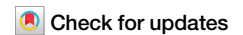


<https://doi.org/10.1038/s42005-025-02464-9>

Mapping bipartite networks into multidimensional hyperbolic spaces



Robert Jankowski^{1,2}, Roya Aliakbarisani^{1,2}, M. Ángeles Serrano^{1,2,3} & Marián Boguñá^{1,2}  

Bipartite networks appear in many real-world contexts, linking entities across two distinct sets. One-mode projections, though widely used, can introduce artificial correlations and inflated clustering, obscuring the true underlying structure. In this work, we propose a geometric model for bipartite networks that leverages the high levels of bipartite four-cycles as a measure of clustering and embeds both node types into a shared similarity space, where the link probability decreases with distance. Additionally, we introduce B-Mercator, an algorithm that infers node positions from the bipartite structure. We evaluate its performance on diverse datasets, illustrating how the resulting embeddings improve downstream tasks such as node classification and distance-based link prediction. These hyperbolic embeddings enable realistic synthetic network generation with node features mirroring real data. By preserving the bipartite structure, our approach avoids projection biases, offering more accurate structural descriptions and providing a robust framework for uncovering hidden geometry in bipartite networks.

Bipartite networks lie at the heart of countless real-world applications, linking authors to the articles they write^{1–4}, users to the products they consume^{5–7}, people to the groups they belong to⁸, or countries to the languages they speak⁹. They also arise naturally in metabolic networks, where metabolites are connected to the chemical reactions or enzymes that transform them^{10–12}, in plant-pollinator networks^{13,14}, and in machine learning applications, where nodes have associated features used to feed graph neural networks¹⁵. By definition, each bipartite system splits its nodes into two disjoint sets, with no edges connecting nodes within the same set. This seemingly simple rule nonetheless yields rich and complex connectivity patterns, enabling researchers to model collaboration, consumption, and association processes across diverse domains. Yet bipartite networks have historically garnered less attention than their unipartite counterparts. Given their ubiquity and straightforward interpretability, a renewed focus on bipartite structures is both timely and necessary to fully capture the multiple facets of interaction present in many real-world complex systems and data structures.

A general practice is to analyze bipartite networks by projecting them onto a single node set, creating a *one-mode* network¹⁶. For instance, in an author–article bipartite network, one might create a unipartite graph of authors by connecting two authors if they have co-authored at least one article^{1–4}. While such one-mode projections allow researchers to employ classic unipartite tools (e.g., clustering coefficients, degree distributions, community-detection algorithms), they also introduce strong

correlations between edges. These correlations arise owing to sets of nodes that share common neighbors in the bipartite structure inducing cliques –or fully connected subgraphs– in the one-mode projection. Hence, the resulting unipartite networks can exhibit inflated clustering and misleading connectivity patterns that do not necessarily reflect the independent pairwise interactions of the underlying bipartite system. Besides, there is an unavoidable loss of information when bipartite networks are projected and it is even possible to obtain the same one-mode projection out of different bipartite networks^{16,17}.

To overcome these limitations and more accurately capture the true structure of bipartite systems, it is critical to develop a modeling framework that treats bipartite networks directly rather than relying solely on their one-mode projections. Thus, we turn our attention to the latent space models, initially developed for social networks¹⁸ and extended to temporal bipartite networks¹⁹ and to growing bipartite generative models²⁰. Within this family of models, we propose to use network geometry to make sense of bipartite networks and to find hyperbolic geometric representations for this class of systems²¹. Hyperbolic geometry is particularly useful for capturing the hierarchical organization of complex networks²². In recent years, network geometry²³ has been extremely successful in explaining undirected²⁴, directed²⁵, weighted²⁶, and multiplex networks^{27,28} in many real systems and has also been extended to bipartite settings, proposing that nodes of both types can indeed lie in a shared latent space with connection probabilities governed by their mutual distances^{12,29}. Despite the fact that bipartite

¹Departament de Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain. ²Universitat de Barcelona Institute of Complex Systems (UBICS), Universitat de Barcelona, Barcelona, Spain. ³ICREA, Passeig Lluís Companys 23, E-08010 Barcelona, Spain.

 e-mail: marian.boguena@ub.edu

networks do not contain triangles by definition (the signature of any metric space), it is still possible to define a clustering coefficient by counting cycles of length four (i.e., squares). Empirical analyses have shown that these four-cycles can be abundant in real bipartite networks, leading to high effective clustering values^{15,30}. This finding suggests that, indeed, bipartite networks may be embedded in a metric space where the likelihood of a connection between two nodes of different types decreases with the distance separating them.

