第一页为封面页

参赛学生姓名：郑朗

中学：北京市鼎石学校

省份：北京

国家/地区：中国

指导老师姓名：Jason Roy

指导老师单位：北京市鼎石学校

论文题目：Fixed points in random cycle factors of d-regular directed graphs

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Fixed points in random cycle factors of d-regular directed graphs

Lang Zheng

August 2026

Abstract

Let G be a directed d-regular graph on n vertices and let σ be a uniformly random cycle factor of G. For such a graph, Christoph, Draganić, Girão, Hurley, Michel, and Müyesser proved that the expected number of cycles in a uniformly random cycle factor is at most $\mathcal{O}((n \log d)/d)$. They proposed the problem of whether or not the expected number of fixed points in a cycle factor of such a graph is at most $\mathcal{O}(n/d)$. This paper gives a self-contained exposition of the entropy method used by Christoph et al. We use the entropy methods to show that the expected number of fixed points is bounded by $\mathcal{O}(n \log d/(d \log \log d))$. This result improves the bound obtained by simply counting every fixed point as a cycle. We extend this bound to cycles of arbitrary length using the same methodology. This strategy allows collectively considering a possible set of fixed points instead of counting individually.

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

1 Introduction

1.1 Background

The cycle structure of a uniformly random permutation is a classical topic in combinatorics and probability. Its origins include de Montmort's study of fixed points in 1708 and Cauchy's formula of permutations given a specific cycle structure. If σ is a uniformly random permutation of n elements, then the expected number of cycles of length k is $1/k$ for $1 \leq k \leq n$. Consequently, its expected number of cycles is

$$H_n = \sum_{k=1}^{n} \frac{1}{k} = \ln(n) + \gamma + \mathcal{O}(\frac{1}{n}).$$

The expected number of fixed points in a random permutation is exactly 1. These two qualities originating from permutations motivates the study of cycles and fixed

points in graphs: how much of these properties survive when the possible images of each vertex are restricted to edges of a graph?

In a directed graph G, v is an out-neighbor of u if the directed edge (u, v) is in G. Similarly, we say that u is an in-neighbor of v. A *d-regular* directed graph is a directed graph where each vertex has in-degree and out-degree exactly d. Loops and directed cycles of length 2 are allowed under this setting, but not multiple edges mapping a vertex to the same out-neighbor. A *cycle factor* of a directed graph G is a spanning directed subgraph that partitions $V(G)$ into disjoint cycles.

For a general d-regular directed graph, the mappings of vertices does not contain the symmetry in completely random permutations. Nevertheless, it is still possible to bound the expected number of cycles with respect to n and d when d is sufficiently large. The counting argument of Vishnoi [Vis12] yields an upper bound of $\mathcal{O}(n/\sqrt{\log d})$ on the expected number of cycles in a uniformly random cycle factor. Applying entropy, Christoph et al. [Chr+25a] improved this to the following bound.

Theorem 1.1 ([Chr+25a]). *For all sufficiently large d, the expected number of cycles in a uniformly random cycle factor of any directed d-regular graph on n vertices is $\mathcal{O}((n \log d)/d)$.*

This bound is asymptotically tight. When $d \mid n$, the disjoint union of n/d copies of K_d° has expected number of $(n/d)H_d = \Theta((n \log d)/d)$ cycles. This establishes the correct order but is not the largest count possible. Indeed, Gajjala [Gaj26] disproved the conjecture of Christoph et al. [Chr+25a, Conjecture 4.1] that this union maximizes the expected number of cycles, constructing counterexamples for every $d \geq 3$ and every multiple $n \geq 2d$ of d.

Entropy provides a way to study such counting problems without listing out all valid permutations. The notion of entropy was introduced in information theory by Shannon [Sha48]. Entropy measures the uncertainty of a random variable. A key development for applying entropy in graph theory was Radhakrishnan's entropy proof of Brégman's permanent bound [Rad97]. Applying entropy and a random reveal order yields a bound on the number of perfect matchings from the chain rule of entropy. The proof of Theorem 1.1 builds on this approach by bounding the entropy loss during the reveal of a cycle factor.

1.2 Fixed points and the main result

A *fixed point* of a cycle factor σ is a vertex i satisfying $\sigma(i) = i$. A fixed point is a cycle of length one. Let

$$S(\sigma) = \{i \in V(G) : \sigma(i) = i\}.$$

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

Only vertices with loops can belong to $S(\sigma)$. In particular, if G has no loops, then all cycle factor contains no fixed points. We define the expected number of fixed points in G by

$$\mathbb{E}[|S(\sigma)|] = \sum_{i \in V(G)} \mathbb{P}(\sigma(i) = i).$$

For K_n°, each vertex has probability $1/n$ of mapping to itself. The entire graph then has expected number of exactly 1 fixed point. Notice that this obeys the bound of n/d since $n = d$ in K_n°. This suggests that the logarithmic factor needed for the expected number of cycles might be unnecessary for fixed points. Accordingly, Christoph et al. [Chr+25a, Problem 4.2] asked whether every directed d-regular graph on n vertices satisfies

$$\mathbb{E}[|S(\sigma)|] = \mathcal{O}(n/d)$$

when σ is chosen uniformly at random from its cycle factors.

Counting each fixed point as a cycle gives the bound $\mathcal{O}((n \log d)/d)$ by Theorem 1.1. Our main contribution is the following improvement toward the proposed $\mathcal{O}(n/d)$ bound.

Theorem 1.2. *For all sufficiently large d, the expected number of fixed points in a uniformly random cycle factor of any directed d-regular graph on n vertices is $\mathcal{O}((n \log d)/(d \log \log d))$.*

Theorem 1.2 improves the bound obtained from Theorem 1.1 by a factor of order $\log \log d$, but leaves a factor of order $\log d / \log \log d$ from the proposed bound.

The proposed bound holds for n/d disjoint copies of K_d°, including the complete graph case when $n = d$. We do not claim that this type of graph is the maximum case for expected number of fixed points. The proposed bound also holds trivially for loopless graphs, and more generally for graphs with at most $\mathcal{O}(n/d)$ looped vertices, since the number of fixed points cannot exceed the number of such vertices.

The entropy method applied to Theorem 1.2 can also be extended to bound the expected number of cycles of arbitrary length. When $r = 1$, this recovers Theorem 1.2. This result shows that improvements obtained for fixed points may be extended to cycles of arbitrary length.

Theorem 1.3. *For all sufficiently large d, the expected number of cycles of length r in a uniformly random cycle factor of any directed d-regular graph on n vertices is $\mathcal{O}((n \log d)/(d \log \log d))$.*

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

1.3 Proof strategy

We give a self-contained account of the entropy argument of Christoph et al. [Chr+25a], with a motivating simplified case and the technical details needed to justify it. Cycle factors correspond to perfect matchings in a regular bipartite graph. Utilizing upper and lower bounds on the number of perfect matchings yield a bound for the entropy deficit during reveal, which in turn bounds the probability of closing a cycle. This yields the optimal order for the expected number of cycles, but does not guarantee a correct constant factor.

