# Stationary diffusion on nonnegatively curved graphs*

Matěj Kripner Václav Rozhoň Robert Šámal
Martin Šindelář Adrián Zámečník

hexagon:2611.00052v1

**Abstract**

Lazy random walk on a bounded-degree graph of nonnegative Ollivier–Ricci curvature has diffusive second moment and Gaussian upper tails when started in stationarity. If the maximum degree is $D$, then $\mathbb{E}d(X_0, X_n)^2 \le 4Dn$, with optimal order of dependence on $D$. The results also hold for stationary random rooted infinite graphs. The proof combines known heat-kernel total-variation decay with a reversible energy identity for the conditional mean displacement. Its exponential form controls variation between starting vertices; martingale concentration then controls the walk itself.

## 1 Introduction

### 1.1 Background and the source problem

A lazy random walk on a cycle travels an expected distance of at most $\sqrt{n/2}$ in $n$ steps, independently of the length of the cycle. Indeed, lift the walk to the integers: its increments have mean zero and variance $1/2$, so Cauchy–Schwarz bounds its expected absolute position by $\sqrt{n/2}$, and projection to the cycle can only decrease distance. On an infinite regular tree of degree at least three, the excess of outward directions instead creates a positive drift away from the starting vertex. Bounded degree by itself therefore does not force diffusive motion. A geometric condition is needed.

Ollivier–Ricci curvature measures how the local geometry affects two nearby walks. Run one lazy step from each of two neighboring vertices. The curvature is nonnegative if the steps can be chosen dependently, while preserving each walk's transition law, so that their expected distance is at most one. Such a joint choice is called a coupling. Iterating these couplings keeps the expected separation from increasing. This discrete transport formulation of curvature was introduced by Ollivier [6, 7]; it is related to path coupling [9] and to the graph-curvature definition of Lin, Lu and Yau [8]. The condition compares *two moving walks*. It does not immediately bound either walk's distance from its fixed starting point.

Throughout the paper, $G$ is a connected undirected graph, $d$ is graph distance, and $P$ is the lazy random-walk kernel: with probability $1/2$ the walk stays put, and otherwise it follows a uniformly chosen incident edge. Loops and multiple edges are permitted; the precise convention is given in Section 2. For probability measures $\rho, \sigma$, let $W_1(\rho, \sigma)$ be the infimum of $\mathbb{E}d(U, V)$ over couplings with these marginals.

**Definition 1.1** (Nonnegative curvature; [1, Introduction, condition 3])**.** The graph has nonnegative Ollivier–Ricci curvature if

$$W_1(P(x,\cdot), P(y,\cdot)) \le 1 \qquad (x \sim y). \tag{1}$$

A second ingredient is the choice of starting vertex. On a finite graph with an edge, the stationary probability is $\pi(x) = \deg(x)/\sum_v \deg(v)$. This is the distribution seen at any fixed time

*This work was produced within the Bolzano project (https://bolzano.app), using AI systems for mathematical research and manuscript preparation. The listed authors contributed to the project and the preparation of this release. The manuscript received only limited human checking.

by a walk in equilibrium; on a regular graph it is uniform. Expectations with subscript $\pi$ average both this starting vertex and the subsequent walk. A bound uniform in the size of the graph is much stronger than the trivial bound by its diameter.

Salez [3] established sublinear stationary displacement under nonnegative curvature and a bounded-degree assumption. Brena, Hutchcroft and Münch [1, Introduction, p. 3] ask whether the square-root scale suggested by Brownian motion on manifolds of nonnegative Ricci curvature holds in this setting. Their stationary question can be formulated as follows.

**Conjecture 1.2** (Stationary first-moment question from [1])**.** For each $D$, there is a constant $C(D)$ such that every finite graph of maximum degree at most $D$ satisfying (1) obeys $\mathbb{E}_\pi d(X_0, X_n) \le C(D)\sqrt{n}$ for every $n \ge 1$.

The source proves a diffusive second-moment bound for transitive graphs, where all vertices have the same rooted geometry, and an averaged bound with a logarithmic loss for general finite graphs [1, Corollary 1.2 and Theorem 1.3]. Removing that loss requires control of how displacement varies between starting vertices. Concentration of a walk around its own conditional mean does not supply such control when those means are different.

## 1.2 Contribution

The stationary question has an affirmative answer, with diffusive moments of every order and a Gaussian upper tail. The quantitative bounds below also apply without transitivity. The proof separates fluctuations of the walk from variation between starting vertices. A reversible energy identity controls the latter: the square of a change in conditional mean absorbs a mixed distance term. Combined with heat-kernel smoothing, this forces the diffusive scale.

1. Theorem 1.3 gives the finite-graph second-moment and tail bounds. Proposition 6.1 shows that the factor $D$ in the second moment has the correct order.

2. Theorem 1.4 extends both estimates to stationary random rooted graphs, including infinite graphs. Curvature supplies the reversibility needed for this extension.

3. Proposition 5.1 bounds the conditional moments at most roots simultaneously at all times, with only an iterated-logarithm loss. The bound from every prescribed root with a constant depending only on $D$ remains open.

**Theorem 1.3.** *Let $G$ be a finite connected graph of maximum degree at most $D \ge 1$ satisfying (1), and start the lazy walk from $\pi$. For every integer $n \ge 1$,*

$$\mathbb{E}_\pi d(X_0, X_n)^2 \le 4Dn, \qquad \mathbb{E}_\pi d(X_0, X_n) \le 2\sqrt{Dn}. \tag{2}$$

*Moreover, for every $u \ge 0$,*

$$\mathbb{P}_\pi\left(d(X_0, X_n) \ge \sqrt{Dn/2} + u\right) \le 3\exp\left(-\frac{u^2}{8n}\right). \tag{3}$$

*A one-vertex graph is interpreted as an absorbing walk and has zero displacement.*

For the infinite version, stationarity concerns the graph as *viewed from the walk*. A random rooted graph $(G, o)$ has a stationary law $\mu$ if moving its root from $o$ to $X_1$ does not change its law. The integer lattice rooted at the origin is an example, since its rerootings are isomorphic. Reversibility means that the doubly rooted graphs $(G, o, X_1)$ and $(G, X_1, o)$ have the same law.

