frac{9}{16}v_G^{01};$$

$$(b)\ v_{G+vw}^0 = \frac{3}{4}v_G^0 + \frac{3}{16}v_G^{01},\ and\ v_{G+vw}^{01} = \frac{9}{16}v_G^{01};$$

and

$$(c)\ q(G + vw) = \frac{3}{4}q(G) + \frac{3}{16}v_G^{01}.$$

证明。 所有关于两个变量 x_v, x_w 的布尔函数构成的集合 $\mathcal{F}_{vw}$ 包含 16 个函数。对于每个与 G 相对应的可解方程组 $\mathcal{S}$，我们考虑方程组 $\mathcal{S} \cup \{f = 0\}$, 其中 $f \in \mathcal{F}_{vw}$. 。因此，$G + vw$ 中的每个方程组都是通过向 G. 中的一个方程组添加这样一个方程构造而成的。如果 $\mathcal{S}$ 是一个可解方程组，且 $x_v = 0$ 在所有解中都成立，那么在 16 个多项式 $f \in \mathcal{F}_{vw}$, 中，恰有 4 个会使扩展后的方程组在所有解中都强制满足 $x_w = 0$ 。当 $x_v = 1$ 在所有解中都成立时，情况也相同。如果 $\mathcal{S}$ 允许 x_v 取两个值（即计入 v_G^{01}），那么在恰好 16 个函数中的 3 个 f 下，所得方程组会在所有解中强制满足 $x_w = 0$ 。因此，综合起来,

$$w^0_{G+vw} = \frac{1}{4}v^0_G + \frac{1}{4}v^1_G + \frac{3}{16}v^{01}_G = \frac{1}{2}v^0_G + \frac{3}{16}v^{01}_G.$$

同理,

$$w^{01}_{G+vw} = \frac{1}{4}v^0_G + \frac{1}{4}v^1_G + \frac{9}{16}v^{01}_G = \frac{1}{2}v^0_G + \frac{9}{16}v^{01}_G.$$

这证明了 (a) 部分。至于 (c) 部分,

$$q(G + vw) = 2w^0 + w^{01} =^{\text{by (a)}} \frac{1}{16}(24v^0_G + 15v^{01}_G) =$$

$$\frac{1}{16}(12(2v^0_G + v^{01}_G) + 3v^{01}_G) =^{\text{by Lemma 9}} \frac{3}{4}q(G) + \frac{3}{16}v^{01}_G.$$

对于 (b) 部分，恰有 9 个函数 f 保持扩展系统中 x_v 的值不变；也就是说，如果在 $\mathcal{S}$. 中 $x_v = 0$ 和 $x_v = 1$ 都可以出现在某些解中，那么在 $\mathcal{S} \cup \{f = 0\}$ 的某些解中也允许它们同时出现。因此，$v^{01}_{G+vw} = \frac{9}{16}v^{01}_G$. 利用恒等式 —

$$q(G + vw) = 2v^0_{G+vw} + v^{01}_{G+vw}$$

并对 v^0_{G+vw} 求解，得到

$$\begin{aligned} 2v^0_{G+vw} &= q(G + vw) - v^{01}_{G+vw} =^{\text{by (c)}} (\frac{3}{4}q(G) + \frac{3}{16}v^{01}_G) - \frac{9}{16}v^{01}_G \\ &= \frac{3}{4}q(G) - \frac{3}{8}v^{01}_G. \end{aligned}$$

因此,

$$v^0_{G+vw} = \frac{3}{8}q(G) - \frac{3}{16}v^{01}_G = \frac{3}{8}(2v^0_G + v^{01}_G) - \frac{3}{16}v^{01}_G = \frac{3}{4}v^0_G + \frac{3}{16}v^{01}_G.$$

这证明了 (b) 部分。证明完毕。

■

对于每个 $n \geq 1$, 令 P_n 为顶点集为 $\{v_0, ..., v_n\}$，且边集为 $\{v_{i-1}v_i, i = 1, ..., n\}$. 的路径。另外，令 S_n 为顶点集为 $\{w_0, ..., w_n\}$，且边集为 $\{w_{i-1}w_n, i = 1, ..., n\}$. 的星形树。因此， w_n 是星形树 S_n. 的中心。

