On the Expressive Power of Ollivier-Ricci Curvature on Graphs

Joshua Southern [*1] **Jeremy Wayland** [*2,3] **Michael Bronstein** [4] **Bastian Rieck** [2,3]

Abstract

Discrete curvature has recently been used in graph machine learning to improve performance, understand message-passing and assess structural differences between graphs. Despite these advancements, the theoretical properties of discrete curvature measures, such as their representational power and their relationship to graph features is yet to be fully explored. This paper studies Ollivier–Ricci curvature on graphs, providing both a discussion and empirical analysis of its *expressivity*, i.e. the ability to distinguish non-isomorphic graphs.

1. Introduction

Curvature is a fundamental concept in differential geometry and topology that allows one to distinguish between different types of manifolds. Among the assortment of curvature constructions and the variety of properties they exhibit, *Ricci curvature* has become one of the most prominent. Roughly speaking, Ricci curvature is based on measuring the differences in the growth of volumes in a space as compared to a 'model' Euclidean space. While originally requiring a smooth manifold, recent work has started to explore how to formulate a theory of Ricci curvature in discrete settings (Coupette et al., 2023; Devriendt & Lambiotte, 2022; Forman, 2003; Liu et al., 2018; Ollivier, 2007; Saucan et al., 2020). We focus on Ollivier-Ricci curvature (OR) (Ollivier, 2007) which, intuitively, quantifies a notion of similarity between node neighbourhoods, the discrete analogue for 'volume'. In the context of graphs, OR curvature provides sophisticated tools to characterise edges by analysing their neighbourhoods. Recent works have shown the benefits of using OR curvature to assess differences between real-world networks (Samal et al., 2018), as well as its ability to reduce over-squashing in GNNs by facilitating *graph rewiring* (Nguyen et al., 2023). However, the representational power of OR curvature on graphs remains largely unexplored.

Our *contributions* are as follows:

- We analyse the expressive power of OR curvature in terms of the Weisfeiler–Le(h)man (WL) heirarchy.
- We outline the use of higher-order neighbourhoods to calculate OR curvature and show that this provides additional expressive power.

2. Background

Ollivier–Ricci Curvature. Ollivier introduced a notion of curvature for metric spaces that measures the Wasserstein distance between Markov chains, i.e. random walks, defined on two nodes (Ollivier, 2007). To define this for graphs, let $G = (V, E)$ be a graph with some metric d_G and a probability measure μ_v at each node $v \in V$. The *Ollivier–Ricci curvature* of *any*[1] pair $i, j \in V \times V$ with $i \neq j$ is then defined as

$$\kappa_{\mathrm{OR}}(i,j) := 1 - \frac{1}{d_G(i,j)} W_1(\mu_i, \mu_j), \qquad (1)$$

where W_1 refers to the first *Wasserstein distance* between μ_i, μ_j. Eq. (1) defines the Ollivier–Ricci (OR) curvature in a general setting outlined by van der Hoorn et al. (2020); this is in contrast to the majority of previous works in the graph setting which specify d_G to be the shortest-path distance and μ_i, μ_j to be uniform probability measures in the 1-hop neighbourhood of the node. Extending the probability measures to act on larger locality scales is known to be beneficial for characterising graphs (Benjamin et al., 2021; Gosztolai & Arnaudon, 2021; Jiradilok & Kamtue, 2021).

Higher-order OR curvature. Defining κ_{OR} in a general framework allows us to alter the probability meausure and metric. In this work, we look to expand our understanding of OR curvature's utility by investigating measures based on *random walk probabilities*. Specifically, for a node x and

[*]Equal contribution [1]Imperial College London [2]Helmholtz Munich [3]Technical University of Munich [4]University of Oxford. Correspondence to: Bastian Rieck <bastian.rieck@helmholtz-munich.de>.

Presented at the 2^{nd} Annual Workshop on Topology, Algebra, and Geometry in Machine Learning (TAG-ML) at the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA.

[1]In contrast to other notions of curvature, Ollivier–Ricci curvature is defined for *both* edges and non-edges.

a positive integer m, we calculate μ_{RW} as

$$\mu_{\text{RW}}(y) := \sum_{k \leq m} \phi_k(x, y), \tag{2}$$

with $\phi_k(x, y)$ denoting the probability of reaching node y in a k-step random walk that starts from node x. Subsequently, we normalise Eq. (2) to ensure a valid probability distribution. By changing the value of m, we can consider up to m-step random walks.

