



OPEN

Generalised popularity-similarity optimisation model for growing hyperbolic networks beyond two dimensions

Bianka Kovács¹, Sámuel G. Balogh^{1✉} & Gergely Palla^{1,2,3}

Hyperbolic network models have gained considerable attention in recent years, mainly due to their capability of explaining many peculiar features of real-world networks. One of the most widely known models of this type is the popularity-similarity optimisation (PSO) model, working in the native disk representation of the two-dimensional hyperbolic space and generating networks with small-world property, scale-free degree distribution, high clustering and strong community structure at the same time. With the motivation of better understanding hyperbolic random graphs, we hereby introduce the d PSO model, a generalisation of the PSO model to any arbitrary integer dimension $d > 2$. The analysis of the obtained networks shows that their major structural properties can be affected by the dimension of the underlying hyperbolic space in a non-trivial way. Our extended framework is not only interesting from a theoretical point of view but can also serve as a starting point for the generalisation of already existing two-dimensional hyperbolic embedding techniques.

Network theory has become an essential and ubiquitous tool for modelling various types of complex systems ranging from the level of interactions within cells to the level of the Internet, economic networks, and the society^{1,2}. In the past decades, a vast number of related studies reported a few universal features that most of the real networks seem to have in common, such as sparsity³, small-world property^{4,5}, inhomogeneous degree distribution^{6,7}, high clustering coefficient⁸ or community structure^{9–11}. Incorporating all, or at least some of these universal properties into a unified modelling framework is, however, a non-trivial issue and still presents a theoretical challenge of high relevance. Along this line, a variety of different network models have been proposed so far, including the celebrated Barabási–Albert (BA) model with preferential attachment¹², the hidden variables formalism^{13–17} or models based on the mechanism of triadic closure, which has been specifically designed for explaining the high clustering of social networks^{18,19}. Besides these examples, a further notable approach is given by hyperbolic network models that are capable of simultaneously explaining many observed network characteristics in a natural manner by assuming that nodes are embedded into a negatively curved hidden metric space^{20–25}.

The random hyperbolic graph (RHG)²⁰, for instance, is a static network model where nodes are placed at random on the hyperbolic disk of constant curvature $K = -\zeta^2$, and the connection probability between any pair of nodes is a decreasing function of their hyperbolic distance. A mathematically equivalent model is given by the \mathbb{S}^1 model²⁶, where the nodes are positioned on a circle and become connected according to a probability depending on the angular distance and a hidden variable drawn from a power-law distribution. By converting the hidden variables into radial coordinates we arrive to the hyperbolic \mathbb{H}^2 model²⁷ that is equivalent to the RHG model; hence, the RHG is often also referred to as the $\mathbb{S}^1/\mathbb{H}^2$ model. (Actually, the most general form of the $\mathbb{S}^1/\mathbb{H}^2$ model described in Ref.²⁷ is somewhat broader than the RHG model as it can accommodate any type of degree distributions.)

In contrast to the RHG, in the popularity-similarity optimisation (PSO) model²¹ the networks are not static but evolve over time via the continual appearance of new nodes on the hyperbolic plane. More precisely, new nodes are placed one by one in the native disk representation of the two-dimensional hyperbolic plane²⁰ with logarithmically increasing radial coordinates and uniformly random angular coordinates. Once a new node appears, it establishes connections to the previous ones with a probability depending on the hyperbolic distance

¹Department of Biological Physics, Eötvös Loránd University, Pázmány P. stny. 1/A, 1117 Budapest, Hungary. ²MTA-ELTE Statistical and Biological Physics Research Group, Pázmány P. stny. 1/A, 1117 Budapest, Hungary. ³Health Services Management Training Centre, Semmelweis University, 1125, Kútvolgyi út 2, Budapest, Hungary. ✉email: balogh@hal.elte.hu

in a similar way as in the RHG model. The tendency to connect to hyperbolically close nodes can be interpreted as an optimisation of a trade-off between the popularity (arising from the node birth time and reflected by the radial coordinate) of a possible candidate and its similarity (the angular distance abstracting the distance in an attribute space) compared to the newly arriving node. In vague terms, the degree of the nodes is determined by the radial coordinate, and owing to an outward shift of the nodes (referred to as the popularity fading, controlled by a parameter β), the degree distribution takes the scaling form of $\mathcal{P}(k) \sim k^{-\gamma}$ with a tuneable decay exponent $\gamma = 1 + \frac{1}{\beta}$. By changing the sharpness of the cutoff in the connection probability as a function of the hyperbolic distance with another parameter T called temperature, the average clustering coefficient \bar{c} of the resulting graphs can be adjusted as well. Although this model has been shown to be capable of generating networks that are small-world, highly clustered and scale-free at the same time, several other variants of the original PSO model have been suggested in order to explain further features of real-world graphs. Examples include the E-PSO model that inherently accounts for the creation of internal links, i.e. connections emerging between old nodes in the network²², or alternatively, the deletion of already existing links²⁸. Or, the nonuniform popularity-similarity optimisation (nPSO) model^{24,29} that allows the generation of networks with an adjustable community structure by assuming a heterogeneous angular node distribution with multiple peaks is also worth mentioning. Nevertheless, it has also been revealed quite recently that both the RHG and the PSO models can generate networks that possess strong community structure despite lacking any explicitly built-in community generating mechanisms^{25,30–34}.

In parallel with the developments of hyperbolic network models, another closely related field given by hyperbolic embedding techniques has also received great attention^{22,27,28,35–38}. Briefly, this tackles the problem of inferring the most plausible coordinates for the network nodes based on the topology of a given network. One of the first methods pointing in this direction was HyperMap²², relying on a maximum likelihood estimation, where we assign hyperbolic coordinates to the nodes of the network by maximising the probability that the network was generated by the E-PSO model. Contrarily, in Refs.^{36,37} an embedding technique based on a nonlinear dimension reduction of the Laplacian matrix was introduced. Along similar lines, a whole set of embedding algorithms were studied in Ref.³⁸, using different pre-weighted matrices encapsulating the network structure and multiple unsupervised dimension reduction techniques borrowed from machine learning. The rationale behind these approaches (usually coined as coalescent embeddings) is that when they are applied to hyperbolic networks, a common node aggregation pattern can be observed that is circularly or linearly ordered (angular coalescence) according to the original angular coordinates on the hyperbolic plane. An embedding algorithm that mixes the coalescent embedding with local angular optimisation based on likelihood maximisation was proposed in Ref.²⁸. A further, very efficient embedding method is given by Mercator²⁷, adopting the Laplacian eigenmaps approach with the coordinates optimised according to the RHG model.

Despite the excellent performance of the above embedding techniques, there is clearly a theoretical limitation behind most of them: they are defined on the hyperbolic disk, that is, in $d = 2$ dimensions. Nevertheless, it has recently been revealed that higher-dimensional hyperbolic embeddings can outperform lower-dimensional ones in link prediction and in graph reconstruction with respect to mean average precision (MAP), for instance, in author collaboration networks³⁹. Moreover, in Ref.⁴⁰ it has also been shown that the presence of additional dimensions can lead to a clearer separation between the communities of a network, quantitatively confirmed by a higher value of the angular separation index (ASI). Furthermore, using higher number of dimensions can be advantageous also from the point of the behaviour of the optimisation procedures behind the embedding algorithms. I.e., in many cases the embedding is based on the maximisation of the likelihood that the network was generated according to a hyperbolic random graph model, and in 2 dimensions, this function can be extremely non-convex with respect to the node coordinates, making e.g. stochastic gradient descent methods inefficient²². However, raising the number of dimensions of the hyperbolic space can lift some of the local maxima, helping the optimisation methods in finding better solutions of the embedding problem. Uncovering the role of the number of dimensions in hyperbolic embeddings is, therefore, of great interest that simultaneously provides a strong motivation for investigating appropriate higher-dimensional hyperbolic network models as well.

The first hyperbolic network model that treats the number of dimensions as a model parameter by placing the nodes in a d -dimensional hyperbolic ball was considered in the Supplementary Information of Ref.⁴¹ in the context of its asymptotic equivalence to the causal sets in the d -dimensional de Sitter spacetime. Then, the RHG model has been extended to $d > 2$ dimensions in Refs.^{42,43}. Here, we introduce the d PSO model, a generalisation of the original two-dimensional popularity-similarity optimisation model to any arbitrary integer dimension of $d \geq 2$, which, we believe, provides a further substantial step towards a comprehensive theoretical characterisation of hyperbolic graphs.

Besides its theoretical relevance, our d PSO model opens up the possibility of systematically generalising already existing embedding techniques to higher dimensions, the necessity of which has explicitly been outlined e.g. for coalescent embeddings in Ref.³⁸. Therein the authors claim that as a supplement to their findings, an additional interesting analysis could be to generate synthetic networks using for instance a three-dimensional PSO model and examine the accuracy of their methods by comparing the obtained embeddings to the original node arrangement. The present work contributes to this issue by thoroughly elaborating the PSO model for $d = 3$ and higher dimensions. Since the higher number of dimensions of the underlying hyperbolic space allows a much richer characterisation of the nodes in general, here we conjecture that the suggested d -dimensional embeddings could provide further and deeper insights into the architecture of the hidden geometry behind the structure of complex networks.

In the present paper, we introduce the d PSO model as a natural generalisation of the well-known two-dimensional PSO model²¹ to hyperbolic spaces of dimension $d > 2$. We show analytically that the degree distribution of d PSO networks can be written as $\mathcal{P}(k) \sim k^{-\gamma}$ in the large k regime, where the degree decay exponent γ is directly related to the dimension d and the popularity fading parameter β as $\gamma = 1 + \frac{1}{(d-1)\beta}$. Besides the scale-free behaviour, the networks generated by the d PSO model can exhibit a large average clustering coefficient \bar{c} and

a strong community structure for a relatively wide range of the parameter settings if the number of dimensions of the underlying hyperbolic space is not extremely high. According to our results, \bar{c} is controlled by an interesting interplay between the dimension d and the temperature T , where for $T < \frac{1}{d-1}$ the average clustering coefficient is a decreasing function of T , whereas at temperatures near and above $T = \frac{1}{d-1}$, \bar{c} becomes independent of T . A further noteworthy feature of the d PSO model is that in dimensions $d > 2$ extremely skewed degree distributions with $\gamma < 2$ become accessible, which can lead to networks displaying a number of exotic properties that are absent in scale-free networks with $2 \leq \gamma$. The rich variety of networks that can be obtained in our proposed framework together with the capability of reproducing the fundamental properties of real networks in a natural way make the d PSO model a very promising candidate upon which higher-dimensional hyperbolic embedding techniques may be developed in the future.