For Theorem 1.2, we instead reveal the entire fixed-point set $S(\sigma)$ first. The entropy chain rule divides total uncertainty into the uncertainty in $S(\sigma)$ and the uncertainty in the remaining matching. We bound the former using binary entropy and the latter using Brégman's inequality after deleting the vertices in $S(\sigma)$. Comparing their sum with Schrijver's lower bound on the number of perfect matchings gives an inequality for $\mathbb{E}[|S(\sigma)|]$. The advantage of this method is that it treats fixed points collectively rather than estimating each loop probability separately. Its limitation is that the upper bound counts all remaining perfect matchings, including those with additional fixed points, and ignores the degree losses after deletion. This results in a logarithmic gap from the proposed bound.

The same approach of collectively revealing a certain set of vertices extends to prove Theorem 1.3. Instead of revealing the fixed vertices, we reveal the set of all cycles of a specific length.

1.4 Organization

Section 2 gives an introduction to entropy, provides entropy tools used later, and shows Cauchy's formula for the expected number of cycles in a random permutation. Section 3 presents the proof of Theorem 1.1 and discusses its limitations. Section 4 proves Theorem 1.2. Section 5 explains the modifications in order to achieve the extension to cycles of arbitrary length. Finally, Section 6 discusses the remaining gap to the proposed $\mathcal{O}(n/d)$ bound and possible refinements of the method.

2 Preliminaries

Write $[n] = \{1, 2, \ldots, n\}$. Unless stated otherwise, this paper considers directed graphs.

Definition 2.1 (Shannon Entropy)**.** For a discrete random variable X taking values

2026 S.-T. Yau High School Science Award
仅用于2025江苏省青少年科学作家论文公示

in a finite set $\mathcal{X}$, the *Shannon entropy* of X is defined as

$$\mathcal{H}(X) = -\sum_{x\in\mathcal{X}} \mathbb{P}(X = x) \log \mathbb{P}(X = x).$$

For $p \in [0, 1]$, write binary entropy as

$$\mathcal{H}_{\mathrm{b}}(p) = -p \log p - (1 - p) \log(1 - p)$$

with the convention that $0 \log 0 = 0$.

This definition guarantees that entropy cannot be negative. Intuitively, we may think of entropy as measuring uncertainty of a random variable. Higher uncertainty in the outcome of a random variable corresponds to higher entropy of that random variable. Unless stated otherwise, we use base 2 logarithms. The base does not affect our asymptotic results because changing bases multiplies by a constant factor.

A natural question is how large entropy can be when the number of possible outcomes is fixed. We expect uncertainty to be greatest when the outcomes are uniformly distributed.

Theorem 2.2. *For a discrete random variable X taking values in $\mathcal{X}$, $\mathcal{H}(X) \le \log |\mathcal{X}|$, with equality if and only if X is uniformly distributed on $\mathcal{X}$.*

Proof. Write $p_x = \mathbb{P}(X = x)$ for each $x \in \mathcal{X}$, and let $|\mathcal{X}| = s$. We use the natural-logarithm identity $\log_2 t = (\ln t)/(\ln 2)$ and the inequality $\ln t \le t - 1$, which holds for all $t > 0$.

Consider the difference between the entropy of X and $\log s$:

$$\mathcal{H}(X) - \log s = \sum_{x\in\mathcal{X}} p_x \log \frac{1}{p_x} - \sum_{x\in\mathcal{X}} p_x \log s = \sum_{x\in\mathcal{X}} p_x \log \left(\frac{1}{sp_x}\right).$$

For each x with $p_x > 0$, apply the inequality $\ln t \le t - 1$ with $t = 1/(sp_x)$. Since $\log t = (\ln t)/(\ln 2)$, we obtain

$$\log \left(\frac{1}{sp_x}\right) = \frac{1}{\ln 2} \ln \left(\frac{1}{sp_x}\right) \le \frac{1}{\ln 2} \left(\frac{1}{sp_x} - 1\right).$$

Multiplying by p_x and summing over the support of X, denoted by $\operatorname{supp}(X)$, we obtain

$$\mathcal{H}(X) - \log s \le \frac{1}{\ln 2} \sum_{x\in\operatorname{supp}(X)} p_x \left(\frac{1}{sp_x} - 1\right) = \frac{1}{\ln 2} \left(\frac{|\operatorname{supp}(X)|}{s} - 1\right) \le 0.$$

Hence $\mathcal{H}(X) \le \log s$. If equality holds, then $|\operatorname{supp}(X)| = s$ and equality must hold in $\ln t \le t - 1$ for every x. This occurs when $t = 1$, or equivalently, when $p_x = 1/s$

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公开

for every $x \in \mathcal{X}$. Conversely, under uniform distribution each x has probability $1/s$. Direct substitution into the definition yields entropy $\log s$. □

This theorem is easily understood intuitively knowing that entropy measures uncertainty. With uniform distribution, each event is equally likely to be chosen, so the uncertainty is as high as possible. When the balance is broken, some events are more likely to happen than others, making it easier to predict and lowering entropy.

Similar to probability, entropy may also be applied to conditional distributions:

1. If E is an event,

$$\mathcal{H}(X \mid E) = -\sum_{x \in \mathcal{X}} \mathbb{P}(X = x \mid E) \log \mathbb{P}(X = x \mid E),$$

2. If Y is a random variable taking values in $\mathcal{Y}$,

$$\mathcal{H}(X \mid Y) = \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) \mathcal{H}(X \mid Y = y).$$

   Note that conditioning on another random variable cannot increase entropy. That is, $\mathcal{H}(X|Y) \leq \mathcal{H}(X)$.

3. Moreover, for random variables $X_1, \ldots, X_n$, it follows from 1 and 2 that:

$$\mathcal{H}(X_1, \ldots, X_n) = \sum_{i=1}^{n} \mathcal{H}(X_i \mid X_1, \ldots, X_{i-1}),$$

   where the n-tuple $(X_1, \ldots, X_n)$ is viewed as a single random variable with the joint distribution of $X_1, \ldots, X_n$.

When conditioning X on another random variable Y, the following property holds:

Theorem 2.3. *$\mathcal{H}(X \mid Y) = 0$ if and only if Y determines X.*

Proof. Entropy is always non-negative, so $\mathcal{H}(X \mid Y)$ is a weighted average of non-negative numbers. If this average is zero, then every term with positive weight must be zero. Thus, if $\mathcal{H}(X \mid Y) = 0$, then $\mathcal{H}(X \mid Y = y) = 0$ for every y within the support of Y.

A random variable has entropy zero if and only if it takes a single value with probability 1. Therefore, for every possible outcome y with positive probability, there exists a value x such that

$$\mathbb{P}(X = x \mid Y = y) = 1.$$

2026 S.-T. Yau High School Science Award
仅用于2026日友桐中学科学奖论文公示

Therefore, for each y in the support of Y, there is a unique value $f(y)$ such that $\mathbb{P}(X = f(y) \mid Y = y) = 1$. Hence $X = f(Y)$, indicating that Y determines X.

Conversely, suppose that $X = f(Y)$. For each y in the support of Y, the conditional random variable $X \mid Y = y$ takes the single value $f(y)$ with probability 1. Hence $\mathcal{H}(X \mid Y = y) = 0$, and averaging over y gives $\mathcal{H}(X \mid Y) = 0$.

□

Indeed, this theorem shows entropy behaves exactly as we expect it. Since entropy measures the uncertainty of an event, if Y determines X, then when Y is given, X is fixed. Hence, $\mathcal{H}(X|Y) = 0$.