**Theorem 1.4.** *Let $\mu$ be a stationary law on connected random rooted graphs of maximum degree at most $D \ge 1$. Suppose (1) holds almost surely at every edge. Then $\mu$ is reversible, and (2)–(3) hold with expectations and probabilities over $(G, o) \sim \mu$ and the walk started at $o$. The graphs may be infinite.*

The averaging is essential to these statements. In particular, (3) is an *annealed* bound: it averages both the rooted graph and the walk. It is not a bound on deviations around the conditional mean at each individually chosen root. The conjecture $\mathbb{E}_x d(x, X_n)^2 \le C(D)n$ from every vertex, stated separately in [1, Remark 1.4], is not resolved here.

### 1.3 The mechanism and related work

The key step controls the distribution, over the root, of the conditional mean displacement. To describe it, run the nonlazy walk at rate one in continuous time, and set $a_t(x) = \mathbb{E}_x d(x, Y_t)$. For a stationary reversible oriented edge $(x, y)$, write $v = a_t(y) - a_t(x)$. Differentiating the stationary average of $a_t^2$ produces a negative energy term $-v^2$. Reversing the initial edge turns the remaining unbounded distance into $d(x, z) - d(y, z)$, whose absolute value is at most one. The only mixed term has the form $-vu$ with $|u| \le 1$, and $-v^2 - vu \le 1/4$. Thus variation between roots can be absorbed by the energy rather than bounded separately. Applying the same identity to $e^{\theta a_t}$ gives Gaussian control of those conditional means. A heat martingale then handles fluctuations of the walk around its mean.

The heat-kernel input is established work. Münch [2, Corollary 3.1.2, preprint numbering] proves that neighboring heat kernels approach each other in total variation at rate $O(\sqrt{D/t})$. Salez [3, Section 3.2] develops the coupling method, and Münch and Salez [5, Theorem 4] give an explicit finite discrete-time estimate. We use Münch's continuous-time form because differentiation leaves no discrete Taylor remainder in the energy identity. An independent coupling estimate needed for the random-rooted reversibility argument is included in Appendix A.

Stationarity of a random rooted graph does not by itself imply reversibility. Salez's zero-entropy theorem [3, Theorem 11] and the entropy theory of Benjamini and Curien [4, Proposition 3.6 and Theorem 4.4] supply the relevant principle. Appendix B gives the required entropy-production argument, counting vertices with their multiplicities under graph isomorphisms. This principle and the heat-kernel estimate are background ingredients; the reversible squared and exponential energy estimates are the displacement mechanism used here.

For general finite graphs, the uniformly averaged second moment in [1, Theorem 1.3] is bounded by $C(D)n(\log n)^{3/2}$. Since the uniform vertex measure is at most $D\pi$, Theorem 1.3 also gives a uniform-vertex bound $4D^2n$, removing this logarithmic loss. Zhou's pointwise estimate [13, Theorem 2.1 and Corollary 2.2] retains a subpolynomial loss in time; it applies to every vertex, a stronger quantifier than stationary averaging. Hutchcroft and Lopez [12, Theorem 2.1] give upper tails relative to the *supremum* of the expected displacement over all starting vertices. A stationary mean does not control that supremum, so their theorem does not directly give (3).

Transition-probability bounds such as the Carne–Varopoulos estimate [10] do not alone settle the question, since a graph may have many vertices at a given distance. Diffusive lower bounds, such as those of Lee and Peres [11] and Münch and Salez [5, Theorem 3 and Corollary 5], address the complementary question of slow escape. In the transitive setting the polynomial-growth theorem of [1] and the rigidity theorem of Brena and Bruè [14, Theorem 1.9] give stronger geometric information. The proof below obtains stationary diffusion directly, without a geometric classification.

## 2 Heat evolution and stationary edge reversal

Let $a(x, y)$ count oriented edges from $x$ to $y$, with each loop contributing two orientations, and put $\deg(x) = \sum_y a(x, y)$. On a graph with at least two vertices define

$$R(x, y) = \frac{a(x, y)}{\deg(x)}, \qquad P = \frac{I + R}{2}, \qquad L = R - I, \qquad S_t = e^{tL}.$$

Thus $R$ is the nonlazy kernel, $P$ is the discrete kernel in the main theorems, and $S_t$ is the heat semigroup of a rate-one Poissonization of $R$. A loop may produce a zero-distance step. On each fixed graph, degree measure is reversible for all these kernels.

Condition (1), the triangle inequality for $W_1$ along a geodesic, and composition of couplings imply $W_1(P^n(x,\cdot), P^n(y,\cdot)) \le d(x,y)$. Hence $P^n$ preserves 1-Lipschitz functions. The identity $S_t = e^{2t(P-I)}$ expresses $S_t$ as a Poisson mixture of these powers, so it also preserves 1-Lipschitz functions.

**Lemma 2.1** (Heat-kernel input)**.** *For every $t > 0$ and neighboring vertices $x, y$,*

$$\|S_t(x,\cdot) - S_t(y,\cdot)\|_{\mathrm{TV}} \le \varepsilon_t := \min\left\{1, \frac{\sqrt{D}}{2\sqrt{t}}\right\}, \qquad \|\rho - \sigma\|_{\mathrm{TV}} = \frac{1}{2}\sum_z |\rho(z) - \sigma(z)|. \tag{4}$$

*Proof.* Apply [2, Theorem 2.5.1 and Corollary 3.1.2] to the jump rates $q(x,y) = R(x,y)$ for $x \neq y$, with $q(x,x) = 0$. These rates are reversible with respect to degree measure, and every positive rate is at least $1/D$. The source allows countable locally finite graphs. Here total jump rates are at most one, so the heat process is nonexplosive. Its infinitesimal curvature condition follows from (1): if $f$ is 1-Lipschitz and $f(y) - f(x) = 1$ on an edge, then $Pf(y) - Pf(x) \le 1$ gives $Lf(y) - Lf(x) \le 0$. The cited result bounds the $\ell^1$ distance of heat rows by $d(x,y)/\sqrt{tq_{\min}}$. Our total variation is half that distance. The bound by one completes (4). □