Methods

The original popularity-similarity optimisation model places the network nodes in the native representation of the hyperbolic plane during the network generation. First, [in the following section](#) we describe the native representation of the d -dimensional hyperbolic space and the corresponding formula of the hyperbolic distance. Next, the network generation algorithm of the d PSO model is introduced by extending the PSO model to the hyperbolic space of any integer dimension $d \geq 2$.

Native representation of the hyperbolic space. The d -dimensional hyperbolic space of constant curvature $K < 0$ is represented in the so-called native representation²⁰ by a d -dimensional ball of infinite radius in the Euclidean space (for which $K = 0$). In this representation the Euclidean angles between hyperbolic lines are equal to their hyperbolic values, and the radial coordinate r of a point (defined as its Euclidean distance from the centre of the ball) is equal to its hyperbolic distance from the ball centre. The hyperbolic distance between two points is measured along their connecting hyperbolic line, which is either an arc going through the given points and intersecting the ball's boundary perpendicularly or – if the ball centre falls on the Euclidean line connecting the two points in question – the corresponding diameter of the ball. The hyperbolic distance x between two points given by the Cartesian coordinate vectors $\underline{u} = (u_1, u_2, \dots, u_d)$ and $\underline{v} = (v_1, v_2, \dots, v_d)$ of norms $\|\underline{u}\| = \sqrt{\sum_{q=1}^d u_q^2} \equiv r_u$ and $\|\underline{v}\| = \sqrt{\sum_{q=1}^d v_q^2} \equiv r_v$ fulfills the hyperbolic law of cosines written as

$$\cosh(\zeta x) = \cosh(\zeta r_u) \cosh(\zeta r_v) - \sinh(\zeta r_u) \sinh(\zeta r_v) \cos(\theta_{u,v}), \quad (1)$$

where $\zeta = \sqrt{-K}$, and $\theta_{u,v} = \arccos(\frac{\underline{u} \cdot \underline{v}}{\|\underline{u}\| \|\underline{v}\|}) = \arccos(\frac{\sum_{q=1}^d u_q v_q}{r_u r_v})$ is the angle between the examined points. Note that in the case of $r_u = 0$ simply $x = r_v$, and if $r_v = 0$ then $x = r_u$. According to Ref.²⁰, for sufficiently large ζr_u and ζr_v with an angular distance $\theta_{u,v}$ larger than $2 \cdot \sqrt{e^{-2\zeta r_u} + e^{-2\zeta r_v}}$ but small enough to use the approximation $\sin(\theta_{u,v}/2) \approx \theta_{u,v}/2$, the hyperbolic distance can be approximated as

$$x \approx r_u + r_v + \frac{2}{\zeta} \cdot \ln \left(\sin \left(\frac{\theta_{u,v}}{2} \right) \right) \approx r_u + r_v + \frac{2}{\zeta} \cdot \ln \left(\frac{\theta_{u,v}}{2} \right). \quad (2)$$

Description of the extended PSO model. The PSO model²¹ generates networks that have a scale-free degree distribution characterised by a degree decay exponent γ that is determined by the radial arrangement of the network nodes. In terms of these, there are two natural possibilities for the extension of the two-dimensional PSO model to any integer number of dimensions $d \geq 2$ that both gives back the original PSO model at $d = 2$: one can either make γ independent of the number of dimensions by introducing a d -dependent multiplier in the radial coordinates of the low-temperature regime that yields highly clustered networks, or not change this coordinate formula compared to the two-dimensional case and make the degree decay exponent γ dependent on the value of d . In the present study, we chose the latter option since it offers the opportunity to expand the range of the achievable γ values below 2 by increasing the number of dimensions. This choice is established in more detail in Sect. S1.4 of the Supplementary Information, where we also introduce the f PSO model that provides the possibility to decrease the degree decay exponent γ below 2 even in the two-dimensional case by treating the multiplying factor in the formula of the initial radial coordinates as a new model parameter.

In the d PSO model, the network nodes appear one by one in the above described native representation of the d -dimensional hyperbolic space and connect to previously appeared nodes with probabilities depending on the hyperbolic distances. The parameters of the model can be listed as follows:

- The curvature $K \in \mathbb{R}^-$ of the hyperbolic space, controlled by $\zeta = \sqrt{-K} > 0$. Changing the value of ζ corresponds to a simple rescaling of the hyperbolic distances. The usual custom is to set the value of ζ to 1 (i.e. K to -1).
- The dimension $2 \leq d \in \mathbb{Z}^+$ of the hyperbolic space.
- The final number of nodes $N \in \mathbb{Z}^+$ in the network.
- The number of connections $m \in \mathbb{Z}^+$ established by each node after the m th one at its appearance. The average degree of the network is approximately $\bar{k} = 2 \cdot m$.
- The popularity fading parameter $\beta \in (0, 1]$, controlling the outward drift of the nodes in the native ball. The exponent γ of the power-law decaying tail of the degree distribution is related to the popularity fading parameter as

$$\gamma = 1 + \frac{1}{(d-1) \cdot \beta}. \quad (3)$$

According to this relation between γ , β and d , in the case of different dimensions different popularity fading parameters are needed to obtain the same degree decay exponent γ . Note that as the dimension d increases, the achievable smallest degree decay exponent ($\gamma_{\min} = 1 + 1/(d-1)$), yielded by $\beta = 1$, decreases, meaning that in higher dimensions the attainable highest degree is larger than in hyperbolic spaces of smaller dimensions. (The details of the derivation of Eq. (3) are given in Sect. Degree distribution and in Sect. S1 of the Supplementary Information.)

- The temperature $0 \leq T$, $T \neq \frac{1}{d-1}$, controlling the average clustering coefficient \bar{c} of the network. As the temperature increases from 0, the average clustering coefficient decreases, and settles to a more or less constant value at $T = \frac{1}{d-1}$. For $\beta \leq \frac{1}{d-1}$ (i.e., for $2 \leq \gamma$), the clustering is asymptotically zero for any $\frac{1}{d-1} < T$, while for larger popularity fading parameters (i.e., for $\gamma < 2$) the lowest possible \bar{c} is an increasing function of β .

During the random graph generation process, initially the network is empty, and at each time step $j = 1, 2, \dots, N$ a new node joins the network as follows:

1. The new node j appears at radial distance r_{jj} from the origin with a position chosen uniformly at random on the surface of the corresponding d -dimensional ball by sampling independently d number of Cartesian coordinates from the standard normal distribution and then normalising the obtained position vector (i.e., dividing all the sampled coordinates u_q , $q \in [1, d]$ by $\sqrt{\sum_{q=1}^d u_q^2}$ and then multiplying all of them by r_{jj}). The radial coordinate r_{jj} is calculated as

$$\begin{aligned} \text{(a)} \quad r_{jj} &= \frac{2}{\zeta} \ln j \text{ if } T < \frac{1}{d-1}, \text{ and} \\ \text{(b)} \quad r_{jj} &= \frac{2T(d-1)}{\zeta} \ln j \text{ if } \frac{1}{d-1} < T, \end{aligned}$$

where the change in the multiplying factor at $T = 1/(d-1)$ was introduced to ensure that the formula of the degree decay exponent γ remains the same (given by Eq. (3)) for all temperature settings. (Note that a similar adjustment of r_{jj} was already introduced at $T = 1$ in the original PSO model of $d = 2^1$.)

2. The radial coordinate of each previously (at time $i < j$) appeared node i is increased according to the formula $r_{ij} = \beta r_{ii} + (1 - \beta)r_{jj}$ in order to simulate popularity fading.
3. The new node j establishes connections with m number of previously appeared nodes. Only single links are permitted. If the number of previously appeared nodes is not larger than m , then node j connects to all of them. Otherwise (i.e., for $m + 1 < j$),

- (a) node j connects to the m hyperbolically closest nodes if $T = 0$, and
- (b) at temperatures $0 < T$, any previous node $i = 1, 2, \dots, j-1$ gets connected to node j with probability

$$p(x_{ij}) = \frac{1}{1 + e^{\frac{\zeta}{2T}(x_{ij} - R_j)}}, \quad (4)$$

where the hyperbolic distance x_{ij} between the node pair $i-j$ can be calculated based on Eq. (1) and the so-called cutoff distance R_j can be obtained by solving the equation $m = \bar{k}_j$, where the expected number \bar{k}_j of the realised connections of the new node j at its arrival time j can be written as

$$\bar{k}_j = \eta(d) \cdot \int_1^j \int_0^\pi \frac{\sin^{d-2} \theta_{ij}}{1 + \left(e^{\frac{\zeta}{2} \cdot (r_{ij} + r_{jj} - R_j)} \cdot \sin\left(\frac{\theta_{ij}}{2}\right) \right)^{\frac{1}{T}}} d\theta_{ij} di \quad (5)$$

using

$$\eta(d) = \begin{cases} \frac{\left(\frac{d-1}{2}\right)! \cdot \frac{d-2}{2}! \cdot 2^{d-2}}{(d-2)! \cdot \pi} & \text{if } d \text{ is even,} \\ \frac{(d-1)!}{\left(\frac{d-1}{2}\right)! \cdot \frac{d-1}{2}! \cdot 2^{d-1}} & \text{if } d \text{ is odd.} \end{cases} \quad (6)$$