2.1 Cycles in random permutations

An important fact we use throughout the paper is that cycle factors, perfect matchings in simple bipartite graphs, and permutations carry the same information. A directed cycle factor on G assigns to each vertex i a unique out-neighbor $\sigma(i)$, giving a permutation σ where each mapping is an edge of G. The same permutation can be viewed as a perfect matching in the simple bipartite graph B_G, whose left and right partite sets are both copies of $V(G)$, with an edge i on the left adjacent to j on the right if and only if $(i, j) \in E(G)$. Indeed, the edges $(i, \sigma(i))$ form a perfect matching of B_G: since σ is a permutation, every vertex on both sides is matched exactly once. Conversely, any perfect matching of B_G assigns to each i a unique out-neighbor $\sigma(i)$, and since every right vertex is matched exactly once, σ is a permutation and hence a cycle factor. Thus the three notions are in bijection. When G is d-regular, B_G is likewise d-regular. We will often switch between these concepts as each offers access to different theorems and simplifications.

Permutations map one element to another. Thus, each permutation can be understood and written as a collection of disjoint cycles. Up to the order in which the cycles are listed and cyclic rotations within each cycle, each permutation corresponds uniquely to such a collection. In cycle notation, $(a_1\, a_2\, \ldots\, a_k)$ denotes the cycle where a_1 maps to a_2, a_2 maps to a_3, finally sending a_k back to a_1.

We list the cycles of a permutation using the following canonical presentation. In cycle notation, permute the elements cyclically until the largest element is first. Then, list the cycles by increasing order of their largest element. Removing the parentheses produces a sequence containing the elements of $[n]$. Conversely, given any such sequence, insert an opening parenthesis before each left-to-right maximum (or record), close each cycle immediately before the next record, and close the final cycle at the end of the sequence. Doing so recovers the original permutation. Thus, there is a bijection between permutations of $[n]$ and orderings of its elements.

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Example 2.4. Consider the permutation $\pi = (1\,4\,2)(3\,6)(5)$ of $[6]$. Rotating each cycle so that its largest element comes first and ordering the cycles by these elements gives

$$\pi = (4\,2\,1)(5)(6\,3).$$

Removing the parentheses produces the sequence

$$\mathbf{4},\ 2,\ 1,\ \mathbf{5},\ \mathbf{6},\ 3.$$

Its left-to-right maxima are 4, 5, and 6: each larger than every entry preceding it, with the first entry always counted as a record. To reverse the construction, start a new cycle at each of these records, recovering $(4\,2\,1)(5)(6\,3)$. Thus, the three cycles of π correspond exactly to the three records in the sequence.

For each $k \in [n]$, define the indicator variable $\mathbb{I}_k$ to be 1 if the entry in position k is a left-to-right maximum and 0 otherwise. Among the first k entries, the maximum has equal probability to occur in each of the k positions. Thus, the probability that the entry in position k is a left-to-right maximum is $1/k$. Therefore, $\mathbb{P}(\mathbb{I}_k = 1) = 1/k$. Since each cycle corresponds to its unique starting left-to-right maxima,

$$C(\pi) = \sum_{k=1}^{n} \mathbb{I}_k.$$

By linearity of expectation,

$$\mathbb{E}[C(\pi)] = \sum_{k=1}^{n} \mathbb{E}[\mathbb{I}_k] = \sum_{k=1}^{n} \frac{1}{k} = \ln(n) + \gamma + \mathcal{O}(\frac{1}{n}).$$

This proves that a random permutation has $\mathcal{O}(\log n)$ cycles on average.

3 Cycle factors in directed graphs via entropy

Because perfect matchings correspond to cycle factors, the same bound may be used for cycle factors of d-regular directed graph on n vertices as perfect matchings in d-regular bipartite graphs on $2n$ vertices. The regular cases below appear as Theorem 3.1 and Theorem 3.2 in [Chr+25a].

Theorem 3.1 ([Brè73]). *Let B be a bipartite graph with left part $\{u_1, \ldots, u_n\}$, where u_i has degree r_i. Then B has at most*

$$\prod_{i=1}^{n} (r_i!)^{1/r_i}$$

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公开

perfect matchings. In particular, every d-regular bipartite graph on $2n$ vertices has at most $(d!)^{n/d}$ perfect matchings.

Theorem 3.2 ([Ego81; Fal81])**.** *Every d-regular bipartite graph on $2n$ vertices has at least $n!d^n/n^n$ perfect matchings.*

3.1 A simplified case

We begin by discussing a simplified version of Theorem 1.1. Let G be a d-regular directed graph on n vertices, labeled by $[n]$. Define $\mathcal{C}$ as the collection of permutations σ of $[n]$ for which $(i, \sigma(i)) \in E(G)$ holds for every $i \in [n]$. By the preceding remark, $\mathcal{C}$ is also the set of cycle factors of G. Choose σ uniformly at random from $\mathcal{C}$, and let τ be an independent uniformly random permutation of $[n]$. In a cycle factor, each vertex i is assigned exactly one out-neighbor, namely $\sigma(i) \in N^+(i)$. Reveal these assignments in the order $\tau(1), \ldots, \tau(n)$: at step j, reveal the value $\sigma(\tau(j))$. Thus, τ maps a reveal position to a vertex, while $\tau^{-1}(i)$ is the step at which the assignment of vertex i is revealed. In particular, if $\tau(j) = i$, then the value $\sigma(i)$, or equivalently the edge $(i, \sigma(i))$, is revealed at step j.

Given permutations σ' and τ and a vertex i, let $s(i, \sigma', \tau)$ be the number of out-neighbors still available to i when its assignment is revealed. An out-neighbor j is unavailable to i if an earlier revealed assignment has already used j. Among the d vertices that map to the possible out-neighbors of i, the position of i in the random order τ is uniformly distributed on $[d]$. Consequently, $s(i, \sigma', \tau)$ is also uniformly distributed on $[d]$.

Revealing which vertex an arbitrary matches to is equivalent to revealing a specific edge going out of that vertex. By saying that an out-going edge of a vertex closes a cycle, we mean it maps back to the vertex beginning the path that vertex is in. For every cycle, exactly one edge $(i, \sigma(i))$ can map back to the unique starting vertex of that path, so exactly one choice of edge closes a cycle. Hence, instead of directly counting the expected number of cycles, we may calculate the expected number of edges that close cycles. *If we assume that the chosen out-neighbor is uniformly distributed among all available out-neighbors*, then the probability that $\sigma(i)$ closes a cycle is equal to the probability of choosing one specific out-neighbor

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

from the remaining $s(i, \sigma', \tau)$. Given that $s(i, \sigma', \tau)$ is uniformly distributed on $[d]$,

$$\mathbb{E}_\sigma[|\sigma|] = \sum_{i=1}^{n} \mathbb{P}(\sigma(i) \text{ closes a cycle}) \tag{1}$$

$$\leq \sum_{i=1}^{n} \frac{1}{d} \sum_{t=1}^{d} \frac{1}{t} \tag{2}$$

$$\leq \frac{n}{d} (\log d + 1), \tag{3}$$

proving the desired bound.