引理 12 令 $\{a_n\}_1^\infty$ 和 $\{b_n\}_1^\infty$ 为递归定义的数列：

$$a_1 = \frac{3}{16},\ b_1 = \frac{9}{16},\ \ and\ for\ n > 1$$
$$a_n = \frac{1}{2}a_{n-1} + \frac{3}{16}b_{n-1},$$
$$b_n = \frac{1}{2}a_{n-1} + \frac{9}{16}b_{n-1}.$$

于是 $v_n^0 = a_n$、$v_n^{01} = b_n$、$q\ (P_n)\ = 2a_n + b_n$，且 $b_n \geq \left(\frac{9}{16}\right)^n$。此外，$w_n^{01} = \left(\frac{9}{16}\right)^n$，且 $q\ (S_n)\ = 2\left(\frac{3}{4}\right)^n - \left(\frac{3}{4}\right)^{2n}$.

证明。 我们通过对 $n.$ 归纳来证明所有结论。对于 $n = 1$，根据断言 10，$v_1^0 = \frac{3}{16}$，$v_1^{01} = \frac{9}{16}$，也就是说，根据引理 9，$q(P_1) = 2v_1^0 + v_1^{01} = 2a_1 + b_1$. 对于 $n > 1$，根据引理 11 (c)，

$$q(P_n) = \frac{3}{4}q(P_{n-1}) + \frac{3}{16}b_{n-1} = \frac{3}{4}(2a_{n-1} + b_{n-1}) + \frac{3}{16}b_{n-1} =$$
$$\frac{3}{2}a_{n-1} + \frac{15}{16}b_{n-1} = 2(\frac{1}{2}a_{n-1} + \frac{3}{16}b_{n-1}) + (\frac{1}{2}a_{n-1} + \frac{9}{16}b_{n-1}) = 2a_n + b_n,$$

并且根据 (a) 部分，$v_n^0 = a_n,\ \ v_n^{01} = b_n$。

对于 $n = 1$，$b_1 = \frac{9}{16}$，并根据引理 11 (a) 和归纳假设 $b_n = v_{P_n}^{01} = \frac{1}{2}v_{P_{n-1}}^0 + \frac{9}{16}v_{P_{n-1}}^{01} > \frac{9}{16}b_{n-1} = \left(\frac{9}{16}\right)^n$.

至于星形树 S_n，对于 $n = 1$，我们有 $q(S_1) = q(P_1) = \frac{15}{16} = 2(\frac{3}{4})^1 - (\frac{3}{4})^2$，并且，$w_1^{01} = \frac{9}{16}$. 对于 $n > 1$，根据引理 11 (c)，

$$\begin{aligned} q(S_n) &= \frac{3}{4}q(S_{n-1}) + \frac{3}{16}w_{n-1}^{01} \\ &= \frac{3}{4}\left(2\left(\frac{3}{4}\right)^{n-1} - \left(\frac{3}{4}\right)^{2n-2}\right) + \frac{3}{16}\left(\frac{9}{16}\right)^{n-1} \\ &= \frac{3}{4}2\left(\frac{3}{4}\right)^{n-1} - \frac{3}{4}\left(\frac{3}{4}\right)^{2n-2} + \frac{3}{16}\left(\frac{3}{4}\right)^{2n-2} \\ &= 2\left(\frac{3}{4}\right)^{n} - \left(\frac{3}{4}\right)^{2n}. \end{aligned}$$

最后，根据 (b) 部分，$w_n^{01} = \frac{9}{16}w_{n-1}^{01} = \left(\frac{9}{16}\right)^n$. ■

引理 13 对于每个 s、$t \geq 1$，均有

$$b_s b_t < b_{s+t}.$$

证明。 固定 $s \geq 1$ 。我们对 $t.$使用归纳法来证明该命题。为了进行归纳步骤，我们证明一个更强的命题

$$b_s b_t < b_{s+t} \quad \text{and} \quad b_s a_t < a_{s+t}.$$

对于 $t = 1$，我们得到

$$b_s b_1 = b_s \frac{9}{16} < \frac{1}{2} a_s + \frac{9}{16} b_s = b_{s+1},$$
$$b_s a_1 = b_s \frac{3}{16} < \frac{1}{2} a_s + \frac{3}{16} b_s = a_{s+1}$$

