A Regularization

We perform a regularization of our graphs (see section 3) which allows for an efficient characterization of local neighborhood growth rates and in turn, the local geometry. In this appendix, we provide more details on the implementation and the theoretical guarantees of the regularization.

Fig. 1 shows the regularization schematically. For each vertex v in the graph, we enforce a uniform node degree of 3 within its 1-hop neighborhood $\mathcal{N}_1(v)$. Hereby auxiliary vertices a are inserted and modified edges reweighed. The regularization algorithm is given in Alg. 1. The regularization allows for a quasi-isometric embedding of any

Algorithm 1 Regularization

Input: $G = \{V(G), E(G)\}$
$\epsilon \leftarrow \max_{e \in E} \omega(e)$
for $v \in V$ **do**
    $\mathcal{N}(v) \leftarrow \{u \in V : v \sim u\}$ ▷ Neighborhood of v.
    if $\deg(v) == 1$ **then** ▷ Leaf: $\mathcal{N}(v) = \{u\}$
        CREATE(a^0, a^1, a^2, a^3) ▷ auxiliary nodes
        $\omega(u,v) \leftarrow \omega(u,v)/2$
        $\omega(v,a^0),\ \omega(v,a^1) \leftarrow \epsilon/4$
        $\omega(a^i, a^{i+k}) \leftarrow \epsilon/4$ for $i = 0,1;\ k = 1,2$
    else if $\deg(v) == 2$ **then** ▷ Chain: $\mathcal{N}(v) = \{u_1, u_2\}$
        CREATE(a^0, a^1, a^2) ▷ auxiliary nodes
        $\omega(u_1, v) \leftarrow \omega(u_1, v)/2$
        $\omega(v, a^i) \leftarrow \omega(u_2, v)/2$ for $i = 0,1$
        $\omega(u_2, a^2) \leftarrow \omega(u_2, v)/2$
        $\omega(a^0, a^1),\ \omega(a^0, a^2),\ \omega(a^1, a^2) \leftarrow \epsilon/4$
    else if $\deg(v) == 3$ **then** ▷ 3-regular: $\mathcal{N}(v) = \{u_1, u_2, u_3\}$
        CONTINUE
    else ▷ Star: $\mathcal{N}(v) = \{u_i\}_{i=1}^{\deg(v)}$
        CREATE($a^1, \dots, a^{\deg(v)}$) ▷ auxiliary nodes
        for $i = 1, \dots, \deg(v)$ **do**
            $\omega(a^i, u^i) \leftarrow \omega(u_i, v)/2$
            $\omega(a^i, a^{i+1}) \leftarrow \epsilon/4$
        end for
    end if
end for

graph into a 3-regular graph. We provide the theoretical reasoning below:

Theorem A.1 (Bermudo et al. (2013)). *$G \hookrightarrow^{\phi} G_3$ is a $(\epsilon+1, \epsilon)$-quasi-isometric embedding, i.e.*

$$d_{G_3}(\phi(x), \phi(y)) \leq (\epsilon + 1) d_G(x,y) + \epsilon \ . \tag{3}$$

From this we can derive the following additive distortion:

$$|d_{G_3} - d_G| \leq |(\epsilon+1)d_G + \epsilon - d_G| = |\epsilon(\underbrace{d_G}_{\leq \operatorname{diam}(G)} + 1)| \leq O(\epsilon) \ ,$$

For the multiplicative distortion we have

$$d_G \leq d_{G_3} \leq (1+\epsilon)d_G + \epsilon \Rightarrow \frac{1}{1+\epsilon} d_G \leq d_{G_3} \leq (\underbrace{d_G}_{\leq \operatorname{diam}(G)} + 1)(1+\epsilon) \leq O(1+\epsilon) \ .$$

This gives $c_A = O(\epsilon)$ and $c_M = O(1+\epsilon)$ as given in the main text.

B Neighborhood growth rates in canonical graphs

We restate the result from the main text:

Theorem B.1 (Neighborhood growth of canonical graphs)**.** *For $3 \leq R \ll \mathrm{diam}(G)$, every R-neighborhood in*

1. *a b-regular tree is exponentially expanding;*
2. *an $(\sqrt{N} \times \sqrt{N})$-lattice is linearly expanding;*
3. *an N-cycle is sublinearly expanding.*

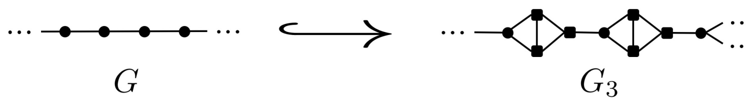

Figure 3: Regularization of N-cycle.