Example 3.3. Suppose $n = 4$ and the chosen cycle factor and reveal order are

$$(\sigma(1), \sigma(2), \sigma(3), \sigma(4)) = (2, 4, 1, 3),$$
$$(\tau(1), \tau(2), \tau(3), \tau(4)) = (3, 1, 4, 2).$$

These tuples list function values, not in cycle notation. The assignments are revealed as follows:

Step j	1	2	3	4
Vertex $\tau(j)$	3	1	4	2
Out-neighbor $\sigma(\tau(j))$	1	2	3	4

For example, $\tau(2) = 1$, so the second step reveals $\sigma(1) = 2$, namely the edge $(1, 2)$. However, this does not yet reveal the outgoing assignment of vertex 2: being revealed as an out-neighbor is different from having one's own assignment revealed.

3.1.1 Counterexample for uniformity

The assumption that the chosen out-neighbor will be uniformly distributed among all available out-neighbors is not true in general. It is usually not true that each edge occurs in the same number of cycle factors.

Consider a graph on 5 vertices labeled 0 to 4, with an edge from i to each of $i - 1$, i, and $i + 1$, where arithmetic is modulo 5. This is a 3-regular directed graph. This graph contains 13 cycle factors. Of all cycle factors, 5 cycle factors contain the loop $(0, 0)$ while only 4 cycle factors contain $(0, 1)$ and only 4 cycle factors contain $(0, 4)$. Thus, in a uniformly random cycle factor, the three outgoing edges of vertex 0 have probabilities 5/13, 4/13, and 4/13, respectively. If vertex 0 is revealed first, all three of its out-neighbors are still available, but its chosen out-neighbor is not uniformly distributed among all available ones.

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

3.2 Bounding non-uniformity by entropy

Recall that maximum entropy is achieved through uniform distribution. Thus, when entropy decreases, the probability of the random variable taking a certain value must increase, breaking this balance. The following lemma gives an upper bound to this probability as entropy deficit grows.

Lemma 3.4 ([Chr+25a]). *Let X be a random variable taking values in a set $\mathcal{X}$ of size s, and let $\ell = \log s - \mathcal{H}(X)$. Then, for every $x \in \mathcal{X}$,*

$$\mathbb{P}(X = x) \leq \frac{2}{s} + \ell.$$

Proof. Since by Theorem 2.2 $\mathcal{H}(X) \leq \log s$ and entropy is non-negative, the bound holds trivially when $s = 1$ or $s = 2$. Assume $s \geq 3$. Fix $x \in \mathcal{X}$, let $p = \mathbb{P}(X = x)$, and let $\mathbb{I}_{X=x}$ be the indicator variable of the event $X = x$. That is, $\mathbb{I}_{X=x}$ is 1 if $X = x$ and 0 otherwise. Using the chain rule, $\mathcal{H}(X, \mathbb{I}_{X=x}) = \mathcal{H}(X) + \mathcal{H}(\mathbb{I}_{X=x} \mid X)$. Since X determines $\mathbb{I}_{X=x}$, $\mathcal{H}(\mathbb{I}_{X=x} \mid X) = 0$. Thus,

$$\mathcal{H}(X) = \mathcal{H}(X, \mathbb{I}_{X=x}).$$

Since $(X, \mathbb{I}_{X=x})$ denotes the joint distribution, changing their order does not change the entropy. Hence,

$$\begin{aligned} \mathcal{H}(X) &= \mathcal{H}(\mathbb{I}_{X=x}) + \mathcal{H}(X|\mathbb{I}_{X=x}) && (4) \\ &= \mathcal{H}_{\mathrm{b}}(p) + (1-p)\mathcal{H}(X \mid X \neq x). && (5) \end{aligned}$$

The random variable X given that $X \neq x$ can take the remaining $s - 1$ values in $\mathcal{X}$. Therefore,

$$\mathcal{H}(X) \leq \mathcal{H}_{\mathrm{b}}(p) + (1-p)\log(s-1).$$

Let $h(q) = \mathcal{H}_{\mathrm{b}}(q) + (1-q)\log(s-1)$. Intending to find a random variable whose entropy is $h(q)$, let Y be a random variable that takes the value x with probability q and each other value in $\mathcal{X}$ with probability $(1-q)/(s-1)$. Then

$$\mathcal{H}(Y) = -q\log(q) - (1-q)\log\left(\frac{1-q}{s-1}\right) = h(q).$$

Hence, $h(q) \leq \log s$. Moreover, taking the derivative gives $h'(q) = \log\left(\frac{1-q}{q(s-1)}\right)$, which is decreasing in q. If $p \leq 2/s$, since ℓ is non-negative,the result holds immediately. Suppose that $p \geq 2/s$. For every $q \in [2/s, p]$,

$$h'(q) \leq h'(2/s) = \log\left(\frac{s-2}{2s-2}\right) \leq -1.$$

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖官方公示

The mean value theorem guarantees that there exists $c \in [2/s, p]$ such that

$$h(p) - h(2/s) = h'(c)\left(p - \frac{2}{s}\right).$$

Since h' is decreasing, $h'(c) \le h'(2/s) \le -1$. Therefore,

$$h(p) \le h(2/s) - \left(p - \frac{2}{s}\right).$$

It follows that

$$\log s - \ell = \mathcal{H}(X) \le h(p) \le h(2/s) - \left(p - \frac{2}{s}\right) \le \log s - \left(p - \frac{2}{s}\right).$$

Rearranging this inequality yields the desired bound. □

3.3 Proof of Theorem 1.1

We use the same notation and setup as the simplified case. We reveal the assigned out-neighbors one by one. To utilize Lemma 3.4, we give a bound to $\mathcal{H}(\sigma)$. Taking the entropy of σ and recalling that σ was chosen uniformly at random from $\mathcal{C}$,

$$\mathcal{H}(\sigma) = \mathcal{H}(\sigma(1), \ldots, \sigma(n)) = \log |\mathcal{C}|.$$

Recall that τ is the auxiliary permutation introduced to specify the reveal order; it is not the permutation σ that corresponds to the cycle factor. At step j, we reveal the outgoing assignment of vertex $\tau(j)$, namely $\sigma(\tau(j))$. Fix an arbitrary reveal order τ. Since τ is a bijection, the tuple

$$\big(\sigma(\tau(1)), \ldots, \sigma(\tau(n))\big)$$

is just a reordering of $\big(\sigma(1), \ldots, \sigma(n)\big)$. Applying the chain rule therefore gives

$$\mathcal{H}(\sigma) = \sum_{j=1}^{n} \mathcal{H}\big(\sigma(\tau(j)) \;\big|\; \sigma(\tau(1)), \ldots, \sigma(\tau(j-1))\big),$$

where the conditioning is empty when $j = 1$. If $\tau^{-1}(i) = j$, let $\tau_{<i} = \{\tau(1), \ldots, \tau(j-1)\}$ be the set of vertices revealed before i, and write $\sigma(\tau_{<i})$ for restricting σ to this set. Since τ is bijective, each term in the sum indexed by j occurs exactly once in the sum indexed by i. We may therefore reindex the sum in terms of i:

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \mathcal{H}\big(\sigma(i) \mid \sigma(\tau_{<i})\big).$$

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Conditioning each term on the event $\sigma(\tau_{<i}) = \pi$ gives

