Sparse expanders have negative curvature

Justin Salez*

February 2, 2021

Abstract

We prove that bounded-degree expanders with non-negative Ollivier-Ricci curvature do not exist, thereby solving a long-standing open problem suggested by Naor and Milman and publicized by Ollivier (2010). In fact, this remains true even if we allow for a vanishing proportion of large degrees, large eigenvalues, and negatively-curved edges. To establish this, we work directly at the level of Benjamini-Schramm limits, and exploit the entropic characterization of the Liouville property on stationary random graphs to show that non-negative curvature and spectral expansion are incompatible "at infinity". We then transfer this result to finite graphs via local weak convergence. The same approach also applies to the Bacry-Emery curvature condition $\mathrm{CD}(0,\infty)$, thereby settling a recent conjecture of Cushing, Liu and Peyerimhoff (2019).

Contents

*CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, 75016 Paris, France

1 Introduction

1.1 Non-negative curvature

The *Ricci curvature* of a manifold is a fundamental concept in Riemannian geometry, see e.g. [28]. In two celebrated works [42, 43], Ollivier proposed a notion of curvature based on optimal transport which applies to arbitrary metric spaces, hence in particular to the discrete setting of graphs. Specifically, let $G = (V_G, E_G)$ be a locally finite connected graph. As usual, write $\deg_G(x)$ for the degree of a vertex x, and $\mathrm{d}_G(x, y)$ for the length of a minimal path from x to y in G. Let also $P_G \colon V_G \times V_G \to [0, 1]$ denote the transition matrix of the lazy simple random walk on G, i.e.

$$P_G(x,y) \quad := \quad \begin{cases} \frac{1}{2\deg_G(x)} & \text{if } \{x,y\} \in E_G; \\ \frac{1}{2} & \text{if } x = y; \\ 0 & \text{else.} \end{cases}$$

The *Ollivier-Ricci curvature* at an edge $\{x, y\} \in E_G$ is defined as

$$\kappa_G(x,y) \quad := \quad 1 - \mathcal{W}_1\left(P_G(x,\cdot), P_G(y,\cdot)\right),$$

where $\mathcal{W}_1$ denotes the L^1-Wasserstein distance on $\mathcal{P}_1(V_G, \mathrm{d}_G)$, see (20) below. Note that the computation of $\kappa_G(x, y)$ amounts to solving a finite-dimensional linear optimization problem, and is therefore amenable to standard algorithmic techniques (see [22] for a beautiful interactive curvature calculator). The Ollivier-Ricci curvature of the whole graph is then defined as

$$\kappa(G) \quad := \quad \inf_{\{x,y\}\in E_G} \kappa_G(x,y).$$

This fundamental geometric quantity measures how distances are contracted, on average, under the action of P_G. When $\kappa(G) \geq 0$, the graph G is called *non-negatively curved*. This is the case, for example, when G is the Cayley graph of an abelian group, as witnessed by the obvious coupling that uses the same random generators for both trajectories. Non-negative curvature is equivalent to the requirement that P_G is a contraction under the Wasserstein metric $\mathcal{W}_1$, and constitutes the essence of the powerful *path coupling method* for bounding mixing times [18]. Consequences in terms of geometry, mixing, and concentration of measure have been massively investigated, and quantified by a variety of functional inequalities. The literature is too vast for an exhaustive account, and we refer the reader to the seminal papers [42, 43, 34, 30], the survey [44], and the more recent works [24, 41, 21, 32, 40] for details, variations, references, and open problems. In particular, the present work was motivated by the following long-standing question, due to Naor and Milman, and publicized by Ollivier [44, Problem T]. Recall that a *family of expanders* is a sequence of finite graphs with uniformly bounded degrees, diverging sizes, and spectral gap bounded away from 0.

Question 1 (Problem T in [44])**.** *Is there a family of non-negatively curved expanders ?*

An instructive special class of graphs for which non-negative curvature is completely understood is that of cubic graphs. Specifically, it was shown in [22] that prism graphs and Möbius ladders are the only cubic graphs with non-negative Ollivier-Ricci curvature. Since these are not expanders, the answer to Question 1 is negative for cubic graphs. To the best of our knowledge, this is the only result in the direction of Question 1, despite the rich body of works on non-negative curvature.

1.2 Main result

In the present paper, we answer Question 1 negatively in full generality, as well as its $\mathrm{CD}(0,\infty)$ analogue raised by Cushing, Liu and Peyerimhoff [23, Conjecture 9.11], see Remark 1 below. Moreover, we show that the answer to Question 1 remains negative even if we significantly relax the required properties. Specifically, denote by $\Delta(G)$ the maximum degree of a finite graph G, and by

$$1 \;=\; \lambda_1(G) \;\geq\; \lambda_2(G) \;\geq \ldots \;\geq\; \lambda_N(G) \;\geq\; 0,$$

the $N = |V_G|$ ordered eigenvalues of its transition matrix P_G. With these notations, Question 1 simply asks whether there exist constants $\Delta \geq 1$, $\rho < 1$ and arbitrary large graphs satisfying

(A) sparsity: $\Delta(G) \leq \Delta$;

(B) spectral expansion: $\lambda_2(G) \leq \rho$;

(C) non-negative curvature: $\kappa(G) \geq 0$.

Our main result says that no large graph can even come close to satisfying these three requirements.

Theorem 2 (Main result)**.** *Fix $\Delta \geq 1$ and $\rho \in (0,1)$. Then, there exists a constant $\varepsilon = \varepsilon_{\Delta,\rho} > 0$ such that* every *finite graph G must satisfy one of the following conditions:*

- *either G is far from satisfying the sparsity requirement (A), in the following sense:*

$$\sum_{x \in V_G} \deg_G(x) \log \deg_G(x) \quad > \quad (\Delta \log \Delta)|V_G|;$$

- *or G is far from satisfying the expansion requirement (B), in the following sense:*

$$\mathrm{card}\{i \colon \lambda_i(G) > \rho\} \quad \geq \quad \varepsilon |V_G|;$$

- *or G is far from satisfying the curvature requirement (C), in the following sense:*

$$\mathrm{card}\{e \in E_G \colon \kappa_G(e) < -\varepsilon\} \quad \geq \quad \varepsilon |E_G|.$$

Note that the conclusion is only meaningful for large graphs, since the second condition is trivially satisfied when $|V_G| \leq \frac{1}{\varepsilon}$. Here is an equivalent – but perhaps more intuitive – formulation.

Theorem 3 (Rephrasing)**.** *Let $G_n = (V_n, E_n), n \geq 1$ be finite graphs with the sparsity property*

$$\sup_{n\geq 1} \left\{ \frac{1}{|V_n|} \sum_{x \in V_n} \deg_{G_n}(x) \log \deg_{G_n}(x) \right\} \quad < \quad \infty. \tag{1}$$

Suppose in addition that the Ollivier-Ricci curvature is almost non-negative on most edges, i.e.

$$\forall \varepsilon > 0, \quad \frac{1}{|E_n|} \operatorname{card}\{e \in E_n \colon \kappa_{G_n}(e) < -\varepsilon\} \quad \xrightarrow[n\to\infty]{} \quad 0. \tag{2}$$

Then, a macroscopic proportion of eigenvalues of the transition matrix must accumulate near 1*:*

$$\forall \rho < 1, \quad \liminf_{n\to\infty} \left\{ \frac{1}{|V_n|} \operatorname{card}\{i \colon \lambda_i(G_n) \geq \rho\} \right\} \quad > \quad 0. \tag{3}$$

Here again, the theorem is only meaningful in the large-size limit $|V_n| \to \infty$, since the conclusion (3) trivially holds otherwise. The high-level message is that on large sparse graphs, non-negative curvature (in an even weak sense) induces extremely poor spectral expansion. This stands in stark contrast with the traditional idea – quantified by a broad variety of functional inequalities over the past decade – that non-negative curvature is associated with *good* mixing behavior.

Remark 1 (Bacry-Emery curvature)**.** *Bacry and Emery [7, 8, 9] developed a different notion of non-negative curvature based on $\Gamma-$calculus and known as the* $\mathrm{CD}(0,\infty)$ *condition, see also [33, 26]. Since this notion is local, our proof also applies, with the role of Theorem 11 being played by a recent result of Hua [27, Theorem 2]. Consequently, there is no family of expanders satisfying* $\mathrm{CD}(0,\infty)$*, as conjectured by Cushing, Liu and Peyerimhoff [23, Conjecture 9.11]. We note that the weaker statement obtained by replacing* $\mathrm{CD}(0,\infty)$ *with* $\mathrm{CD}(0,n)$ *was recently established by Münch [39]. We warmly thank David Cushing, Shiping Liu and Florentin Münch for pointing this out.*

Remark 2 (Laziness)**.** *The literature actually contains a whole family of variants $(\kappa_\alpha)_{\alpha\in[0,1)}$ of the Ollivier-Ricci curvature κ, obtained by replacing the matrix P_G with its $\alpha-$idle version:*

$$P_G^{(\alpha)} \quad := \quad (2-2\alpha)P_G + (2\alpha - 1)\,\mathrm{Id}.$$

There is even a continuous-time version $\kappa_\star := \lim_{\alpha\to 1} \frac{\kappa_\alpha}{1-\alpha}$, proposed in [34] and largely adopted since then. In fact, it was later shown (see [19, Remark 5.4]) that $\frac{\kappa_\alpha}{1-\alpha} \leq \kappa_\star \;=\; 2\kappa$, where $\kappa = \kappa_{1/2}$ is the version considered in the present paper. Consequently, our result is stated in the strongest possible form, and applies to all versions of the Ollivier-Ricci curvature.

