Journal de l'École polytechnique
Mathématiques

Florentin Münch & Justin Salez

Mixing time and expansion of non-negatively curved Markov chains

MIXING TIME AND EXPANSION OF
NON-NEGATIVELY CURVED MARKOV CHAINS

by Florentin Münch & Justin Salez

Abstract. — We establish three remarkable consequences of non-negative curvature for sparse Markov chains. First, their conductance decreases logarithmically with the number of states. Second, their displacement is at least diffusive until the mixing time. Third, they never exhibit the cutoff phenomenon. The first result provides a nearly sharp quantitative answer to a classical question of Ollivier, Milman & Naor. The second settles a conjecture of Lee and Peres for graphs with non-negative curvature. The third offers a striking counterpoint to the recently established cutoff for non-negatively curved chains with uniform expansion.

Résumé (Temps de mélange et expansion des chaînes de Markov en courbure positive)

Nous établissons trois conséquences remarquables de la courbure positive pour les chaînes de Markov. D'abord, la conductance de ces chaînes décroît logarithmiquement avec la taille de l'espace. Ensuite, leur déplacement est diffusif jusqu'au temps de mélange. Enfin, le phénomène de cutoff ne peut pas se produire. Le premier résultat fournit une réponse quantitative presqu'optimale à une question classique d'Ollivier, Milman et Naor. Le second confirme une conjecture de Lee et Peres, dans le cas particulier des graphes à courbure positive. Le troisième offre un contraste frappant avec les résultats positifs récents concernant le cutoff pour les chaînes de Markov ayant à la fois une courbure positive et une expansion uniforme.

Contents

1. Introduction

In Riemannian geometry, a lower bound on the Ricci curvature classically implies an array of powerful estimates for the underlying manifold, including diameter bounds, volume growth, comparison principles, splitting theorems, spectral estimates, and concentration inequalities [10]. Over the past decade, those remarkable implications have motivated the development of non-smooth analogues of curvature that

can be applied to discrete geometries [27, 15, 9, 8, 21, 11, 20, 19]. In particular, Ollivier [21] proposed a transportation-based definition that makes sense on arbitrary metric spaces, hence in particular on graphs and Markov chains. Informally, a metric space has non-negative Ollivier-Ricci curvature if balls are at least as close to each other as their centers are. The simplest example of a finite non-negatively curved graph is a cycle. It is classical that this graph has poor expansion, that the random walk on it exhibits a diffusive behavior, and that its mixing time is of the same order as the inverse spectral gap. The aim of the present paper is to show that those three properties are in fact shared by all sparse Markov chains with non-negative curvature. Before we state our results in full generality, let us describe their content in the simple but important special case of random walk on graphs.

Non-negatively curved graphs. — Let $G = (V, E)$ be a finite simple graph, and let P denote the random-walk transition matrix of G. Thus, P acts on any function $f: V \to \mathbb{R}$ as follows:

$$(Pf)(x) = \frac{1}{\deg(x)} \sum_{y \sim x} f(y),$$

where the notation $y \sim x$ indicates that $\{x, y\} \in E$. Following Ollivier [21, 22], we say that G has *non-negative curvature* if P contracts the Lipschitz norm, i.e.,

$$\|Pf\|_{\text{LIP}} \leqslant \|f\|_{\text{LIP}},$$

where $\|f\|_{\text{LIP}} := \max_{y \sim x} |f(y) - f(x)|$. This fundamental property is satisfied by many natural families of graphs, including all Abelian Cayley graphs and, more generally, all Cayley graphs whose generating set is conjugacy-invariant. Additional details, including a more effective formulation in terms of couplings, will be provided in the next section when we discuss curvature for general Markov chains.

Expansion. — Our first result concerns the expansion of graphs with non-negative curvature. Write ∂A for the edge-boundary of a set $A \subseteq V$, and $\deg(A)$ for the sum of the degrees of all vertices in A. With this notation, the *conductance* (also known as Cheeger constant, or bottleneck ratio) of G is

$$\Phi := \min\left\{ \frac{|\partial A|}{\deg(A)} : A \subseteq V, 0 < \deg(A) \leqslant |E| \right\}.$$

Sequences of bounded-degree graphs whose size diverges but whose conductance remains bounded away from zero are famously known as *expanders*. Whether such graphs can have non-negative curvature is an important question, which explicitly appears in a survey by Ollivier [22, Probl. T], and is therein attributed to Milman and Naor. The problem has remained open until very recently, when a negative answer was given by the second author [26]. Specifically, the latter used the notion of entropy for graph limits to prove that non-negative curvature and expansion are incompatible "at infinity", and the conclusion was then transferred to finite graphs using a compactness argument. A clear drawback of this approach is its non-quantitative nature. In particular, the second author asked for a direct, quantitative relation between

volume, degree and expansion on non-negatively curved graphs. This is precisely the content of our first main result.

Theorem 1 (Poor expansion). — *If G has non-negative curvature, then*

$$\Phi \leqslant c\sqrt{\frac{d\log d}{\log n}},$$

where n is the number of vertices, d the maximum degree, and c a universal constant.

In other words, large graphs can not simultaneously enjoy non-negative curvature and uniform expansion unless their maximum degree grows at least like $\log n/\log\log n$. We note that this is sharp up to the $\log\log n$ correction. Indeed, a celebrated result of Alon and Roichman asserts that random Cayley graphs with logarithmic degrees have uniform expansion with high probability [1], and specializing this result to random Cayley graphs of Abelian groups produces examples of non-negatively curved graphs with logarithmic degrees and uniform expansion.

Mixing times. — Our second result is a complete determination of the order of magnitude of the mixing time of all vertex-transitive graphs with bounded degrees and non-negative curvature. Suppose that G is vertex-transitive, with degree d and volume n. Fix an arbitrary origin $x \in V$ (the choice is irrelevant, by transitivity), and consider the lazy simple random walk on G started at x, i.e., the Markov chain $(X_t)_{t\geqslant 0}$ on V with initial condition $X_0 = x$ and transition matrix $(P+I)/2$. The *mixing time* of G is a fundamental graph-theoretical parameter, defined as follows [13]:

$$\mathrm{t_{MIX}} := \min\left\{t \in \mathbb{N}\colon \max_{A \subseteq V}\left|\mathbb{P}(X_t \in A) - \frac{|A|}{n}\right| \leqslant \frac{1}{4}\right\}.$$

An important, closely related quantity is the so-called *relaxation time*

$$\mathrm{t_{REL}} := \frac{1}{1-\lambda_2},$$

where $1 = \lambda_1 > \lambda_2 \geqslant \cdots \geqslant \lambda_n$ denote the ordered eigenvalues of P. It is classical that $\mathrm{t_{MIX}} \geqslant \mathrm{t_{REL}}$, and that this inequality can be off by a factor as large as $\log n$, as is the case for expanders (see [13]).