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \sum_{\pi} \mathbb{P}\big(\sigma(\tau_{<i}) = \pi\big)\mathcal{H}\big(\sigma(i) \mid \sigma(\tau_{<i}) = \pi\big),$$

where π ranges over all possible partial maps. Since σ was chosen uniformly at random, $\mathbb{P}\big(\sigma(\tau_{<i}) = \pi\big) = |\{\sigma' \in \mathcal{C} : \sigma'(\tau_{<i}) = \pi\}|/|\mathcal{C}|$. We reindex the sum by all full cycle factor $\sigma' \in \mathcal{C}$. For each such σ', condition on $\sigma'(\tau_{<i}) = \pi$. Since any fixed history π appears in the reindexed sum once for every σ' that extends it, the number of occurrences of π in this sum is exactly the numerator $|\{\sigma' \in \mathcal{C} : \sigma'(\tau_{<i}) = \pi\}|$. Thus, dividing by $|\mathcal{C}|$ and factoring out of the sum gives

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathcal{H}\big(\sigma(i) \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})\big).$$

Given the history $\sigma(\tau_{<i}) = \sigma'(\tau_{<i})$, the value $\sigma(i)$ must be one of the available out-neighbors because it is the vertex chosen. Hence its conditional entropy is at most $\log s(i, \sigma', \tau)$. Define the entropy deficit

$$\ell(i, \sigma', \tau) = \log s(i, \sigma', \tau) - \mathcal{H}\big(\sigma(i) \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})\big).$$

This notation is meant to resemble Lemma 3.4. Indeed, later on we use the upper bound of $\ell(i, \sigma', \tau)$ and Lemma 3.4 to bound the probability that an edge closes a cycle. With this in mind, we write $\mathcal{H}(\sigma)$ by substituting:

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \big(\log s(i, \sigma', \tau) - \ell(i, \sigma', \tau)\big).$$

The preceding identity holds for every fixed reveal order τ. We now average over a uniformly random τ, chosen independently of σ. Independence ensures that fixing the reveal order does not change the distribution of σ, so the left-hand side remains $\mathcal{H}(\sigma)$. Thus, we obtain

$$\mathcal{H}(\sigma) = \mathbb{E}_{\tau}\left[\sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \big(\log s(i, \sigma', \tau) - \ell(i, \sigma', \tau)\big)\right].$$

By linearity of expectation,

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{E}_{\tau}\big[\log s(i, \sigma', \tau)\big] - \mathcal{L}$$

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

where

$$\mathcal{L} = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{E}_\tau \big[\ell(i, \sigma', \tau)\big].$$

Since $s(i, \sigma', \tau)$ is uniformly distributed on $[d]$,

$$\mathbb{E}_\tau[\log s(i, \sigma', \tau)] = \frac{1}{d} \sum_{t=1}^{d} \log t.$$

Substituting,

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \frac{1}{d} \sum_{t=1}^{d} \log t - \mathcal{L}.$$

We can factor out $1/|\mathcal{C}|$ in the first expression, and since the second sum is independent of σ', it is equivalent to removing the sum and multiplying by $|\mathcal{C}|$. Canceling out and using $\sum_{t=1}^{d} \log t = \log d!$,

$$\mathcal{H}(\sigma) = \frac{n}{d} \sum_{t=1}^{d} \log t - \mathcal{L} = \frac{n}{d} \log(d!) - \mathcal{L}.$$

By Theorem 3.2, $|\mathcal{C}| \geq n!(d/n)^n \geq (d/e)^n$. Using $|\mathcal{C}| \geq (d/e)^n$ and Stirling's approximation $d! \leq (d/e)^d(ed)$,

$$n \log\left(\frac{d}{e}\right) \leq \log |\mathcal{C}| \leq n \log\left(\frac{d}{e}\right) + \frac{n}{d} \log(ed) - \mathcal{L}.$$

Thus, $\mathcal{L} \leq (n/d) \log(ed)$.

Forming a cycle requires exactly one edge to close it. Counting the expected number of edges closing a cycle and taking the expectation,

$$\mathbb{E}_\sigma[|\sigma|] = \sum_{i=1}^{n} \mathbb{P}\big(\sigma(i) \text{ closes a cycle}\big).$$

Conditioning on the event that the previously revealed vertices follows a specific permutation π, where π runs through all possible permutations,

$$\mathbb{E}_\sigma[|\sigma|] = \sum_{i=1}^{n} \sum_{\pi} \mathbb{P}\big(\sigma(\tau_{<i}) = \pi\big) \mathbb{P}\big(\sigma(i) \text{ closes a cycle} \mid \sigma(\tau_{<i}) = \pi\big).$$

Again, since the permutations were chosen uniformly at random, we reindex the sum by all full cycles $\sigma' \in \mathcal{C}$ and divide by $|\mathcal{C}|$:

$$\mathbb{E}_\sigma[|\sigma|] = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{P}\big(\sigma(i) \text{ closes a cycle} \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})\big).$$

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

For a fixed history, if the edge from i closes a cycle, it must use the unique available out-neighbor that completes the corresponding cycle. Thus,

$$\mathbb{P}\big(\sigma(i) \text{ closes a cycle } \big|\, \sigma(\tau_{<i}) = \sigma'(\tau_{<i})\big) \le \max_{j \in N^+(i)} \mathbb{P}\big(\sigma(i) = j \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})\big)$$

Lemma 3.4 therefore gives

$$\mathbb{P}\big(\sigma(i) \text{ closes a cycle } \big|\, \sigma(\tau_{<i}) = \sigma'(\tau_{<i})\big) \le \frac{2}{s(i,\sigma',\tau)} + \ell(i,\sigma',\tau).$$

Summing over all vertices and histories gives

$$\mathbb{E}_\sigma(|\sigma|) \le \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \left(\frac{2}{s(i,\sigma',\tau)} + \ell(i,\sigma',\tau) \right).$$

Again, since this bound holds for every τ, we take the expectation over τ:

$$\begin{aligned}\mathbb{E}_\sigma(|\sigma|) &\le \mathbb{E}_\tau \left(\sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \left(\frac{2}{s(i,\sigma',\tau)} + \ell(i,\sigma',\tau) \right) \right) \\ &= \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{E}_\tau \left(\frac{2}{s(i,\sigma',\tau)} \right) + \mathcal{L}.\end{aligned}$$

Since $s(i,\sigma',\tau)$ is uniformly distributed on $[d]$, this simplifies to

$$\begin{aligned}\mathbb{E}_\sigma(|\sigma|) &\le \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \frac{1}{d} \sum_{t=1}^{d} \frac{2}{t} + \mathcal{L} \\ &\le \frac{2n}{d}(\log d + 1) + \mathcal{L} \\ &\le \frac{4n}{d}(\log d + 1).\end{aligned}$$

3.3.1 Limitations

A natural question arising from Theorem 1.1 is whether the bound $\mathcal{O}((n/d)\log d)$ can be improved. In expectation, this bound is asymptotically tight. To see this, suppose that $d \mid n$, and let the graph be the disjoint union of n/d copies of K_d°. A uniformly random cycle factor on each copy corresponds to a uniformly random permutation of d elements, whose expected number of cycles is $\Theta(\log d)$. By linearity of expectation, the expected number of cycles in the entire graph is therefore $\Theta((n/d)\log d)$.