Remark 3 (Eigenvectors)**.** *Our proof will actually reveal more than (3): not only are there many eigenvalues near* 1*, but the corresponding eigenvectors furthermore charge most vertices significantly. In other words, the poor spectral expansion of non-negatively curved graphs is not restricted to any specific region: it applies everywhere. See Remark 6 for a precise statement.*

1.3 Proof outline

Proof outline. The most natural route towards Question 1 would consist in looking for a quantitative upper-bound on the spectral gap of a finite non-negatively curved graph, in terms of its size and maximum degree. Interestingly, we do *not* pursue this approach here. Neither do we try to obtain asymptotic estimates along a sequence of sparse graphs $(G_n)_{n\geq 1}$ with non-negative curvature. Instead, we work directly at the elegant level of *local weak limits* of finite graphs, and exploit their built-in *stationarity* to prove that non-negative curvature and spectral expansion are incompatible "at infinity". This relies on the central concept of *asymptotic entropy*, and its classical relations with the Liouville property and the spectral radius. We then transfer this incompatibility result to finite graphs via a relative-compactness argument. As far as we know, the idea of using local weak limits as a tool to deduce generic bounds on the mixing parameters of sparse Markov chains have not received much attention. We firmly believe that this viewpoint will have many applications.

Further questions. The surprising "deg log deg" requirement (1) is used to define the asymptotic entropy on which our whole argument relies. We do not know whether it is necessary for the conclusion (3) to hold, or whether it can be further relaxed. Note that some degree restriction is necessary, since the complete graph satisfies $\lambda_2(G) = \kappa(G) = 1/2$, regardless of its size. Also, a drawback of our approach – as of any limit argument – is its non-quantitative nature. It would be interesting to find an explicit upper-bound (vanishing as $n \to \infty$) on the spectral gap of a non-negatively curved graph with n vertices and maximum degree Δ, i.e. to estimate

$$\gamma_\Delta(n) \quad := \quad \max\{1 - \lambda_2(G) \colon |V_G| = n, \Delta(G) \leq \Delta, \kappa(G) \geq 0\}.$$

Organization of the paper. The remainder of the paper is organized as follows: Section 2 offers a brief, self-contained introduction to the framework of random rooted graphs. In particular, we recall the definition of local weak convergence (Section 2.1), introduce the key notions of *unimodularity*, *stationarity* and *tightness* (Section 2.2), and gather important results on the *asymptotic entropy* of random walks on stationary graphs (Section 2.3). Section 3 is devoted to the proof of the main result, which is reduced (in Section 3.1) to the following two main steps:

1. Proving that non-negative curvature implies zero-entropy (Section 3.2).

2. Proving that zero-entropy causes poor spectral expansion (Section 3.3).

Acknowledgment. The author warmly thanks Itai Benjamini, David Cushing, Nicolas Curien, Shiping Liu, Russell Lyons, Florentin Münch and Pierre Pansu for many wonderful comments, connections and references. This work was partially supported by Institut Universitaire de France.

2 Random rooted graphs

In this section, we provide a self-contained introduction to the framework of *local weak convergence*. This limit theory for sparse graphs was introduced by Benjamini and Schramm [14] and developed further by Aldous and Steele [2] and Aldous and Lyons [1]. The limit points are *random rooted graphs* enjoying a powerful form of *stationarity*. They describe the "internal" geometry of large graphs, as seen from a uniformly chosen vertex. Local weak limits are often much more convenient to work with than the finite-graph sequences that they approximate, and have been shown to capture the asymptotic behavior of a number of natural graph parameters, see, e.g. [35, 17, 16, 3]. The present paper can be viewed as another illustration of the strength of this modern viewpoint.

2.1 Local weak convergence

The space of rooted graphs. All graphs considered in this paper will be simple, undirected, countable, and locally finite. A *rooted graph* is a pair (G, o), where G is a graph and o is a distinguished vertex, called the *root*. Two rooted graphs (G, o) and (G', o') are *isomorphic*, written $G \simeq G'$, if there is a bijection $\phi \colon V_G \to V_{G'}$ which preserves the root ($\phi(o) = o'$) and the edges:

$$\forall x, y \in V_G, \quad \{x, y\} \in E_G \iff \{\phi(x), \phi(y)\} \in E_{G'}.$$

We let $\mathscr{G}_\bullet$ denote the set of connected rooted graphs, considered up to the isomorphism relation $\simeq$. To lighten the exposition, we will use the same notation (G, o) for the rooted graph and its equivalence class. We write $\mathcal{B}_t(G, o)$ for the *ball of radius t around the root* in G, i.e. the (finite) rooted subgraph of G induced by the set $\{x \in V_G \colon \mathrm{d}_G(o, x) \le t\}$. We equip $\mathscr{G}_\bullet$ with the *local metric* $\mathrm{d}_{\mathrm{LOC}} \colon \mathscr{G}_\bullet \times \mathscr{G}_\bullet \to [0, 1]$, defined by

$$\mathrm{d}_{\mathrm{LOC}}((G, o), (G', o')) \;:=\; \frac{1}{1+r}, \quad \text{with} \quad r \;=\; \sup\{t \ge 0 \colon \mathcal{B}_t(G, o) \simeq \mathcal{B}_t(G', o')\}.$$

In words, two elements of $\mathscr{G}_\bullet$ are "close" to each other if one has to look "far away" from the root to distinguish them apart. It can be shown that $(\mathscr{G}_\bullet, \mathrm{d}_{\mathrm{LOC}})$ is a complete separable metric space. We equip it with its Borel $\sigma-$algebra, and call $\mathscr{G}_\bullet-$valued random variables *random rooted graphs*.

Local weak convergence. Write $\mathcal{P}(\mathscr{G}_\bullet)$ for the space of Borel probability measures on $\mathscr{G}_\bullet$, equipped with the usual topology of weak convergence. If G is an arbitrary finite graph, define its *local profile* $\mathcal{L}_G \in \mathcal{P}(\mathscr{G}_\bullet)$ to be the empirical distribution of all possible *rootings* of G, i.e.

$$\mathcal{L}_G \;:=\; \frac{1}{|V_G|} \sum_{x \in V_G} \delta_{(G,x)}, \tag{4}$$

where (G, x) is here implicitly restricted to the connected component of x if G is not connected. Finally, if $G_n = (V_n, E_n), n \geq 1$ are finite graphs whose local profiles $(\mathcal{L}_{G_n})_{n\geq 1}$ admit a limit $\mathcal{L}$ in $\mathcal{P}(\mathscr{G}_\bullet)$, we call $\mathcal{L}$ the *local weak limit* of the sequence $(G_n)_{n\geq 1}$, and write simply

$$G_n \xrightarrow[n\to\infty]{} \mathcal{L}.$$

In words, $\mathcal{L}$ is the law of a random rooted graph which describes how the deterministic graph G_n asymptotically looks when seen from a uniformly chosen root. More formally,

$$\frac{1}{|V_n|}\sum_{x\in V_n} f(G_n, x) \xrightarrow[n\to\infty]{} \mathcal{L}\left[f(G,o)\right] \triangleq \int_{\mathscr{G}_\bullet} f\,\mathrm{d}\mathcal{L}, \tag{5}$$

for each continuous, bounded observable $f\colon \mathscr{G}_\bullet \to \mathbb{R}$. The left-hand side can be thought of as a spatial average of "local contributions" from the various vertices of G_n. In short, local weak convergence allows one to conveniently replace the asymptotic analysis of such averages with the direct computation of an expectation at the root of a certain random graph.

Local observables. The class of continuous functions on $\mathscr{G}_\bullet$ clearly contains (but is not restricted to) all $t-$*local* observables ($t \geq 0$), where $f\colon \mathscr{G}_\bullet \to \mathbb{R}$ is called $t-$*local* if the value $f(G, o)$ is determined by the (isomorphic class of the) finite ball $\mathcal{B}_t(G, o)$. Here is a short list of examples, which will be used throughout the paper without notice:

- The root degree $(G, o) \mapsto \deg_G(o)$ is $1-$local.
- The minimum curvature at o, $(G, o) \mapsto \min_{x\sim o} \kappa_G(o, x)$ is $2-$local.
- For each $t \geq 0$, the return probability $(G, o) \mapsto P_G^t(o, o)$ is $t-$local (in fact, $(\lfloor t/2\rfloor + 1)-$local).
- For each $t \geq 0$, the $t-$step entropy $(G, o) \mapsto -\sum_{x\in V_G} P_G^t(o, x)\log P_G^t(o, x)$ is $t-$local.