Theorem 2 (Mixing times). — *All vertex-transitive graphs with non-negative curvature satisfy*

$$\mathrm{t_{MIX}} \;\asymp_d\; \mathrm{t_{REL}} \;\asymp_d\; \frac{1}{\Phi^2},$$

where the notation $a \asymp_d b$ means that the ratio a/b is bounded from above and below by positive constants that depend only on the degree d.

This has the following remarkable consequence. For a sequence of graphs $(G_n)_{n\geqslant 1}$, the condition

$$\frac{\mathrm{t_{MIX}}(G_n)}{\mathrm{t_{REL}}(G_n)} \xrightarrow[n\to\infty]{} +\infty$$

is known as the *product condition*. It is well-known to be necessary (see [13, Prop. 18.4]) for the occurrence of the so-called *cutoff phenomenon*, a celebrated but still mysterious phase transition in the approach to equilibrium of certain Markov chains (see [13, Chap. 18] for the precise definition). Thus, Theorem 2 implies that vertex-transitive graphs with fixed degree and non-negative curvature never exhibit cutoff. This stands in stark contrast with recent results due to the second author, showing that many non-negatively curved graphs with logarithmic degree do exhibit cutoff [25]. Interestingly, the conclusion of Theorem 2 is known to hold for fixed-degree Cayley graphs of *moderate growth* [6]. This geometric condition was later shown to be equivalent to the much simpler requirement that the diameter is algebraically large in the volume [5] (see the recent paper [28] for an extension to vertex-transitive graphs). This raises the following question. A positive answer would be surprisingly strong, but we have not been able to produce any counter-example.

QUESTION 1 (Moderate growth?) — Do all non-negatively curved graphs with degree at most d satisfy

$$\mathrm{diam}(G) \geqslant \varepsilon_d n^{\varepsilon_d},$$

where n is the number of vertices, and $\varepsilon_d > 0$ a constant depending only on d?

Indeed, this question was answered affirmatively in case of a modified Barky Emery curvature dimension condition in [2], and later by the first author in case of the weaker, unmodified Bakry Emery curvature dimension condition [17].

Diffusivity. — Finally, our last result concerns the speed of random walk on vertex-transitive graphs with non-negative curvature. Many infinite graphs such as the line $\mathbb{Z}$ are known to exhibit a *diffusive* behavior, in the sense that the typical graph distance between X_t and X_0 grows like $\sqrt{t}$. On a finite graph, the distance to the starting point can of course no longer grow indefinitely with time, but one may still hope for a diffusive behavior on *appropriate time-scales*. This vague statement was recently given a powerful rigorous content by Lee and Peres [12], who showed that the simple random walk on any finite vertex-transitive graph satisfies the diffusive lower-bound

$$\mathbb{E}\big[\mathrm{dist}(X_0, X_t)\big] \geqslant c\sqrt{\frac{t}{d}}$$

for all $t \in [d, \mathrm{t_{REL}}]$, where $c > 0$ is a universal constant. The graph $\mathbb{Z}_2^d \times \mathbb{Z}_n$ shows that this lower-bound is sharp. However, the authors conjectured that the time-scale on which the diffusive behavior remains valid should actually be much longer, namely, of order $\mathrm{t_{MIX}}$ [12, Conj. 2.5]. Our second result confirms this prediction in the case of non-negatively curved graphs.

THEOREM 3 (Diffusive lower-bound). — *If G is vertex-transitive and non-negatively curved, then*

$$\mathbb{E}\big[\mathrm{dist}(X_0, X_t)\big] \geqslant c\sqrt{\frac{t}{d}},$$

for $t \in [d, \mathrm{t_{MIX}}]$, where c is a universal constant.

We emphasize that our estimates are not restricted to simple random walks on graphs. Analogous results will be stated for general Markov chains with non-negative curvature. In particular, neither reversibility, nor even the symmetry of the support of P are actually required for a version of Theorem 1 to hold. Ollivier curvature with respect to a directed metric has been explored before in [30, 23, 24, 7]. However the specific consequences of non-negative curvature seem to be unexplored to the best of our knowledge. Our general results are exposed in Section 2 below, and are proved in Section 3.

2. Main results

In the remainder of the paper, we consider an arbitrary, irreducible stochastic matrix P on a finite state space V. A natural measure of the "distance" from a state $x \in V$ to a state $y \in V$ is the minimum number of transitions needed for the chain to move from x to y, namely

$$\mathrm{dist}(x, y) := \min\{k \in \mathbb{N} : P^k(x, y) > 0\}.$$

This quantity is not necessarily symmetric, but it clearly satisfies the two other axioms of a distance. We may then use optimal transport to extend this notion to probability measures as follows: write $\mathcal{P}(V)$ for the set of probability measures on V, and define $W \colon \mathcal{P}(V) \times \mathcal{P}(V) \to \mathbb{R}_+$ by

$$W(\mu, \nu) := \inf_{\substack{X \sim \mu \\ Y \sim \nu}} \mathbb{E}\big[\mathrm{dist}(X, Y)\big],$$

where the infimum runs over all possible random pairs (X, Y) whose marginals are μ and ν. Again, this quantity is not necessarily symmetric, but it always satisfies the two other axioms of a distance. Due to Kantorovich duality [29, Th. 5.10 & Particular Case 5.4], we can write

$$W(\mu, \nu) = \sup\left\{\nu f - \mu f : \|f\|_{\mathrm{LIP}} \leqslant 1\right\}$$

with

$$\|f\|_{\mathrm{LIP}} := \sup_{y \sim x} f(y) - f(x).$$

Finally, we say that P has *non-negative curvature* if it is a contraction under W, i.e.,

$$(1) \qquad\qquad \forall \mu, \nu \in \mathcal{P}(V), \quad W(\mu P, \nu P) \leqslant W(\mu, \nu).$$

Ollivier curvature with a non-symmetric distance has been studied in [30, 23, 7, 24]. Due to Kantorovich duality and as in the introduction, non-negative curvature is equivalent to $\|Pf\|_{\mathrm{LIP}} \leqslant \|f\|_{\mathrm{LIP}}$. By convexity, it is in fact sufficient to check property (1) on Dirac masses $\mu = \delta_x, \nu = \delta_y, x, y \in V$. Moreover, by the triangle inequality, we may further restrict our attention to the case where y is a neighbor of x (by which we mean that $\mathrm{dist}(x, y) = 1$ and which we denote by $y \sim x$), i.e.,

$$(2) \qquad\qquad \forall y \sim x, \quad W\left(P(x, \cdot), P(y, \cdot)\right) \leqslant 1.$$

This local condition is easily verified in practice. For example, it holds for random walks on Abelian groups and, more generally, random walks with a conjugacy-invariant support, as we now explain.