The k-dimensional Weisfeiler-Le(h)man test. The WL hierarchy has become a standard way of measuring the expressivity of GNNs (Morris et al., 2021a;b). It is based on the idea of labelling all k-tuples of vertices in graphs. Initially, two k-tuples $\mathbf{v} = (v_1, \ldots, v_k)$ and $\mathbf{w} = (w_1, \ldots, w_k)$ are assigned the same label if the map $v_j \mapsto w_j$ induces a homomorphism between the subgraphs induced by $\mathbf{v}$ and $\mathbf{w}$, respectively. For subsequent iterations, tuples are relabelled according to a neighbourhood relation between them. This process is performed iteratively and results in a *graph colouring* or *vertex partition*. If two graphs give rise to different colour sequences, the graphs are guaranteed to be non-isomorphic (the other direction does not hold, as seen in Fig. 1). There are classes of graphs that cannot be distinguished by k-WL but that can be distinguished by $(k + 1)$-WL (Morris et al., 2021a). We thus obtain a *hierarchy* of tests for graph isomorphism, with higher orders substantially increasing in computational complexity.

3. Expressivity of Ollivier-Ricci Curvature

To understand the expressive power of OR curvature, we explore its ability to distinguish non-isomorphic graphs in comparison to k-WL. Specifically, we study regular graphs that 1-WL fails to distinguish, while also analysing strongly-regular graphs that 3-WL fails to distinguish.

3.1. Distinguishing Beyond 1-WL

1-WL cannot distinguish any n-sized r-regular graph as the degree is constant across nodes. Figure 1 shows three examples of regular graphs that cannot be distinguished by 1-WL. Feng et al. (2022) show that using 2-hop message-passing, Example 1 *can* be distinguished with a graph diffusion kernel but not with the shortest path kernel, whereas Example 2 can be distinguished with the shortest path kernel and not with the graph diffusion kernel. In comparison to 2-hop message passing, standard κ_{OR} (uniform measure on 1-hop neighbourhoods) has a smaller field of view. κ_{OR} does, however, consider pairs of neighbourhoods which can improve representational power. For example, κ_{OR} can bound the number of triangles within a locally finite graph (Jost & Liu, 2011), allowing it to differentiate the pairs of graphs which have different numbers of triangles and thus improve upon

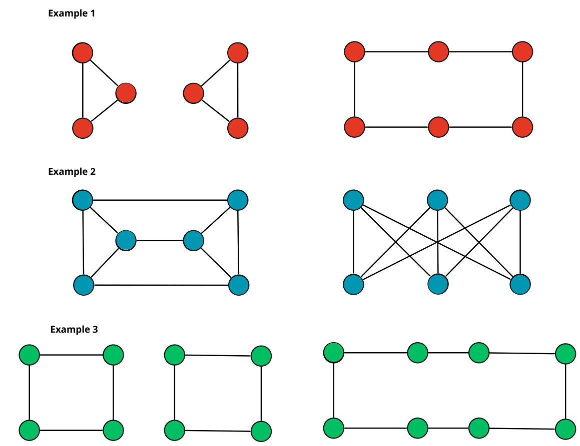

Figure 1: Examples of pairs of regular graphs that cannot be distinguished by 1-WL, the first-order Weisfeiler–Le(h)man test for graph isomorphism.

2-hop message-passing with a specific kernel choice. Example 3 cannot be distinguished by 2-hop message-passing (regardless of kernel choice) nor standard κ_{OR}. However, by changing the probability measure used by κ_{OR}, we can shift the focus towards even larger substructures. This allows a random walk measure with $m = 2$ to distinguish the graphs in Example 3.

3.2. Distinguishing Beyond 3-WL

Strongly-regular graphs are often used to assess the expressive power of graph learning algorithms, constituting a class of graphs that are particularly hard to distinguish. A strongly-regular graph is defined using four parameters $\{n, d, \alpha, \beta\}$ where n is the number of nodes in the graph, d is the degree of each node, α is the number of common neighbours between adjacent vertices and β is the number of common neighbours between non-adjacent vertices. Let $i, j \in V \times V$ and define N_i to be the set of vertices which are adjacent to node i, not including j or any nodes adjacent to j. Suppose there is a maximal matching M of size m between N_i and N_j. Following Bonini et al. (2020), the standard OR curvature of a strongly-regular graph is then given by

$$\kappa_{\text{OR}}(i, j) = \frac{\alpha + 2}{d} - \frac{\mid N_i \mid - m}{d}. \tag{3}$$

Theorem 1 (Expressivity of OR curvature)**.** *Ollivier–Ricci curvature* can *distinguish the Rook and Shrikhande graphs, which are strongly-regular graphs with the same intersection array (Southern et al., 2023).*

The Rook and Shrikhande graphs *cannot* be distinguished by 3-WL (Bouritsas et al., 2022; Morris et al., 2021a). For strongly regular graphs with a girth of 3, the OR curvature does not depend only on the graph parameters (Bonini et al., 2020). It is therefore sensitive to differences in the first-hop peripheral subgraphs (Feng et al., 2022) of the Rook and Shrikhande graphs, thus distinguishing them.