As detailed in Sect. S2 of the Supplementary Information, at $\beta \neq 1/(d-1)$ the cutoff distance can be approximated as

$$R_j \approx \begin{cases} \frac{2}{\zeta \cdot (d-1)} \cdot \ln \left(\frac{m \cdot \sin((d-1) \cdot T \cdot \pi) \cdot (1 - (d-1) \cdot \beta)}{\eta(d) \cdot \pi \cdot 2^{d-1} \cdot T \cdot (j^{3-2-d} - j^{(d-1) \cdot (\beta-2)})} \right) & \text{if } 0 < T < \frac{1}{d-1}, \\ \frac{2 \cdot T}{\zeta} \cdot \ln \left(\frac{m \cdot (1 - (d-1) \cdot \beta)}{\eta(d) \cdot \int_0^\pi \sin^{d-2}(\phi) \cdot \sin^{-1/T}(\phi/2) d\phi \cdot (j^{3-2-d} - j^{(d-1) \cdot (\beta-2)})} \right) & \text{if } \frac{1}{d-1} < T, \end{cases} \quad (7)$$

whereas at $\beta = 1/(d-1)$ the formula

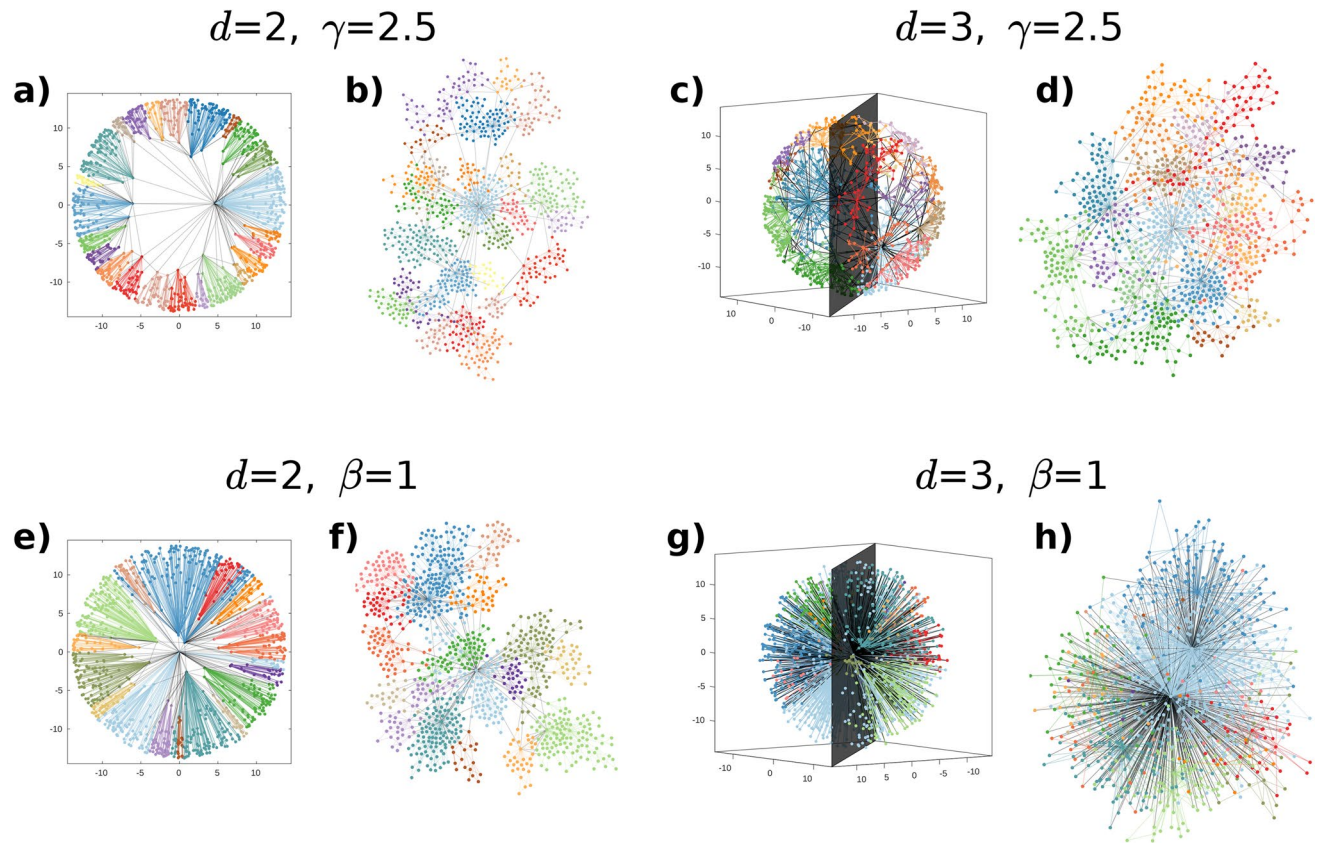


Figure 1. Layouts of networks generated by the d PSO model in 2- and 3-dimensional hyperbolic spaces of curvature $K = -1$. The colouring of the nodes and the links indicates communities found by the Louvain algorithm. In panels **a**, **c**, **e**, **g** we display the network in the native d -dimensional hyperbolic ball, whereas panels **b**, **d**, **f**, **h** show the standard Euclidean layout of the same graph for comparison. In the top row, we present networks of the same degree decay exponent $\gamma = 2.5$, generated on the 2-dimensional hyperbolic plane, setting the popularity fading parameter β to $2/3$ (panels **a**, **b**), and in the 3-dimensional hyperbolic space, setting the popularity fading parameter β to $1/3$ (panels **c**, **d**). In the bottom row, we show networks of the same popularity fading parameter $\beta = 1$, corresponding to the smallest degree exponent achievable in a given dimension, namely $\gamma = 2.0$ in the 2-dimensional case (panels **e**, **f**) and $\gamma = 1.5$ in the 3-dimensional case (panels **g**, **h**). For each network, we set the number of nodes N to 1000, the expected average degree $2m$ to 4 and the temperature T to 0. The average clustering coefficient \bar{c} of the displayed networks and the modularity Q for their displayed partitions are the following: $\bar{c} = 0.729$ and $Q = 0.882$ for panels **a**, **b**; $\bar{c} = 0.644$ and $Q = 0.845$ for panels **c**, **d**; $\bar{c} = 0.788$ and $Q = 0.829$ for panels **e**, **f**; while $\bar{c} = 0.964$ and $Q = 0.354$ for panels **g**, **h**.

$$R_j \approx \begin{cases} \frac{2}{\zeta \cdot (d-1)} \cdot \ln \left(\frac{m \cdot \sin((d-1) \cdot T \cdot \pi) \cdot j^{2 \cdot d-3}}{\eta(d) \cdot \pi \cdot 2^{d-1} \cdot T \cdot \ln j} \right) & \text{if } 0 < T < \frac{1}{d-1}, \\ \frac{2 \cdot T}{\zeta} \cdot \ln \left(\frac{m \cdot j^{2 \cdot d-3}}{\eta(d) \cdot \int_0^\pi \sin^{d-2}(\phi) \cdot \sin^{-1/T}(\phi/2) \, d\phi \cdot \ln j} \right) & \text{if } \frac{1}{d-1} < T. \end{cases} \quad (8)$$

can be used, where $\int_0^\pi \sin^{d-2}(\phi) \cdot \sin^{-1/T}(\phi/2) \, d\phi$ can be calculated numerically.

Results

We generated networks with the above-described d PSO model using numerous parameter settings. As an illustration, in Fig. 1 we show layouts of d PSO networks of size $N = 1000$ both in the native representation of the d -dimensional hyperbolic space and according to a standard force-directed layout algorithm in the Euclidean plane. We analyzed the major structural properties of d PSO networks from various points of view. The results along with their detailed explanations are presented in the following subsections, focusing on a few fundamental network characteristics such as the degree distribution (Sect. Degree distribution), the average clustering coefficient (Sect. Clustering coefficient) and the community structure (Sect. Finding and evaluating communities). Further plots regarding simulation results are shown in Sect. S3 of the Supplementary Information. Moreover, in Sect. S4 of the Supplementary Information we also study a three-dimensional extension of the nonuniform popularity-similarity optimisation model (nPSO)^{24,29} that samples the angular coordinates of the network nodes from a multimodal distribution to allow control over the number and the size of the communities.

Degree distribution. Our analytical calculations show that when generating networks according to the d PSO model presented in the [Methods section](#), we obtain degree distributions that follow a scaling form of

$$\mathcal{P}(k) \sim k^{-\gamma}, \quad (9)$$

where the degree decay exponent γ , given by Eq. (3), depends on the popularity fading parameter β and also the number of dimensions d , but is independent of the temperature T . This is a direct consequence of the fact that the probability for node i (appearing at time i) to attract a link from node j (appearing at time $j > i$) in the above model can be written as

$$\Pi_{d\text{PSO}}(i, j) = m \cdot \frac{i^{-(d-1)\beta}}{\int_1^j \ell^{-(d-1)\beta} d\ell}, \quad (10)$$

with m denoting the number of connections established by node j at its appearance. (The derivation of the above formula is given in Sect. S1 of the Supplementary Information.) Indeed, by following the idea in Ref.²¹, it can be shown that Eq. (10) is equivalent to imposing an extended preferential attachment rule (EPA) that defines the connection probability between an existing node i with a degree $k_i(j)$ and the newly appearing node j as

$$\Pi_{\text{EPA}}[k_i(j)] = m \frac{k_i(j) - m + A}{(m + A)j}, \quad (11)$$

where $A = (\gamma - 2)m$ is a parameter called initial attractiveness. According to Ref.⁴⁴, in networks generated by this EPA rule, the expected degree of node i becomes

$$\overline{k_i(j)} = m + A \left[\left(\frac{i}{j} \right)^{-\alpha} - 1 \right] \quad (12)$$

at time j , and the degree distribution develops into a scale-free form of $\mathcal{P}(k) \sim k^{-\gamma}$, where the exponents α and γ are connected by $\alpha(\gamma) = \frac{1}{\gamma-1}$. We can relate Eqs. (11) and (12) to Eq. (10) by identifying α as $\alpha = (d-1)\beta$, verifying that

$$\Pi_{d\text{PSO}}(i, j) = \Pi_{\text{EPA}}[\overline{k_i(j)}]. \quad (13)$$

Finally, since γ can be expressed from $\alpha(\gamma)$ as $\gamma = 1 + \frac{1}{\alpha}$, we arrive at the formula in Eq. (3) for the decay exponent of the degree distribution in Eq. (9). Thus, networks that grow according to the rules outlined in the [Methods section](#) inherit the scale-free property from the original PSO model, albeit using the same β parameter leads to a smaller γ exponent for greater values of d . Note that by setting d to 2, one can easily recover the results of the original PSO model in Ref.²¹, where the degree decay exponent has been found to be $\gamma = 1 + \frac{1}{\beta}$. The analytical results are in perfect agreement with the numerical simulations, as indicated by Fig. 2, where we display the complementary cumulative distribution function (CCDF) of the node degrees for several networks obtained from the d PSO model with different combinations of the d and the β parameters at different values of the temperature T .