Example 1 (Random walks on groups). — *Suppose that V is a group, and fix $\mu \in \mathcal{P}(V)$. By definition, the* random walk *on V with increment distribution μ is the Markov chain whose transitions correspond to left-multiplication by a μ−distributed element, i.e., $P(x,y) := \mu(yx^{-1})$. This chain has non-negative curvature as soon as the set $S := \{z \in V : \mu(z) > 0\}$ is* conjugacy-invariant, *i.e.,*

$$(3) \qquad\qquad \forall z \in V, \quad zSz^{-1} = S.$$

Indeed, this assumption implies that $\mathrm{dist}(zx, zy) = \mathrm{dist}(x,y)$ for all $x, y, z \in V$. In particular, if Z denotes a random variable with law μ, then the "obvious" coupling of $P(x,\cdot)$ and $P(y,\cdot)$ given by $X := Zx$ and $Y := Zy$ verifies (2). Note that the condition (3) trivially holds if the group is Abelian. An emblematic non-Abelian example is the transposition walk on the symmetric group [3].

To avoid periodicity issues, we now assume that P is lazy, i.e., $P(x,x) \geqslant 1/2$ for all $x \in V$. This is more than enough to guarantee that the chain mixes, i.e.,

$$\forall x, y \in V, \quad P^t(x,y) \xrightarrow[t\to\infty]{} \pi(y),$$

where $\pi = \pi P$ denotes the unique invariant distribution. Quantifying the speed at which this convergence to equilibrium occurs is a fundamental question, with many applications [13, 16]. Formally, this amounts to estimating the so-called *mixing time*:

$$t_{\mathrm{MIX}} := \min\{t \geqslant 0 : d_{\mathrm{TV}}(t) \leqslant 1/4\}, \quad \text{where} \quad d_{\mathrm{TV}}(t) := \max_{x \in V} \big\|P^t(x,\cdot) - \pi\big\|_{\mathrm{TV}}.$$

Here $\|\mu - \nu\|_{\mathrm{TV}}$ denotes the *total-variation distance* between $\mu, \nu \in \mathcal{P}(V)$, defined as

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A \subseteq V} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{x \in V} |\mu(x) - \nu(x)| = \inf_{\substack{X \sim \mu \\ Y \sim \nu}} \mathbb{P}(X \neq Y),$$

where the infimum in the last expression runs over all possible couplings (X, Y) of μ and ν. Thus, a natural way to estimate mixing times is to exhibit good couplings, and this is precisely where curvature enters the play. Indeed, an elementary but crucial reformulation of the non-negative curvature assumption (1) is that the trajectories $(X_t)_{t \geqslant 0}$ and $(Y_t)_{t \geqslant 0}$ emanating from any two states $X_0 = x$ and $Y_0 = y$ can be coupled in such a way that their distance $t \mapsto \mathrm{dist}(X_t, Y_t)$ forms a super-martingale. When combined with an appropriate diffusive estimate for super-martingales, this observation turns out to imply the following $O(1/\sqrt{t})$ decay for the total-variation distance between the laws of X_t and Y_t.

Theorem 4 (Total-variation decay). — *If P is lazy and non-negatively curved, then,*

$$\big\|P^t(x,\cdot) - P^t(y,\cdot)\big\|_{\mathrm{TV}} \leqslant \mathrm{dist}(x,y) \sqrt{\frac{10}{(t+1)P_{\min}}},$$

for all $x, y \in V$ and all $t \geqslant 0$, where $P_{\min}$ denotes the smallest non-zero entry of P.

Variants of this result have appeared in a number of works, under various forms [4, 14, 18, 26]. However, all proofs use the fact that the increments of the process $t \mapsto \mathrm{dist}(X_t, Y_t)$ are uniformly bounded (by 2), and this property may dramatically fail in our more general setup where the metric is directed. Nevertheless, the conclusion turns out to remain valid, and a proof is presented in Section 3.3. The most "obvious" application of Theorem 4 consists in taking a maximum over all states $x, y \in V$ to obtain the following mixing-time estimate, which is new in our directed setup.

COROLLARY 1 (Diameter bound). — *If P is lazy and non-negatively curved, then*

$$\mathrm{t}_{\mathrm{MIX}} \leqslant \frac{160\,(\mathrm{diam})^2}{P_{\min}},$$

where $\mathrm{diam} := \max_{x,y} \mathrm{dist}(x, y)$ *denotes the diameter of the state space.*

While interesting in its own right, this estimate is actually not the key to the new results mentioned in the Introduction. Our main finding is that a significantly finer estimate can be deduced from Theorem 4 provided we replace the *worst-case* mixing time by its *average* version:

$$\mathrm{t}_{\mathrm{MIX}}^{\sharp} := \min\{t \geqslant 0 : d_{\mathrm{TV}}^{\sharp}(t) \leqslant 1/4\}, \quad \text{where} \quad d_{\mathrm{TV}}^{\sharp}(t) := \sum_{x \in V} \pi(x) \big\| P^t(x, \cdot) - \pi \big\|_{\mathrm{TV}}.$$

REMARK 1 (Transitive chains). — Obtaining a bound on $\mathrm{t}_{\mathrm{MIX}}^{\sharp}$ rather than $\mathrm{t}_{\mathrm{MIX}}$ is not a huge drawback. For example, we have $\mathrm{t}_{\mathrm{MIX}}^{\sharp} = \mathrm{t}_{\mathrm{MIX}}$ for all random walks on groups and, more generally, for all *transitive* chains (P is transitive if for each $x, y \in V$, there is a bijection $f \colon V \to V$ which maps x to y and preserves the transition kernel, i.e., $P(f(u), f(v)) = P(u, v)$ for all $u, v \in V$).