The implications of this section are that (i) OR curvature can distinguish graph pairs that are 3-WL indistinguishable without needing more than 1-hop of neighborhood information. (ii) Using OR curvature with a higher-order measure allows us to distinguish some graphs that cannot be distinguished using standard OR curvature.

3.3. Other Metrics and Future Work

Recent work has explored expressivity beyond WL for GNN expressiveness evaluation, including diameter counting, counting substructures (Chen et al., 2020) and biconnectivity (Zhang et al., 2023). Interestingly, OR curvature has been shown to provide an upper bound for the graph diameter (Paeng, 2012) as well as a lower bound for the number of triangles (Jost & Liu, 2011). Additionally, the resistance distance, which is another form of discrete curvature (Devriendt & Lambiotte, 2022), was shown to solve biconnectivity (Zhang et al., 2023). We leave a detailed exploration of such relations to future work.

4. Experiments

In the previous section, we discussed the theoretical expressivity of OR curvature. Subsequently, we will show empirical experiments to evaluate these claims.

4.1. Distinguishing Graphs

We explore the ability of our method to distinguish strongly-regular graphs in a subset of data sets, i.e. `sr16622`, `sr251256`, `sr261034`, `sr281264`, and `sr291467`. These data sets are challenging to classify since they cannot be described in terms of the 1-WL test (Bodnar et al., 2021). Additionally, we evaluate discrete curvature on the BREC dataset, which was recently introduced to evaluate GNN expressiveness (Wang & Zhang, 2023). The dataset consists of different categories of graph pairs (Basic, Regular, and Extension), which are distinguishable by 3-WL but not by 1-WL, as well as Strongly-Regular (STR) and CFI graph pairs which are indistinguishable using 3-WL. We explore the ability of OR curvature to distinguish these graph pairs and compare them to substructure counting, S_3 and S_4, which involves enumerating all 3-node/4-node substructures around nodes in combination with the WL algorithm. These approaches, unlike OR curvature, have limited practical applications due to their high computational complexity.

We calculate Wasserstein distances based on histograms of OR curvature measurements between the pairs of graphs. Subsequently, we count all non-zero distances ($> 1 \times 10^{-8}$ to correct for precision errors). Our main observations from Table 1 and Table 2 are that OR curvature *can* distinguish graphs which are 3-WL indistinguishable and using a higher-order random walk measure *can* improve discriminative power: we observe improvements in success rate from $m = 1$ to $m = 2$ on the Regular, STR, Extension, and CFI graph pairs, respectively. Additionally, the approach performs competitively and sometimes better than S_4, which has been shown to be extremely effective in graph learning tasks (Bouritsas et al., 2022). Despite its prowess, S_4 is computationally expensive, making it an *infeasible* measure in many applications. OR curvature, by contrast, even with higher-order neighbourhoods, scales significantly better.

Table 1: Success rate (↑) of distinguishing pairs of graphs in the BREC dataset when using different probability measures in the OR curvature calculation.

Method	Basic (56)	Regular (50)	STR (50)	Extension (97)	CFI (97)
1-WL	0.00	0.00	0.00	0.00	0.00
3-WL	1.00	1.00	0.00	1.00	0.59
S_3	0.86	0.96	0.00	0.05	0.00
S_4	1.00	0.98	1.00	0.84	0.00
RW, $m = 1$	1.00	0.96	0.06	0.87	0.00
RW, $m = 2$	1.00	1.00	0.14	0.97	0.01
RW, $m = 3$	1.00	1.00	0.14	0.99	0.04
RW, $m = 4$	1.00	1.00	0.14	1.00	0.09
RW, $m = 5$	1.00	1.00	0.14	1.00	0.19
RW, $m = 6$	1.00	1.00	0.14	1.00	0.19

Table 2: Success rate (↑) of distinguishing pairs of strongly-regular graphs when using Ollivier-Ricci curvature.

Method	sr16622	sr251256	sr261034	sr281264	sr291467
1-WL	0.00	0.00	0.00	0.00	0.00
3-WL	0.00	0.00	0.00	0.00	0.00
RW, $m = 1$	1.00	0.00	0.00	1.00	0.00
RW, $m = 2$	1.00	0.00	0.78	1.00	0.00

5. Conclusion

Our work provides the first thorough analysis of the expressivity of Ollivier-Ricci curvature on graphs. We have shown not only that OR curvature can distinguish graphs outside the capabilities of 3-WL, but also that its expressivity can be improved by varying its probability measure, thus motivating the use of higher-order OR curvature. Finally, we demonstrate OR curvature's practical utility by performing experiments that distinguish pairs of graphs in various strongly regular graph datasets and the BREC dataset. Future work could explore expressivity beyond WL, showing how OR curvature can improve diameter approximation, counting substructures, and biconnectivity solutions as well as exploring the utility of other random walk measures.