Clustering coefficient. Based on simulations, we have found that the average clustering coefficient \bar{c} of d PSO networks displays an unusually rich behaviour depending on the specific choices of the model parameters. First, in Fig. 3 we show the measured \bar{c} as a function of the rescaled temperature $T \cdot (d-1)$ for different settings of the further model parameters. A rather simple observation regarding this figure is that \bar{c} is an increasing function of m . This is reasonable in the light of the role of this model parameter outlined in the [Methods section](#): m is related to the expected average degree of the d PSO networks as $k = 2 \cdot m$, meaning that higher values of m correspond to higher average degrees.

Besides, according to Fig. 3, the average clustering coefficient is always maximal at $T = 0$, i.e. when newly appearing nodes connect only to the hyperbolically closest existing nodes. If, however, the temperature increases, the cutoff in the connection probability becomes milder. This implies that rather distant nodes also become likely to be connected and consequently, the value of \bar{c} decreases. Nevertheless, above a certain point $T_c = \frac{1}{d-1}$, hereinafter referred to as the critical temperature, the average clustering coefficient remains more or less constant. This can be attributed to the fact that at the critical temperature we change the formula of the radial coordinates from $r_{ij} = \frac{2}{\zeta} \ln j$ to $r_{ij} = \frac{2T(d-1)}{\zeta} \ln j$, meaning that the radial coordinates become an increasing function of the temperature above T_c . According to the approximating formula given by Eq. (2), the larger radial coordinates obviously yield larger hyperbolic distances between all node pairs. Therefore, despite the continuous slowing down of the decay in the connection probability as a function of the hyperbolic distance, above $T_c = \frac{1}{d-1}$ the increase in the temperature does not increment further the number of nodes that are reachable for a newly coming node, and thus, the average clustering coefficient becomes settled to a constant value.

Finally, it can be clearly seen in Fig. 3 that the average clustering coefficient is an increasing function of the popularity fading parameter β . This is especially striking in the higher-dimensional cases, where the range of the degree decay exponents that are achievable extends to lower values. When $\gamma = 1 + \frac{1}{(d-1)\beta}$ is decreased, the degree distribution of the emerging network decays more slowly and, consequently, the largest occurring degrees increase. This is realised by the increase in the preference of the newly coming nodes for connecting to the early-appeared ones, overshadowing the attractiveness of the angular neighbours located at larger radii.

Now let us concentrate on the subgraph of nodes that appeared before a given time point. Since each node has to create m links at its appearance, the ratio between the connected and the non-connected node pairs in

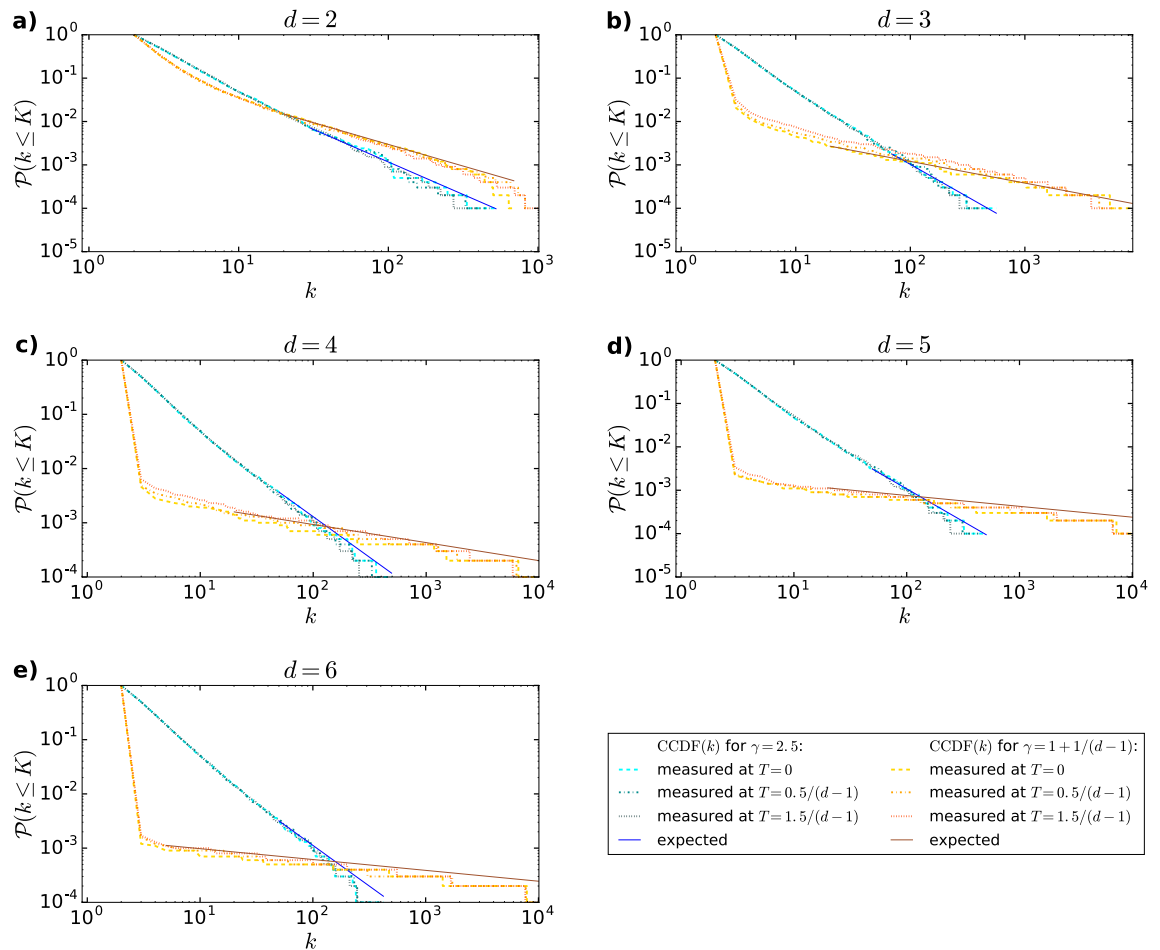


Figure 2. Degree distribution of networks generated by the *dPSO* model at different parameter settings. Panels **a–e** correspond to different dimensions d from 2 to 6. We display in all the panels the complementary cumulative distribution function (CCDF) of the node degrees for networks of two different degree decay exponents: $\gamma = 2.5$ (blue) and $\gamma = 1 + 1/(d - 1)$ (orange). In both cases, we generated a network using the temperature setting $T = 0$ (dashed lines), $T = 0.5/(d - 1)$ (dash-dotted lines) or $T = 1.5/(d - 1)$ (dotted lines). The curvature of the hyperbolic space, the number of nodes and the half of the expected average degree were the same for all networks, namely $K = -\zeta^2 = -1$, $N = 10,000$ and $m = 2$. All the indicated simulation results match well the curve $\mathcal{P}(k \leq K) \sim k^{-(\gamma-1)}$ (shown by solid lines) that was expected based on the analytical calculations.

this subgraph increases towards earlier times (i.e., as the number of nodes in the subgraph decreases). Therefore, if the degree decay exponent γ of the network is decreased, the set of the primarily attractive inner nodes tightens and the new nodes tend to connect into a more densely connected group of nodes, which increases the number of triangles and, accordingly, the average clustering coefficient. It is important to note that at extremely small values of γ , the high density of the connections among the most popular inner nodes moderates the effect of the temperature increase on \bar{c} by limiting the probability to connect to nodes that are not connected to each other. Due to this, the networks of small enough degree decay exponents are highly clustered not only at small temperatures, but in the high-temperature regime as well.

To provide further insight into how the model parameters affect the triangle formation in *dPSO* networks, in Fig. 4 we show the average clustering coefficient as a function of the degree decay exponent γ at different settings of the temperature T and the dimension d . This figure again proves that the critical temperature $T_c = \frac{1}{d-1}$ separates two distinct regimes, where \bar{c} shows a fundamentally different nature.

Below the critical temperature $T_c = \frac{1}{d-1}$ (Fig. 4a,b), the value of \bar{c} measured at a given degree decay exponent γ is affected by the number of dimensions of the underlying hyperbolic space: *dPSO* networks created in higher-dimensional spaces using smaller popularity fading parameters display smaller values of \bar{c} compared to networks that have the same degree decay exponent γ but were generated in lower-dimensional hyperbolic spaces with higher popularity fading parameters. This can be explained roughly by considering that at not too high temperatures, the newly appearing nodes tend to connect to already existing nodes that are relatively similar to them, where high similarity means small angular distance²¹. If the number of dimensions or, equivalently, the number of independent angular coordinates characterising the node attributes is increased, the number of possible coordinate combinations that define the set of positions of similar attributes for a given new node also