For the rest of the main proof, write $\mu$ for a stationary reversible root law, including the finite law $\pi$. In the setting of Theorem 1.4, Lemmas B.1 and B.2 in Appendix B establish these properties before any displacement estimate is used. Stationarity and reversibility for $P$ imply both properties for $R$, by subtracting the diagonal part of $P = (I + R)/2$. Write $\mathbb{E}_e$ for the expectation over the oriented $R$-edge: first sample $(G, x) \sim \mu$, then $y \sim R_G(x,\cdot)$. Its two root marginals are $\mu$, and its law is invariant under $(x,y) \mapsto (y,x)$. Graph arguments will be suppressed in the formulas.

Define

$$a_t(x) = \sum_z S_t(x,z)d(x,z), \qquad B(t) = \int_0^t \varepsilon_s\, ds \le \sqrt{Dt}. \tag{5}$$

The last inequality follows by integrating $\varepsilon_s \le \sqrt{D}/(2\sqrt{s})$. The Poisson step count gives $0 \le a_t(x) \le t$, uniformly in the graph and root.

**Lemma 2.2** (Reversible energy identity)**.** *Fix $t \ge 0$. For an oriented edge put*

$$a = a_t(x), \quad b = a_t(y), \quad v = b - a, \quad g(z) = d(x,z) - d(y,z), \quad \nu_x = S_t(x,\cdot), \quad u = \nu_x(g).$$

*Then $|g| \le 1$ and $|u| \le 1$. For a continuously differentiable function $\phi$ on $[0,\infty)$,*

$$\frac{d}{dt}\mathbb{E}_\mu \phi(a_t) = \frac{1}{2}\mathbb{E}_e\left[\phi'(a)(\nu_y - \nu_x)(g) - (\phi'(b) - \phi'(a))(v + u)\right]. \tag{6}$$

*Proof.* The backward heat equation, applied with the starting vertex fixed, gives

$$a_t'(x) = La_t(x) + \beta_t(x), \qquad \beta_t(x) = \sum_y R(x,y)S_t(y,\cdot)(d(x,\cdot) - d(y,\cdot)). \tag{7}$$

Indeed, expanding $La_t(x) + \beta_t(x)$ cancels the terms involving $d(y,z)$ and leaves

$$\sum_{y,z} R(x,y)S_t(y,z)d(x,z) - a_t(x).$$

For unbounded distance, write $a_t(x) = e^{-t}\sum_{k\ge 0} t^k r_k(x)/k!$, where $r_k(x) = \mathbb{E}_x d(x, Z_k)$ for the $R$-walk. Since $0 \le r_k \le k$, this series may be differentiated uniformly on bounded time intervals. Coupling successive endpoints on the same walk gives $|r_{k+1} - r_k| \le 1$, hence $|a_t'(x)| \le 1$. These bounds justify both the heat equation and differentiation under $\mathbb{E}_\mu$.

Edge reversal gives the two identities

$$\begin{aligned}\mathbb{E}_\mu \phi'(a_t) L a_t &= -\frac{1}{2}\mathbb{E}_e[(\phi'(b)-\phi'(a))v],\\ \mathbb{E}_\mu \phi'(a_t)\beta_t &= \frac{1}{2}\mathbb{E}_e[\phi'(a)\nu_y(g) - \phi'(b)\nu_x(g)]\\ &= \frac{1}{2}\mathbb{E}_e[\phi'(a)(\nu_y-\nu_x)(g) - (\phi'(b)-\phi'(a))u].\end{aligned}$$

Their sum is (6). All quantities are integrable: $a_t \le t$, and $g, u$ are bounded. The same calculation is valid with expectation over a random rooted graph, since it uses only the symmetry and marginals of the oriented edge law. □

## 3 Diffusive second moments

**Proposition 3.1.** *For the rate-one heat walk from a stationary reversible root,*

$$\mathbb{E}_\mu a_t \le B(t), \qquad \mathbb{E}_\mu a_t^2 \le B(t)^2 + \frac{t}{4}, \qquad \mathbb{E}_\mu d(Y_0, Y_t)^2 \le \left(D + \frac{5}{4}\right) t. \tag{8}$$

*Proof.* Apply Lemma 2.2 first to $\phi(a) = a$. Since $|(\nu_y - \nu_x)(g)| \le 2\varepsilon_t$, it gives $|\frac{d}{dt}\mathbb{E}_\mu a_t| \le \varepsilon_t$. The mean starts at zero, proving the first bound.

Next let $A(t) = \mathbb{E}_\mu a_t^2$ and take $\phi(a) = a^2$ in the same identity:

$$A'(t) = \mathbb{E}_e[a(\nu_y - \nu_x)(g) - v^2 - vu]. \tag{9}$$

The first term is at most $2\varepsilon_t \mathbb{E}_\mu a_t$; the remaining terms satisfy

$$-v^2 - vu = -(v + u/2)^2 + u^2/4 \le \frac{1}{4}.$$

Consequently $A'(t) \le 2\varepsilon_t B(t) + 1/4 = (B(t)^2)' + 1/4$. Integrating from zero proves the second bound.