References

Benjamin, S., Mantri, A., and Perian, Q. On the Wasserstein distance between k-step probability measures on finite graphs, 2021.

Bodnar, C., Frasca, F., Wang, Y., Otter, N., Montufar, G. F., Lió, P., and Bronstein, M. Weisfeiler and Lehman go topological: Message passing simplicial networks. In Meila, M. and Zhang, T. (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 1026–1037. PMLR, 2021.

Bonini, V., Carroll, C., Dinh, U., Dye, S., Frederick, J., and Pearse, E. Condensed ricci curvature of complete and strongly regular graphs. *Involve, a Journal of Mathematics*, 13(4):559–576, nov 2020. doi: 10.2140/involve.2020.13.559.

Bouritsas, G., Frasca, F., Zafeiriou, S. P., and Bronstein, M. Improving graph neural network expressivity via subgraph isomorphism counting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2022. doi: 10.1109/TPAMI.2022.3154319.

Chen, Z., Chen, L., Villar, S., and Bruna, J. Can graph neural networks count substructures?, 2020.

Coupette, C., Dalleiger, S., and Rieck, B. Ollivier–Ricci curvature for hypergraphs: A unified framework. In *International Conference on Learning Representations (ICLR)*, 2023.

Devriendt, K. and Lambiotte, R. Discrete curvature on graphs from the effective resistance, March 2022.

Feng, J., Chen, Y., Li, F., Sarkar, A., and Zhang, M. How powerful are k-hop message passing graph neural networks, 2022.

Forman. Bochner's method for cell complexes and combinatorial ricci curvature. *Discrete & Computational Geometry*, 29(3):323–374, Feb 2003. ISSN 1432-0444. doi: 10.1007/s00454-002-0743-x.

Gosztolai, A. and Arnaudon, A. Unfolding the multiscale structure of networks with dynamical ollivier-ricci curvature. *Nature Communications*, 12(1):4561, Jul 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-24884-1.

Jiradilok, P. and Kamtue, S. Transportation distance between probability measures on the infinite regular tree, 2021.

Jost, J. and Liu, S. Ollivier's ricci curvature, local clustering and curvature dimension inequalities on graphs, 2011.

Liu, S., Münch, F., and Peyerimhoff, N. Bakry–émery curvature and diameter bounds on graphs. *Calculus of Variations and Partial Differential Equations*, 57(2), mar 2018. doi: 10.1007/s00526-018-1334-x.

Morris, C., Lipman, Y., Maron, H., Rieck, B., Kriege, N. M., Grohe, M., Fey, M., and Borgwardt, K. Weisfeiler and Leman go machine learning: The story so far, 2021a.

Morris, C., Ritzert, M., Fey, M., Hamilton, W. L., Lenssen, J. E., Rattan, G., and Grohe, M. Weisfeiler and leman go neural: Higher-order graph neural networks, 2021b.

Nguyen, K., Nguyen, T., Nong, H., Nguyen, V., Ho, N., and Osher, S. Revisiting over-smoothing and over-squashing using ollivier-ricci curvature, 2023.

Ollivier, Y. Ricci curvature of metric spaces. *Comptes Rendus Mathematique*, 345(11):643–646, 2007. ISSN 1631-073X. doi: 10.1016/j.crma.2007.10.041.

Paeng, S.-H. Volume and diameter of a graph and ollivier's ricci curvature. *European Journal of Combinatorics*, 33 (8):1808–1819, 2012. ISSN 0195-6698. doi: https://doi.org/10.1016/j.ejc.2012.03.029.

Samal, A., Sreejith, R. P., Gu, J., Liu, S., Saucan, E., and Jost, J. Comparative analysis of two discretizations of ricci curvature for complex networks. *Scientific Reports*, 8(1), 2018. doi: 10.1038/s41598-018-27001-3.

Saucan, E., Samal, A., and Jost, J. A simple differential geometry for complex networks, 2020.

Southern, J., Wayland, J., Bronstein, M., and Rieck, B. Curvature filtrations for graph generative model evaluation, 2023.

van der Hoorn, P., Lippner, G., Trugenberger, C., and Krioukov, D. Ollivier curvature of random geometric graphs converges to ricci curvature of their riemannian manifolds, 2020.

Wang, Y. and Zhang, M. Towards better evaluation of gnn expressiveness with brec dataset, 2023.

Zhang, B., Luo, S., Wang, L., and He, D. Rethinking the expressive power of gnns via graph biconnectivity, 2023.