因为对于所有 $s \geq 1.$，$a_s > 0$ 。假设对于所有 $k < t$，$b_s b_k < b_{s+k}$ 和 $b_s a_k < a_{s+k}$ 已得到证明。那么

$$b_s b_t = b_s(\frac{1}{2} a_{t-1} + \frac{9}{16} b_{t-1}) < \frac{1}{2} a_{s+t-1} + \frac{9}{16} b_{s+t-1} = b_{s+t},$$
$$b_s a_t = b_s(\frac{1}{2} a_{t-1} + \frac{3}{16} b_{t-1}) < \frac{1}{2} a_{s+t-1} + \frac{3}{16} b_{s+t-1} = a_{s+t}.$$

■

4.2 具有极端一致性概率的树

下一个定理刻画了能够达到极端一致性概率的树。

定理 14 对于任意一棵具有 $n \geq 3$ 条边、既不是路径也不是星形树的树 T_n，它满足

$$q(S_n) < q(T_n) < q(P_n).$$

证明。 我们对边的数量进行归纳法来证明该定理。在归纳步骤中，证明一个更强的命题会更方便：对于任意一棵具有 $n \geq 3$条边、既不是路径也不是星形树的树 T_n，它满足$q(S_n) < q(T_n) < q(P_n)$；此外，对于每棵具有 $n \geq 1$ 条边的树（包括星形树和路径）的每个顶点 v，都有

$$\left(\frac{9}{16}\right)^n \leq v^{01} \leq b_n. \tag{8}$$

预备知识与引理中的值。设 P_n 为顶点集为 $\{v_0, ..., v_n\}$，且边集为 $\{v_{i-1}v_i, i = 1, ..., n\}$，的路径，并设 S_n 为中心为 $w_n.$的星形树。我们回顾一下，根据引理 12,

$$q(S_n) = 2\left(\frac{3}{4}\right)^n - \left(\frac{3}{4}\right)^{2n}, w_n^{01} = (\frac{9}{16})^n,$$
$$q(P_n) = 2a_n + b_n, \text{ and } b_n \geq \left(\frac{9}{16}\right)^n.$$

根据递推关系（见引理 12），我们得到具体值 $a_2 = \frac{51}{16^2}$，$b_2 = \frac{105}{16^2}$，$a_3 = \frac{723}{16^3}$，$b_3 = \frac{1353}{16^3}$.

基础情形 $n = 3$. 当 $n = 3$ 时，不存在既不是路径也不是星形树的树，因此陈述中的不等式归结为证明$q(S_3) < q(P_3)$。

利用上述数字

$$q(S_3) = \frac{2727}{16^3} < 2\frac{723}{16^3} + \frac{1353}{16^3} = \frac{2799}{16^3} = q(P_3)$$

如所需。

关于 (8)，需要证明它对所有具有 $n = 1$，2，和 3 条边的树 T 都成立。若 $n = 1$ $T = P_1$ $v_0^{01} = v_1^{01} = b_1 = \frac{9}{16}$. $n = 2$ $T =$，则，且 P_2 $v_0^{01} = v_2^{01} = b_2 \geq \frac{9}{16}^2$. v_1 对于，，以及 ( ) 该顶点可视为星形树 S_2的中心 w_2，因此 $w_2^{01} = (\frac{9}{16})^2$. 对于 $n = 3$, $T = S_3$或 $T = P_3$.由对称性，只需检验 P_3，中的顶点 v_2 和 v_3，以及 S_3.的中心和一个悬挂顶点