For the last bound, fix a root $x$ and set $f(z) = d(x, z)$. The backward heat martingale $M_s = S_{t-s} f(Y_s)$, $0 \le s \le t$, has jumps of absolute value at most one, because $S_{t-s} f$ is 1-Lipschitz. Its predictable quadratic variation has rate

$$\sum_y R(Y_s, y)\big(S_{t-s}f(y) - S_{t-s}f(Y_s)\big)^2 \le 1.$$

Therefore $\mathrm{Var}_x\, d(x, Y_t) \le t$. One can first use $f \wedge K$ and then let $K \to \infty$; the Poisson step count has finite second moment and dominates the distance. Averaging the conditional variance identity gives

$$\mathbb{E}_\mu d(Y_0, Y_t)^2 \le A(t) + t \le B(t)^2 + \frac{5t}{4} \le \left(D + \frac{5}{4}\right) t.$$

□

*Proof of the moment bounds in Theorems 1.3 and 1.4.* Use one nonlazy $R$-walk $(Z_j)_{j\ge 0}$ from the stationary root. Independently of the walk and of one another, choose $K \sim \mathrm{Bin}(n, 1/2)$ and $N \sim \mathrm{Pois}(n/2)$. Then $Z_K$ has the $P^n$ law and $Z_N$ has the $S_{n/2}$ law. Pathwise,

$$d(Z_0, Z_K) \le d(Z_0, Z_N) + |K - N|. \tag{10}$$

Both clocks have mean $n/2$, and $\mathbb{E}(K-N)^2 = n/4 + n/2 = 3n/4$. Using $(r+s)^2 \le 2r^2 + 2s^2$ and Proposition 3.1 yields

$$\mathbb{E}_\mu d(X_0, X_n)^2 \le 2\left(D + \frac{5}{4}\right)\frac{n}{2} + 2\frac{3n}{4} = \left(D + \frac{11}{4}\right) n \le 4Dn.$$

Cauchy–Schwarz gives the first moment. Finite stationary walks are reversible by degree detailed balance; the random-rooted case uses Appendix B, as explained above. Restoring any mass on a one-vertex graph contributes zero displacement. □

## 4 Exponential energy and Gaussian tails

**Proposition 4.1.** *For every $t \geq 0$ and $\theta \geq 0$,*

$$\mathbb{E}_\mu e^{\theta a_t} \leq \exp\left(\theta B(t) + \frac{\theta^2 t}{8}\right). \tag{11}$$

*In particular, $\mu\{a_t \geq B(t) + u\} \leq e^{-2u^2/t}$ for $t > 0$ and $u \geq 0$.*

*Proof.* Set $\phi(a) = e^{\theta a}$ and $F(t) = \mathbb{E}_\mu \phi(a_t)$. The first term of (6) is bounded above by $\theta \varepsilon_t F(t)$. Define, including when $a = b$,

$$h(a, b) = \theta^2 \int_0^1 e^{\theta((1-s)a+sb)}\, ds.$$

Then $\phi'(b) - \phi'(a) = h(a, b)v$, and convexity of the exponential gives $0 \leq h(a, b) \leq \theta^2(\phi(a) + \phi(b))/2$. Completing the square,

$$-\frac{1}{2}h(a, b)(v^2 + vu) \leq \frac{1}{8}h(a, b)u^2.$$

Its edge expectation is at most $\theta^2 F(t)/8$, because $|u| \leq 1$ and both edge marginals are $\mu$. Hence $F'(t) \leq (\theta\varepsilon_t + \theta^2/8)F(t)$. Integration with $F(0) = 1$ proves (11). Markov's inequality with $\theta = 4u/t$ proves the tail assertion. □

**Lemma 4.2** (Fluctuations at a fixed root)**.** *For every fixed rooted graph satisfying (1), $t \geq 0$, and $\theta \geq 0$,*

$$\mathbb{E}_x \exp\{\theta(d(x, Y_t) - a_t(x))\} \leq \exp\{t(e^\theta - 1 - \theta)\}. \tag{12}$$

*Proof.* Use the backward heat martingale from the proof of Proposition 3.1, first with bounded terminal function $f \wedge K$. Each possible jump $w$ lies in $[-1, 1]$, and

$$e^{\theta w} - 1 - \theta w \leq (e^\theta - 1 - \theta)w^2.$$

For $w \geq 0$ this follows term by term from the exponential series. For $w < 0$ the integral Taylor remainder is at most $\theta^2 w^2/2$, which is bounded by the right-hand side. To make the exponential estimate explicit, write $H_s = S_{t-s}(f \wedge K)$ and $w_y = H_s(y) - H_s(x)$. The backward equation $\partial_s H_s = -LH_s$ gives

$$(\partial_s + L)e^{\theta H_s}(x) = e^{\theta H_s(x)} \sum_y R(x, y)(e^{\theta w_y} - 1 - \theta w_y) \leq (e^\theta - 1 - \theta)e^{\theta H_s(x)}.$$

Take expectations along $Y_s$ and integrate this differential inequality from 0 to $t$. Since $H_0(x) = S_t(f \wedge K)(x)$, the result is (12) for $f \wedge K$. Let $K \to \infty$. The terminal distance is dominated by the Poisson step count, which has every exponential moment, so dominated convergence proves the claimed formula. □

Combining the conditional bound with Proposition 4.1 gives

$$\mathbb{E}_\mu e^{\theta d(Y_0, Y_t)} \leq \exp\left\{\theta B(t) + t\left(\frac{\theta^2}{8} + e^\theta - 1 - \theta\right)\right\}. \tag{13}$$

For $0 \leq \theta \leq 1$, the exponential series yields $e^\theta - 1 - \theta \leq (e - 2)\theta^2$, and $1/8 + e - 2 < 1$. Chernoff's bound with $\theta = v/(2t)$ consequently gives

$$\mathbb{P}_\mu(d(Y_0, Y_t) \geq B(t) + v) \leq e^{-v^2/(4t)} \qquad (0 \leq v \leq 2t). \tag{14}$$

The range matters: an unbounded Poisson clock does not have a Gaussian far tail. The bounded discrete-time walk permits the full tail estimate in the main theorems.