2.2 Tightness, unimodularity and stationarity

Tightness. One of the many reasons for the success of the local weak convergence framework (compared to other limit theories for sparse graphs) is the fact that every "reasonable" sequence of sparse graphs admits a local weak limit. The following tightness criterion, due to Benjamini, Lyons and Schramm, gives an honest mathematical content to this vague claim. Note, of course, that passing to sub-sequences is unavoidable.

Theorem 4 (Tightness, see Theorem 3.1 in [12])**.** *Let $G_n = (V_n, E_n), n \geq 1$ be finite graphs so that*

$$\sup_{n\geq 1}\left\{\frac{1}{|V_n|}\sum_{x\in V_n}\phi\left(\deg_{G_n}(x)\right)\right\} < \infty,$$

for some function $\phi\colon \mathbb{Z}_+ \to \mathbb{R}_+$ satisfying $\phi(d) \gg d$ as $d \to \infty$. Then, $(G_n)_{n\geq 1}$ has a subsequence which admits a local weak limit.

In particular, this criterion applies to the sequence $(G_n)_{n\geq 1}$ in Theorem 3, with $\phi(d) = d \log d$. This will ensure that we can "pass to the limit" and study the question of existence of non-negatively curved expanders directly at the level of local weak limits.

Unimodularity. Local weak limits of finite graphs happen to enjoy a powerful distributional invariance, which is directly inherited from the fact that the root is equally likely to be any vertex under the local profile (4). More precisely, a measure $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ is called *unimodular* if it satisfies

$$\mathcal{L}\left[\sum_{x\in V_G} f(G,o,x)\right] \quad = \quad \mathcal{L}\left[\sum_{x\in V_G} f(G,x,o)\right], \tag{6}$$

for every Borel function $f\colon \mathscr{G}_{\bullet\bullet} \to [0,\infty]$, where $\mathscr{G}_{\bullet\bullet}$ denotes the analogue of the space $\mathscr{G}_\bullet$ with two distinguished roots instead of one. Thinking of $f(G,o,x)$ as an amount of mass sent from o to x, the identity (6) expresses the fact that the expected masses received and sent by the root coincide. This *Mass Transport Principle* is clearly satisfied when $\mathcal{L}$ is the local profile of a finite graph, and is preserved under weak convergence. Thus, we obtain the following fundamental result.

Theorem 5 (Inherited unimodularity)**.** *All local weak limits of finite graphs are unimodular.*

Whether the converse holds is a notoriously hard open problem with deep implications, see [1, 25, 12]. Let us here record a first simple consequence of unimodularity, which will be useful.

Lemma 6 (Everything shows at the root, see Lemma 2.3 in [1])**.** *Suppose that $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ is unimodular, and let $B \subseteq \mathscr{G}_\bullet$ be a Borel set such that $\mathcal{L}(B) = 1$. Then we also have,*

$$\mathcal{L}\left(\{\forall x \in V_G,\ (G,x) \in B\}\right) \quad = \quad 1.$$

Proof. Just apply the Mass Transport Principle with $f(G,o,x) = \mathbf{1}_{(G,o)\notin B}$. ☐

Stationarity. Under a mild integrability condition and a trivial change of measure, unimodularity can be rephrased as *reversibility* under a natural Markov chain on $\mathscr{G}_\bullet$. We will here only need the weaker notion of *stationarity*. Specifically, we say that a law $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ is *stationary* if it is invariant for the Markov chain on $\mathscr{G}_\bullet$ which, at each step, keeps the underlying graph as it is and moves the root according to the transition matrix P_G. In other words, $\mathcal{L}$ is stationary if

$$\mathcal{L}\left[\sum_{x\in V_G} P_G^t(o,x) h(G,x)\right] \quad = \quad \mathcal{L}\left[h(G,o)\right], \tag{7}$$

for every Borel function $h\colon \mathscr{G}_\bullet \to [0,\infty]$ and every $t \geq 0$ (equivalently, for $t=1$). The relation with unimodularity is summed up in the following classical lemma (see, e.g. [10]).

Lemma 7 (Degree-biasing)**.** *Let $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ be a unimodular law with* $\deg(\mathcal{L}) := \mathcal{L}[\deg_G(o)] < \infty$. *Then, the law $\widehat{\mathcal{L}} \in \mathcal{P}(\mathscr{G}_\bullet)$ defined by the following change of measure is stationary:*

$$\mathrm{d}\widehat{\mathcal{L}}(G,o) \quad := \quad \frac{\deg_G(o)}{\deg(\mathcal{L})}\,\mathrm{d}\mathcal{L}(G,o). \tag{8}$$

Proof. Apply the Mass Transport Principle to $\mathcal{L}$ with $f(G,o,x) = h(G,o)\mathbf{1}_{\{x,o\}\in E_G}$. ☐

Remark 4 (Mutual absolute continuity)**.** *It follows from (8) that the original law $\mathcal{L}$ and its degree-biased version $\widehat{\mathcal{L}}$ are mutually absolutely continuous. In other words, we have*

$$\mathcal{L}(B) = 1 \quad \Longleftrightarrow \quad \widehat{\mathcal{L}}(B) = 1,$$

for any Borel set $B \subseteq \mathscr{G}_\bullet$, allowing us to transfer results from one law to the other.

2.3 Spectral radius, entropy and the Liouville property

Stationarity is a powerful property, because it enables the development of an *ergodic theory* of random rooted graphs. See the inspiring works [37] on Galton-Watson trees, [10] on random rooted graphs, and [11] on general random environments. In particular, a classical application of Kingman's sub-additive ergodic theorem allows one to define the (quenched) *asymptotic entropy* of random walks on stationary random graphs, as recalled in the following lemma.

Lemma 8 (Entropy)**.** *Let $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ be stationary with $\mathcal{L}[\log \deg_G(o)] < \infty$. Then the limit*

$$\mathscr{H}(G,o) \quad := \quad \lim_{t\to\infty} \frac{1}{t} \sum_{x\in V_G} P_G^t(o,x) \log \frac{1}{P_G^t(o,x)},$$

exists $\mathcal{L}-$almost-surely and in $L^1(\mathscr{G}_\bullet, \mathcal{L})$, and does not depend on the choice of the root o.

We will henceforth simply write $\mathscr{H}(G)$ instead of $\mathscr{H}(G,o)$, and call this the *entropy* of G.

Proof. Let (G,o) have law $\mathcal{L}$, and conditionally on (G,o), let $X = (X_t)_{t\geq 0}$ be a lazy simple random walk on G starting from $X_0 = o$. For $0 \leq s \leq t$, define a non-negative random variable $Z_{s,t}$ by

$$Z_{s,t} \quad := \quad \log \frac{1}{P_G^{t-s}(X_s, X_t)}.$$

Note that $Z_{t,s} \overset{d}{=} Z_{0,t-s}$. Indeed, for any Borel function $f\colon \mathbb{R}_+ \to \mathbb{R}_+$, we have by definition

$$\begin{aligned}
\mathbb{E}\left[f(Z_{s,t})\right] &= \mathbb{E}\left[\sum_{x,y\in V_G} P_G^s(o,x)P_G^{t-s}(x,y) f\left(\log\frac{1}{P_G^{t-s}(x,y)}\right)\right] \\
&= \mathbb{E}\left[\sum_{y\in V_G} P_G^{t-s}(o,y) f\left(\log\frac{1}{P_G^{t-s}(o,y)}\right)\right] \\
&= \mathbb{E}\left[f(Z_{0,t-s})\right],
\end{aligned}$$

where the second line uses the stationarity (7) with $h(G,o) = \sum_y P_G^{t-s}(o,y) f\left(\log\frac{1}{P_G^{t-s}(o,y)}\right)$. Moreover, the trivial inequality $P_G^t(o,y) \geq P_G^s(o,x)P_G^{t-s}(x,y)$ readily implies the sub-additive property

$$Z_{0,t} \leq Z_{0,s} + Z_{s,t}. \tag{9}$$

Finally, the assumption $\mathcal{L}[\log \deg_G(o)] < \infty$ ensures that $\mathbb{E}[Z_{0,1}] < \infty$. Consequently, Kingman's sub-additive ergodic theorem (see, e.g. [38, Theorem 14.44]) guarantees the existence of a non-negative, integrable random variable Z_∞ such that almost-surely and in L^1,

$$\frac{Z_{0,t}}{t} \xrightarrow[t\to\infty]{} Z_\infty.$$

Averaging this convergence over the random walk X (i.e., taking conditional expectation given the random rooted graph) yields the existence of the limit $\mathscr{H}(G,o)$. By Lemma 6, the same is true if o is replaced by any $x \in V_G$. Moreover, the sub-additive property (9) with $s = 1$ shows that

$$\mathscr{H}(G,o) \leq \sum_{x\in V_G} P_G(o,x)\mathscr{H}(G,x),$$

$\mathcal{L}-$almost-surely. Since $\theta \mapsto (\theta - a)_+$ is monotone and convex for $a \geq 0$, this inequality implies

$$\forall a \geq 0, \quad (\mathscr{H}(G,o) - a)_+ \leq \sum_{x\in V_G} P_G(o,x)\left(\mathscr{H}(G,x) - a\right)_+.$$

But the two sides have the same law by stationarity, so they must coincide $\mathcal{L}-$almost-surely. The fact that this is true for all $a \geq 0$ deterministically forces the equality $\mathscr{H}(G,x) = \mathscr{H}(G,o)$ for all neighbours x of o, and hence for all $x \in V_G$ by Lemma 6. □

The Liouville property. One of the interests of asymptotic entropy lies in its relation with the Liouville property. A function $f\colon V_G \to \mathbb{R}$ is called *harmonic* on G if $P_G f = f$, where

$$\forall x \in V_G, \quad (P_G f)(x) := \sum_{y\in V_G} P_G(x,y) f(y). \tag{10}$$

This is trivially the case, in particular, when f is constant. The graph G has the *Liouville property* if it admits no non-constant bounded harmonic function. For stationary random graphs, this functional-analytic property turns out to admit the following simple entropic characterization.