对于 v_3，它成立

$$v_3^{01} = b_3 \geq \left(\frac{9}{16}\right)^3$$

由引理 11 可得 $v_2^{01} =^{(b)} \frac{9}{16} b_2 <^{(a)} \frac{1}{2} a_2 + \frac{9}{16} b_2 = b_3$. 此外 $\frac{9}{16} b_2 = \frac{9}{16} \frac{105}{16^2} > \left(\frac{9}{16}\right)^3$，因此 v_2^{01}也满足 (8)。对于 w_3， S_3 的中心，由引理 12

$$w_3^{01} = \left(\frac{9}{16}\right)^3 < \frac{1353}{16^3} = b_3.$$

最后，将证明 (8) 对 S_3 的一个悬挂顶点 w 成立：将 S_3 视为 $P_2 + v_1 w$，其中 v_1 是 P_2. 的中间顶点。由 P_2 中的引理 11 可知，它是 $v_1^0 = \frac{63}{16^2}$ $v_1^{01} = \frac{81}{16^2}$. $w_{P_2+v_1w}^{01} = \frac{1}{2} v_1^0 + \frac{9}{16} v_1^{01} = \left(\frac{9}{16}\right)^3 + \frac{8 \cdot 63}{16^3} = \frac{1233}{16^3} <$，于是，当.....时，( ) 对所有 $\frac{1353}{16^3} = b_3$. $n = 3$. 顶点都成立

归纳步骤。现在设 T 为一棵有 $n > 3$ 条边、既不是路径也不是星形树的树，并且 vw 为一条悬挂边，满足 $\deg_T(w) = 1$，且 $T - vw$不是一条路径。根据归纳假设，有 $q(T - vw) < q(P_{n-1})$ $v_{T-vw}^{01} \leq b_{n-1}$，于是得到

$$\begin{aligned} q(T) = q((T - vw) + vw) &=^{\text{Lemma 11(c)}} \frac{3}{4} q(T - vw) + \frac{3}{16} v_{T-vw}^{01} \\ &< \frac{3}{4} q(P_{n-1}) + \frac{3}{16} b_{n-1} = \frac{3}{4}(2a_{n-1} + b_{n-1}) + \frac{3}{16} b_{n-1} \\ &= 2(\frac{1}{2} a_{n-1} + \frac{3}{16} b_{n-1}) + (\frac{1}{2} a_{n-1} + \frac{9}{16} b_{n-1}) = 2a_n + b_n = q(P_n). \end{aligned}$$

对于左侧不等式，存在一条悬挂边 vw，其位于 T 中，且 $\deg_T(w) = 1$, 使得 $T - vw$ 不是星形树。根据归纳假设， $q(T - vw) >$

$q(S_{n-1})$，并且 $v^{01}_{T-vw} \geq \left(\frac{9}{16}\right)^{n-1}$.。将其与引理 11 结合，可得

$$\begin{aligned} q(T) &= \frac{3}{4}q(T-vw) + \frac{3}{16}v^{01}_{T-vw} \\ &> \frac{3}{4}\left(2\left(\frac{3}{4}\right)^{n-1} - \left(\frac{9}{16}\right)^{n-1}\right) + \frac{3}{16}\left(\frac{9}{16}\right)^{n-1} \\ &= 2\left(\frac{3}{4}\right)^{n} - \left(\frac{9}{16}\right)^{n} = q(S_n). \end{aligned}$$

现在我们来证明 (8)。首先，令 $\deg_T(w) = 1$，并设 v 是与 w 相邻的 T 的一个顶点。反设 $b_n < w^{01}$. 成立。于是 $2w^0+w^{01} = q(T) < q(P_n) = 2a_n + b_n$，这意味着 $w^0 < a_n$.。进而得到

$$\begin{aligned} b_n < w^{01} &\Longrightarrow \frac{1}{2}a_{n-1} + \frac{9}{16}b_{n-1} < \frac{1}{2}v^0_{T-vw} + \frac{9}{16}v^{01}_{T-vw}, \text{ and} \\ w^0 < a_n &\Longrightarrow \frac{1}{2}v^0_{T-vw} + \frac{3}{16}v^{01}_{T-vw} < \frac{1}{2}a_{n-1} + \frac{3}{16}b_{n-1}. \end{aligned}$$