*Proof of the tail bound in Theorems 1.3 and 1.4.* Use the independent clocks $K, N$ in (10). For $|\theta| \le 1$, their exact moment generating functions give

$$\begin{aligned}\log \mathbb{E}e^{\theta(K-N)} &= n \log\cosh(\theta/2) + \frac{n}{2}(e^{-\theta} - 1 + \theta) \\ &\le \frac{n\theta^2}{8} + \frac{n}{2}(e-2)\theta^2 \le \frac{n\theta^2}{2}.\end{aligned}$$

Here $\log\cosh z \le z^2/2$ follows by integrating $|\tanh z| \le |z|$. Applying Markov's inequality with $\theta = \pm w/n$ proves

$$\mathbb{P}(|K - N| \ge w) \le 2e^{-w^2/(2n)} \qquad (0 \le w \le n). \tag{15}$$

If $0 \le u \le n$, split the excess in (10) into two parts of size $u/2$. Use (14) at $t = n/2$ and $v = u/2$, and (15) at $w = u/2$. Since $B(n/2) \le \sqrt{Dn/2}$, the union bound gives

$$\mathbb{P}_\mu\left(d(X_0, X_n) \ge \sqrt{Dn/2} + u\right) \le e^{-u^2/(8n)} + 2e^{-u^2/(8n)}.$$

No independence between the heat displacement and the clock difference is required. If $u > n$, the event is empty because $d(X_0, X_n) \le n$. □

**Corollary 4.3.** *Under either main theorem, for every real $p \ge 1$,*

$$\left(\mathbb{E}d(X_0, X_n)^p\right)^{1/p} \le \sqrt{Dn/2} + \left[3(8n)^{p/2}\Gamma(p/2+1)\right]^{1/p} \le C\left(\sqrt{Dn} + \sqrt{pn}\right),$$

*where $C$ is universal.*

*Proof.* Integrate (3) using $\mathbb{E}V^p = \int_0^\infty pu^{p-1}\mathbb{P}(V > u)\,du$ for $V = (d(X_0, X_n) - \sqrt{Dn/2})_+$. The change of variables $s = u^2/(8n)$ gives the displayed gamma factor. Minkowski's inequality adds the deterministic shift. For completeness, if $r = p/2 \ge 1/2$, then $s^r e^{-s/2} \le (2r/e)^r$, so $\Gamma(r+1) = \int_0^\infty s^r e^{-s}\,ds \le 2(2r/e)^r$. Taking $p$th roots gives the last inequality with a universal constant. □

## 5 Conditional moments at typical roots

Let $b_n(x) = \mathbb{E}_x d(x, X_n)$ for the discrete lazy walk, and put $c_D = \sqrt{D/2} + \sqrt{3}/2$. Taking conditional expectations in (10), then using Cauchy–Schwarz for the clocks, gives $b_n(x) \le a_{n/2}(x) + \sqrt{3n/4}$. Proposition 4.1 therefore implies the following root-distribution bound for every $u \ge 0$:

$$\mu\{b_n \ge c_D\sqrt{n} + u\} \le e^{-4u^2/n}. \tag{16}$$

This controls how the conditional mean varies between roots. The next proposition turns it into a simultaneous statement in time.

**Proposition 5.1.** *For every $0 < \delta < 1$, outside a set of rooted graphs of $\mu$-probability at most $\delta$, the bounds*

$$b_n(o) \le (2+\sqrt{2})\sqrt{n}\left[c_D + \frac{1}{2}\sqrt{\log\frac{\pi^2(1+\lfloor\log_2 n\rfloor)^2}{6\delta}}\right], \tag{17}$$

$$\mathbb{E}_o d(o, X_n)^2 \le Cn\left[D + \log(1/\delta) + \log\log(n+3)\right] \tag{18}$$

*hold for all integers $n \ge 1$, with a universal $C$.*

*Proof.* Wasserstein contraction gives subadditivity at a fixed root:

$$\begin{aligned}b_{m+n}(x) &= W_1(\delta_x, \delta_x P^{m+n}) \\ &\le W_1(\delta_x, \delta_x P^m) + W_1(\delta_x P^m, (\delta_x P^n)P^m) \le b_m(x) + b_n(x).\end{aligned}$$

In (16) set $n = 2^j$ and $u = \frac{1}{2}\sqrt{2^j \log(\pi^2(j+1)^2/(6\delta))}$. The sum of the resulting exceptional probabilities is $\delta$. On the complementary event, write $n$ in binary and apply subadditivity. For $J = \lfloor \log_2 n \rfloor$, the logarithmic factor is bounded by its value at $J$, while $\sum_{j=0}^{J} 2^{j/2} \le (2+\sqrt{2})\sqrt{n}$. This proves (17). No monotonicity of $b_n$ is used.

Finally the discrete backward martingale $P^{n-j} f(X_j)$, with $f(z) = d(x,z)$, has conditional variance per step at most $1/2$: its conditional variance is at most the mean squared difference from $P^{n-j} f(X_{j-1})$, and a lazy step moves distance one with probability at most $1/2$. Each smoothed function remains 1-Lipschitz. Summing martingale variances gives $\mathrm{Var}_x\, d(x, X_n) \le n/2$. Adding the square of (17) proves (18). □

For a finite graph, (16) also gives, for every vertex and time,

$$b_n(x) \le \sqrt{n}\left(c_D + \frac{1}{2}\sqrt{\log(1/\pi(x))}\right).$$

Indeed, if this inequality failed, a slightly smaller threshold than $b_n(x)$ in (16) would contain $x$ but have probability strictly smaller than $\pi(x)$. With $N$ vertices, $\pi(x) \ge 1/(ND)$, so the conditional second moment is at most $Cn(D + \log N)$. The dependence on $N$ distinguishes this consequence from the open pointwise degree-only conjecture.