Theorem 9 (Entropic characterization of the Liouville property)**.** *The equivalence*

$$\mathscr{H}(G) = 0 \iff G \text{ has the Liouville property},$$

holds almost-surely under any stationary law $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ *with* $\mathcal{L}[\log \deg_G(o)] < \infty$*.*

This remarkable result has a long history: it originates with the pioneering works of Avez [5, 6, 4], and was then made famous in a celebrated paper of Kaimanovich and Vershik [31]. In the present setting of stationary random graphs, the implication $\Longrightarrow$ was established by Benjamini and Curien [10], and refined by Benjamini, Duminil-Copin, Kozma and Yadin [11]. The converse $\Longleftarrow$ was proved by Carrasco Piaggio and Lessa [20] (see also [13]), but under an additional growth assumption. Since this is the implication that we are going to use, we need to give more details.

Proof of Theorem 9. Fix a connected graph G, and let $X = (X_t)_{t\geq 0}$ denote a lazy simple random walk on G starting at some fixed vertex $o \in V_G$. Write $\mathbf{P}^G$ for its law, which is a probability measure on the product space $V_G^{\mathbb{Z}_+}$. On this space, let $\mathcal{I}$ denote the $\sigma-$field of all events which are invariant under the natural shift $(x_t)_{t\geq 0} \mapsto (x_{t+1})_{t\geq 0}$. Then [38, Proposition 14.12] states that

$$G \text{ has the Liouville property} \iff \mathcal{I} \text{ is } \mathbf{P}^G-\text{trivial}.$$

On the other hand, writing $\mathcal{T} = \bigcap_{t=0}^\infty \sigma(x_t, x_{t+1}, \ldots)$ for the tail $\sigma-$field on $V_G^{\mathbb{Z}_+}$, we have

$$\mathcal{I} \text{ is } \mathbf{P}^G-\text{trivial} \iff \mathcal{T} \text{ is } \mathbf{P}^G-\text{trivial},$$

by Theorem [38, Theorem 14.18] and because X is lazy. Finally, the equivalence

$$\mathcal{L}\left(\mathcal{T} \text{ is } \mathbf{P}^G-\text{trivial}\right) = 1 \iff \mathcal{L}(\mathscr{H}(G) = 0) = 1,$$

was proved in [10, Theorem 3.2] for any stationary law $\mathcal{L}$ with $\mathcal{L}[\log \deg_G(o)] < \infty$. Thus,

$$\mathcal{L}(G \text{ has the Liouville property}) = 1 \iff \mathcal{L}\left(\mathscr{H}(G) = 0\right) = 1, \tag{11}$$

and this annealed statement will actually suffice for the present paper. However, deducing the quenched claim is easy, as we now explain. Define the events $A := \{G \text{ has the Liouville property}\}$ and $B := \{\mathscr{H}(G) = 0\}$, and let $A\Delta B$ denote their symmetric difference. We want to show that

$$\mathcal{L}(A\Delta B) \;=\; 0, \tag{12}$$

for any stationary law $\mathcal{L}$ with $\mathcal{L}[\log \deg_G(o)] < \infty$. We already know this if A, B are $\mathcal{L}-$trivial, thanks to (11). Moreover, the events A, B are clearly *root-invariant*, in the sense that

$$(G, o) \in A \implies \{\forall x \in V_G, (G, x) \in A\}.$$

Consequently, (12) holds under the extra assumption that *root-invariant events are* $\mathcal{L}-$*trivial.* But this is known as *ergodicity*, and any stationary law can be decomposed as a mixture of ergodic laws, by [1, Theorem 4.7]. Thus, (12) extends to all stationary laws $\mathcal{L}$ with $\mathcal{L}[\log \deg_G(o)] < \infty$. ☐

Spectral radius. The entropy $\mathscr{H}(G)$ is related to several other fundamental graph-theoretical quantities, such as the *speed*, *growth*, or *spectral radius*, see [38]. Let us recall the last notion. Fix a rooted graph $(G,o) \in \mathscr{G}_\bullet$. For any $t,s \geq 0$, we trivially have $P_G^{t+s}(o,o) \geq P_G^t(o,o)P_G^s(o,o)$. By Fekete's lemma, we deduce that the limit

$$\varrho(G,o) \quad := \quad \lim_{t\to\infty} \left(P_G^t(o,o)\right)^{\frac{1}{t}}, \tag{13}$$

exists in $(0,1]$. Moreover, the connectivity of G together with the trivial inequality

$$P_G^{t+2s}(o,o) \quad \geq \quad P_G^s(o,x)P_G^t(x,x)P_G^s(x,o),$$

shows that $\varrho(G,o)$ does not depend on the choice of the root o. Thus, we will henceforth simply write $\varrho(G)$, and call this quantity the *spectral radius* of G.

Lemma 10 (Spectral radius vs entropy)**.** *The inequality*

$$\mathscr{H}(G) \quad \geq \quad 2\log\frac{1}{\varrho(G)},$$

holds almost-surely under any stationary law $\mathcal{L}$ with $\mathcal{L}[\log \deg_G(o)] < \infty$.

Proof. For any rooted graph (G,o) and any $t \geq 0$, we have by concavity

$$\begin{aligned} \log\left(P_G^{2t}(o,o)\right) \quad &= \quad \log\left(\sum_{x\in V_G} P_G^t(o,x)P_G^t(x,o)\right) \\ &\geq \quad \sum_{x\in V_G} P_G^t(o,x)\log P_G^t(x,o) \\ &= \quad \sum_{x\in V_G} P_G^t(o,x)\log P_G^t(o,x) + \sum_{x\in V_G} P_G^t(o,x)\log\left(\frac{\deg_G(o)}{\deg_G(x)}\right), \end{aligned}$$

where the last line uses the reversibility $\deg_G(o)P_G^t(o,x) = \deg_G(x)P_G^t(x,o)$. Dividing by $-2t$ and taking the limit as $t\to\infty$ in $L^1(\mathscr{G}_\bullet,\mathcal{L})$ yields the claim, provided we can show that

$$\frac{1}{t}\sum_{x\in V_G} P_G^t(o,x)\log\left(\frac{\deg_G(o)}{\deg_G(x)}\right) \quad \xrightarrow[t\to\infty]{L^1(\mathscr{G}_\bullet,\mathcal{L})} \quad 0.$$

But this follows from the crude bound

$$\begin{aligned} \mathcal{L}\left[\left|\sum_{x\in V_G} P_G^t(o,x)\log\left(\frac{\deg_G(o)}{\deg_G(x)}\right)\right|\right] \quad &\leq \quad \mathcal{L}\left[\sum_{x\in V_G} P_G^t(o,x)\left(\log\deg_G(o) + \log\deg_G(x)\right)\right] \\ &= \quad 2\mathcal{L}\left[\log\deg_G(o)\right], \end{aligned}$$

where the second line simply uses the stationarity property (7) with $h(G,o) = \log\deg_G(o)$. ☐

Remark 5 (Unimodular analogues)**.** *By Lemma 7 and Remark 4, all results in this section also apply to any unimodular law $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ with $\mathcal{L}[\deg_G(o)\log\deg_G(o)] < \infty$.*

3 Proof of the main result

We are now ready to prove our main result. We work with the formulation given in Theorem 3. Section 3.1 below reduces it to two key results, which are then proved in Sections 3.2 and 3.3.