将上述不等式相加并整理，得到 $b_{n-1} < v^{01}_{T-vw}$，这与归纳假设矛盾。

另一方面，

$$w^{01} = \frac{1}{2}v^0_{T-vw} + \frac{9}{16}v^{01}_{T-vw} > \frac{9}{16}v^{01}_{T-vw} \geq \frac{9}{16}\left(\frac{9}{16}\right)^{n-1} = \left(\frac{9}{16}\right)^{n}$$

根据归纳假设。现在考虑一个满足 $\deg_T(w) > 1$. 的顶点 w。如果 T 是一棵星形树，那么 w 是它的中心，并且 $w^{01} = (\frac{9}{16})^n$.。—另一种情况下，存在一个满足 $\deg_T(z) > 1$，且与 w 相邻的顶点 z。令 T_w 和 T_z 分别为 $T - wz$ 中包含顶点 w 和 z 的子树。通过与引理 11 类似的推理，我们有

$$w^{01}_T = w^{01}_{T_w} z^{01}_{T_z}\frac{9}{16} + w^{01}_{T_w} z^0_{T_z}\frac{1}{2} = w^{01}_{T_w}\left(\frac{1}{2}z^0_{T_z} + \frac{9}{16}z^{01}_{T_z}\right) = w^{01}_{T_w} y^{01}_{T_z+yz},$$

其中 y 是 $T_z + yz$. 中一个待处理的新顶点。根据归纳假设，
$\left(\frac{9}{16}\right)^s \leq w^{01}_{T_w} \leq b_s$ 和 $\left(\frac{9}{16}\right)^{t+1} \leq y^{01}_{T_z+yz} \leq b_{t+1}$，其中 s 和 t 分别是 T_w 和 T_z，中的边数。显然，$s + t = n - 1$. 因此 $w^{01}_T = w^{01}_{T_w} y^{01}_{T_z+yz} \geq \left(\frac{9}{16}\right)^n$，而由引理 13 可得 $w^{01}_T \leq \overline{b_s b_{t+1}} < b_n$。证明完毕。 ■

4.3 路径的概率

在本小节中，我们推导出 $q(P_n)$。的显式公式。

二次方程 $x^2 = \frac{17}{16}x - \frac{3}{16}$ 的根将记为 $s = \frac{17+\sqrt{97}}{32}$、和 $t = \frac{17-\sqrt{97}}{32}$.

引理 15 设 $\{d_n\}_1^\infty$ 是一个递归定义的数列，其定义为 $d_n = \frac{17}{16}d_{n-1} - \frac{3}{16}d_{n-2}$ $n \geq 3$. 其中 $d_1 = a$ $d_2 = b$ 则满足 的特解为

$$d_n = \frac{b - at}{s(s-t)}s^n + \frac{-b + as}{t(s-t)}t^n.$$

证明。 递推关系的通解为

$$d_n = c_1 s^n + c_2 t^n,$$

其中， s 和 t 是特征方程的根。常数c_1 和 c_2 满足以下系统

$$a = c_1 s^1 + c_2 t^1,\ b = c_1 s^2 + c_2 t^2$$

使用标准方法可得

$$c_1 = \frac{b - a \cdot t}{s(s-t)} \text{ and } c_2 = \frac{-b + a \cdot s}{t(s-t)}$$

结论得证。 ■

由引理 12， $q(P_n) = 2a_n + b_n$. 直接计算得出 $q(P_1) = 15\,16$，且 $q(P_2) = \frac{207}{256}$. 现在陈述本小节的主要结果:

定理 16 设 $c = \frac{q_{(P_2)} - q\ P_1 \cdot t\ ()() \cdot}{s\ s-t}$ 那么，对于 $n \geq 1$,

$$q(P_n) = cs^n + (1-c)t^n \approx 1.16 \cdot s^n - 0.16 \cdot t^n,$$

其中 $s = \frac{17+\sqrt{97}}{32}$, $t = \frac{17-\sqrt{97}}{32}$.