## 6 Sharp degree dependence

**Proposition 6.1.** *For every integer $D \ge 1$, any bounds $\mathbb{E}_\pi d(X_0, X_n) \le C_1(D)\sqrt{n}$ and $\mathbb{E}_\pi d(X_0, X_n)^2 \le C_2(D)n$, uniform over the graphs in Theorem 1.3, must satisfy*

$$C_1(D) \ge \frac{1 - e^{-1}}{2}\sqrt{D}, \qquad C_2(D) \ge \frac{(1-e^{-1})^2}{4} D.$$

*Proof.* Use the $D$-dimensional cube $\{0,1\}^D$ with Hamming distance. Couple two lazy walks by the same holding or coordinate-flipping action. Each action preserves their distance, so (1) holds. A fixed coordinate flips independently at each time with probability $1/(2D)$. Its probability of differing from its initial value after $n$ steps is $[1 - (1 - 1/D)^n]/2$. Thus, from every initial vertex,

$$\mathbb{E} d(X_0, X_n) = \frac{D}{2}(1 - (1 - 1/D)^n).$$

At $n = D$ this is at least $(1 - e^{-1})D/2$. Dividing by $\sqrt{D}$ proves the first assertion; Jensen's inequality proves the second. □

## A A self-contained coupling estimate

Set

$$Q = \frac{I + P}{2}, \qquad \alpha = \frac{1}{4D}.$$

The extra holding probability supplies the mass needed for the coupling switch below. This auxiliary kernel is used only to establish sublinear entropy and reversibility of the rooted law. Thus $Q(x,x) \ge 3/4$, its nonholding probability is at most $1/4$, and $Q(x,a) \ge 1/(4D)$ for each distinct neighbor $a$ of $x$. By the triangle inequality for $W_1$, applied along a geodesic, (1) implies $W_1(P(x,\cdot), P(y,\cdot)) \le d(x,y)$ for all $x, y$. Mixing such a coupling with the point mass at $(x,y)$ gives the same inequality for $Q$.

**Lemma A.1.** *For distinct vertices $x, y$ with $r = d(x,y)$, there is a coupling $(U,V)$ of $Q(x,\cdot)$ and $Q(y,\cdot)$ satisfying*

$$\mathbb{E} d(U,V) \le r, \qquad \mathbb{P}(d(U,V) \le r-1) \ge \alpha, \qquad |d(U,V) - r| \le 2.$$

*Proof.* Choose a coupling $\gamma$ of transportation cost at most $r$, and a neighbor $a$ of $x$ on a geodesic to $y$. Every coupling of these marginals has $\gamma(x,y) \geq Q(x,x) + Q(y,y) - 1 \geq 1/2$. For every entry $(a,b)$ of positive mass with $d(a,b) \geq r$, replace its mass $\beta_b = \gamma(a,b)$ by the switch

$$(a,b),(x,y) \longmapsto (a,y),(x,b).$$

The total mass removed from $(x,y)$ is at most $Q(x,a) \leq 1/4$, so all switches are feasible and preserve the marginals. Since $d(a,y) = r-1$ and $d(x,b) \leq r+1$, each switch changes the cost by at most $\beta_b((r-1)+(r+1)-r-r) = 0$. Afterwards every entry in row $a$ has distance at most $r-1$. This row has mass $Q(x,a) \geq \alpha$. Finally each coordinate moves by distance at most one, giving the last assertion. All one-step supports are finite, so the proof also applies on infinite locally finite graphs. □

**Lemma A.2.** *Suppose a nonnegative integer-valued process $Z_k$, absorbed at zero, starts at $r \geq 1$. Before absorption its increments $\xi_k$ satisfy*

$$|\xi_k| \leq 2, \qquad \mathbb{E}(\xi_k \mid \mathcal{F}_k) \leq 0, \qquad \mathbb{E}(\xi_k^2 \mid \mathcal{F}_k) \geq \alpha.$$

*If $\tau$ is its absorption time, then*

$$\mathbb{P}(\tau > n) \leq \min\{1, 9r/\sqrt{\alpha n}\} \qquad (n \geq 1).$$

*Proof.* For an integer $R > r$, stop at $T = \inf\{k : Z_k = 0 \text{ or } Z_k \geq R\}$. The function $v(z) = z(2(R+2) - z)$ satisfies, before $T$,

$$\mathbb{E}[v(Z_{k+1}) - v(Z_k) \mid \mathcal{F}_k] = (2(R+2) - 2Z_k)\mathbb{E}(\xi_k \mid \mathcal{F}_k) - \mathbb{E}(\xi_k^2 \mid \mathcal{F}_k) \leq -\alpha.$$

The stopped process stays in $[0, R+2]$, where $v$ is nonnegative. Summing to a deterministic cutoff and then using monotone convergence yields $\mathbb{E}T \leq v(r)/\alpha \leq 2(R+2)r/\alpha$. The stopped supermartingale $Z$ is bounded, and consequently $\mathbb{E}Z_T \leq r$ and $\mathbb{P}(Z_T \geq R) \leq r/R$. Hence

$$\mathbb{P}(\tau > n) \leq \frac{r}{R} + \frac{2(R+2)r}{\alpha n}. \tag{19}$$

If $\alpha n \geq 1$ and $r < \lceil\sqrt{\alpha n}\rceil$, take $R = \lceil\sqrt{\alpha n}\rceil$. Then $R+2 \leq 4\sqrt{\alpha n}$, and (19) is at most $9r/\sqrt{\alpha n}$. In the other cases that expression is at least one, and the assertion follows from the trivial probability bound. □

**Proposition A.3.** *For every graph satisfying the hypotheses above, finite or infinite,*

$$\|Q^n(x,\cdot) - Q^n(y,\cdot)\|_{\mathrm{TV}} \leq \min\{1, 18\sqrt{D}\, d(x,y)/\sqrt{n}\} \qquad (n \geq 1). \tag{20}$$

*Proof.* On a fixed graph, choose the coupling of Lemma A.1 at each distinct ordered vertex pair, and identical transitions on the diagonal. The vertex set is countable, so these choices define a Markov coupling. Its distance is absorbed at zero, has nonpositive drift and jumps of size at most two, and has conditional second moment of its increments at least $\alpha$ before absorption. Lemma A.2 bounds the probability that the walks have not coalesced. The coupling inequality for total variation gives (20). No choice of coupling measurable as a function of the graph is needed: the resulting inequality holds on each fixed graph. □

# B Entropy and reversibility

Rooted graphs are considered up to rooted isomorphism, with the Borel structure generated by finite rooted neighborhoods. We give the details needed to use stationary edge reversal without presupposing reversibility. All vertex sums count actual vertices, even when different vertices give isomorphic rerootings. Such sums are Borel: restrict first to finite balls, use the monotone class theorem, and then take increasing limits. The event that the graph has one vertex is invariant under rerooting and contributes zero displacement. We may condition on its complement, so that all degrees lie between 1 and $D$.