Throughout the paper, we let $X = (X_t)_{t \geqslant 0}$ denote a Markov chain with transition matrix P starting from stationarity ($X_0 \sim \pi$). Our main new estimate on $\mathrm{t}_{\mathrm{MIX}}^{\sharp}$ depends on two statistics of this chain. The first one is the *mean displacement* in t steps:

$$\mathbb{E}\big[\mathrm{dist}(X_0, X_t)\big] = \sum_{x,y \in V} \pi(x) P^t(x, y)\,\mathrm{dist}(x, y).$$

The second is the *escape probability* in t steps, i.e., the conductance of P^t:

$$\Phi(P^t) = \min\Big\{\mathbb{P}\left(X_t \notin A | X_0 \in A\right) : A \subseteq V, 0 < \pi(A) \leqslant \frac{1}{2}\Big\}$$
$$= \min\Big\{\frac{1}{\pi(A)} \sum_{x \in A} \sum_{y \in A^c} \pi(x) P^t(x, y) : A \subseteq V, 0 < \pi(A) \leqslant \frac{1}{2}\Big\}.$$

THEOREM 5 (Main estimate). — *If P is lazy and non-negatively curved, then*

$$(4) \qquad \mathrm{t}_{\mathrm{MIX}}^{\sharp} \leqslant \frac{160}{P_{\min}} \inf_{t \geqslant 1}\Big\{\frac{\mathbb{E}[\mathrm{dist}(X_0, X_t)]}{\Phi(P^t)}\Big\}^2,$$

where we recall that $P_{\min}$ denotes the smallest non-zero entry of P.

Theorem 5 has a number of notable consequences, which we now enumerate. The simplest one is an "average" version of Corollary 1, obtained by sending $t \to \infty$ in the infimum (4):

$$(5) \qquad\qquad t_{\mathrm{MIX}}^{\sharp} \leqslant \frac{640 \, (\mathrm{diam}^{\sharp})^2}{P_{\min}},$$

where $\mathrm{diam}^{\sharp} := \sum_{x,y \in V} \pi(x)\pi(y) \, \mathrm{dist}(x,y)$ denotes the *effective diameter*. Note that the latter can be significantly smaller than the true diameter appearing in Corollary 1 (consider, e.g., the biased random walk on a segment). A much more refined consequence of Theorem 5 is obtained by taking $t = 1$ in the infimum (4): writing $\Phi = \Phi(P)$, we readily obtain the following surprising bound.

Corollary 2 (Conductance bound). — *If P is lazy and non-negatively curved, then*

$$t_{\mathrm{MIX}}^{\sharp} \leqslant \frac{40}{P_{\min}\Phi^2}.$$

This offers a considerable improvement over (5) in situations where the effective diameter diverges while the conductance remains bounded away from 0 (consider, e.g., random walk on a random Abelian Cayley graph with logarithmic degree). More importantly, by virtue of an elementary combinatorial lower-bound on $t_{\mathrm{MIX}}^{\sharp}$ (see, e.g., [13, §7.1.1]), Corollary 2 implies the quantitative non-existence of non-negatively curved expanders promised in Theorem 1. For general chains, we will show that $t_{\mathrm{MIX}}^{\sharp}$ can be bounded below by $\mathrm{diam}^{\sharp}$, leading to the following result.

Corollary 3 (Poor expansion). — *If P is non-negatively curved, then*

$$\Phi \leqslant \frac{19}{\sqrt{P_{\min} \, \mathrm{diam}^{\sharp}}}.$$

Thus, non-negatively curved chains which are large ($\mathrm{diam}^{\sharp} \gg 1$) and sparse ($P_{\min}$ bounded away from 0) must have poor expansion ($\Phi \ll 1$). This constitutes a precise quantitative answer to the Markov-chain generalization of the question of Milman, Naor and Ollivier [22, Probl. T]. Note that there are examples of sparse chains with non-negative curvature and arbitrarily many states (consider, e.g., a biased random walk on a segment). However, the fact that their effective diameter is bounded forces their stationary measure to concentrate on a bounded number of states.

Corollary 2 is sharp in the important case where P is transitive, reversible and sparse. Indeed, we have the classical lower-bound $t_{\mathrm{MIX}} \geqslant t_{\mathrm{REL}}$, where $t_{\mathrm{REL}} := (1 - \lambda_2)^{-1}$ denotes the inverse spectral gap of P (see, e.g., [13]), and the first author established in [18] the Buser inequality

$$t_{\mathrm{REL}} \geqslant \frac{P_{\min}}{12\Phi^2},$$

for any non-negatively curved, reversible chain. When combined with Corollary 2, this yields the following result, of which Theorem 2 is clearly a special case.

Corollary 4 (No cutoff for sparse chains). — *Fix $p \in (0,1)$. Then, any lazy reversible transitive chain with non-negative curvature and $P_{\min} \geqslant p$ satisfies*

$$\mathrm{t_{MIX}} \; \asymp_p \; \mathrm{t_{REL}} \; \asymp_p \; \frac{1}{\Phi^2},$$

where the notation $a \asymp_p b$ means that the ratio a/b is bounded from above and below by positive constants that depend only on p. In particular, no family of such chains can exhibit cutoff.

An important observation here is that the transitivity of the chain is only used to ensure that $\mathrm{t_{MIX}} = \mathrm{t_{MIX}^{\sharp}}$. Consequently, Corollary 4 extends to any collection of chains which are "spatially homogeneous" in the mild sense that $\mathrm{t_{MIX}} \asymp \mathrm{t_{MIX}^{\sharp}}$.

Finally, a last notable consequence of Theorem 5 is that the expected displacement of the chain over short time-scales is already substantial. More precisely, assuming that P is reversible, we have

$$\Phi(P^t) \geqslant \frac{1 - \lambda_2^t}{2} \geqslant \frac{1 - e^{-t/\mathrm{t_{REL}}}}{2},$$

and the right-hand side is at least $(1 - e^{-1})/2$ for all $t \geqslant \mathrm{t_{REL}}$, yielding the following estimate.