3.1 Setting the stage

Let $G_n = (V_n, E_n)$, $n \geq 1$ be finite graphs satisfying the assumptions of Theorem 3, i.e.

$$\sup_{n\geq 1}\left\{\frac{1}{|V_n|}\sum_{x\in V_n}\deg_{G_n}(x)\log\deg_{G_n}(x)\right\} \quad < \quad \infty; \tag{14}$$

$$\forall\varepsilon>0, \quad \frac{1}{|E_n|}\operatorname{card}\{e\in E_n\colon \kappa_{G_n}(e)<-\varepsilon\} \quad \xrightarrow[n\to\infty]{} \quad 0. \tag{15}$$

Recall that our goal is to establish

$$\forall\rho\in(0,1), \quad \liminf_{n\to\infty}\left\{\frac{1}{|V_n|}\operatorname{card}\{i\colon \lambda_i(G_n)>\rho\}\right\} \quad > \quad 0. \tag{16}$$

By (14) and Theorem 4, we may assume, upon extracting a subsequence if necessary, that

$$G_n \quad \xrightarrow[n\to\infty]{} \quad \mathcal{L}, \tag{17}$$

for some $\mathcal{L}\in\mathcal{P}(\mathscr{G}_\bullet)$. Note that $\mathcal{L}$ is automatically unimodular by Theorem 5, and such that

$$\mathcal{L}\left[\deg_G(o)\log\deg_G(o)\right] \quad < \quad \infty. \tag{18}$$

Just like the degree, the curvature is a local notion, hence it also "passes to the limit", i.e

$$\mathcal{L}\left(\kappa(G)\geq 0\right) \quad = \quad 1. \tag{19}$$

Proof. As already mentioned, the observable $f\colon (G,o)\mapsto \min_{x\sim o}\kappa_G(o,x)$ is $2-$local, hence continuous on $\mathscr{G}_\bullet$. By the Portmanteau Theorem, we deduce that for any $\varepsilon>0$,

$$\begin{aligned}
\mathcal{L}\left(f<-\varepsilon\right) \quad &\leq \quad \liminf_{n\to\infty}\mathcal{L}_{G_n}\left(f<-\varepsilon\right)\\
&= \quad \liminf_{n\to\infty}\left\{\frac{1}{|V_n|}\operatorname{card}\{o\in V_n\colon f(G_n,o)<-\varepsilon\}\right\}\\
&\leq \quad \liminf_{n\to\infty}\left\{\frac{2}{|V_n|}\operatorname{card}\{e\in E_n\colon \kappa_{G_n}(e)<-\varepsilon\}\right\}\\
&= \quad \mathcal{L}\left[\deg_G(o)\right]\liminf_{n\to\infty}\left\{\frac{1}{|E_n|}\operatorname{card}\{e\in E_n\colon \kappa_{G_n}(e)<-\varepsilon\}\right\},
\end{aligned}$$

where the last inequality follows from the observation that $\frac{2|E_n|}{|V_n|}\to\mathcal{L}[\deg_G(o)]$, by the continuity and uniform integrability of $(G,o)\mapsto\deg_G(o)$. Sending $\varepsilon\to 0$ yields $\mathcal{L}(f<0)=0$, by (15). To conclude, we simply apply Lemma 6 to the event $B=\{f\geq 0\}$. □

The first crucial step in our proof consists in deducing from (19) that the entropy is zero under $\mathcal{L}$. This is the content of the following theorem, which will be proved in Section 3.2.

Theorem 11 (Non-negative curvature implies zero-entropy)**.** *The implication*

$$\kappa(G) \geq 0 \implies \mathscr{H}(G) = 0$$

holds almost-surely under any stationary law $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ *satisfying* $\mathcal{L}\left[\log \deg_G(o)\right] < \infty$.

In view of Remark 4, this result also applies to any unimodular law $\mathcal{L} \in \mathcal{P}(\mathscr{G}_\bullet)$ satisfying $\mathcal{L}\left[\deg_G(o) \log \deg_G(o)\right] < \infty$, hence in particular to the limit $\mathcal{L}$ in (17). Combining this with Lemma 10, we immediately deduce that our local weak limit satisfies

$$\mathcal{L}(\rho(G) = 1) \quad = \quad 1.$$

It turns out that this simple condition suffices to guarantee (16). This is the content of the following second result, established in Section 3.3 below, and which completes the proof of our main result.

Theorem 12 (Zero-entropy implies poor spectral expansion)**.** *Let* $G_n = (V_n, E_n), n \geq 1$ *be finite graphs having local weak limit* $\mathcal{L}$, *and suppose that* $\mathcal{L}\left(\rho(G) = 1\right) = 1$. *Then, for any* $\rho < 1$,

$$\liminf_{n\to\infty} \left\{ \frac{1}{|V_n|} \operatorname{card}\left\{i \colon \lambda_i(G_n) > \rho\right\} \right\} \quad > \quad 0.$$

In fact, a stronger statement about eigenvectors will be derived, as claimed in Remark 3.

3.2 Non-negative curvature implies zero entropy

Consider a connected graph G and two vertices $x, y \in V_G$. The proof of Theorem 11 relies on the following intuitive idea: if G has non-negative curvature and bounded degrees, then it takes time $O(\mathrm{d}_G^2(x,y))$ for two random walks starting at x and y to meet. This classical observation constitutes the very essence of the path coupling method of Bordewich and Dyer [18]. It was later re-discovered and further developed by Münch [40]. We will here prove a refinement that does not require bounded degrees, see Corollary 16 below. Write $\mathcal{B}_x, \mathcal{B}_y$ for the balls of radius 1 around x and y, and recall that the Wassertein distance $\mathcal{W}_1\left(P_G(x,\cdot), P_G(y,\cdot)\right)$ is defined as

$$\mathcal{W}_1\left(P_G(x,\cdot), P_G(y,\cdot)\right) \quad = \quad \inf_{\pi} \left\{ \sum_{u \in \mathcal{B}_x} \sum_{v \in \mathcal{B}_y} \pi(u,v)\, \mathrm{d}_G(u,v) \right\}, \tag{20}$$

where the infimum runs over all probability distributions $\pi \in \mathcal{P}(\mathcal{B}_x \times \mathcal{B}_y)$ with marginals $P_G(x,\cdot)$ and $P_G(y,\cdot)$. By compactness, the above infimum is actually achieved, and the minimizers will be called *optimal couplings*. As in [18, 40], our first task consists in showing that an optimal coupling can always be chosen so as to assign a "decent" probability to the "good" set

$$\Gamma \quad := \quad \{(u,v) \in \mathcal{B}_x \times \mathcal{B}_y \colon \mathrm{d}_G(u,v) < \mathrm{d}_G(x,y)\}\,.$$

The argument crucially uses the laziness of P_G but is otherwise rather general.

Lemma 13 (Good optimal couplings)**.** *If $x \neq y$, then there is an optimal coupling π such that*

$$\pi\,(\Gamma) \quad \geq \quad \frac{1}{2}\max\left\{\frac{1}{\deg_G(x)}, \frac{1}{\deg_G(y)}\right\}.$$

Proof. By compactness, we can find an optimal coupling π which, among all optimal couplings, maximizes $\pi(\Gamma)$. Suppose for a contradiction that this "doubly optimal" coupling satisfies

$$\pi\,(\Gamma) \quad < \quad \frac{1}{2\deg_G(x)}. \tag{21}$$

The set $A := \{u \in \mathcal{B}_x \colon (u,y) \in \Gamma\}$ is not empty, since it contains the first vertex on a geodesic from x to y. Thus, $\pi(A \times \mathcal{B}_y) \geq 1/(2\deg_G(x))$. In view of (21), this forces $\pi((A \times \mathcal{B}_y) \setminus \Gamma) > 0$, i.e.

$$\exists (x_0, y_0) \in (A \times \mathcal{B}_y) \setminus \Gamma, \quad \pi(x_0,y_0) \quad \geq \quad \varepsilon, \tag{22}$$

for some $\varepsilon > 0$. On the other hand, we have $\pi(A \times \{y\}) + \pi(A^c \times \{y\}) = P_G(y,y) \;=\; \frac{1}{2}$. This forces $\pi(A^c \times \{y\}) > 0$, because $\pi(A \times \{y\}) \leq \pi(\Gamma) < \frac{1}{2}$. In other words,

$$\exists x_1 \in A^c, \quad \pi(x_1, y) \quad \geq \quad \varepsilon, \tag{23}$$

provided $\varepsilon > 0$ is chosen small enough. We now use the vertices x_0, y_0, x_1 found at (22)-(23) to construct a new coupling $\widehat{\pi}$ which contradicts the optimality of π. For all $(u,v) \in \mathcal{B}_x \times \mathcal{B}_y$, we set

$$\widehat{\pi}(u,v) \quad := \quad \begin{cases} \pi(u,v) & \text{if } u \notin \{x_0,x_1\} \text{ and } b \notin \{y_0,y\}; \\ \pi(u,v) - \varepsilon & \text{if } (u,v) = (x_0,y_0) \text{ or } (u,v) = (x_1,y); \\ \pi(u,v) + \varepsilon & \text{if } (u,v) = (x_0,y) \text{ or } (u,v) = (x_1,y_0). \end{cases}$$

By construction, $\widehat{\pi}$ is non-negative on $\mathcal{B}_x \times \mathcal{B}_y$ and has the same marginals as π. Thus, it is a coupling of $P_G(x,\cdot), P_G(y,\cdot)$. This coupling is moreover optimal, since

$$\begin{aligned} \sum_{u \in \mathcal{B}_x}\sum_{v \in \mathcal{B}_y} \mathrm{d}_G(u,v)\,(\widehat{\pi}(u,v) - \pi(u,v)) \quad &= \quad \varepsilon\,(\mathrm{d}_G(x_0,y) + \mathrm{d}_G(x_1,y_0) - \mathrm{d}_G(x_0,y_0) - \mathrm{d}_G(x_1,y)) \\ &\leq \quad \varepsilon\,(\mathrm{d}_G(x,y) - 1 + \mathrm{d}_G(x_1,y_0) - \mathrm{d}_G(x,y) - \mathrm{d}_G(x_1,y)) \\ &\leq \quad 0, \end{aligned}$$

where the first inequality uses $x_0 \in A$ and $(x_0, y_0) \notin \Gamma$, while the second uses the triangle inequality $\mathrm{d}_G(x_1, y_0) \le \mathrm{d}_G(x_1, y) + \mathrm{d}_G(y, y_0)$. Finally, since Γ contains (x_1, y) but not $(x_0, y_0), (x_1, y)$, we have

$$\widehat{\pi}(\Gamma) \quad \geq \quad \pi(\Gamma) + \varepsilon,$$

contradicting the definition of π. Thus, (21) can not be true, and the claim follows by symmetry. □

We will also need the following technical lemma, which is of independent interest and quantifies the intuition that non-negative super-martingales that "move a lot" must "quickly" hit zero.