## B.1 Saturation and entropy

Stationarity for $P$ implies stationarity for $Q = (I + P)/2$. Write $Y$ for the $Q$-walk. We will repeatedly use the following consequence of stationarity. If a rooted Borel event $A$ has full measure, then almost surely it holds at every vertex of the sampled graph. Indeed, for each $n$,

$$0 = \mu(A^c) = \mathbb{E}_\mu \sum_x Q_G^n(o, x)\mathbf{1}_{A^c}(G, x).$$

Intersect over $n$. Every vertex is reached with positive probability at some finite time, so no rerooting can belong to $A^c$. We call this the saturation property.

Set

$$H_n(G, x) = -\sum_z Q_G^n(x, z) \log Q_G^n(x, z), \qquad h_n = \mathbb{E}_\mu H_n(G, o).$$

The convention is $0 \log 0 = 0$. Since there are at most $(D+1)^n$ possible endpoints, all these entropies are finite and bounded by $n \log(D + 1)$.

**Lemma B.1.** *For $N \geq 1$,*

$$h_N \leq \log(D + 1)\big(1 + 36D^{3/2}\sqrt{N}\big).$$

*In particular, $h_N = o(N)$.*

*Proof.* On a fixed rooted graph let $A = Y_1$ and $Z = Y_{n+1}$. Conditional on $A = a$, the law of $Z$ is $\nu_a = Q_G^n(a, \cdot)$. The entropy chain rule gives

$$H_{n+1}(G, o) - \sum_a Q_G(o, a)H_n(G, a) = I(A; Z),$$

Here $I(A; Z) = H(Z) - H(Z \mid A)$ denotes mutual information on this fixed graph, with conditional entropy obtained by averaging the entropy of the conditional law. Averaging and using stationarity identifies the left side with $h_{n+1} - h_n$.

For $S = \{o\} \cup N_G(o)$ put $j(z) = \min_{a \in S} \nu_a(z)$ and $c = \sum_z j(z)$. Then (20) gives

$$1 - c \leq \sum_{a \sim o} \sum_z (\nu_o(z) - \nu_a(z))_+ = \sum_{a \sim o} \|\nu_o - \nu_a\|_{\mathrm{TV}} \leq \frac{18D^{3/2}}{\sqrt{n}}.$$

The channel from $A$ to $Z$ has common part $j$ of mass $c$. More explicitly, independently of $A$ choose a Bernoulli variable $L$ with $\mathbb{P}(L = 1) = 1 - c$. If $L = 0$, draw $Z$ from $j/c$, independently of $A$; if $L = 1$, use the residual law $(\nu_A - j)/(1 - c)$. The chain rule implies

$$I(A; Z) \leq I(A; Z, L) \leq (1 - c)H(A) \leq (1 - c) \log(D + 1).$$

When $c = 0$ or $c = 1$, the same estimate follows directly. Consequently $0 \leq h_{n+1} - h_n \leq 18D^{3/2} \log(D + 1)/\sqrt{n}$. Sum this inequality, starting with $h_1 \leq \log(D + 1)$. □

## B.2 Entropy and reversal of the rooted edge

The next argument is a counting-measure version of the Radon–Nikodym cocycle method in [4, Section 4].

**Lemma B.2.** *For a stationary law of graphs with degrees in $\{1, \ldots, D\}$, sublinear $Q$-entropy $h_n = o(n)$ implies reversibility for $Q$ and for $P$.*

*Proof.* On graphs with two ordered roots define the sigma-finite counting measures

$$\mathcal{M}(F) = \mathbb{E}_\mu \sum_y F(G, o, y), \qquad \check{\mathcal{M}}(F) = \mathbb{E}_\mu \sum_y F(G, y, o).$$

They are finite on sets where the roots have bounded distance. They are equivalent: if $\mathcal{M}(A) = 0$, saturation applied to $\sum_y \mathbf{1}_A(G, x, y)$ shows that this sum vanishes at every $x$ almost surely, and hence $\check{\mathcal{M}}(A) = 0$. Reversing the set proves the converse. Let $J = d\check{\mathcal{M}}/d\mathcal{M}$ be a positive finite Borel version outside its null exceptional set.

We record why $J$ is a cocycle. On triples use $\mathcal{T}(F) = \mathbb{E}_\mu \sum_{y,z} F(G, o, y, z)$, and write $S_{ij}$ for the exchange of coordinates $i, j$. Summing over the unused root shows

$$\frac{d(S_{12})_*\mathcal{T}}{d\mathcal{T}}(x, y, z) = J(x, y), \qquad \frac{d(S_{13})_*\mathcal{T}}{d\mathcal{T}}(x, y, z) = J(x, z), \qquad (S_{23})_*\mathcal{T} = \mathcal{T},$$

where the graph argument is suppressed. For an involution $S$ with density $j_S = dS_*\mathcal{T}/d\mathcal{T}$, the image of $f\mathcal{T}$ has density $(f \circ S)j_S$. Applying successively $S_{12}, S_{23}, S_{12}$ gives densities $J(x, y)$, then $J(x, z)$, then $J(y, z)J(x, y)$. Since this composition is $S_{13}$,

$$J(x, z) = J(x, y)J(y, z) \tag{21}$$

for almost every triple. On the diagonal, $J(x, x) = 1$. Counting and then saturating these full-measure events makes the identities valid simultaneously for all vertices of almost every sampled graph.