Corollary 5 (Fast escape). — *If P is lazy, reversible and non-negatively curved, then*

$$\forall t \geqslant \mathrm{t_{REL}}, \quad \mathbb{E}\big[\mathrm{dist}(X_0, X_t)\big] \geqslant \frac{\sqrt{\mathrm{t_{MIX}^{\sharp}} P_{\min}}}{41},$$

where we recall that $\mathbb{E}\big[\mathrm{dist}(X_0, X_t)\big] = \sum_{x,y} \pi(x) P^t(x,y) \, \mathrm{dist}(x,y)$.

For reversible transitive chains, Lee and Peres [12] proved the diffusive lower-bound

$$\mathbb{E}\big[\mathrm{dist}(X_0, X_t)\big] \geqslant c\sqrt{t P_{\min}},$$

for all t such that $P_{\min}^{-1} \leqslant t \leqslant \mathrm{t_{REL}}$, where $c > 0$ is a universal constant. They conjectured that this diffusive lower-bound should remain valid until the mixing time [12, Conj. 2.5]. Corollary 5 readily implies that this is true in the non-negatively curved case, and Theorem 3 follows as a special case.

3. Proofs

Section 3.1 below is devoted to the proof of our main result, namely the relation between conductance, displacement and mixing times (Theorem 5). The latter exploits the diffusive total-variation decay of non-negatively curved chains (Theorem 4), which will be proved independently in Section 3.3. Once Theorem 5 is established, all announced corollaries follow effortlessly, except for Corollary 3: the latter requires a lower bound on the average mixing time in terms of the effective diameter, which we prove in Section 3.2.

3.1. MIXING TIME VS. CONDUCTANCE. — In this section, we prove Theorem 5. We will make crucial use of Theorem 4, as well as the following L^1 version of Cheeger's inequality. An important remark is that the latter holds without any assumption on the transition matrix P, and will thus also apply to powers of P.

LEMMA 1 (L^1 analogue of Cheeger's inequality). — *If $f\colon V \to \mathbb{R}$ satisfies $\pi f = 0$, then,*

$$\sum_{x \in V} \pi(x) |f(x)| \leqslant \frac{1}{\Phi(P)} \sum_{x,y \in V} \pi(x)P(x,y) |f(y) - f(x)|.$$

Proof. — Upon replacing f with $-f$, we may assume that $\pi(f \geqslant 0) \geqslant 1/2$. For any $t \geqslant 0$, we may take $A = \{f \geqslant t\}$ in the definition of $\Phi(P)$ to obtain

$$\Phi(P) \sum_{x \in V} \pi(x)\mathbf{1}_{(f(x) \geqslant t)} \leqslant \sum_{x,y \in V} \pi(x)P(x,y)\mathbf{1}_{(f(y) \leqslant t < f(x))}.$$

Integrating over $t \in \mathbb{R}_+$ and interchanging the sum and integral, we obtain

$$\Phi(P) \sum_{x \in V} \pi(x)f_+(x) \leqslant \sum_{x,y \in V} \pi(x)P(x,y)(f_+(x) - f_+(y))_+,$$

where $a_+ := \max(0, a)$ denotes the positive part of a. Now, since f is centered under π, the left-hand side does not change if we replace $f_+(x)$ by $|f(x)|/2$. Similarly, since any gradient is centered under the measure $(x, y) \mapsto \pi(x)P(x,y)$, the right-hand side does not change if we replace $(f_+(x) - f_+(y))_+$ by $|f_+(x) - f_+(y)|/2$. Finally, observe that $|f_+(x) - f_+(y)| \leqslant |f(x) - f(y)|$. $\qquad\square$

Proof of Theorem 5. — Fix $s \geqslant 1$. Lemma 1 applied to P^s instead of P gives

$$\sum_{x \in V} \pi(x) |f(x)| \leqslant \frac{1}{\Phi(P^s)} \sum_{x,y \in V} \pi(x)P^s(x,y)|f(x) - f(y)|,$$

for any centered observable $f\colon V \to \mathbb{R}$. Now, fix $t \in \mathbb{N}$ and $z \in V$, and let us apply this to the observable $f(x) := P^t(x,z) - \pi(z)$, which is centered because $\pi P^t = \pi$. We readily obtain

$$\sum_{x \in V} \pi(x) \left| P^t(x,z) - \pi(z) \right| \leqslant \frac{1}{\Phi(P^s)} \sum_{x,y \in V} \pi(x)P^s(x,y) \left| P^t(x,z) - P^t(y,z) \right|.$$

We may now sum over all $z \in V$ and use Theorem 4 to get

$$d_{\mathrm{TV}}^\sharp(t) \leqslant \frac{1}{\Phi(P^s)} \sum_{x,y \in V} \pi(x)P^s(x,y) \left\| P^t(x,\cdot) - P^t(y,\cdot) \right\|_{\mathrm{TV}}$$

$$\leqslant \sqrt{\frac{10}{(t+1)P_{\min}}} \, \frac{\mathbb{E}\left[\mathrm{dist}(X_0, X_s)\right]}{\Phi(P^s)}.$$

Finally, choosing t so that the right-hand side is smaller than $1/4$ shows that

$$\mathrm{t}_{\mathrm{MIX}}^\sharp \leqslant \frac{160}{P_{\min}} \left(\frac{\mathbb{E}\left[\mathrm{dist}(X_0, X_s)\right]}{\Phi(P^s)} \right)^2.$$

The result follows by taking an infimum over all $s \geqslant 1$. $\qquad\square$

3.2. Effective diameter vs. conductance. — Here we prove Corollary 3. We will use the following concentration inequality.

Lemma 2 (Concentration inequality). — *For any $f\colon V \to \mathbb{R}$ and $a > 0$,*

$$\pi\left(f \geqslant \pi f + a\right) \leqslant \frac{1}{a\Phi} \min\left(\max_{y \sim x}(f(y) - f(x))_+, \max_{y \sim x}(f(x) - f(y))_+\right),$$

and the same holds for $\pi(f \leqslant \pi f - a)$.

We remark that if "$\sim$" is symmetric, then the minimum is the Lipschitz constant of f.