证明。 设$\{a_n\}_1^\infty$和$\{b_n\}_1^\infty$为引理 12 中定义的数列。

现在将推导出 $\{b_n\}_1^\infty$的二阶递推关系。

$$\begin{aligned}
\frac{1}{2}a_{n-1} &= b_n - \frac{9}{16}b_{n-1}, \\
a_n &= b_n - \frac{9}{16}b_{n-1} + \frac{3}{16}b_{n-1} = b_n - \frac{6}{16}b_{n-1}, \\
a_{n-1} &= b_{n-1} - \frac{6}{16}b_{n-2}, \\
b_n &= \frac{1}{2}(b_{n-1} - \frac{6}{16}b_{n-2}) + \frac{9}{16}b_{n-1} = \frac{17}{16}b_{n-1} - \frac{3}{16}b_{n-2}.
\end{aligned}$$

因此 $\{b_n\}$满足与引理 15 中相同的递推关系，所以

$$b_n = \frac{b_2 - b_1 \cdot t}{s(s-t)}s^n + \frac{-b_2 + b_1 \cdot s}{t(s-t)}t^n$$

类似地，可以验证 $\{a_n\}_1^\infty$遵循相同的递推关系

$$a_n = \frac{17}{16}a_{n-1} - \frac{3}{16}a_{n-2}$$

因此，再由引理 15,

$$a_n = \frac{a_2 - a_1 \cdot t}{s(s-t)}s^n + \frac{-a_2 + a_1 \cdot s}{t(s-t)}t^n.$$

由引理 12，$q(P_n) = 2a_n + b_n$，并代入 a_n 和 b_n 的表达式，可得

$$\begin{aligned} q(P_n) &= 2a_n + b_n \\ &= \frac{(2a_2 + b_2) - (2a_1 + b_1) \cdot t}{s(s-t)}s^n + \frac{-(2a_2 + b_2) + (2a_1 + b_1) \cdot s}{t(s-t)}t^n \\ &= \frac{q(P_2) - q(P_1) \cdot t}{s(s-t)}s^n + \frac{-q(P_2) + q(P_1) \cdot s}{t(s-t)}t^n. \end{aligned}$$

要完成证明，还需证明

$$\frac{q(P_2) - q(P_1) \cdot t}{s(s-t)} + \frac{-q(P_2) + q(P_1) \cdot s}{t(s-t)} = 1.$$

经简单变形并应用韦达公式 $s + t = \frac{17}{16}$， $st = \frac{3}{16}$，得到

$$\frac{q(P_2) - q(P_1) \cdot t}{s(s-t)} + \frac{-q(P_2) + q(P_1) \cdot s}{t(s-t)} = \frac{-q(P_2) + q(P_1)(s+t)}{st} = \frac{-\frac{207}{256} + \frac{15}{16}\frac{17}{16}}{\frac{3}{16}} = \frac{15 \cdot 17 - 207}{48} = 1.$$

因此，$q(P_n) = cs^n + (1-c)t^n$. 证明至此完成。 ■

4.4 概率最大的无环图

本节中，我们考虑不含孤立顶点的无环图（即森林），因为孤立顶点对应于未参与任何方程的变量。显然，我们有 $n > m$；也就是说，顶点数多于边数。我们刻画使一致性概率最大化的无环图。接下来的两个引理为证明极值图的唯一性提供了关键要素。

引理 17 设 C_1、...、C_t 是图 G. 的连通分量，则

$$q(G) = \prod_{i=1}^{k} q(C_i).$$

证明。 连通分量 C_1、.....、 C_t 对应于变量集合互不相交的方程组。因此，它们的解彼此独立。结论得证。

■

引理18 设 $k \geq 1$，且 $d \geq k+2$. 那么

$$q(P_{k+1}) \cdot q(P_{d-1}) > q(P_d) \cdot q(P_k).$$

证明。 由于 $q(P_n) = cs^n + (1-c)t^n$，我们需要证明

$$\left(cs^{k+1} + (1-c)t^{k+1}\right)\left(cs^{d-1} + (1-c)t^{d-1}\right) > (cs^d + (1-c)t^d)(cs^k + (1-c)t^k).$$