Define $\Delta(x, y) = \deg(x)J(x, y)/\deg(y)$. It obeys the same cocycle identity and $\Delta(x, y)\Delta(y, x) = 1$. Let $\Lambda_n$ be the law of $(G, o, Y_n)$ and $\check{\Lambda}_n$ its reversal. Their densities with respect to $\mathcal{M}$ are $Q_G^n(x, y)$ and $J(x, y)Q_G^n(y, x)$. Degree detailed balance on each fixed graph implies

$$\deg(x)Q_G^n(x, y) = \deg(y)Q_G^n(y, x), \qquad \check{\Lambda}_n = \Delta\Lambda_n.$$

Both endpoint laws have first-root marginal $\mu$, by stationarity. It follows that

$$\sum_y Q_G^n(x, y)\Delta(x, y) = 1 \tag{22}$$

for every $x, n$, after intersecting and saturating the relevant full-measure events.

For a possible one-step transition, $Q_G(x, y), Q_G(y, x) \geq 1/(4D)$. Apply (22) at $x$ and at $y$, and use reciprocity, to obtain $(4D)^{-1} \leq \Delta(x, y) \leq 4D$. Thus its one-step logarithm is integrable. Set

$$\eta = -\mathbb{E}_\mu \log \Delta(o, Y_1) = \mathbb{E}_\mu[\Delta(o, Y_1) - 1 - \log \Delta(o, Y_1)] \geq 0.$$

The equality uses (22). The cocycle identity and stationarity along the walk give $-\mathbb{E}_\mu \log \Delta(o, Y_n) = n\eta$. On the other hand, the $o$ term of (22) at $Y_n$ gives

$$\Delta(o, Y_n) \geq Q_G^n(Y_n, o) = \frac{\deg(o)}{\deg(Y_n)} Q_G^n(o, Y_n) \geq D^{-1} Q_G^n(o, Y_n).$$

Taking negative logarithms and expectations gives $n\eta \leq h_n + \log D$. Sublinearity forces $\eta = 0$. Since $u - 1 - \log u$ is nonnegative and vanishes only at $u = 1$, we have $\Delta(o, Y_1) = 1$ almost surely. Hence $\Lambda_1 = \check{\Lambda}_1$, proving $Q$-reversibility. The $Q$ endpoint law is half the diagonal law plus half the $P$ endpoint law; the diagonal law is symmetric, so $P$ is reversible as well. □

*Remark* B.3 (Why vertices must be counted with multiplicity). Consider the three-leaf star rooted according to its degree-biased stationary probability. The center has probability $1/2$ and each leaf has probability $1/6$. On the isomorphism type whose first root is the center and whose second root is a leaf, $\mathcal{M}$ has mass $3/2$, whereas $\check{\mathcal{M}}$ has mass $1/2$. Thus $J = 1/3$, and the degree correction gives $\Delta = 3J = 1$, as reversibility requires. Counting only one representative of the three isomorphic leaves would give the wrong Radon–Nikodym derivative. The counting measures in the proof avoid this problem.

# References


[1] C. Brena, T. Hutchcroft and F. Münch, *Transitive graphs of non-negative Ollivier–Ricci curvature have polynomial growth*, arXiv:2512.03968v2 (2026), https://arxiv.org/abs/2512.03968v2.

[2] F. Münch, *Non-negative Ollivier curvature on graphs, reverse Poincaré inequality, Buser inequality, Liouville property, Harnack inequality and eigenvalue estimates*, J. Math. Pures Appl. **170** (2023), 231–257, https://doi.org/10.1016/j.matpur.2022.12.007; preprint https://arxiv.org/abs/1907.13514v1.

[3] J. Salez, *Sparse expanders have negative curvature*, Geom. Funct. Anal. **32** (2022), 1486–1513; preprint numbering from https://arxiv.org/abs/2101.08242v2.

[4] I. Benjamini and N. Curien, *Ergodic theory on stationary random graphs*, Electron. J. Probab. **17** (2012), no. 93, 1–20, https://doi.org/10.1214/EJP.v17-2401.

[5] F. Münch and J. Salez, *Mixing time and expansion of non-negatively curved Markov chains*, J. Éc. polytech. Math. **10** (2023), 575–590, https://doi.org/10.5802/jep.226.

[6] Y. Ollivier, *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256** (2009), 810–864, https://doi.org/10.1016/j.jfa.2008.11.001.

[7] Y. Ollivier, *A survey of Ricci curvature for metric spaces and Markov chains*, in Probabilistic Approach to Geometry, Adv. Stud. Pure Math. **57**, Math. Soc. Japan, Tokyo, 2010, 343–381, https://doi.org/10.2969/aspm/05710343.

[8] Y. Lin, L. Lu and S.-T. Yau, *Ricci curvature of graphs*, Tohoku Math. J. **63** (2011), 605–627, https://doi.org/10.2748/tmj/1325886283.

[9] R. Bubley and M. Dyer, *Path coupling: a technique for proving rapid mixing in Markov chains*, in Proc. 38th Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 1997, 223–231, https://doi.org/10.1109/SFCS.1997.646111.

[10] T. K. Carne, *A transmutation formula for Markov chains*, Bull. Sci. Math. (2) **109** (1985), 399–405.

[11] J. R. Lee and Y. Peres, *Harmonic maps on amenable groups and a diffusive lower bound for random walks*, Ann. Probab. **41** (2013), 3392–3419, https://doi.org/10.1214/12-AOP779.

[12] T. Hutchcroft and I. M. Lopez, *A relation between isoperimetry and total variation decay with applications to graphs of non-negative Ollivier–Ricci curvature*, arXiv:2411.04988v1 (2024), https://arxiv.org/abs/2411.04988v1.

[13] C. Zhou, *Pointwise subexponential growth and near-diffusive displacement on bounded-degree graphs with non-negative Ollivier–Ricci curvature*, arXiv:2607.27162v1 (2026), https://arxiv.org/abs/2607.27162v1.

[14] C. Brena and E. Bruè, *Ollivier–Ricci curvature on groups of polynomial growth*, arXiv:2608.14232v1 (2026), https://arxiv.org/abs/2608.14232v1.