Proof. — Upon replacing f by $f - \pi f$ if necessary, we may assume that f is centered under π, i.e., $\pi f = 0$. Now, we use Markov's inequality and Lemma 1 to write

$$\pi\left(f \geqslant a\right) \leqslant \frac{1}{a} \sum_{x \in V} \pi(x) f_+(x) = \frac{1}{2a} \sum_{x \in V} \pi(x)|f(x)|$$

$$\leqslant \frac{1}{2a\Phi} \sum_{x,y \in V} \pi(x)P(x,y)\,|f(x) - f(y)|$$

$$= \frac{1}{a\Phi} \sum_{x,y \in V} \pi(x)P(x,y)\,(f(x) - f(y))_+$$

$$\leqslant \frac{1}{a\Phi} \max_{y \sim x}(f(x) - f(y))_+.$$

Note that, since the function $(x,y) \mapsto f(x) - f(y)$ is centered under the measure $(x,y) \mapsto \pi(x)P(x,y)$, the integral of its positive part equals that of its negative part. This establishes the first claim, and the second is obtained by replacing f with $-f$. $\square$

We will use this lemma to prove the following lower-bound on the mixing time.

Lemma 3 (Diameter lower-bound). — *For any lazy chain P, we have*

$$t_{\mathrm{MIX}}^\sharp \geqslant \mathrm{diam}^\sharp - \frac{4}{\Phi}.$$

Proof. — Let us first note that for a lazy chain, Lemma 2 holds with the better constant 2Φ instead of Φ (just apply the lemma to the non-lazy chain $2P - I$). Now, fix $x \in V$ and $t \in \mathbb{N}$, and write $B_x(t) := \{y \in V \colon \mathrm{dist}(x,y) \leqslant t\}$. By definition,

$$\|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} = \max_{A \subseteq V} |P^t(x,A) - \pi(A)|$$

$$\geqslant 1 - \pi\left(B_x(t)\right)$$

$$= 1 - \mathbb{P}\left(\mathrm{dist}(x,Y) \leqslant t\right),$$

where Y denotes a π−distributed random variable. Averaging over $x \in V$, we obtain

$$d_{\mathrm{TV}}^\sharp(t) \geqslant 1 - \mathbb{P}\left(\mathrm{dist}(X,Y) \leqslant t\right),$$

where X is π−distributed and independent of Y. The claim follows if we can show

$$(6) \qquad \mathbb{P}\left(\mathrm{dist}(X,Y) \leqslant \mathrm{diam}^\sharp - 4/\Phi\right) \leqslant \frac{1}{2}.$$

Now, for fixed $x \in V$, the function $f \colon y \mapsto \mathrm{dist}(x, y)$ satisfies $f(z) \leqslant f(y) + 1$ whenever $z \sim y$, by the triangle inequality. Thus, Lemma 2 ensures that

$$\mathbb{P}\Big(\mathrm{dist}(X, Y) \leqslant \mathbb{E}\big[\mathrm{dist}(X, Y)|X\big] - 2/\Phi\Big) \leqslant \frac{1}{4}.$$

On the other hand, by the triangle inequality again, the function $f \colon x \mapsto \mathbb{E}[\mathrm{dist}(x, Y)]$ satisfies $f(x) \leqslant f(y) + 1$ whenever $y \sim x$, and $\pi f = \mathbb{E}[\mathrm{dist}(X, Y)] = \mathrm{diam}^{\sharp}$. Thus, Lemma 2 yields

$$\mathbb{P}\left(\mathbb{E}\big[\mathrm{dist}(X, Y)|X\big] \leqslant \mathrm{diam}^{\sharp} - 2/\Phi\right) \leqslant \frac{1}{4}.$$

Combining those two estimates readily yields (6).   □

Proof of Corollary 3. — If P is lazy and non-negatively curved, then Corollary 2 and Lemma 3 give

$$\mathrm{diam}^{\sharp} \leqslant \frac{40}{\Phi^2 P_{\min}} + \frac{4}{\Phi} \leqslant \frac{41}{\Phi^2 P_{\min}},$$

because $P_{\min}, \Phi \leqslant \frac{1}{2}$. If P is not lazy, we apply the above result to $(P + I)/2$. The latter is still non-negatively curved, but its conductance and minimal entry are half those of P, so we loose a factor of 8 and obtain

$$\mathrm{diam}^{\sharp} \leqslant \frac{328}{\Phi^2 P_{\min}}.$$

This readily implies the claim, because $\sqrt{328} < 19$.   □

3.3. Diffusive total-variation decay. — In this section, we prove Theorem 4. Fix two distinct states $x \neq y$, and recall that

$$W\left(P(x, \cdot), P(y, \cdot)\right) = \inf_{\chi}\Big\{\textstyle\sum_{u, v \in V} \chi(u, v)\, \mathrm{dist}(u, v)\Big\},$$

where the infimum runs over all probability distributions $\chi \in \mathcal{P}(V^2)$ with marginals $P(x, \cdot)$ and $P(y, \cdot)$. Minimizers are called optimal couplings. As in [4, 18], our first task consists in showing that they can be chosen so as to assign a "decent" probability to the "good" set

$$\Gamma := \big\{(u, v) \in V^2 \colon \mathrm{dist}(u, v) < \mathrm{dist}(x, y)\big\}.$$

Lemma 4 (Good optimal couplings). — *If P is lazy and $x \neq y$, then there is an optimal coupling χ of $P(x, \cdot), P(y, \cdot)$ such that $\chi(\Gamma) \geqslant P_{\min}$.*

Proof. — By compactness, we can find an optimal coupling χ which, among all optimal couplings, maximizes $\chi(\Gamma)$. Suppose for a contradiction that this "doubly optimal" coupling satisfies $\chi(\Gamma) < P_{\min}$. The set $A := \{u \sim x \colon (u, y) \in \Gamma\}$ is not empty, since it contains the first vertex on a geodesic from x to y. Thus, $\chi(A \times V) = P(x, A) \geqslant P_{\min} > \chi(\Gamma)$. This forces $\chi((A \times V) \smallsetminus \Gamma) > 0$, i.e.,

$$(7) \qquad\qquad \exists (x_0, y_0) \in (A \times V) \smallsetminus \Gamma, \quad \chi(x_0, y_0) \geqslant \varepsilon,$$

for some $\varepsilon > 0$. On the other hand, we have $\chi(A \times \{y\}) + \chi(A^c \times \{y\}) = P(y, y) \geqslant 1/2$. This forces $\chi(A^c \times \{y\}) > 0$, because $\chi(A \times \{y\}) \leqslant \chi(\Gamma) < P_{\min} \leqslant 1/2$. In other words,

$$(8) \qquad \exists x_1 \in A^c, \quad \chi(x_1, y) \geqslant \varepsilon,$$

provided $\varepsilon > 0$ is chosen small enough. We now use the vertices x_0, y_0, x_1 found at (7)–(8) to construct a new coupling $\widetilde{\chi}$ which contradicts the optimality of χ. For all $(u, v) \in V^2$, we set

