On the Number of Sources in a Random Orientation of a Graph

March 30, 2026

Abstract

We record a short chain of observations about random orientations of a fixed connected simple graph. The original quantity of interest is the number of sources. We first pass to the easier variable counting vertices of even indegree, whose law is completely determined by parity. We then use the directed handshaking lemma to transfer this information back to the source count, obtaining a deterministic lower bound. Finally, we identify the exact defect in that bound and show that it vanishes identically for graphs of maximum degree at most 2. The point of the note is not only the inequalities themselves, but also the mechanism behind them: parity gives a linear shadow of the orientation, while indegree conservation controls how much of that shadow can fail to come from sources.

MS: Note on provenance. This documents an experimental mathematical writing session using LLMs. We are reproving a result on random orientations of finite connected graphs: our earlier result (Pelekis and Schauer, 2013, https://link.springer.com/chapter/10.1007/978-1-4614-6825-7_4.) This is hopefully elementary enough to be able to follow along.

Starting from basic prompts, the discussion develops the distribution of vertex indegrees, proves that indegree parity is uniformly distributed subject to the directed handshaking constraint, derives bounds for the number of source vertices in terms of the number of even-indegree vertices. This is the generated document.

1 Introduction

Let $G = (V, E)$ be a connected simple graph with $n = |V|$ and $m = |E|$. We orient each edge independently and uniformly at random. For the resulting random orientation, let

$$S = \#\{v \in V : d^-(v) = 0\}$$

be the number of sources.

The variable S is locally rigid and nonlinear: a vertex is a source only when *all* incident edges point outward. A more tractable statistic is

$$E_0 = \#\{v \in V : d^-(v) \text{ is even}\},$$

the number of vertices of even indegree. Parity linearizes the problem. Reversing one edge toggles the indegree parity at exactly its two endpoints, and for a connected graph this is the only global obstruction. As a consequence, E_0 has a universal distribution depending only on n and the parity of m.

The second step is to return from parity to sources. The directed handshaking lemma,

$$\sum_{v \in V} d^-(v) = m,$$

forces a budget constraint: an odd-indegree vertex costs at least one unit of indegree, while an even non-source costs at least two. This yields a deterministic lower bound on S in terms of E_0. The sharpness question then becomes structural: when can indegree mass concentrate at vertices of indegree at least 3? The answer is exact, and in particular it shows that graphs of maximum degree at most 2 are the regime in which parity already determines the source count.

2 Parity of indegrees

For each orientation of G, define the parity vector

$$\eta = (\eta(v))_{v \in V} \in \mathbf{F}_2^V, \qquad \eta(v) = d^-(v) \pmod 2.$$

Then E_0 is precisely the number of zeros of η.

Lemma 2.1 (Parity hyperplane)**.** *For every orientation of G one has*

$$\sum_{v \in V} \eta(v) \equiv m \pmod 2.$$

Moreover, as the orientation varies over all 2^m orientations of G, the parity vector η is uniformly distributed on the affine hyperplane

$$H_m := \Big\{ x \in \mathbf{F}_2^V : \sum_{v \in V} x(v) = m \pmod 2 \Big\}.$$

Equivalently, each element of H_m occurs with probability $2^{-(n-1)}$.

Proof. [MS: I asked it to replace the linear algebra proof by a symmetry argument.] The parity constraint is immediate, since each edge contributes exactly one to the total indegree:

$$\sum_{v \in V} \eta(v) \equiv \sum_{v \in V} d^-(v) = m \pmod 2.$$

Thus η takes values only in H_m.

It remains to show that all vectors in H_m occur equally often. Since all 2^m orientations are equally likely, it suffices to prove that for any two vectors $x, y \in H_m$ there is a bijection between the set of orientations with parity vector x and the set of orientations with parity vector y.

Let

$$D := \{v \in V : x(v) \neq y(v)\}.$$

Because x and y have the same total parity, the set D has even cardinality. Pair its vertices arbitrarily as

$$D = \{a_1, b_1, \ldots, a_k, b_k\}.$$

Since G is connected, for each i choose a path P_i from a_i to b_i.

Now fix an orientation ω. For a path P, let $R_P(\omega)$ denote the orientation obtained by reversing every edge of P. This operation toggles the indegree parity at the two endpoints of P and leaves the indegree parity of every other vertex unchanged. Indeed, each internal vertex of P is incident with exactly two edges of the path, so reversing the path changes its indegree by 0 or ± 2, which is invisible modulo 2, whereas at each endpoint exactly one incident edge is reversed.