Lemma 14 (Non-negative super-martingales quickly hit zero)**.** *Let $\tau := \inf\{t \ge 0 \colon Z_t = 0\}$ be the hitting time of zero by a non-negative super-martingale $Z = (Z_t)_{t \ge 0}$. Suppose that $Z_0 = z$, and that all increments $(Z_{t+1} - Z_t)_{t \ge 0}$ are upper-bounded by a constant K. Then,*

$$\mathbb{P}\left(\tau \geq t\right) \quad \leq \quad z\left(\frac{2a + K - z}{a^2}\right) + \mathbb{P}\left(\tau \geq t, \sum_{s=0}^{t-1} W_s < a^2\right),$$

for all $t \in \mathbb{Z}_+, a > 0$, where $W_s = \mathbb{E}\left[(Z_{s+1} - Z_s)^2 | \mathscr{F}_s\right]$ and $(\mathscr{F}_s)_{s \ge 0}$ is the underlying filtration.

Proof. First note that the process Z is trivially square-integrable, because $Z_t \in [0, z + Kt]$ for each $t \ge 0$. Now fix $t \ge 0$ and $a > 0$, and consider the bounded stopping time

$$\sigma \quad := \quad \inf\left\{s \geq 0 \colon Z_s \geq a\right\} \wedge t.$$

Using the Optional Stopping Theorem, the non-negativity of Z and the definition of σ, we have

$$\begin{aligned} z \quad &\geq \quad \mathbb{E}\left[Z_{\sigma \wedge \tau}\right] \\ &\geq \quad \mathbb{E}\left[Z_{\sigma \wedge \tau} \mathbf{1}_{(\sigma < \tau \wedge t)}\right] \\ &\geq \quad a\mathbb{P}\left(\sigma < \tau \wedge t\right). \end{aligned}$$

On the other hand, observe that for all $s \ge 0$, we may rewrite W_s as

$$W_s \quad = \quad \mathbb{E}\left[Z_{s+1}^2 - Z_s^2 | \mathscr{F}_s\right] + 2Z_s\mathbb{E}[Z_s - Z_{s+1} | \mathscr{F}_s].$$

Note that the second conditional expectation is non-negative by assumption. Moreover, we have $Z_s \le a$ on the event $\{\sigma > s\}$, which is in $\mathscr{F}_s$. Thus,

$$W_s \mathbf{1}_{\sigma > s} \quad \leq \quad \mathbb{E}\left[\left(Z_{s+1}^2 - Z_s^2\right)\mathbf{1}_{\sigma > s} | \mathscr{F}_s\right] + 2a\mathbb{E}\left[\left(Z_s - Z_{s+1}\right)\mathbf{1}_{\sigma > s} | \mathscr{F}_s\right].$$

Taking expectations and summing over all $s \ge 0$, we obtain

$$\begin{aligned} \mathbb{E}\left[\sum_{s=0}^{\sigma-1} W_s\right] \quad &\leq \quad \mathbb{E}\left[Z_\sigma^2\right] - 2a\mathbb{E}[Z_\sigma] - z^2 + 2az \\ &\leq \quad (K + a - z)z, \end{aligned}$$

where the second inequality follows from the observations that $Z_\sigma \le K + a$ and $\mathbb{E}[Z_\sigma] \le z$. Let us now use these two estimates to conclude. By union bound, we have

$$\begin{aligned}
\mathbb{P}(\tau \ge t) &\le \mathbb{P}(\sigma < \tau \wedge t) + \mathbb{P}(\sigma \wedge \tau \ge t) \\
&\le \mathbb{P}(\sigma < \tau \wedge t) + \mathbb{P}\left(\tau \ge t, \sum_{s=0}^{\sigma-1} W_s \ge \sum_{s=0}^{t-1} W_s\right) \\
&\le \mathbb{P}(\sigma < \tau \wedge t) + \mathbb{P}\left(\sum_{s=0}^{\sigma-1} W_s \ge a^2\right) + \mathbb{P}\left(\tau \ge t, \sum_{s=0}^{t-1} W_s < a^2\right) \\
&\le \frac{z}{a} + \frac{(K+a-z)z}{a^2} + \mathbb{P}\left(\tau \ge t, \sum_{s=0}^{t-1} W_s < a^2\right).
\end{aligned}$$

This is exactly the claimed bound. ☐

Combining these two lemmas, we may now deduce the following estimate, which exploits non-negative curvature to control the action of P_G on the variations of bounded observables.

Proposition 15 (Variational estimate via non-negative curvature)**.** *Let G be a connected graph with $\kappa(G) \ge 0$. Then, for any $f\colon V_G \to [-1,1]$, any vertices $x, y \in V_G$, and any $a > 0, t \in \mathbb{Z}_+$,*

$$|P_G^t f(x) - P_G^t f(y)| \le \frac{8\mathrm{d}_G(x,y)}{a} + 2\mathbb{P}\left(\sum_{s=0}^{t-1} \frac{1}{\deg_G(X_s)} < 2a^2\right),$$

where X denotes a lazy random walk on G starting from x.

Proof. Let (X, Y) be the Markov chain on $V_G \times V_G$ which, from any state $(x, y) \in V_G \times V_G$, draws the next state according to the "good" optimal coupling of $P_G(x,\cdot), P_G(y,\cdot)$ described in Lemma 13. We use the standard notations $\mathbb{P}_{(x,y)}(\cdot), \mathbb{E}_{(x,y)}[\cdot]$ to specify the choice of the initial state. Since the two coordinates X, Y are marginally distributed as lazy random walks on G, we have

$$\begin{aligned}
\left|P_G^t f(x) - P_G^t f(y)\right| &= \left|\mathbb{E}_{x,y}\left[f(X_t)\right] - \mathbb{E}_{x,y}\left[f(Y_t)\right]\right| \\
&\le \mathbb{E}_{x,y}\left[|f(X_t) - f(Y_t)|\right] \\
&\le 2\mathbb{P}_{x,y}(X_t \ne Y_t) \\
&\le 2\mathbb{P}_{x,y}(\tau > t),
\end{aligned}$$

where $\tau = \inf\{t \ge 0\colon X_t = Y_t\}$ denotes the meeting time of the two walkers. Note that τ is also the hitting time of zero by the non-negative process $Z = (Z_t)_{t \ge 0}$ defined as follows:

$$\forall t \ge 0, \quad Z_t := \mathrm{d}_G(X_t, Y_t).$$

We claim that Z is a super-martingale w.r.t. the natural filtration $(\mathscr{F}_t)_{t\geq 0}$ associated with (X, Y). Indeed, by the Markov property and the optimality of the chosen couplings, this claim reduces to

$$\mathcal{W}_1\left(P_G(x, \cdot), P_G(y, \cdot)\right) \quad \leq \quad \mathrm{d}_G(x, y),$$

for all $x, y \in V_G$. But this inequality readily follows from the assumption $\kappa_G(x, y) \geq 0$ in the case $\{x, y\} \in E_G$, and it then automatically extends to all $x, y \in V_G$ by the triangle inequality of $\mathcal{W}_1(\cdot, \cdot)$ (see, e.g., [45]). On the other hand, Lemma 13 ensures that on the event $\{\tau > t\}$,

$$\mathbb{E}_{x,y}\left[(Z_{t+1} - Z_t)^2 | \mathscr{F}_t\right] \quad \geq \quad \frac{1}{2\deg_G(X_t)}.$$

Finally, note that the distance between the two walkers can not increase by more than 2 at each step. Thus, we may invoke Lemma 14 to conclude that

$$\begin{aligned}
\mathbb{P}_{x,y}\left(\tau \geq t\right) \quad & \leq \quad 2\mathrm{d}_G(x, y)\left(\frac{a+1}{a^2}\right) + \mathbb{P}_{x,y}\left(\sum_{s=0}^{t-1} \frac{1}{\deg_G(X_s)} < 2a^2\right) \\
& \leq \quad \frac{4\mathrm{d}_G(x, y)}{a} + \mathbb{P}_{x,y}\left(\sum_{s=0}^{t-1} \frac{1}{\deg_G(X_s)} < 2a^2\right),
\end{aligned}$$

where the second line follows from the first if $a \geq 1$, and is trivial otherwise. □

In particular, this applies to any bounded harmonic function f, after a trivial normalization. Since $P_G^t f = f$ for all $t \geq 0$, we may send $t \to \infty$ and then $a \to \infty$ in the resulting estimate to obtain the following key result, which ensures that non-negatively curved graphs satisfy the Liouville property, provided they have a "decent proportion" of vertices with "reasonable" degree.