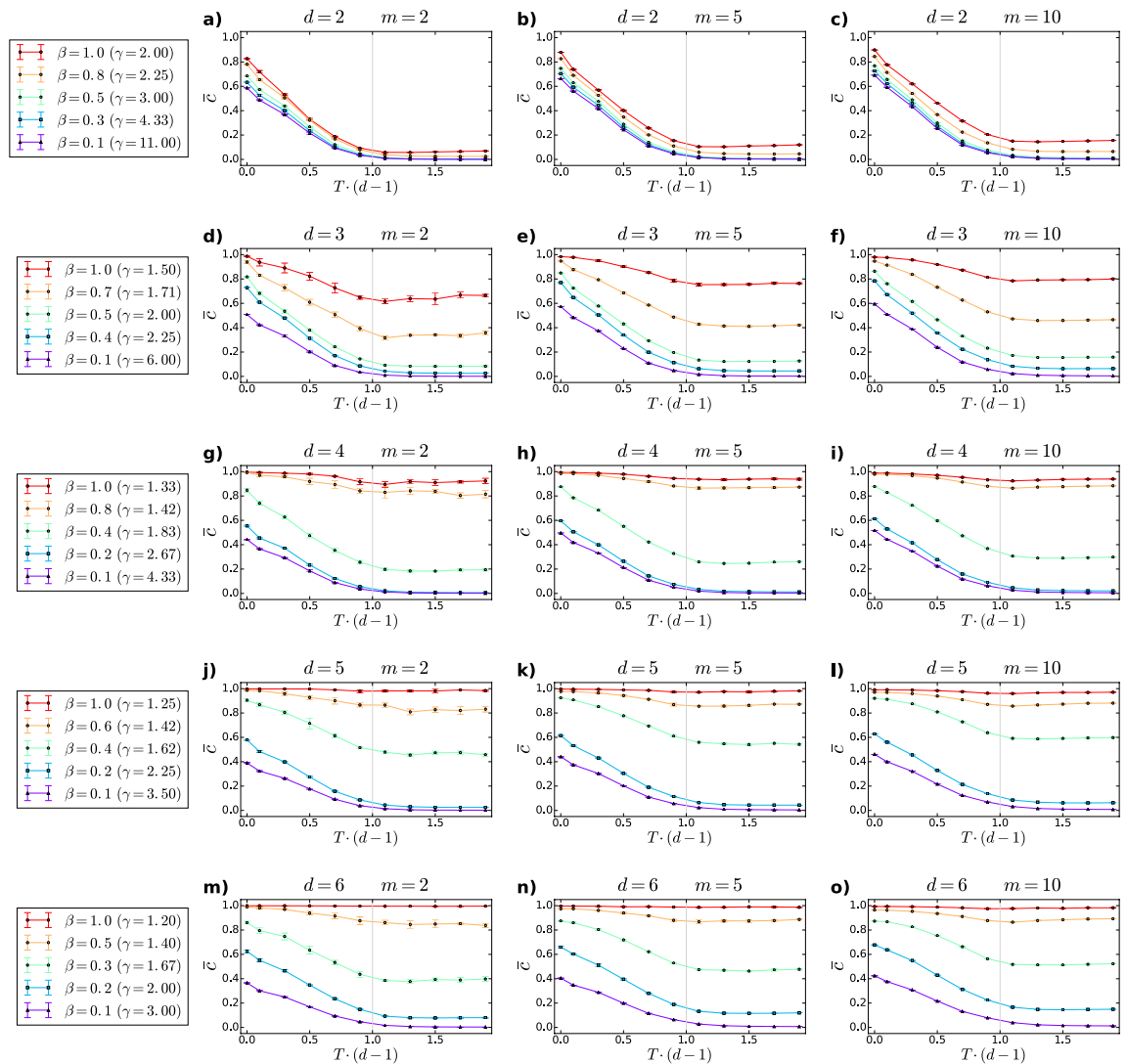


Figure 3. Average clustering coefficient \bar{c} of d PSO networks as a function of the rescaled temperature $T \cdot (d-1)$. Each row of panels (i.e., panels a–c, d–f, g–i, j–l and m–o) was created using a given dimension d , and each column of subplots presents the results obtained with a given value of the expected average degree $2m$, as written in the panel titles. The different curves of each panel correspond to different values of the popularity fading parameter β (yielding different degree decay exponents γ), listed for each dimension in the leftmost panel of the corresponding row. The displayed data points were obtained by averaging over 5 d PSO networks generated independently with a given set of model parameters, setting the number of nodes to 10,000 and the curvature of the hyperbolic space to -1 in each case. The error bars show the standard deviations measured among the 5 networks. The grey vertical lines indicate the critical point $T_c = 1/(d-1)$.

increases. Therefore, any two nodes that can be considered to be similar from a third node's point of view are less and less likely to have angular coordinates that are relatively close to each other as well with the increase in the number of dimensions. Consequently, the selected nodes to which the new node connects tend to share fewer links in higher dimensions; thus, the number of triangles in the emerging network is reduced.

As we approach the critical point from below, the role of the number of dimensions in local triangle formation gradually weakens. Near to and above $T = T_c$ (Fig. 4c,d), where connections between rather distant nodes are likely to occur too, the value of \bar{c} is independent of the individual values of the d and the β parameters that produce the same degree decay exponent γ . Here, it seems that the randomising influence of the high temperature on link formation is so strong that the similar effect of the large number of dimensions is negligible compared to it.

Finally, let us consider the case where both the degree decay exponent γ and the average clustering coefficient \bar{c} are fixed by some external condition (e.g., when modelling real networks), and the goal is to provide parameter settings using which the d PSO model is capable of reproducing these fixed values. As described above, γ is controlled by $(d-1) \cdot \beta$, whereas \bar{c} can be regulated by tuning $(d-1) \cdot T$. This would allow in principle multiple solutions with different d values, where of course the popularity fading parameter β and the temperature T have to be adjusted correspondingly to leave γ and \bar{c} unaltered. However, under certain circumstances this freedom in the choice of d can become restricted. On the one hand, extreme low required γ values impose a lower boundary

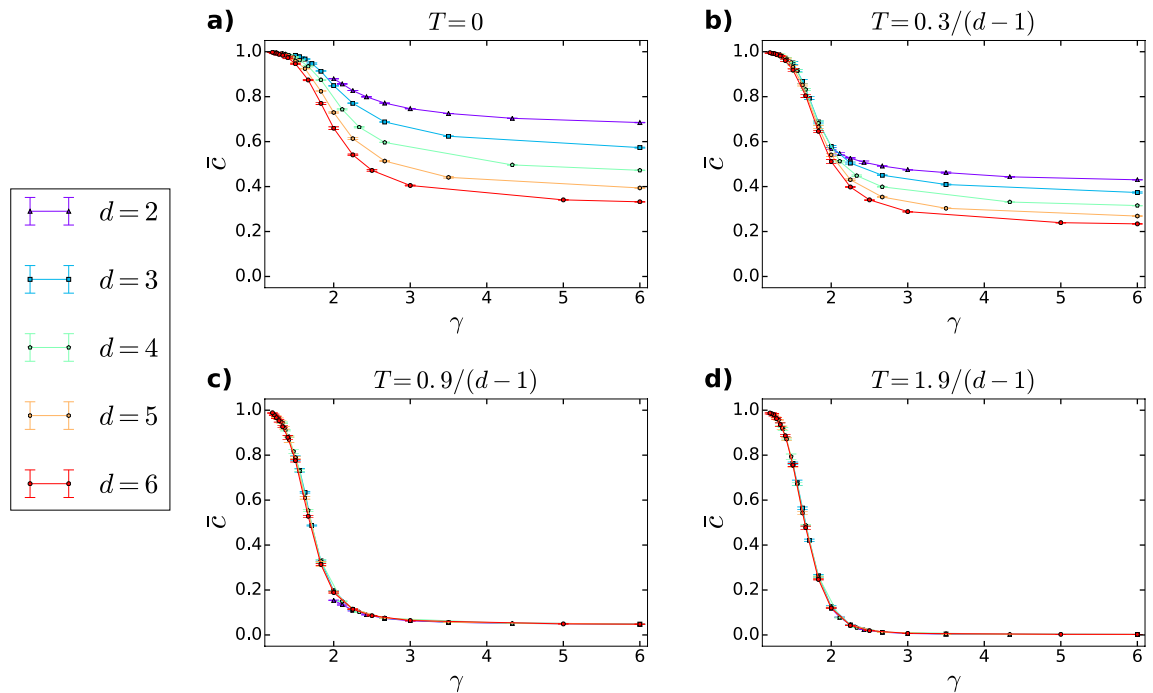


Figure 4. Average clustering coefficient \bar{c} measured in d -dimensional PSO networks as a function of the degree decay exponent γ at different values of the temperature T and the dimension d . We plotted the average clustering coefficient averaged over 5 networks for each parameter setting, with the error bars indicating the standard deviations among the 5 networks. The size of the networks was $N = 10,000$, the expected average degree was $k = 2m = 10$, and each network was generated in a hyperbolic space of curvature $K = -1$. The curves of different colours correspond to different values of the dimension d of the hyperbolic space, listed in the legend. The popularity fading parameter was always set to $\beta = \frac{1}{(d-1)(\gamma-1)}$. The different panels **a–d** were created using different values of the temperature T , specified in the panel title.

on the number of dimensions as $d \geq 1 + \frac{1}{\gamma-1}$. On the other hand, high required \bar{c} values can impose an upper bound on d at a given γ . Nevertheless, allowing the number of dimensions d to take also other values than $d = 2$ is certainly increasing our freedom in the choice of the model parameters in general.

Finding and evaluating communities. Communities are very important structural units in complex networks at the “mesoscopic” scale, without a widely accepted unique definition, but usually associated with subgraphs with a larger internal and a smaller external link density. The automated extraction of communities based solely on the network topology is a challenging problem, with an immense number of different solutions proposed in the literature^{9–11}. An interesting related feature of two-dimensional hyperbolic networks is that they also contain communities for the major part of the parameter space, in spite of the lack of any explicit built-in community formation mechanism in the graph generation algorithms²⁵. Motivated by this and the overall importance of communities in network science, here we examine also the community structure of d -dimensional PSO networks.

Along this line, we apply three independent and well-established community finding algorithms to locate the modules, namely the asynchronous label propagation algorithm (we used the Python function ‘`asyn_lpa_communities`’ available in the ‘`networkx.algorithms.community.label_propagation`’ package)⁴⁵, the Infomap method (we used the Python package available at <https://pypi.org/project/infomap/>)⁴⁶ and the Louvain algorithm (we used the Python implementation available at <https://github.com/taynaud/python-louvain>)⁴⁷. The basic idea of asynchronous label propagation is to simulate the diffusion of community labels along the examined network, where the regular updating of the labels based on the neighbouring nodes brings a rapid consensus among the members of a dense group on a unique label. In contrast, the Infomap algorithm provides an information-theoretic approach for finding communities, taking advantage of the fact that communities can actively help in achieving the most parsimonious description of the trajectory of an infinitely long random walk on the network. The algorithm itself searches for the minimum of the so-called map equation, which expresses the code length for an average movement in the above-mentioned random walk process. Finally, the Louvain method performs a fast and efficient heuristic maximisation of modularity, which corresponds to the most widely used quality measure for communities⁴⁸, expressed in general as

$$Q = \frac{1}{2E} \sum_{i=1}^N \sum_{j=1}^N [A_{ij} - P_{ij}] \delta_{c_i, c_j}, \quad (14)$$

where N is the number of nodes in the network, A_{ij} denotes an element of the adjacency matrix ($A_{ij} \equiv A_{ji} = 1$ if i is connected to j , and otherwise $A_{ij} \equiv A_{ji} = 0$), P_{ij} gives the connection probability between nodes i and j in a random null model, E stands for the total number of links in the network, c_i is the community to which node i belongs and the Kronecker delta δ_{c_i, c_j} ensures that non-zero contribution can come only from node pairs of the same community. A natural choice for the null model is given by the configuration model, yielding $P_{ij} = \frac{k_i k_j}{2E}$. In the Louvain approach, Q is optimised in a hierarchical manner, where after finding the local maximum at a given organisation level of the network, in the next step we move up to the next level by aggregating the current communities into single nodes.