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

By construction, $\widetilde{\chi}$ is non-negative and has the same marginals as χ. Thus, it is a coupling of $P(x, \cdot)$ and $P(y, \cdot)$. This coupling is moreover optimal, since

$$\sum_{u,v \in V} \operatorname{dist}(u, v) \left(\widetilde{\chi}(u, v) - \chi(u, v) \right)$$
$$= \varepsilon \left(\operatorname{dist}(x_0, y) + \operatorname{dist}(x_1, y_0) - \operatorname{dist}(x_0, y_0) - \operatorname{dist}(x_1, y) \right)$$
$$\leqslant \varepsilon \left(\operatorname{dist}(x, y) - 1 + \operatorname{dist}(x_1, y_0) - \operatorname{dist}(x, y) - \operatorname{dist}(x_1, y) \right)$$
$$\leqslant 0,$$

where we have successively used $x_0 \in A$, $(x_0, y_0) \notin \Gamma$, and the triangle inequality $\operatorname{dist}(x_1, y_0) \leqslant \operatorname{dist}(x_1, y) + \operatorname{dist}(y, y_0)$. Finally, Γ contains (x_1, y) but not $(x_0, y_0), (x_1, y)$, so we have $\widetilde{\chi}(\Gamma) \geqslant \chi(\Gamma) + \varepsilon$, contradicting the double optimality of χ. $\qquad \square$

Our second ingredient is a diffusive hitting-time estimate for super-martingales. Results of this sort are standard, but usually require a uniform bound on the increments, a property which fails in our directed setting ($\operatorname{dist}(x, x') = \operatorname{dist}(y, y') = 1$ no longer implies $\operatorname{dist}(x', y') \leqslant \operatorname{dist}(x, y) + 2$).

Lemma 5 (Hitting-time estimate). — *Let $(Z_t)_{t \geqslant 0}$ be a discrete-time, $\mathbb{N}-$valued super-martingale with $Z_0 = z_0 \in \mathbb{N}$, and set $\tau := \min\{t \geqslant 0 : Z_t = 0\}$. Assume that almost-surely,*

$$\mathbb{P}\left(Z_{t+1} \neq Z_t | \mathcal{F}_t \right) \geqslant p\mathbf{1}_{(\tau > t)},$$

for all $t \geqslant 0$, where $(\mathcal{F}_t)_{t \geqslant 0}$ denotes the underlying filtration. Then, for all $t \geqslant 1$,

$$\mathbb{P}(\tau \geqslant t) \leqslant z_0 \sqrt{\frac{10}{pt}}.$$

Proof. — Our starting point is the following easily verified inequality: for all $x \in \mathbb{R}$,

$$e^{-x/2} \geqslant 1 - \frac{x}{2} + \frac{1 \wedge x^2}{10}.$$

In particular, for all $t \in \mathbb{N}$ and $\lambda \in [0, 1]$, we have on the event $\{\tau > t\}$,

$$\mathbb{E}\left[e^{-\frac{\lambda}{2}(Z_{t+1}-Z_t)}|\mathcal{F}_t\right] \geqslant \mathbb{E}\left[1 - \frac{\lambda}{2}(Z_{t+1}-Z_t) + \frac{1 \wedge (\lambda Z_{t+1} - \lambda Z_t)^2}{10}\Big|\mathcal{F}_t\right]$$
$$\geqslant 1 + \frac{p\lambda^2}{10},$$

where the second line uses our assumptions on Z. It inductively follows that for $t \in \mathbb{N}$,

$$\mathbb{E}\left[e^{-\frac{\lambda}{2}(Z_{t\wedge\tau}-Z_0)}\left(1+\frac{p\lambda^2}{10}\right)^{-t\wedge\tau}\right] \geqslant 1.$$

Since $Z_{t\wedge\tau} - Z_0 \geqslant -Z_0 = -z_0$, we deduce that

$$\mathbb{E}\left[\left(1+\frac{p\lambda^2}{10}\right)^{-t\wedge\tau}\right] \geqslant e^{-\lambda z_0/2} \geqslant 1 - \frac{\lambda z_0}{2}.$$

In particular,

$$\lambda z_0 \geqslant 2\mathbb{E}\left[1 - \left(1+\frac{p\lambda^2}{10}\right)^{-t\wedge\tau}\right]$$
$$\geqslant 2\left(1 - \left(1+\frac{p\lambda^2}{10}\right)^{-t}\right)\mathbb{P}(\tau \geqslant t)$$
$$\geqslant 2\left(1 - \left(1+\frac{pt\lambda^2}{10}\right)^{-1}\right)\mathbb{P}(\tau \geqslant t).$$

The result now readily follows by choosing $\lambda = \sqrt{10/pt}$. Note that this choice satisfies $\lambda \in [0, 1]$ only when $10/pt \leqslant 1$, but the conclusion trivially holds when $10/pt > 1$. $\quad\square$

Proof of Theorem 4. — For each $(x, y) \in V$, let $\chi_{x,y}$ denote an optimal coupling of $P(x, \cdot)$ and $P(y, \cdot)$ satisfying the condition in Lemma 4, and let us define a transition matrix K on V^2 by

$$K((x, y), (u, v)) := \chi_{x,y}(u, v).$$

Finally, consider a Markov chain $(X_t, Y_t)_{t\geqslant 0}$ with transition matrix K, and let us use the notation $\mathbb{E}_{(x,y)}[\cdot]$ to indicate that $(X_0, Y_0) = (x, y)$. By construction, we have for all $x, y \in V$,

$$\mathbb{E}_{(x,y)}\left[\text{dist}(X_1, Y_1)\right] = W\left(P(x, \cdot), P(y, \cdot)\right) \leqslant \text{dist}(x, y);$$
$$\mathbb{P}_{(x,y)}\left[\text{dist}(X_1, Y_1) < \text{dist}(x, y)\right] \geqslant P_{\min}\mathbf{1}_{(x\neq y)}.$$

By the Markov property, the first condition implies that the process $Z := (Z_t)_{t\geqslant 0}$ defined by $Z_t := \text{dist}(X_t, Y_t)$ is a super-martingale with respect to the natural filtration $(\mathcal{F}_t)_{t\geqslant 0}$ of $(X_t, Y_t)_{t\geqslant 0}$, and the second implies that $\mathbb{P}(Z_{t+1} \neq Z_t|\mathcal{F}_t) \geqslant P_{\min}\mathbf{1}_{Z_t\neq 0}$. Thus, Lemma 5 applies and yields

$$\mathbb{P}_{x,y}(X_t \neq Y_t) \leqslant \text{dist}(x, y)\sqrt{\frac{10}{(t+1)P_{\min}}}.$$

The result follows, since $\mathbb{P}_{(x,y)}\left((X_t, Y_t) = (\cdot, \cdot)\right)$ is a coupling of $P^t(x, \cdot)$ and $P^t(y, \cdot)$.