Corollary 16 (Liouville property and non-negative curvature)**.** *Let G be a connected graph with $\kappa(G) \geq 0$. Fix $o \in V_G$ and suppose that the simple random walk X on G starting from o satisfies*

$$\mathbb{P}\left(\sum_{t=0}^{\infty} \frac{1}{\deg_G(X_t)} = \infty\right) \quad = \quad 1. \tag{24}$$

Then, G has the Liouville property.

A simple situation where the above condition trivially holds is that where G has bounded degrees. In that case, the Liouville property was recently established by Jost, Münch, and Rose [29]. Our relaxation allows for arbitrary large degrees, as long as the random walk can avoid them from times to times. This is the case under any stationary law by Birkhoff's Ergodic Theorem, allowing us to prove Theorem 11.

Proof of Theorem 11. Let (G,o) have law $\mathcal{L}$ and, conditionally on (G,o), let X be a lazy random walk starting from the root. Then the process $Z=(Z_t)_{t\geq 0}$ defined by

$$\forall t \geq 0, \quad Z_t \quad := \quad \frac{1}{\deg_G(X_t)}$$

is stationary, in the usual sense that its law is invariant under the shift $(z_t)_{t\geq 0} \mapsto (z_{t+1})_{t\geq 0}$ on $[0,1]^{\mathbb{Z}_+}$. Thus, Birkhoff's Ergodic Theorem (see, e.g. [38, Theorem 14.43]) ensures that

$$\frac{1}{t}\sum_{s=0}^{t-1} Z_s \quad \xrightarrow[t\to\infty]{} \quad \mathbb{E}[Z_1|\mathscr{I}],$$

almost-surely, where $\mathscr{I}$ is the invariant $\sigma-$algebra. Since Z_1 is almost-surely positive, we deduce

$$\sum_{s=0}^{\infty} Z_s \quad = \quad \infty,$$

almost-surely. In other words, the random graph (G,o) satisfies (24) almost-surely. By the above corollary, this implies that G has the Liouville property almost-surely on the event $\{\kappa(G)\geq 0\}$. By Theorem 9, we conclude that $\mathscr{H}(G)=0$ almost-surely on the same event. □

3.3 Zero entropy implies poor spectral expansion

This final section is devoted to proving Theorem 12, which relates the eigenvalues of finite graphs to the spectral radius of their local weak limits. If G is a finite graph, the $N=|V_G|$ eigenvalues $\lambda_1(G)\geq \ldots \geq \lambda_N(G)$ of its transition matrix P_G can be conveniently encoded into a probability measure $\mu_G \in \mathcal{P}([0,1])$, called the *empirical eigenvalue distribution* of the matrix P_G:

$$\mu_G \quad := \quad \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(G)}.$$

It turns out that the large-size asymptotics of this fundamental object can be understood directly at the level of local weak limits. When P_G is replaced with the more standard adjacency matrix, this classical observation is the starting point of a rich and well-established theory, see the comprehensive introductory survey [15] by Bordenave, and the references therein.

Local spectral measures. The transition kernel P_G of a graph G can be viewed as a linear operator acting via (10) on the Hilbert space

$$\ell^2(G) \quad := \quad \left\{ f\in\mathbb{C}^{V_G} \colon \sum_{o\in V_G} \deg_G(o)|f(o)|^2 < \infty \right\},$$

with inner product $\langle f, g\rangle = \sum_{o \in V_G} \deg_G(o)\overline{f(o)}g(o)$. The stochasticity, laziness and reversibility

$$\sum_{y \in V_G} P_G(x,y) = 1, \qquad P_G(x,x) \geq 1/2, \qquad \deg_G(x) P_G(x,y) = \deg_G(y) P_G(y,x),$$

easily (and classically) imply that P_G is a positive contraction on $\ell^2(G)$, i.e.

$$\forall f \in \ell^2(G), \qquad 0 \ \leq \ \langle f, P_G f\rangle \ \leq \langle \ f, f\rangle.$$

In particular, for each $o \in V_G$, the spectral theorem for self-adjoint operators ensures the existence of a *local spectral measure* $\mu_{(G,o)} \in \mathcal{P}([0,1])$, characterized by the moment identity

$$\forall t \geq 0, \qquad \int_0^1 \lambda^t \mu_{(G,o)}(\mathrm{d}\lambda) \ = \ P_G^t(o,o). \tag{25}$$

As we will now see, $\mu_{(G,o)}$ can be interpreted as the local contribution of o to the spectrum of P_G. Local spectral measures are a powerful tool to investigate the mixing properties of graphs, see [36].

The finite case. When G is finite with N vertices, there is an orthonormal basis $(\phi_1, \ldots, \phi_N)$ of $\ell^2(G)$ consisting of eigenvectors of P_G with eigenvalues $\lambda_1(G), \ldots, \lambda_N(G)$, and we easily find

$$\mu_{(G,o)} \ = \ \sum_{i=1}^N \deg_G(o) |\phi_i(o)|^2 \delta_{\lambda_i(G)}. \tag{26}$$

Thus, the local spectral measure $\mu_{(G,o)}$ is a mixture of Dirac masses located at the various eigenvalues of P_G, and weighted by the squared amplitudes of the corresponding eigenvectors at o. Moreover, thanks to the orthonormality of $(\phi_1, \ldots, \phi_N)$, the identity (26) readily implies

$$\mu_G \ = \ \frac{1}{|V_G|} \sum_{o \in V_G} \mu_{(G,o)}. \tag{27}$$

In other words, the empirical eigenvalue distribution of a finite graph G coincides with the spatial average of its local spectral measures.

Spectral continuity. In light of (5), it is tempting to pass to the limit in the formula (27) along a convergent sequence of finite graphs $(G_n)_{n \geq 1}$. This is made rigorous by the following continuity principle. As usual, $\mathcal{P}([0,1])$ is here equipped with the topology of weak convergence.

Lemma 17 (Spectral continuity)**.** *The map $(G,o) \mapsto \mu_{(G,o)}$ is continuous on $\mathscr{G}_\bullet$. In particular, if a sequence of graphs $(G_n)_{n \geq 1}$ admits a local weak limit $\mathcal{L}$, then*

$$\mu_{G_n}(\mathrm{d}\lambda) \ \xrightarrow[n \to \infty]{} \ \mu_{\mathcal{L}}(\mathrm{d}\lambda) \ := \ \mathcal{L}\left[\mu_{(G,o)}(\mathrm{d}\lambda)\right].$$

Proof. For each fixed $t \geq 0$, the observable $(G, o) \mapsto P^t_G(o, o)$ is clearly $t-$local, hence continuous. In particular, via the identity (25), the convergence $(G_n, o_n) \to (G, o)$ in $\mathscr{G}_\bullet$ implies

$$\forall t \geq 0, \quad \int_0^1 \lambda^t \, \mu_{(G_n,o_n)}(\mathrm{d}\lambda) \quad \xrightarrow[n\to\infty]{} \quad \int_0^1 \lambda^t \, \mu_{(G,o)}(\mathrm{d}\lambda). \tag{28}$$

Since convergence in $\mathcal{P}([0,1])$ is equivalent to the convergence of moments, we conclude that $\mu_{(G_n,o_n)} \xrightarrow[n\to\infty]{} \mu_{(G,o)}$, and the continuity is proved. Similarly, the second claim is obtained by applying (5) to the $t-$local observable $f\colon (G, o) \mapsto P^t_G(o, o)$, for each $t \geq 1$. □

Corollary 18 (Unit spectral radius implies poor spectral expansion)**.** *Let $G_n = (V_n, E_n), n \geq 1$ be finite graphs having a local weak limit $\mathcal{L}$ such that $\mathcal{L}(\rho(G) = 1) = 1$. Then, for any $0 \leq \rho < 1$,*

$$\liminf_{n\to\infty} \mu_{G_n}\left([\rho, 1]\right) \quad > \quad 0. \tag{29}$$

Moreover, we have the refinement

$$\sup_{n\geq 1} \frac{\left|\left\{x \in V_n \colon \mu_{(G_n,x)}([\rho, 1]) \leq \varepsilon\right\}\right|}{|V_n|} \quad \xrightarrow[\varepsilon\to 0]{} \quad 0. \tag{30}$$

Proof. Fix $0 \leq \rho < 1$. By the second part of Lemma 17 and the Portmanteau Theorem, we have

$$\liminf_{n\to\infty} \mu_{G_n}([\rho, 1]) \quad \geq \quad \mathcal{L}\left[\mu_{(G,o)}((\rho, 1])\right]. \tag{31}$$