展开各因子并减去对应项后得到

$$c(1-c)(s^{d-1}t^{k+1} + s^{k+1}t^{d-1}) > c(1-c)(s^d t^k + s^k t^d).$$

由于 $c > 1$，故 $c(1-c) < 0$，于是

$$\begin{aligned} s^{d-1}t^{k+1} + s^{k+1}t^{d-1} &< s^d t^k + s^k t^d, \\ s^k t^{d-1}(s-t) &< s^{d-1}t^k(s-t), \\ t^{d-1-k} &< s^{d-1-k}. \end{aligned}$$

由于 $s > t$，且 $d > k+1.$，不等式成立 ■

定理19 设 G 是一个有 n 个顶点和 $m < n$ 条边且没有孤立顶点的无环图。那么，$\max q(G)$ 在某个图上达到，该图的所有连通分量 T_i 都是路径，且这些连通分量所含的顶点数 $n(T_i)$ 尽可能相等；即，

$$|n(T_i) - n(T_j)| \leq 1 \ for \ all \ i, j.$$

证明。 设 G 是具有最大一致性概率的无环图。由于 G 是没有孤立顶点的森林，因此每个连通分量都是至少含有一条边的树。根据定理14，每个连通分量 C_i 都是一条路径；根据引理17， $q(G) = \prod q(C_i)$。如果 G 包含两个连通分量 $_i$ 和 P_d，其中 $d \geq k+2$， 那么用 P_{d-1} 和 P_{k+1} 替换这两条路径所得到的图将具有更高的一致性概率，见引理18。因此，在 G中，任意两个连通分量的阶至多相差 1。

■

4.5 环的一致性概率

通常，含有 n 条边的环记作 C_n. 。我们提出以下猜想

猜想 20 环 C_n 的一致性概率为

$$q(C_n) = s^n + t^n - \frac{1}{4^n} - \frac{1}{8^{n-1}},$$

其中 $s = \frac{17+\sqrt{97}}{32}, t = \frac{17-\sqrt{97}}{32}$.

作为支持性证据，我们使用计算机计算了 $q(C_n)$ 在 $n = 3,...,7,$ 时的取值，在每种情况下，计算结果都与公式给出的值一致。

利益冲突

本稿件为原创，未经发表，目前也未在其他地方考虑发表。所有作者均已批准投稿。所有作者均无须披露的利益冲突。

参考文献

[1] D. Achlioptas。2000年。“每次设定 2 个变量，为随机 3-SAT 得到一个新的下界（扩展摘要）”。第 *32* 届 *ACM* 计算理论研讨会论文集：28–37。[2] G. V. Bard。2009年。代数密码分析。施普林格。[3] V. Chv´atal 和 B. Reed。*1992*年。“Mick 得到了一些机会（胜算站在他这边）”。第 *33* 届计算机科学基础年度研讨会：620–627。IEEE 计算机学会出版社。[4] J. Ding、A. Sly 和 N. Sun。2015年。“大 k 情形下可满足性猜想的证明”，第 *47* 届 *ACM* 计算理论年度研讨会 *STOC 2015* 论文集：59–68。[5] N. E´en，N. S¨orensson。*MiniSat* 主页，http://minisat.se。[6]A. Goerdt。1992年。“不可满足性阈值”。数学基础与计算机科学，第 17 届国际研讨会，I.M. Havel 和 V. Koubek 编，《计算机科学讲义》629：264–274，施普林格。[7] S. Janson、Y.C. Stamatiou 和 M. Vamvakari。2000年。“随机 3-SAT 不可满足性阈值的界定”。随机结构与算法。17：103-116。[8] B.A LaMacchia 和 A.M Odlyzko。1991年。“在有限域上求解大型稀疏线性系统”。载于 A. J. Menezes 和 S. A. Vanstone（编），密码学进展 – *CRYPTO '90* （《计算机科学讲义》，537：109–133）。施普林格，海德堡。

[9]I. Semaev。2009年。“有限域上的稀疏代数方程”。*SIAM*计算杂志，*39(2)*：388–409。[10] I. Semaev。2013年。“改进的一致-粘合算法”。计算机科学数学 7(3)：321–339。[11]G. B. Sorkin。2001年。“关于随机可满足性的一些注记”。载于 SAGA 2001。《计算机科学讲义》，2264：117-130。施普林格。