$\quad\square$

References

[1] N. Alon & Y. Roichman – "Random Cayley graphs and expanders", *Random Structures Algorithms* **5** (1994), no. 2, p. 271–284.

[2] F. Bauer, P. Horn, Y. Lin, G. Lippner, D. Mangoubi & S.-T. Yau – "Li-Yau inequality on graphs", *J. Differential Geom.* **99** (2015), no. 3, p. 359–405.

[3] N. Berestycki & B. Şengül – "Cutoff for conjugacy-invariant random walks on the permutation group", *Probab. Theory Relat. Fields* **173** (2019), no. 3-4, p. 1197–1241.

[4] M. Bordewich & M. Dyer – "Path coupling without contraction", *J. Discrete Algorithms* **5** (2007), no. 2, p. 280–292.

[5] E. Breuillard & M. C. H. Tointon – "Nilprogressions and groups with moderate growth", *Adv. Math.* **289** (2016), p. 1008–1055.

[6] P. Diaconis & L. Saloff-Coste – "Moderate growth and random walk on finite groups", *Geom. Funct. Anal.* **4** (1994), no. 1, p. 1–36.

[7] M. Eidi & J. Jost – "Ollivier ricci curvature of directed hypergraphs", *Sci. Rep.* **10** (2020), no. 1, article no. 12466.

[8] M. Erbar & J. Maas – "Ricci curvature of finite Markov chains via convexity of the entropy", *Arch. Rational Mech. Anal.* **206** (2012), no. 3, p. 997–1038.

[9] R. Forman – "Bochner's method for cell complexes and combinatorial ricci curvature", *Discrete Comput. Geom.* **29** (2003), no. 3, p. 323–374.

[10] J. Jost – *Riemannian geometry and geometric analysis*, seventh ed., Universitext, Springer, Cham, 2017.

[11] J. Jost & F. Münch – "Characterizations of Forman curvature", 2021, `arXiv:2110.04554`.

[12] J. R. Lee & Y. Peres – "Harmonic maps on amenable groups and a diffusive lower bound for random walks", *Ann. Probab.* **41** (2013), no. 5, p. 3392–3419.

[13] D. A. Levin & Y. Peres – *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2017.

[14] Y. Lin & S. Liu – "Equivalent properties of CD inequality on graph", 2015, `arXiv:1512.02677`.

[15] Y. Lin & S.-T. Yau – "Ricci curvature and eigenvalue estimate on locally finite graphs", *Math. Res. Lett.* **17** (2010), no. 2, p. 343–356.

[16] R. Montenegro & P. Tetali – "Mathematical aspects of mixing times in Markov chains", *Found. Trends Theor. Comput. Sci.* **1** (2006), no. 3, p. x+121.

[17] F. Münch – "Li-Yau inequality under $CD(0, n)$ on graphs", 2019, `arXiv:1909.10242`.

[18] _____, "Non-negative Ollivier curvature on graphs, reverse Poincaré inequality, Buser inequality, Liouville property, Harnack inequality and eigenvalue estimates", 2019, `arXiv:1907.13514`.

[19] F. Münch & R. K. Wojciechowski – "Ollivier Ricci curvature for general graph Laplacians: Heat equation, Laplacian comparison, non-explosion and diameter bounds", *Adv. Math.* **356** (2019), article no. 106759 (45 pages).

[20] L. Najman & P. Romon (eds.) – *Modern approaches to discrete curvature*, Lect. Notes in Math., vol. 2184, Springer, Cham, 2017.

[21] Y. Ollivier – "Ricci curvature of Markov chains on metric spaces", *J. Functional Analysis* **256** (2009), no. 3, p. 810–864.

[22] _____, "A survey of Ricci curvature for metric spaces and Markov chains", in *Probabilistic approach to geometry*, Adv. Stud. Pure Math., vol. 57, Math. Soc. Japan, Tokyo, 2010, p. 343–381.

[23] R. Ozawa, Y. Sakurai & T. Yamada – "Geometric and spectral properties of directed graphs under a lower Ricci curvature bound", *Calc. Var. Partial Differential Equations* **59** (2020), no. 4, article no. 142 (39 pages).

[24] _____, "Heat flow and concentration of measure on directed graphs with a lower Ricci curvature bound", *Potential Anal.* (2022), p. 1–15.

[25] J. Salez – "Cutoff for non-negatively curved Markov chains", 2021, `arXiv:2102.05597`.

[26] _____, "Sparse expanders have negative curvature", *Geom. Funct. Anal.* **32** (2022), no. 6, p. 1486–1513.

[27] M. Schmuckenschläger – "Curvature of nonlocal Markov generators", in *Convex geometric analysis (Berkeley, CA, 1996)*, Math. Sci. Res. Inst. Publ., vol. 34, Cambridge Univ. Press, Cambridge, 1999, p. 189–197.

[28] R. Tessera & M. C. H. Tointon – "A finitary structure theorem for vertex-transitive graphs of polynomial growth", *Combinatorica* **41** (2021), no. 2, p. 263–298.

[29] C. Villani – *Optimal transport: old and new*, Grundlehren Math. Wiss., vol. 338, Springer-Verlag, Berlin, 2009.

[30] T. Yamada – "The Ricci curvature on directed graphs", 2016, arXiv:1602.07779.

Florentin Münch, Max-Planck-Institut für Mathematik in den Naturwissenschaften
Inselstr. 22, 04103 Leipzig, Germany
E-mail : muench@mis.mpg.de
Url : https://sites.google.com/view/florentin-muench/home/

Justin Salez, Université Paris-Dauphine & PSL, CEREMADE
Place du Maréchal de Lattre de Tassigny, F-75775 Paris Cedex 16, France
E-mail : justin.salez@dauphine.psl.eu
Url : https://www.ceremade.dauphine.fr/~salez/