Proof. The regularization introduces edge weights (see Alg. 1), however, due to the periodic structure of the canonical graphs, those weights are uniform (for lattices and trees) or up to an additive error of $\frac{\epsilon}{4}$ uniform (for cycles). Therefore, we can renormalize the edge weights and analyze the regularized graphs as unweighted graphs. For cycles, the residual additive error does not affect the neighbor count and can therefore be neglected.

Consider first (c) an N-cycle. Due to the periodic structure of the *chains* in the regularized graph (see Fig. 3), there are always either two or three vertices at a distance r from the root, in particular, we have

$$|\mathcal{N}_R(v)| = 1 + \sum_{r=1}^{R} \alpha(r) \; ,$$
$$\alpha(r) = \begin{cases} 2, & \mathrm{mod(r,3)} = 0 \\ 3, & \text{else} \end{cases} \; .$$

It is clear that $\sum_{r=1}^{R} \alpha(r) < 3R$, i.e. the growth is sublinear.

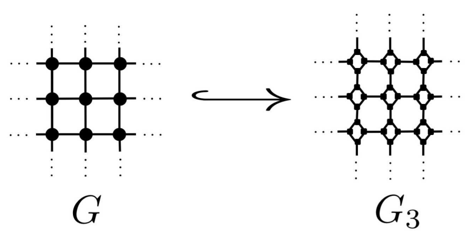

Figure 4: Regularization of $(\sqrt{N} \times \sqrt{N})$-cycle.

Next, consider (b) a $(\sqrt{N} \times \sqrt{N})$-lattice. From the periodicity of the regularized graph (see Fig. 4), we see that the number of nodes at distance r from the root grows linearly in r. In particular, we have

$$|\mathcal{N}_R(v)| = |v| + \sum_{r=1}^{R} \deg(v)\alpha(r) \; ,$$
$$\alpha(r) = \begin{cases} r-1, & \mathrm{mod(r,3)} = 1 \\ r, & \text{else} \end{cases} \; .$$

Since $\deg(v) = 4$, we have $\deg(v)\alpha(r) \geq 3r$, i.e. the lattice expands linearly. Finally, consider (a) a b-ary tree. We first consider the case $b = 3$, i.e., a ternary tree. Note that this structure is invariant under our regularization. Since every node has exactly two children, we get the following growth rate:

$$|\mathcal{N}_R(v)| = 1 + \sum_{r=1}^{R} 3 \cdot 2^{r-1} = \gamma_{EE} \; ,$$

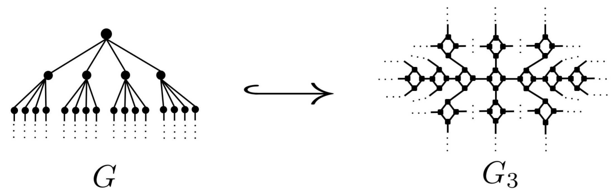

Figure 5: Regularization of b-ary tree (here $b = 4$).

i.e., the ternary tree expands exponentially. Now consider general b-ary trees with $b > 3$. The building blocks of the periodic structure unfolding are b-rings (see Fig. 5). On each level, the nodes have either two children or two nodes have a combined three children, if the b-rings close. This results in the following growth rate:

$$|\mathcal{N}_R(v)| = |v| + \sum_{r=1}^{R} b \cdot \alpha(r) \;,$$
$$\alpha(r) = \begin{cases} \frac{3}{2} \cdot 2^{r-1}, & \mathrm{mod}(\lceil \frac{\mathrm{r}}{2} \rceil, \mathrm{b}) = \mathrm{b} - 1 \\ 2^r, & \text{else} \end{cases} \;.$$

Since $3 < b$ and $2^{r-1} < \alpha(r)$, all b-regular trees expand exponentially. □

C Comparison with other discrete curvatures

We restate the result from the main text (Thm. 3.3):

Theorem C.1. *At any node v, we have*

1. $\mathrm{Ric_O}(v) < 0$ *and* $\mathrm{Ric_F}(v) < 0$ *in a b-regular tree,*
2. $\mathrm{Ric_O} \leq 0$ *and* $\mathrm{Ric_F}(v) < 0$ *in an* $(\sqrt{N} \times \sqrt{N})$*-lattice, and*
3. $\mathrm{Ric_O} = 0$ *and* $\mathrm{Ric_F} = 0$ *in an N-cycle.*