On the other hand, comparing (25) with the definition of the spectral radius, we see that $\rho(G)$ is exactly the supremum of the support of $\mu_{(G,o)}$, for any $(G, o) \in \mathscr{G}_\bullet$. In other words,

$$\mu_{(G,o)}((\rho, 1]) > 0 \quad \Longleftrightarrow \quad \rho(G) > \rho.$$

In particular, since $\mathcal{L}(\rho(G) = 1) = 1$, the right-hand side of (31) is positive, as desired. To prove the second claim, note that the continuity of $(G, o) \mapsto \mu_{(G,o)}$ implies that the event $F_\varepsilon = \left\{\mu_{(G,o)}([\rho, 1]) \leq \varepsilon\right\}$ is closed in $\mathscr{G}_\bullet$. Consequently, the convergence $G_n \to \mathcal{L}$ implies

$$\limsup_{n\to\infty} \mathcal{L}_{G_n}(F_\varepsilon) \quad \leq \quad \mathcal{L}(F_\varepsilon),$$

and the right-hand side tends to $\mathcal{L}(F_0) \leq \mathcal{L}(\rho(G) \leq \rho) = 0$ as $\varepsilon \to 0$. The limsup can then be replaced with a sup, since for each $n \geq 1$, $\mathcal{L}_{G_n}(F_\varepsilon)$ decreases monotonically to 0 with ε. □

Remark 6 (Corollary 18 vs Theorem 12)**.** *The statement (29) asserts that a macroscopic proportion of eigenvalues of G_n accumulate in $[\rho, 1]$, which is exactly the conclusion of Theorem 12. The refinement (30), on the other hand, constitutes a rigorous formalization of the "delocalization" announced in Remark 3. To see this, recall that for any graph G with N vertices, we have by (26),*

$$\mu_{(G,x)}([\rho, 1]) \quad = \quad \sum_{i=1}^{N} \deg_G(x)|\phi_i(x)|^2 \mathbf{1}_{\lambda_i(G)\geq\rho}.$$

In words, the number $\mu_{(G,x)}([\rho, 1]) \in [0, 1]$ measures the cumulative squared amplitude at x of all the basis eigenvectors corresponding to "bad" eigenvalues (those in $[\rho, 1]$). In particular, the set $\{x \in V_G \colon \mu_{(G,x)}([\rho, 1]) \le \varepsilon\}$ represents the region where these "bad" eigenvectors have a small cumulative squared amplitude. The statement (30) asserts that the relative size of this region can be made arbitrarily small by choosing ε small, uniformly in n. Thus, bad eigenvectors have their cumulative mass "spread out" across most vertices.

References

[1] David Aldous and Russell Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007.

[2] David Aldous and J. Michael Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

[3] Venkat Anantharam and Justin Salez. The densest subgraph problem in sparse random graphs. *Ann. Appl. Probab.*, 26(1):305–327, 2016.

[4] A. Avez. Harmonic functions on groups. In *Differential geometry and relativity*, pages 27–32. Mathematical Phys. and Appl. Math., Vol. 3. 1976.

[5] André Avez. Entropie des groupes de type fini. *C. R. Acad. Sci. Paris Sér. A-B*, 275:A1363–A1366, 1972.

[6] André Avez. Théorème de Choquet-Deny pour les groupes à croissance non exponentielle. *C. R. Acad. Sci. Paris Sér. A*, 279:25–28, 1974.

[7] D. Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[8] Dominique Bakry. Étude des transformations de Riesz dans les variétés riemanniennes à courbure de Ricci minorée. In *Séminaire de Probabilités, XXI*, volume 1247 of *Lecture Notes in Math.*, pages 137–172. Springer, Berlin, 1987.

[9] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[10] Itai Benjamini and Nicolas Curien. Ergodic theory on stationary random graphs. *Electron. J. Probab.*, 17:no. 93, 20, 2012.

[11] Itai Benjamini, Hugo Duminil-Copin, Gady Kozma, and Ariel Yadin. Disorder, entropy and harmonic functions. *Ann. Probab.*, 43(5):2332–2373, 2015.

[12] Itai Benjamini, Russell Lyons, and Oded Schramm. Unimodular random trees. *Ergodic Theory Dynam. Systems*, 35(2):359–373, 2015.

[13] Itai Benjamini, Elliot Paquette, and Joshua Pfeffer. Anchored expansion, speed and the Poisson-Voronoi tessellation in symmetric spaces. *Ann. Probab.*, 46(4):1917–1956, 2018.

[14] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13, 2001.

[15] Charles Bordenave. Spectrum of random graphs. In *Advanced topics in random matrices*, volume 53 of *Panor. Synthèses*, pages 91–150. Soc. Math. France, Paris, 2017.

[16] Charles Bordenave, Marc Lelarge, and Justin Salez. The rank of diluted random graphs. *Ann. Probab.*, 39(3):1097–1121, 2011.

[17] Charles Bordenave, Marc Lelarge, and Justin Salez. Matchings on infinite graphs. *Probab. Theory Related Fields*, 157(1-2):183–208, 2013.

[18] Magnus Bordewich and Martin Dyer. Path coupling without contraction. *J. Discrete Algorithms*, 5(2):280–292, 2007.

[19] D. P. Bourne, D. Cushing, S. Liu, F. Münch, and N. Peyerimhoff. Ollivier-Ricci idleness functions of graphs. *SIAM J. Discrete Math.*, 32(2):1408–1424, 2018.

[20] Matías Carrasco Piaggio and Pablo Lessa. Equivalence of zero entropy and the Liouville property for stationary random graphs. *Electron. J. Probab.*, 21:Paper No. 55, 24, 2016.

[21] D. Cushing, S. Kamtue, J. Koolen, S. Liu, F. Münch, and N. Peyerimhoff. Rigidity of the Bonnet-Myers inequality for graphs with respect to Ollivier Ricci curvature. *Adv. Math.*, 369:107188, 53, 2020.

[22] David Cushing, Riikka Kangaslampi, Valtteri Lipiäinen, Shiping Liu, and George W. Stagg. The graph curvature calculator and the curvatures of cubic graphs. *Experimental Mathematics*, 0(0):1–13, 2019.

[23] David Cushing, Shiping Liu, and Norbert Peyerimhoff. Bakry-Émery curvature functions on graphs. *Canad. J. Math.*, 72(1):89–143, 2020.

[24] Ronen Eldan, James R. Lee, and Joseph Lehec. Transport-entropy inequalities and curvature in discrete-space Markov chains. In *A journey through discrete mathematics*, pages 391–406. Springer, Cham, 2017.

[25] Gábor Elek. On the limit of large girth graph sequences. *Combinatorica*, 30(5):553–563, 2010.

[26] Max Fathi and Yan Shu. Curvature and transport inequalities for Markov chains in discrete spaces. *Bernoulli*, 24(1):672–698, 2018.

[27] Bobo Hua. Liouville theorem for bounded harmonic functions on manifolds and graphs satisfying non-negative curvature dimension condition. *Calc. Var. Partial Differential Equations*, 58(2):Paper No. 42, 8, 2019.

[28] Jürgen Jost. *Riemannian geometry and geometric analysis.* Universitext. Springer, Cham, seventh edition, 2017.

[29] Jürgen Jost, Florentin Münch, and Christian Rose. Liouville property and non-negative ollivier curvature on graphs, 2019.

[30] Aldéric Joulin and Yann Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.*, 38(6):2418–2442, 2010.

[31] V. A. Kaĭmanovich and A. M. Vershik. Random walks on discrete groups: boundary and entropy. *Ann. Probab.*, 11(3):457–490, 1983.

[32] Mark Kempton, Gabor Lippner, and Florentin Münch. Large scale Ricci curvature on graphs. *Calc. Var. Partial Differential Equations*, 59(5):Paper No. 166, 17, 2020.

[33] Bo'az Klartag, Gady Kozma, Peter Ralli, and Prasad Tetali. Discrete curvature and abelian groups. *Canad. J. Math.*, 68(3):655–674, 2016.

[34] Yong Lin, Linyuan Lu, and Shing-Tung Yau. Ricci curvature of graphs. *Tohoku Math. J. (2)*, 63(4):605–627, 2011.

[35] Russell Lyons. Asymptotic enumeration of spanning trees. *Combin. Probab. Comput.*, 14(4):491–522, 2005.

[36] Russell Lyons and Shayan Oveis Gharan. Sharp bounds on random walk eigenvalues via spectral embedding. *Int. Math. Res. Not. IMRN*, (24):7555–7605, 2018.

[37] Russell Lyons, Robin Pemantle, and Yuval Peres. Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems*, 15(3):593–619, 1995.

[38] Russell Lyons and Yuval Peres. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016.

[39] Florentin Münch. Li-Yau inequality under $CD(0, n)$ on graphs. *arXiv e-prints*, page arXiv:1909.10242, September 2019.

[40] Florentin Münch. Non-negative Ollivier curvature on graphs, reverse Poincaré inequality, Buser inequality, Liouville property, Harnack inequality and eigenvalue estimates. *arXiv e-prints*, page arXiv:1907.13514, July 2019.

[41] Florentin Münch and Radosław K. Wojciechowski. Ollivier Ricci curvature for general graph Laplacians: heat equation, Laplacian comparison, non-explosion and diameter bounds. *Adv. Math.*, 356:106759, 45, 2019.

[42] Yann Ollivier. Ricci curvature of metric spaces. *C. R. Math. Acad. Sci. Paris*, 345(11):643–646, 2007.

[43] Yann Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256(3):810–864, 2009.

[44] Yann Ollivier. A survey of Ricci curvature for metric spaces and Markov chains. In *Probabilistic approach to geometry*, volume 57 of *Adv. Stud. Pure Math.*, pages 343–381. Math. Soc. Japan, Tokyo, 2010.

[45] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.