Define

$$\Phi(\omega) := R_{P_k} \circ \cdots \circ R_{P_1}(\omega).$$

Then Φ toggles the indegree parity exactly at the vertices of D. Consequently, if $\eta(\omega) = x$, then

$$\eta(\Phi(\omega)) = y.$$

Moreover, Φ is a bijection, because each path-reversal is an involution.

Thus the number of orientations with parity vector x equals the number of orientations with parity vector y. Hence all vectors in H_m occur equally often. Since $|H_m| = 2^{n-1}$, each element of H_m occurs with probability $2^{-(n-1)}$. □

Theorem 2.2 (Distribution of the number of even-indegree vertices)**.** *For* $0 \le k \le n$*,*

$$\mathbf{P}(E_0 = k) = \begin{cases} \frac{\binom{n}{k}}{2^{n-1}}, & k \equiv n - m \pmod 2, \\ 0, & k \not\equiv n - m \pmod 2. \end{cases}$$

Equivalently, E_0 *is distributed as a* $\mathrm{Bin}(n, \frac{1}{2})$ *random variable conditioned to have parity* $n - m$ *modulo* 2*.*

Proof. By Lemma 2.1, the parity vector is uniform on H_m. If $E_0 = k$, then exactly k coordinates of η are zero and $n - k$ are one. Such a vector lies in H_m precisely when

$$n - k \equiv m \pmod 2,$$

that is, when $k \equiv n - m \pmod 2$. For admissible k, the number of such parity vectors is $\binom{n}{k}$, and each has probability $2^{-(n-1)}$. □

Corollary 2.3. *One has* $\mathbf{E}[E_0] = n/2$*.*

Proof. The admissible parity class contains exactly half of all subsets of V, and the mean size of a uniformly random subset of V is $n/2$. Equivalently, each fixed vertex has even indegree with probability $1/2$. □

3 Sources and an indegree budget

We now compare S with E_0.

Proposition 3.1 (Deterministic sandwich)**.** *For every orientation of* G*,*

$$\frac{n + E_0 - m}{2} \le S \le E_0.$$

In particular,

$$\mathbf{E}[S] \ge \frac{3n - 2m}{4}.$$

Proof. The upper bound $S \le E_0$ is immediate: every source has indegree 0, hence even.

For the lower bound, let $O = n - E_0$ be the number of odd-indegree vertices. Partition the vertex set into three classes:

- sources, of indegree 0;
- even non-sources, of indegree at least 2;
- odd-indegree vertices, of indegree at least 1.

The directed handshaking lemma gives

$$m = \sum_{v \in V} d^-(v) \geq 0 \cdot S + 2(E_0 - S) + 1 \cdot O.$$

Since $O = n - E_0$, this becomes

$$m \geq 2(E_0 - S) + (n - E_0) = n + E_0 - 2S,$$

which rearranges to the stated bound.

Taking expectations and using $\mathbf{E}[E_0] = n/2$ yields

$$\mathbf{E}[S] \geq \frac{n + \mathbf{E}[E_0] - m}{2} = \frac{3n - 2m}{4}.$$

□

The lower bound has a transparent interpretation: parity tells us how many vertices are even and how many are odd, while the handshaking lemma bounds how much indegree mass can be spent on these classes. A source is the cheapest possible even-indegree vertex.

4 The exact defect

The lower bound in Proposition 3.1 is sharp exactly when no indegree mass is hidden at vertices of indegree at least 3.

Proposition 4.1 (Defect formula)**.** *For an orientation of G, let*

$$N_j = \#\{v \in V : d^-(v) = j\} \qquad (j \geq 0).$$

Then

$$S - \frac{n + E_0 - m}{2} = \sum_{j \geq 3} \left\lfloor \frac{j-1}{2} \right\rfloor N_j.$$

In particular, the lower bound in Proposition 3.1 is an equality if and only if every vertex has indegree 0*,* 1*, or* 2*.*

Proof. We write

$$S = N_0, \qquad E_0 = \sum_{r \geq 0} N_{2r}, \qquad n = \sum_{j \geq 0} N_j, \qquad m = \sum_{j \geq 0} j N_j.$$

Split the indegree sum according to the three minimal contributions used in Proposition 3.1. Vertices of indegree 1 contribute exactly the minimal odd cost. Vertices of indegree 2 contribute exactly the minimal non-source even cost. Every vertex of indegree $j \geq 3$ contributes an excess beyond that minimal cost. Concretely,

$$m = (n - E_0) + 2(E_0 - S) + \sum_{\substack{j \geq 3 \\ j \text{ odd}}} (j-1) N_j + \sum_{\substack{j \geq 4 \\ j \text{ even}}} (j-2) N_j.$$

Rearranging gives

$$S = \frac{n + E_0 - m}{2} + \sum_{\substack{j \geq 3 \\ j \text{ odd}}} \frac{j-1}{2} N_j + \sum_{\substack{j \geq 4 \\ j \text{ even}}} \frac{j-2}{2} N_j,$$

which is exactly the asserted formula.