Before proving the theorem, recall the node-based curvature notions for $v \in V(G)$:

$$\mathrm{Ric_O}(v) = \frac{1}{\deg(v)} \sum_{(u,v)} \mathrm{Ric_O}(u,v) = \frac{1}{\deg(v)} \sum_{(u,v)} 1 - W_1(m_u, m_v)$$
$$\mathrm{Ric_F}(v) = \frac{1}{\deg(v)} \sum_{(u,v)} \mathrm{Ric_F}(u,v) = 4 - \deg(v) - \sum_{(u,v)} \frac{\deg(\mathrm{u})}{\deg(v)} \;.$$

Furthermore recall the following curvature inequalities for $\mathrm{Ric_O}$:

Lemma C.2. *(Jost and Liu, 2014)* $\mathrm{Ric_O}$ *fulfills the following inequalities:*

1. *If (u, v) is an edge in a tree, then* $\mathrm{Ric_O}(u, v) \leq 0$.
2. *For any edge u, v in a graph, we have*
$$-2\left(1 - \frac{1}{\deg(u)} - \frac{1}{\deg(v)}\right)_+ \leq \mathrm{Ric_O}(u,v) \leq \frac{\#(u,v)}{\max\{\deg(u), \deg(v)\}} \;,$$
where $\#(u, v)$ denotes the number of common neighbors (or joint triangles) of u and v.

Proof. (Thm. 3.3) Consider first (3) an N-cycle. By Lem. C.2(2) we have for any (u, v) on the right hand side $\mathrm{Ric_O}(u, v) \leq 0$, since a cycle has no triangles. Furthermore, the left hand side gives $\mathrm{Ric_O}(u, v) \geq 0$, since $\deg(v) = \deg(u) = 2$. This implies $\mathrm{Ric_O}(v) = 0$. We also have

$$\mathrm{Ric_F}(v) = 4 - \deg(v) - \left(\frac{\deg(u_1)}{2} + \frac{\deg(u_2)}{2}\right) = 0 \;,$$

since $\deg(v) = \deg(u_i) = 2$.

Next, consider (1) a b-ary tree. By Lem. C.2(1), we have $Ric_O(u, v) \leq 0$ and therefore $Ric_O(v) \leq 0$. Moreover, we have

$$\mathrm{Ric}_F(v) = 4 - (b+1) - \frac{(b+1)(b+1)}{b+1} \leq 0 \ ,$$

since by construction $b \geq 2$.

Finally, consider (2) an $(\sqrt{N} \times \sqrt{N})$-lattice. Since the lattice has no triangles, Lem. C.2(2) gives $\mathrm{Ric}_{\mathrm{O}}(u, v) \leq 0$ for any edge (u, v) and therefore $\mathrm{Ric}_{\mathrm{O}}(v) \leq 0$. In addition,

$$\mathrm{Ric}_F(v) = 4 - \deg(v) - \left(\frac{\deg(u_1)}{4} + \frac{\deg(u_2)}{4} + \frac{\deg(u_3)}{4} + \frac{\deg(u_4)}{4} \right) \leq 0 \ ,$$

since $\deg(v) = \deg(u_i) = 4$ for all i. □

D Embeddability measures

We evaluate the quality of embeddings using two computational distortion measures, following the workflow in (Gu et al., 2019; Nickel and Kiela, 2017). First, we report the average distortion

$$D_{avg} = \sum_{1 \leq i \leq j \leq n} \left| \left(\frac{d_{\mathcal{M}}(x_i, x_j)}{d_G(x_i, x_j)} \right)^2 - 1 \right| \ . \tag{4}$$

Secondly, we report MAP scores that measure the preservation of nearest-neighbor structures:

$$\mathrm{MAP} = \frac{1}{|V|} \sum_{u \in V} \frac{1}{\deg(u)} \sum_{i=1}^{|\mathcal{N}_1(u)|} \frac{|\mathcal{N}_1(u) \cap R_{u,i}|}{|R_{u,i}|} \ . \tag{5}$$

Here, $R_{u,i}$ denotes the smallest set of nearest neighbors required to retrieve the i^{th} neighbor of u in the embedding space M. One can show that for isometric embeddings, $\mathrm{D_{avg}} = 0$ and $\mathrm{MAP} = 1$.