As demonstrated by Fig. 1, a community in a hyperbolic network arises from some inner nodes that serve as community cores and the outer nodes of the corresponding angular sector that are held together by their common preference toward the same attractive centres. As it was detailed in Ref.²⁵ in the case of the two-dimensional PSO model, the emergence of a strong community structure can be achieved under two conditions: the existence of inner nodes that are distant from each other enough to provide well-separated attractive centres for the different angular regions, and the localisation of the connections. The distance between the inner community cores can be increased by accelerating their outward drift that simulates the popularity fading via decreasing the popularity fading parameter β . The strong localisation of the connections can be ensured primarily by setting the temperature T to a small value and thus making the cutoff in the connection probability sharp as a function of the hyperbolic distance. However, it has to be also taken into consideration that if the ratio between the number N of nodes and the number m of connections established by each node at its appearance is smaller, then, to create all the m number of links, the nodes are forced more often to connect even to farther nodes. As a consequence, a small temperature in itself is not always enough to make the connections localised, but it has to be complemented with a relatively large N/m ratio in order to make the connections more strongly determined by the hyperbolic distances and create hereby a more clear separation between the angular regions with respect to the links, thus supporting community formation.

In Fig. 5, we show the highest modularity Q (calculated with the Python function ‘modularity’ available in the ‘networkx.algorithms.community.quality’ package) obtained among the applied three community finding methods as a function of the rescaled temperature $T \cdot (d - 1)$ for d PSO networks generated at different parameter settings. In each panel, the bundle of curves (showing Q for networks with different degree decay exponents) seems to form a fork-like pattern, in which the curves are more distant from each other in the low-temperature regime and approach each other at high values of $T \cdot (d - 1)$, where Q becomes more or less constant and independent from the rescaled temperature.

In Fig. 6, we show the achieved highest modularity Q as a function of the degree decay exponent γ for different number of dimensions at a low temperature (Fig. 6a), a moderate temperature (Fig. 6b) and a high temperature (Fig. 6c). For all of the examined temperatures, Q starts at low values and shows first a strong, then a mild increase toward the higher values of γ , i.e. as the largest node degrees decrease. The effect of the temperature can be observed by comparing the three panels, where we can see that the constant value to which Q settles in the large γ regime is higher if T is lower, as expected. Furthermore, according to Fig. 6c, when the rescaled temperature $T \cdot (d - 1)$ is high enough, the $Q(\gamma)$ curves seem to collapse onto a universal curve for all the examined dimensions, while at smaller temperatures (Fig. 6a,b), the modularity measured at a given degree decay exponent γ is a decreasing function of the dimension d , similarly to what has been seen on the local scale for the average clustering coefficient in Fig. 4.

Discussion

The PSO model²¹ is arguably one of the most successful hyperbolic network models since it offers a quite natural way to reproduce the major structural properties of real-world networks. The original formulation of this approach is, however, given only for the two-dimensional hyperbolic disk. Therefore, a question arising naturally in this context is how it can reasonably be extended to higher-dimensional hyperbolic spaces. In the present paper, we studied this issue in detail and introduced the d PSO model as a d -dimensional generalisation of the original PSO model.

The obtained analytical results show that the scale-free property of d PSO networks is inherited from the original PSO model, meaning that the tail of the degree distribution decays as a power-law, $\lim_{k \rightarrow \infty} \mathcal{P}(k) \sim k^{-\gamma}$, for any number of dimensions $d \geq 2$. The exponent γ is determined by the popularity fading parameter β and the number of dimensions d via the formula $\gamma = 1 + \frac{1}{(d-1)\beta}$, which can be basically explained as follows. The solid angle subtended by the entire surface of the d -dimensional ball representing the d -dimensional hyperbolic space is an increasing function of the number of dimensions, meaning that in the case of a higher d , the uniform distribution of the same number of nodes on the surface results in larger angular distances between the nearest neighbours. Therefore, at a given popularity fading parameter (controlling the attractiveness of the early-appeared nodes arising from their relatively low radial coordinate), in higher dimensions we see an increased propensity of the newly coming nodes to connect to the innermost nodes of any angular position instead of the angular neighbours located at larger radii. Due to this, the same popularity fading parameter β yields larger maximum degree, and thus, smaller degree decay exponent γ for larger values of d . To compensate for the decrease in the attractiveness of the angular neighbours caused by the increase in the number of dimensions and keep the degree decay exponent γ at a given value, one has to reduce the distinguished attractiveness of the inner nodes and limit the angular range in which they are preferred. For this, the advantage of the inner nodes in their radial position has to be decreased by shifting them more outwards, which can be achieved by setting the popularity fading parameter β to a smaller value and enhancing thereby the process of popularity fading.

It is worth emphasising that the smallest attainable degree decay exponent (obtained at $\beta = 1$) is 2 in the original PSO model²¹, whereas in d PSO networks

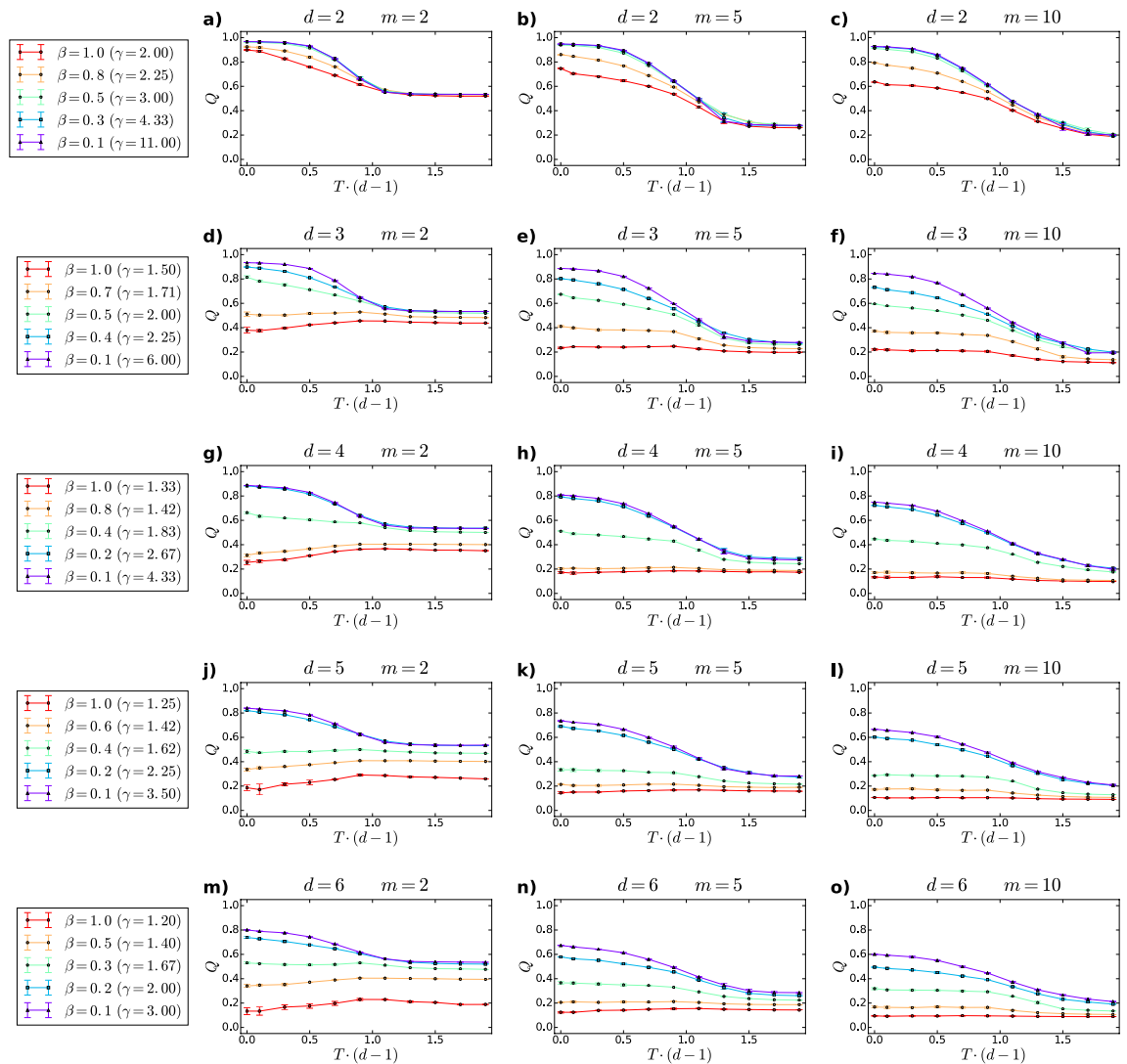


Figure 5. The highest modularity Q achieved among the communities obtained by the asynchronous label propagation, the Louvain and the Infomap algorithms in d PSO networks, as a function of the rescaled temperature $T \cdot (d - 1)$. The dimension d is constant across the panel rows (i.e., in panels **a–c**, **d–f**, **g–i**, **j–l** and **m–o**), whereas the expected average degree $2m$ is constant across the panel columns, as indicated by the panel titles. The different curves in a given subplot correspond to different values of the popularity fading parameter β (yielding different degree decay exponents γ), listed for each dimension in the leftmost panel of the corresponding row. The displayed data points were obtained by averaging over 5 d PSO networks of $N = 10,000$ nodes at curvature $K = -\zeta^2 = -1$, the error bars indicate the standard deviations.

$$\gamma_{\min}(d) = 1 + \frac{1}{d-1}, \quad (15)$$

which can be decreased below the two-dimensional limit $\gamma_{\min}(2) = 2$ by simply increasing the number of dimensions d above 2. Note that the original formulation of the d -dimensional RHG model as it is introduced in Refs.^{42,43} does not allow the generation of such heavy-tailed networks with $\gamma < 2$. Nevertheless, with an appropriate modification of the radial coordinate distribution $\rho(r)$, one can obtain a wide range of degree distributions within this model as well, similarly to what has been shown for the $\mathbb{S}^1/\mathbb{H}^2$ model in Ref.²⁷. The analysis of scale-free networks with a degree decay exponent smaller than two is in itself an interesting topic since such networks exhibit many exotic features that can not be observed in scale-free networks with $2 \leq \gamma$, including the divergence of the average degree or the presence of macroscopic hubs being connected to a finite fraction of the nodes even in the thermodynamic limit $N \rightarrow \infty$ ^{49,50}.