However, tightness in expectation does not rule out stronger existential results. Christoph et al. [Chr+25b] conjectured that every d-regular directed graph on n vertices contains a cycle factor with $\mathcal{O}(n/d)$ cycles. This is consistent with the example above: choosing a single cycle through all vertices of each copy gives a

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

cycle factor with exactly n/d cycles, even though a uniformly random cycle factor has $\Theta((n/d)\log d)$ cycles on average.

4 Fixed points

Christoph et al. [Chr+25a] posed the question of whether a uniformly random cycle factor of a directed d-regular graph on n vertices has an expected number of fixed points bounded by $\mathcal{O}(n/d)$. The expected number of fixed points in a uniformly random permutation of n elements is 1, so the proposed bound holds for complete directed graphs with loops. For general regular graphs, however, the problem is more difficult.

A natural attempt at this problem is to utilize the correspondence between cycle factors and perfect matchings in the bipartite presentation. Let G be a d-regular directed graph on n vertices, and let B_G be its simple bipartite presentation, with partite sets L and R. If v is a fixed point, then $L_v R_v$ is an edge of the corresponding perfect matching in B_G. Deleting L_v and R_v yields a bipartite graph B'_G whose perfect matchings correspond to the perfect matchings of B_G that contain $L_v R_v$. Thus, provided that v has a loop, the probability that v is a fixed point is $\operatorname{pm}(B'_G)/\operatorname{pm}(B_G)$ where $\operatorname{pm}(B_G)$ denotes the number of perfect matchings in B_G. If v has no loop, this probability is zero. Linearity of expectation therefore shows that the expected number of fixed points equals the sum of these vertex-wise probabilities. However, combining Theorems 3.1 and 3.2 in this direct way is too weak to prove the desired $\mathcal{O}(n/d)$ bound. In fact, it does not improve upon the $\mathcal{O}((n/d)\log d)$ bound that follows from Theorem 1.1 by every fixed point being a cycle.

Using entropy, we prove Theorem 1.2. Its bound is weaker than the original conjecture but improves upon the bound implied by Theorem 1.1. We will use a lower bound of Schrijver [Sch98] on the number of perfect matchings, which is stronger than the estimate in Theorem 3.2.

Theorem 4.1 ([Sch98])**.** *Every d-regular bipartite graph on $2n$ vertices has at least $((d-1)^{d-1}/d^{d-2})^n$ perfect matchings.*

The following lemma will be used in the proof of Theorem 1.2. It allows us to obtain an upper bound on $\mathcal{H}(S(\sigma))$, the entropy of a random set of fixed points, with respect to p.

Lemma 4.2. *Let S be a random subset of $[n]$, and let $p = \mathbb{E}[|S|]/n$.*

$$\mathcal{H}(S) \leq n\mathcal{H}_{\mathrm{b}}(p).$$

2026 S. T. (Qiu Hing) School Science Award
仅用于2026丘成桐中学科学奖公示

Proof. Represent S by its indicator vector $X = (X_1, X_2, \ldots, X_n)$, where $X_i = 1$ if and only if $i \in S$. Then S and X determine each other, $|S| = \sum_{i=1}^{n} X_i$, and $\mathbb{E}[|S|] = \sum_{i=1}^{n} \mathbb{E}[X_i]$. Using the chain rule,

$$\mathcal{H}(S) = \mathcal{H}(X) = \sum_{i=1}^{n} \mathcal{H}(X_i \mid X_1, \ldots, X_{i-1}).$$

Conditioning cannot increase entropy, so

$$\mathcal{H}(S) \leq \sum_{i=1}^{n} \mathcal{H}(X_i).$$

Each X_i is a Bernoulli random variable; write $p_i = \mathbb{E}[X_i]$. Its entropy is $\mathcal{H}_{\mathrm{b}}(p_i)$. Thus, $\mathcal{H}(S) \leq \sum_{i=1}^{n} \mathcal{H}_{\mathrm{b}}(p_i)$. Also, $\sum_{i=1}^{n} p_i = \mathbb{E}[|S|]$. Since the binary entropy function is concave, Jensen's inequality yields

$$\frac{1}{n} \sum_{i=1}^{n} \mathcal{H}_{\mathrm{b}}(p_i) \leq \mathcal{H}_{\mathrm{b}}\left(\frac{1}{n} \sum_{i=1}^{n} p_i\right) = \mathcal{H}_{\mathrm{b}}\left(\frac{\mathbb{E}[|S|]}{n}\right).$$

Thus,

$$\mathcal{H}(S) \leq n\mathcal{H}_{\mathrm{b}}\left(\frac{\mathbb{E}[|S|]}{n}\right).$$

□

Proof of Theorem 1.2. Throughout this proof, logarithms and entropy are measured using *natural logarithm*. Changing the base multiplies every entropy by the same constant, so Lemma 4.2 remains valid. Let G be a directed d-regular graph on n vertices with vertex set V. A cycle factor is a permutation σ on the set of vertices such that $i \to \sigma(i)$ is an edge of G for every i. Its fixed points are the vertices satisfying $\sigma(i) = i$.

Let B_G be the bipartite presentation of G. Denote the set of perfect matchings of B_G by $\mathcal{M}$, and let σ be a uniformly random element of $\mathcal{M}$. Let $S(\sigma)$ be the random set of fixed points of σ.

Since σ is chosen uniformly at random,

$$\mathcal{H}(\sigma) = \log |\mathcal{M}|.$$

By Theorem 4.1, $|\mathcal{M}| \geq ((d-1)^{d-1}/d^{d-2})^n$. Therefore,

$$\mathcal{H}(\sigma) \geq n \log \frac{(d-1)^{d-1}}{d^{d-2}}.$$

Fix $\sigma' \in \mathcal{M}$, and let $S(\sigma')$ be its fixed-point set. Then $\mathbb{P}(S(\sigma) = S(\sigma')) > 0$.

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

Let B'_G be obtained from B_G by deleting both copies of every vertex in $S(\sigma')$. Conditional on $S(\sigma) = S(\sigma')$, the matching edges at these vertices are fixed, so the remaining matching in B'_G determines σ uniquely. Let $\mathcal{M}(B')$ denote the set of all perfect matchings of B'_G. Not every matching in $\mathcal{M}(B'_G)$ need be compatible with $S(\sigma) = S(\sigma')$, since some may contain additional fixed points. Nevertheless, the number of possible values of σ under this conditioning is at most $|\mathcal{M}(B'_G)|$. The entropy bound in Theorem 2.2 therefore gives

$$\mathcal{H}(\sigma \mid S(\sigma) = S(\sigma')) \le \log |\mathcal{M}(B'_G)|.$$

Write r_i for the degree in B'_G of the left copy of $i \in V \setminus S(\sigma')$. Since B'_G has a perfect matching and is a subgraph of B, we have $1 \le r_i \le d$. Theorem 3.1 gives

$$|\mathcal{M}(B')| \le \prod_{i \in V \setminus S(\sigma')} (r_i!)^{1/r_i} \le \prod_{i \in V \setminus S(\sigma')} (d!)^{1/d} = (d!)^{(n - |S(\sigma')|)/d}.$$

For the second inequality, note that $(r!)^{1/r}$ is increasing in the positive integer r. Taking logarithms yields

$$\mathcal{H}(\sigma \mid S(\sigma) = S(\sigma')) \le \log |\mathcal{M}(B'_G)| \le \frac{n - |S(\sigma')|}{d} \log(d!).$$