The right-hand side vanishes if and only if $N_j = 0$ for all $j \geq 3$, equivalently if and only if every vertex has indegree at most 2. □

Remark 4.2. *The defect formula turns the lower bound from an estimate into a structural statement. The gap is not mysterious: it is exactly the contribution of vertices where indegree mass concentrates beyond the minimal levels detected by parity.*

5 Restricted degree and exactness

The previous proposition immediately identifies the degree range in which the lower bound is universally sharp.

Corollary 5.1 (Maximum degree at most two)**.** *If $\Delta(G) \le 2$, then for every orientation of G,*

$$S = \frac{n + E_0 - m}{2}.$$

Hence for connected simple graphs of maximum degree at most 2—that is, for paths and cycles—the number of sources is completely determined by the number of even-indegree vertices.

Proof. If $\Delta(G) \le 2$, then every vertex has indegree at most 2 in every orientation. Proposition 4.1 therefore shows that the defect vanishes identically. □

Corollary 5.2 (Paths and cycles)**.** *Let P_n be the path on n vertices and C_n the cycle on n vertices.*

$$\text{for } P_n: \quad S = \frac{E_0 + 1}{2}, \qquad \mathbf{P}(S = k) = \frac{\binom{n}{2k-1}}{2^{n-1}},$$

$$\text{for } C_n: \quad S = \frac{E_0}{2}, \qquad \mathbf{P}(S = k) = \frac{\binom{n}{2k}}{2^{n-1}}.$$

Proof. For P_n one has $m = n - 1$, so Corollary 5.1 gives $S = (n + E_0 - (n-1))/2 = (E_0 + 1)/2$. The distribution follows by substituting $E_0 = 2k - 1$ into Theorem 2.2. For C_n one has $m = n$, so similarly $S = E_0/2$, and the distribution follows by substituting $E_0 = 2k$. □

Once degree 3 is allowed, the lower bound ceases to be universally exact, but the defect is still very simple in the subcubic case.

Corollary 5.3 (Subcubic graphs)**.** *If $\Delta(G) \le 3$, then for every orientation of G,*

$$S = \frac{n + E_0 - m}{2} + N_3.$$

Thus the lower bound in Proposition 3.1 is an equality if and only if there are no vertices of indegree 3. Equivalently, in a simple subcubic graph it is an equality if and only if there are no degree-3 sinks.

Proof. If $\Delta(G) \le 3$, then $N_j = 0$ for all $j \ge 4$, and the defect formula reduces to

$$S - \frac{n + E_0 - m}{2} = N_3.$$

In a subcubic graph, a vertex has indegree 3 if and only if it has degree 3 and all incident edges point inward, i.e. if and only if it is a degree-3 sink. □

Remark 5.4 (Threshold phenomenon)**.** *The degree bound $\Delta \le 2$ is the exact threshold for universal sharpness of the lower bound. Below that threshold, indegree concentrations of size at least 3 are impossible, so parity together with the indegree sum determines the source count exactly. At degree 3, a single sink can absorb three incoming edges and create slack. This is the first point at which the parity view and the source view diverge.*

6 The story of the result

The mathematical mechanism can be summarized in four steps.

1. The original variable S is difficult because it is an extremal local condition.
2. Passing to parity produces a linear shadow of the orientation, encoded by the parity vector $\eta \in \mathbf{F}_2^V$.
3. Connectedness forces this parity vector to be uniformly distributed on a codimension-one affine hyperplane, giving the exact law of E_0.
4. The directed handshaking lemma turns parity information back into a quantitative statement about sources by limiting how much indegree mass can be hidden among non-source vertices.

What the defect formula adds is that it identifies exactly what parity misses: vertices of indegree at least 3. Thus the lower bound is not merely a convenient estimate. It is the exact relation once indegree concentration is forbidden, and the entire failure of exactness is accounted for by a concrete, explicitly measurable obstruction.

For exposition, this is the natural narrative arc. One starts with a hard statistic, replaces it by a solvable parity shadow, and then recovers the original statistic through a mass-balance principle. The sharpness analysis closes the loop by showing precisely when no information was lost in passing to parity.