In terms of d PSO networks, we have found that such extremely skewed degree distributions lead to further unexpected results. As indicated in Figs. 4 and 6, d PSO networks of $\gamma < 2$ are characterised by high average clustering coefficient \bar{c} , but remarkably at the same time relatively small modularity Q at basically any temperature T . Intuitively, this behaviour can be understood if we consider that on the one hand, since the largest hubs are formed from the first few nodes of the network generation process, they are densely connected to each other.

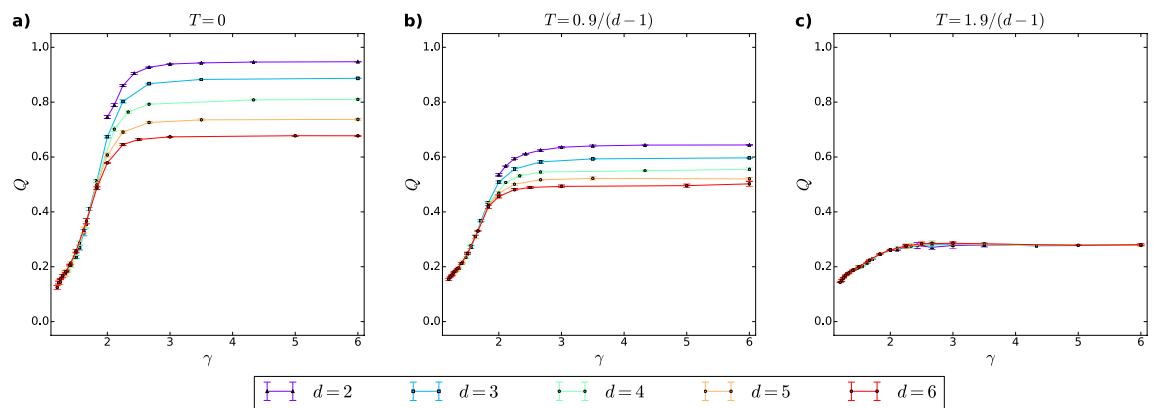


Figure 6. The highest modularity Q achieved between the asynchronous label propagation, the Louvain and the Infomap algorithms in d PSO networks as a function of the degree decay exponent γ at different values of the temperature T and the dimension d . The panels **a–c** refer to different values of the temperature T , given in the title of the subplots. The curves of different colours correspond to different number of dimensions d , as listed below the panels. The popularity fading parameter was calculated as $\beta = \frac{1}{(d-1)(\gamma-1)}$. We always set the curvature $K = -\zeta^2$ of the hyperbolic space to -1 , the network size N to 10,000 and the expected average degree $2m$ to 10. We searched for communities once with all three community detection methods on 5 d PSO networks and plotted the obtained highest modularity averaged over the 5 networks for each parameter setting, with the error bar indicating the standard deviation among the 5 networks.

Thus, in the presence of extremely large hubs to which most of the nodes connect at their appearance, triangles are formed with large probability, resulting in a large \bar{c} . On the other hand, these large hubs make the partitioning of the network into disjunct communities with high modularity practically impossible since the community of any such hub has a macroscopic number of links pointing outside of the given community, resulting in low Q values.

In addition, we found that the number of dimensions d along with the temperature T play a joint role in controlling the average clustering coefficient \bar{c} of d PSO networks as well. More precisely, as it can be seen in Fig. 3, we can distinguish two phases separated by the critical point $T_c = \frac{1}{d-1}$, where the average clustering coefficient of the networks behaves in a fundamentally different way. At temperatures $T < T_c$, \bar{c} is a decreasing function of the temperature; however, at temperatures near and above T_c , \bar{c} tends to become independent from T . Besides, based on Fig. 4, at low temperatures we can state that the average clustering coefficient measured at a given degree decay exponent γ decreases with the number of dimensions d , in perfect accordance with the results found in Ref.⁵¹ for the RHG model. Meanwhile, near the critical temperature \bar{c} begins to show a universal decay with γ , irrespectively of the separate values of the popularity fading parameter β and the dimension d .

Remarkably, a similar separation of phases with the same critical temperature has also been observed in the d -dimensional RHG model^{42,43}. Although the d -dimensional PSO and RHG models share some other common characteristics, such as the Fermi-Dirac connection probability or the uniform angular distribution of the nodes, they fairly differ from each other. One of the most significant differences is apparently the process of network construction: the d -dimensional RHG model generates static hyperbolic graphs with a fixed number of nodes, while the d PSO model simulates the growth of networks in the hyperbolic space, thus providing a dynamic network model. The latter can be advantageous for embeddings by naturally allowing the calculation of local likelihoods instead of the optimisation of global values²². Another difference is that the radial coordinates r of the nodes are sampled from a predefined distribution $\rho(r)$ in the d -dimensional RHG model⁴³, whereas in the case of d PSO networks a logarithmically increasing radial sequence of the nodes is imposed. This implies that in the d -dimensional RHG model two distinct nodes can have the same radial coordinates, a phenomenon which is impossible in the d PSO model. A somewhat less beneficial property of d -dimensional RHG networks is that they are not necessarily sparse: when $\gamma = 2$ and the temperature T is below $1/(d-1)$, the expected average degree of the network scales with the system size as $\langle k \rangle \sim \ln^2(N)$ ⁴³. In contrast to this, the expected average degree of d PSO networks is given by $\langle k \rangle = 2m$; therefore, their sparseness can always be assured by controlling the parameter m .

As it has already been reported in Ref.²⁵, the original, two-dimensional PSO model²¹ is capable of generating networks with strong communities for a wide range of the parameter settings, despite the fact that it does not include any explicitly built-in community structure generating mechanism. Based on the high modularity values measured on d PSO networks (Figs. 5 and 6), here we conclude that the emergence of a strong community structure is not a peculiar feature of the two-dimensional case, but it can be observed even if the number of dimensions is larger than 2, provided that the degree decay exponent γ is not too small and the temperature T is not too high. Nevertheless, at such settings of γ and T , the modularity Q obtained at a given γ decreases as d is increased, similarly to the average clustering coefficient \bar{c} .

In conclusion, motivated by the fact that there is evidently no particular reason to assume that the underlying hyperbolic space of complex networks is certainly two-dimensional, here we proposed a generalisation of the two-dimensional popularity-similarity optimisation model of network growth, namely the d PSO model, in which

the dimension d acts as an additional degree of freedom. A further motivation for higher-dimensional hyperbolic models is given by higher-dimensional embedding methods, where the behaviour of the likelihood optimisation can be more favourable compared to the traditional approaches working in two dimensions. In our studies, we found that d PSO networks are suitable for capturing many essential characteristics of real-world networks such as the scale-free property, the high value of the average clustering coefficient or the strong community structure. Namely, d PSO networks generated in lower-dimensional hyperbolic spaces using not too high temperatures and not too low degree decay exponents simultaneously exhibit all the above-mentioned features of real-world networks. Nevertheless, as the number of dimensions d increases, both the maximal average clustering coefficient and the maximal modularity that can be obtained at a given degree decay exponent γ (by setting the temperature T to 0) decreases. This implies that d PSO networks generated in high-dimensional spaces (e.g., at $d > 10$) can be characterised by neither a clustering nor a community structure of similar strength that can be observed in various real-world examples; hence, a reasonable upper limit can be found on the number of dimensions of the hyperbolic space that underlies real networks. By following similar considerations regarding the clustering coefficient of RHG networks, in Ref.⁵² the authors claim that in terms of both the modelling and the embedding techniques, the most suitable choice for the number of dimensions of the underlying hyperbolic space is $d = 2$. Here, we complement this result by showing that in general, the d PSO model with d slightly above 2 also performs excellently on the modelling ground (still yielding relatively high values of the average clustering coefficient and the modularity), which, along with the recent success of $d > 2$ hyperbolic embedding techniques for instance in link prediction and in graph reconstruction³⁹ or the separation of communities⁴⁰, confirms the relevance of low-dimensional hyperbolic spaces (e.g. with $d = 3$ or $d = 4$) in the theory of complex networks. In Ref.³⁸, several different methods have already been provided for assigning angular coordinates to the network nodes in hyperbolic spaces of arbitrary curvature $K = -\zeta^2$ and number of dimensions d . Using the angular positions yielded by one of these methods, in order to create an embedding that corresponds to our d PSO model, the radial coordinate of the node having the ℓ th ($\ell = 1, 2, \dots, N$) largest degree (with ties in the order of node degrees broken arbitrarily) has to be calculated as $r_{\ell N} = \beta \cdot (2/\zeta) \cdot \ln \ell + (1 - \beta) \cdot (2/\zeta) \cdot \ln N$, where the popularity fading parameter β is determined by the degree decay exponent γ and the number of dimensions d as $\beta = \frac{1}{(d-1) \cdot (\gamma-1)}$.

Data availability

All data generated during the current study are available from the corresponding author upon request.

Code availability

The code used for generating networks with the d -dimensional PSO model will be available at <https://github.com/BianKov/dPSO> upon publication.