By definition, $\mathcal{H}(\sigma \mid S(\sigma))$ is the weighted average of $\mathcal{H}(\sigma \mid S(\sigma) = S(\sigma'))$ over the distinct possible fixed-point sets $S(\sigma')$, with each set weighted by its probability $\mathbb{P}(S(\sigma) = S(\sigma'))$. Applying the preceding bound to each term gives

$$\begin{aligned}
\mathcal{H}(\sigma \mid S(\sigma)) &= \sum_{S(\sigma')} \mathbb{P}(S(\sigma) = S(\sigma')) \mathcal{H}(\sigma \mid S(\sigma) = S(\sigma')) \\
&\le \sum_{S(\sigma')} \mathbb{P}(S(\sigma) = S(\sigma')) \frac{n - |S(\sigma')|}{d} \log(d!) \\
&= \frac{\log(d!)}{d} \left(n \sum_{S(\sigma')} \mathbb{P}(S(\sigma) = S(\sigma')) - \sum_{S(\sigma')} \mathbb{P}(S(\sigma) = S(\sigma')) |S(\sigma')| \right) \\
&= \frac{\log(d!)}{d} \left(n - \mathbb{E}[|S(\sigma)|] \right).
\end{aligned}$$

Here the sums range over the distinct sets $S(\sigma')$ arising from $\sigma' \in \mathcal{M}$, with each set counted once, not once per matching. The final equality uses the fact that these probabilities sum to 1 and the definition $\mathbb{E}[|S(\sigma)|] = \sum_{S(\sigma')} \mathbb{P}(S(\sigma) = S(\sigma')) |S(\sigma')|$. Also, Lemma 4.2 gives $\mathcal{H}(S(\sigma)) \le n \mathcal{H}_{\mathrm{b}}(\mathbb{E}[|S(\sigma)|]/n)$. Since σ determines $S(\sigma)$, the

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

chain rule gives $\mathcal{H}(\sigma) = \mathcal{H}(S(\sigma)) + \mathcal{H}(\sigma \mid S(\sigma))$. Let

$$p = \frac{\mathbb{E}[|S(\sigma)|]}{n}.$$

Applying the preceding bounds and dividing by n gives

$$\log \frac{(d-1)^{d-1}}{d^{d-2}} \leq \mathcal{H}_{\mathrm{b}}(p) + (1-p)\frac{1}{d}\log(d!).$$

To isolate the term containing p, expand the last term and then move the corresponding terms to the opposite sides:

$$\begin{aligned}
\log \frac{(d-1)^{d-1}}{d^{d-2}} &\leq \mathcal{H}_{\mathrm{b}}(p) + \frac{1}{d}\log(d!) - p\frac{1}{d}\log(d!), \\
\log \frac{(d-1)^{d-1}}{d^{d-2}} + p\frac{1}{d}\log(d!) &\leq \mathcal{H}_{\mathrm{b}}(p) + \frac{1}{d}\log(d!), \\
p\frac{1}{d}\log(d!) &\leq \mathcal{H}_{\mathrm{b}}(p) + \frac{1}{d}\log(d!) - \log\frac{(d-1)^{d-1}}{d^{d-2}}. \qquad (6)
\end{aligned}$$

We proceed by bounding $(1/d)\log(d!)$ and $\log((d-1)^{d-1}/d^{d-2})$. Using Stirling's approximation in the form

$$d! = \sqrt{2\pi d}\left(\frac{d}{e}\right)^{d}\left(1 + \mathcal{O}\left(\frac{1}{d}\right)\right),$$

and using $\log(1+x) = \mathcal{O}(x)$ as $x \to 0$, we obtain

$$\begin{aligned}
\frac{1}{d}\log(d!) &= \frac{1}{d}\left[\frac{1}{2}\log(2\pi d) + d\log d - d + \log\left(1 + \mathcal{O}\left(\frac{1}{d}\right)\right)\right] \\
&= \log d - 1 + \frac{1}{2d}\log(2\pi d) + \frac{1}{d}\mathcal{O}\left(\frac{1}{d}\right) \\
&= \log d - 1 + \frac{1}{2d}\log(2\pi d) + \mathcal{O}\left(\frac{1}{d^2}\right) \qquad (7) \\
&= \log d - 1 + \mathcal{O}\left(\frac{\log d}{d}\right). \qquad (8)
\end{aligned}$$

In the final line, we used $\log(2\pi d) = \mathcal{O}(\log d)$ and absorbed the smaller $\mathcal{O}(1/d^2)$ term into $\mathcal{O}((\log d)/d)$. Also,

$$\log \frac{(d-1)^{d-1}}{d^{d-2}} = (d-1)\log(d-1) - (d-2)\log d. \qquad (9)$$

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Since $\log(d-1) = \log d + \log(1-1/d)$,

$$\log \frac{(d-1)^{d-1}}{d^{d-2}} = (d-1)\log d - (d-1)\left(\frac{1}{d} + \frac{1}{2d^2} + \mathcal{O}\left(\frac{1}{d^3}\right)\right) - (d-2)\log d \quad (10)$$

$$= \log d - 1 + \frac{1}{2d} + \mathcal{O}\left(\frac{1}{d^2}\right). \quad (11)$$

The two expansions have the same leading term, $\log d - 1$. Subtracting them therefore gives

$$\frac{1}{d}\log(d!) - \log\frac{(d-1)^{d-1}}{d^{d-2}} = \frac{\log(2\pi d) - 1}{2d} + \mathcal{O}\left(\frac{1}{d^2}\right) = \mathcal{O}\left(\frac{\log d}{d}\right).$$

Substituting this back to the right-hand-side of (6), and the expansion of $(1/d)\log(d!)$ into its left-hand side, yields

$$p\left(\log d - 1 + \mathcal{O}\left(\frac{\log d}{d}\right)\right) \le \mathcal{H}_{\mathrm{b}}(p) + \mathcal{O}\left(\frac{\log d}{d}\right).$$

Move the error term on the left to the right. Since $0 \le p \le 1$, combining the two error terms thus gives

$$p(\log d - 1) \le \mathcal{H}_{\mathrm{b}}(p) + \mathcal{O}\left(\frac{\log d}{d}\right).$$

For $p \in [0,1]$, the binary entropy satisfies $\mathcal{H}_{\mathrm{b}}(p) \le p(-\log p + 1)$. Hence

$$p(\log d - 1) \le p(-\log p + 1) + \frac{C\log d}{d}$$

for some constant $C > 0$, independent of n, d, and p. Moving the $p(-\log p + 1)$ to the left and combining logarithms gives

$$\begin{aligned} p(\log d - 1) - p(-\log p + 1) &\le \frac{C\log d}{d}, \\ p(\log d + \log p - 2) &\le \frac{C\log d}{d}, \\ p\big(\log(dp) - 2\big) &\le \frac{C\log d}{d}. \end{aligned}$$

Let $t = dp = d\,\mathbb{E}[|S(\sigma)|]/n$. Then

$$t(\log t - 2) \le C\log d.$$

Set

$$T = 4(C+1)\frac{\log d}{\log\log d}.$$

2026 S.-T. Yau High School Science Award
仅用于2026年丘成桐中学科学奖论文公示

For all sufficiently large d,

$$\log T - 2 = \log\log d - \log\log\log d + \log(4(C+1)) - 2 \geq \frac{1}{2}\log\log d,$$

so

$$T(\log T - 2) \geq 2(C+1)\log d > C\log d.$$

Since $x(\log x - 2)$ is increasing for $x > e$, it follows that $t \leq T$ for all sufficiently large d. Therefore,

$$\mathbb{E}[|S(\sigma)|] = \frac{nt}{d} = \mathcal{O}\left(\frac{n\log d}{d\log\log d}\right).$$