Received: 25 August 2021; Accepted: 14 December 2021

Published online: 19 January 2022

References

- Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
- Mendes, J. F. F. & Dorogovtsev, S. N. *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford Univ. Press, 2003).
- Del Genio, C. I., Gross, T. & Bassler, K. E. All scale-free networks are sparse. *Phys. Rev. Lett.* **107**, 178701. <https://doi.org/10.1103/PhysRevLett.107.178701> (2011).
- Milgram, S. The small world problem. *Psychol. Today* **2**, 60–67 (1967).
- Kochen, M. (ed.) *The Small World* (Ablex, 1989).
- Faloutsos, M., Faloutsos, P. & Faloutsos, C. On power-law relationships of the internet topology. *Comput. Commun. Rev.* **29**, 251–262 (1999).
- Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
- Watts, D. J. & Strogatz, S. H. Collective dynamics of “small-world” networks. *Nature* **393**, 440–442 (1998).
- Fortunato, S. Community detection in graphs. *Phys. Rep.* **486**, 75 – 174 (2010). <http://www.sciencedirect.com/science/article/pii/S0370157309002841>.
- Fortunato, S. & Hric, D. Community detection in networks: A user guide. *Phys. Rep.* **659**, 1 – 44 (2016). <http://www.sciencedirect.com/science/article/pii/S0370157316302964>.
- Cherifi, H., Palla, G., Szymanski, B. & Lu, X. On community structure in complex networks: challenges and opportunities. *Appl. Netw. Sci.* **4**, 117 (2019).
- Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999). <https://science.sciencemag.org/content/286/5439/509>.
- Caldarelli, G., Capocci, A., De Los Rios, P. & Muñoz, M. A. Scale-free networks from varying vertex intrinsic fitness. *Phys. Rev. Lett.* **89**, 258702 (2002).
- Servedio, V. D. P., Caldarelli, G. & Buttà, P. Vertex intrinsic fitness: How to produce arbitrary scale-free networks. *Phys. Rev. E* **70**, 056126 (2004).
- Boguñá, M. & Pastor-Satorras, R. Class of correlated random networks with hidden variables. *Phys. Rev. E* **68**, 036112. <https://doi.org/10.1103/PhysRevE.68.036112> (2003).
- Garlaschelli, D., Ahnert, S. E., Fink, T. M. A. & Caldarelli, G. Low-temperature behaviour of social and economic networks. *Entropy* **15**, 3148–3169 (2013). <http://www.mdpi.com/1099-4300/15/8/3148>.
- Balogh, S. G., Palla, G. & Pollner, P. Generalised thresholding of hidden variable network models with scale-free property. *Sci. Rep.* **9**, 11273 (2019).
- Holme, P. & Kim, B. J. Growing scale-free networks with tunable clustering. *Phys. Rev. E* **65**, 026107. <https://doi.org/10.1103/PhysRevE.65.026107> (2002).
- Toivonen, R., Onnela, J.-P., Saramäki, J., Hyvönen, J. & Kaski, K. A model for social networks. *Phys. A Stat. Mech. Appl.* **371**, 851–860 (2006). <https://www.sciencedirect.com/science/article/pii/S0378437106003931>.
- Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A. & Boguñá, M. Hyperbolic geometry of complex networks. *Phys. Rev. E* **82**, 036106. <https://doi.org/10.1103/PhysRevE.82.036106> (2010).

21. Papadopoulos, F., Kitsak, M., Serrano, M. Á., Boguñá, M. & Krioukov, D. Popularity versus similarity in growing networks. *Nature* **489**, 537 EP. <https://doi.org/10.1038/nature11459> (2012).
22. Papadopoulos, F., Psomas, C. & Krioukov, D. Network mapping by replaying hyperbolic growth. *IEEE/ACM Trans. Netw.* **23**, 198–211 (2015).
23. Zuev, K., Boguñá, M., Bianconi, G. & Krioukov, D. Emergence of soft communities from geometric preferential attachment. *Sci. Rep.* **5**, 9421 (2015).
24. Muscoloni, A. & Cannistraci, C. V. A nonuniform popularity-similarity optimization (npso) model to efficiently generate realistic complex networks with communities. *New J. Phys.* **20**, 052002 (2018).
25. Kovács, B. & Palla, G. The inherent community structure of hyperbolic networks. *Sci. Rep.* **11**, 16050. <https://doi.org/10.1038/s41598-021-93921-2> (2021).
26. Serrano, M. A., Krioukov, D. & Boguñá, M. Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.* **100**, 078701. <https://doi.org/10.1103/PhysRevLett.100.078701> (2008).
27. García-Pérez, G., Allard, A., Serrano, M. Á. & Boguñá, M. Mercator: uncovering faithful hyperbolic embeddings of complex networks. *New J. Phys.* **21**, 123033 (2019). <https://doi.org/10.1088%2F1367-2630%2F21123033>
28. Kovács, B. & Palla, G. Optimisation of the coalescent hyperbolic embedding of complex networks. *Sci. Rep.* **11**, 8350. <https://doi.org/10.1038/s41598-021-87333-5> (2021).
29. Muscoloni, A. & Cannistraci, C. V. Leveraging the nonuniform pso network model as a benchmark for performance evaluation in community detection and link prediction. *New J. Phys.* **20**, 063022 (2018).
30. Wang, Z., Li, Q., Jin, F., Xiong, W. & Wu, Y. Hyperbolic mapping of complex networks based on community information. *Phys. A Stat. Mech. Appl.* **455**, 104 – 119 (2016). <http://www.sciencedirect.com/science/article/pii/S0378437116001813>.
31. Wang, Z., Wu, Y., Li, Q., Jin, F. & Xiong, W. Link prediction based on hyperbolic mapping with community structure for complex networks. *Phys. A Stat. Mech. Appl.* **450**, 609 – 623 (2016). <http://www.sciencedirect.com/science/article/pii/S0378437116000182>.
32. Wang, Z., Sun, L., Cai, M. & Xie, P. Fast hyperbolic mapping based on the hierarchical community structure in complex networks. *J. Stat. Mech. Theory Exp.*, <https://doi.org/10.1088/1742-5468/ab3bc8> (2019).
33. Faqeeh, A., Osat, S. & Radicchi, F. Characterizing the analogy between hyperbolic embedding and community structure of complex networks. *Phys. Rev. Lett.* **121**, 098301. <https://doi.org/10.1103/PhysRevLett.121.098301> (2018).
34. Wang, Z., Li, Q., Xiong, W., Jin, F. & Wu, Y. Fast community detection based on sector edge aggregation metric model in hyperbolic space. *Phys. A Stat. Mech. Appl.* **452**, 178 – 191 (2016). <http://www.sciencedirect.com/science/article/pii/S0378437116000595>.
35. Boguñá, M., Papadopoulos, F. & Krioukov, D. Sustaining the internet with hyperbolic mapping. *Nat. Commun.* **1**, 62 (2010).
36. Alanis-Lobato, G., Mier, P. & Andrade-Navarro, M. Efficient embedding of complex networks to hyperbolic space via their laplacian. *Sci. Rep.* **6**, 301082 (2016).
37. Alanis-Lobato, G., Mier, P. & Andrade-Navarro, M. A. Manifold learning and maximum likelihood estimation for hyperbolic network embedding. *Appl. Netw. Sci.* **1**, 10 (2016).
38. Muscoloni, A., Thomas, J. M., Ciucci, S., Bianconi, G. & Cannistraci, C. V. Machine learning meets complex networks via coalescent embedding in the hyperbolic space. *Nat. Commun.* **8**, 1615. <https://doi.org/10.1038/s41467-017-01825-5> (2017).
39. Nickel, M. & Kiela, D. Poincaré embeddings for learning hierarchical representations. In Guyon, I. et al. (eds.) *Advances in Neural Information Processing Systems*, vol. 30 (Curran Associates, Inc., 2017). <https://proceedings.neurips.cc/paper/2017/file/59dfa2df42d9e3d41f5b02bfc32229dd-Paper.pdf>.
40. Muscoloni, A. & Cannistraci, C. V. Angular separability of data clusters or network communities in geometrical space and its relevance to hyperbolic embedding (2019). Preprint at [arXiv:1907.00025](https://arxiv.org/abs/1907.00025) [cs.LG].
41. Krioukov, D. et al. Network cosmology. *Sci. Rep.* **2**, 793. <https://doi.org/10.1038/srep00793> (2012).
42. Yang, W. & Rideout, D. High dimensional hyperbolic geometry of complex networks. *Mathematics* **8** (2020). <https://www.mdpi.com/2227-7390/8/11/1861>.
43. Kitsak, M., Aldecoa, R., Zuev, K. & Krioukov, D. Random hyperbolic graphs in $d + 1$ dimensions (2020). Preprint at [arXiv:2010.12303](https://arxiv.org/abs/2010.12303) [physics.soc-ph].
44. Dorogovtsev, S. N., Mendes, J. F. F. & Samukhin, A. N. Structure of growing networks with preferential linking. *Phys. Rev. Lett.* **85**, 4633–4636. <https://doi.org/10.1103/PhysRevLett.85.4633> (2000).
45. Raghavan, U. N., Albert, R. & Kumara, S. Near linear time algorithm to detect community structures in large-scale networks. *Phys. Rev. E* **76**, 036106. <https://doi.org/10.1103/PhysRevE.76.036106> (2007).
46. Rosvall, M. & Bergstrom, C. T. Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLOS ONE* **6**, 1–10. <https://doi.org/10.1371/journal.pone.0018209> (2011).
47. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008> (2008).
48. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).
49. Seyed-allaei, H., Bianconi, G. & Marsili, M. Scale-free networks with an exponent less than two. *Phys. Rev. E* **73**, 046113. <https://doi.org/10.1103/PhysRevE.73.046113> (2006).
50. Timár, G., Dorogovtsev, S. N. & Mendes, J. F. F. Scale-free networks with exponent one. *Phys. Rev. E* **94**, 022302. <https://doi.org/10.1103/PhysRevE.94.022302> (2016).
51. García-Pérez, G., Boguñá, M. & Serrano, M. Á. Multiscale unfolding of real networks by geometric renormalization. *Nat. Phys.* **14**, 583–589. <https://doi.org/10.1038/s41567-018-0072-5> (2018).
52. Boguñá, M. et al. Network geometry. *Nat. Rev. Phys.* **3**, 114–135. <https://doi.org/10.1038/s42254-020-00264-4> (2021).

Acknowledgements

The research was partially supported by the Hungarian National Research, Development and Innovation Office (grant no. K 128780, NVKP_16-1-2016-0004), by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101021607, and the Thematic Excellence Programme (Tématerületi Kiválósági Program, 2020-4.1.1.-TKP2020) of the Ministry for Innovation and Technology in Hungary, within the framework of the Digital Biomarker thematic programme of the Semmelweis University.

Author contributions

B.K. and G.P. developed the concept of the study, B.K. constructed the generative steps of the d PSO, the f PSO and the three-dimensional nPSO model, B.K. carried out the analytical calculations for the degree distribution of d PSO networks in the $T = 0$ case, B.K. and S.G.B. performed the analytical calculations for the degree distribution of d PSO networks in the $0 < T$ case, B.K. performed the numerical analyses and prepared the figures, B.K., S.G.B. and G.P. analysed and interpreted the results, G.P., S.G.B. and B.K. wrote the paper. All authors reviewed the manuscript.

Competing Interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-04379-1>.

Correspondence and requests for materials should be addressed to S.G.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022