□

5 Extension to cycles of arbitrary length

We follow a similar structure as the proof of Theorem 1.2. Instead of revealing all fixed points, we reveal the random set $\mathcal{R}$ for all cycles of length r in σ. The number of possible cycles of length r in the graph satisfies

$$N_r \leq \frac{nd^{r-1}}{r}.$$

Adapting the argument of Lemma 4.2 used in the original proof, we obtain

$$\mathcal{H}(\mathcal{R}) \leq N_r \mathcal{H}_b\left(\frac{m}{N_r}\right) \leq m\log\left(\frac{eN_r}{m}\right),$$

where $m = \mathbb{E}[|\mathcal{R}|]$. Revealing these cycles reveals on average mr vertices. Applying Theorem 3.1 to bound conditional entropy yields

$$\mathcal{H}(\sigma \mid \mathcal{R}) \leq \frac{n - rm}{d}\log(d!).$$

Finally, we compare these two bounds with the same lower bound on the total number of cycle factors. The resulting inequality gives

$$\mathcal{O}_r\left(\frac{n\log d}{d\log\log d}\right).$$

Thus, the original argument extends to arbitrary cycle length r, with the main difference being the random set revealed and the number of vertices removed by each revealed cycle.

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

6 Concluding remarks

The remaining question, conjectured by Christoph et al. [Chr+25a], is whether the bound in Theorem 1.2 can be improved to $\mathcal{O}(n/d)$. Theorem 1.2 reduces the gap to a factor of order $\log d / \log\log d$. It would also be interesting to determine whether the entropy argument can be refined to remove this factor or whether a completely different approach is required. Theorem 3.1 as used in the proof disregards the fact the remaining vertices in B' must be derangements. Applying a stronger bound could improve this result substantially.

References

[Brè73] L. M. Brègman. "Some properties of nonnegative matrices and their permanents". In: *Doklady Akademii Nauk SSSR* 211.1 (1973). In Russian. English translation in Soviet Mathematics Doklady 14 (1973), 945–949, pp. 27–30.

[Chr+25a] Micha Christoph et al. *Cycle-factors of regular graphs via entropy*. Aug. 23, 2025. DOI: 10.48550/arXiv.2507.19417. URL: http://arxiv.org/abs/2507.19417.

[Chr+25b] Micha Christoph et al. *New bounds for linear arboricity and related problems*. July 28, 2025. DOI: 10.48550/arXiv.2507.20500. URL: http://arxiv.org/abs/2507.20500.

[Ego81] G. P. Egorychev. "The solution of van der Waerden's problem for permanents". In: *Advances in Mathematics* 42.3 (1981), pp. 299–305. DOI: 10.1016/0001-8708(81)90044-X.

[Fal81] D. I. Falikman. "Proof of the van der Waerden conjecture regarding the permanent of a doubly stochastic matrix". In: *Mathematical Notes of the Academy of Sciences of the USSR* 29.6 (1981). Translated from Matematicheskie Zametki 29, no. 6 (1981), 931–938, pp. 475–479. DOI: 10.1007/BF01163285.

[Gaj26] Rishikesh Gajjala. *Counterexamples to an Extremal Conjecture for Random Cycle-Factors*. May 7, 2026. DOI: 10.48550/arXiv.2604.26101. URL: http://arxiv.org/abs/2604.26101.

[Rad97] Jaikumar Radhakrishnan. "An Entropy Proof of Bregman's Theorem". In: *Journal of Combinatorial Theory, Series A* 77.1 (1997), pp. 161–164. DOI: 10.1006/jcta.1996.2727.

2026 S.-T. Yau High School Science Award 仅用于2026丘成桐中学生论文公示

[Sch98] Alexander Schrijver. "Counting 1-Factors in Regular Bipartite Graphs". In: *Journal of Combinatorial Theory, Series B* 72 (1998), pp. 122–135. DOI: 10.1006/jctb.1997.1798.

[Sha48] Claude E. Shannon. "A Mathematical Theory of Communication". In: *The Bell System Technical Journal* 27.3 (1948), pp. 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[Vis12] Nisheeth K. Vishnoi. "A Permanent Approach to the Traveling Salesman Problem". In: *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*. IEEE, 2012, pp. 76–80. DOI: 10.1109/FOCS.2012.81.

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Acknowledgements

I would like to sincerely thank my advisor Mr. Roy for his support throughout the project. He inspired me to pursue a topic in graph theory. The specific details of the research question gradually came along the process. The basis of my study was based on the book The Probabilistic Method by Alon and Spencer. This book introduced me to basic tools in probabilistic tools, including linearity of expectation and the notion of using probability to prove combinatorical facts. Without this book, I would not have been introduced to modern approaches to graph theory. The original ideas were mostly expository, aiming to understand the entropy technique used by Christoph et al. Afterwards, I realized I could try to apply this technique to their question on fixed points proposed in the same paper. This would turn out to be the main theorem and contribution of this paper. A related paper due to the same authors on Posa rotations was also read during this study. In fact, another idea coming into this topic was to investigate if Posa Rotation could become useful in the fixed point or related problems on cycle factors. However, due to time constraints, I was only able to mostly focus on the entropy technique.

The expository section was written first in parallel of working on the fixed-point problem. The essay was written entirely by me. Mr. Roy answered some of my questions on math notation and understanding without directly editing the essay.

The first attempt at the fixed-point problem failed and is documented in the essay. This attempt has the most natural idea and intuition: deleting fixed points yield a new perfect matching. However, as noted in the essay, the current bounds of Shrijver and Bregman-Minc are not strong enough to prove it. At one point, I thought I had a proof using this method, but after proofreading found an algebraic mistake. My advisor encouraged me to continue trying this problem. Afterwards, I thought a similar technique using entropy as the Christoph et al. article could also apply to this problem. This led to the proof of the main theorem of the paper on fixed points. A lot of difficulties, mainly in algebra manipulation, emerged in the process of trying to prove this theorem. Using log 2 makes the proof a lot more difficult compared to natural logarithms. Using natural logarithm and noticing that it does not affect asymptotic results simplified the process.

AI was used in the writing process. Deepseek-V4.1-Flash was used for formatting for Latex. The formatting of example 2.4 and 3.3 used Deepseek-V4.1-Flash but the content was written and checked by me. DeepSeek V4.1-Flash was also used for proofreading the paper. It was used for checking grammar, formatting, and syntactical errors. However, the edits after checking are done by me. Deepseek-V4.1-Flash was also used for polishing the language, but the edits are carried out and reviewed by me. Finally, Deepseek-V4.1-Flash was used for some literature search. I asked the model to give known bounds on perfect matchings in regular bipartite graph. The model provided the Shrijver bound. The specifics of this bound was verified and cited by me in the paper.

Additionally, AI found and proved Theorem 1.3 after it was asked to extend the result of theorem 1.2.

2026 Siyuan High School Science Award
仅用于2026上海市科学奖论文公示