Assuming that a bipartite network can be embedded in a similarity (or metric) space, a logical step is to seek methods for inferring the positions of its nodes within this space from real data. Previous approaches typically rely on one-mode projections of bipartite networks –using, for instance, the D-Mercator tool^{31,32}– to embed each type of node. However, such projections often introduce artificial correlations and inflated clustering, reducing embedding accuracy. Embeddings of one-mode projections have been used in ref. 12 to infer the coordinates of one node type, which are then used to estimate the coordinates of the other type. Another approach¹⁵ transfers coordinates from a unipartite layer –when such a layer exists– and assumes that the same coordinates apply in the bipartite representation. However, none of these methods provides a model-based bipartite embedding. Finally,²⁹ uses the number of common neighbors as a proxy for similarity distance between nodes, but these estimates are noisy and do not support a full embedding of the network.

Here, following results in refs. 12,29, we introduce the bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model in which nodes of both types lie in the same D -dimensional similarity space. Using this model, we propose B-Mercator, an algorithm designed specifically for bipartite networks that enables the creation of multidimensional hyperbolic maps of real bipartite datasets. It is important to emphasize that choosing an appropriate similarity-space dimension D is crucial: while many networks are well captured with $D = 1$, others require higher D to prevent distinct communities from being artificially mixed^{32,33}. To illustrate its capabilities, we embedded and analyzed three datasets: Unicodelang⁹, which links countries to the languages spoken in those regions, Metabolic³⁴, in which metabolites are connected through chemical reactions, and Flavor³⁵, which connects food ingredients to their corresponding flavor compounds. Furthermore, we show how B-Mercator can be applied to supervised machine learning tasks, including node classification and link prediction, yielding a significant performance gain with respect to state of the art methods, especially when a strong correlation exists between nodes' labels and their feature distributions.

Results

The bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model and B-Mercator

The bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model generates bipartite networks with type-A and type-B nodes by assigning them coordinates on the surface of a D -sphere and expected degrees. Pairs of type-A and type-B nodes are then connected with a probability that depends on their spherical distance, rescaled by the product of their expected degrees. Figure 1 shows two examples generated by the bipartite- $\mathbb{S}^D$ model for $D = 1$ and $D = 2$, where node size represents expected degree. The model has a purely geometric counterpart, the bipartite- $\mathbb{H}^{D+1}$ model, in which expected degrees are mapped to a radial coordinate so that nodes are placed in the $(D + 1)$ -sphere in the hyperbolic space and the connection probability becomes a function of the hyperbolic distance only. Thus, in the hyperbolic representation, angular coordinates on the D -sphere encode similarity, whereas the radial coordinate encodes popularity as reflected by the expected degree. In both representations, the model includes an inverse temperature β_b that modulates noise and therefore controls the coupling between network topology and geometry.

Given a real bipartite network, we treat it as a likely realization of the bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model. B-Mercator then finds an embedding of both node types on the surface of a D -sphere $\mathbb{S}^D$ and their expected degrees (or, equivalently, in $\mathbb{H}^{D+1}$) that maximizes the likelihood under the bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model. A full description of the model and the technical details of the B-Mercator embedding algorithm are provided in Methods, Sections IIIA and IIIB.

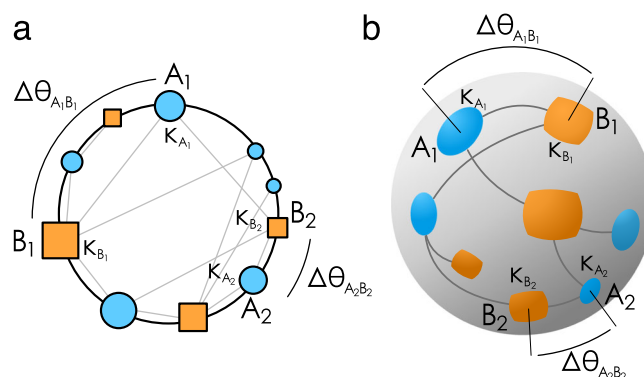


Fig. 1 | Multidimensional hyperbolic bipartite model. Schematic representation of the bipartite- $\mathbb{S}^D$ model in dimension (a) $D = 1$ and (b) $D = 2$. Nodes A are shown as circles whereas nodes B are shown as squares whose sizes are proportional to the nodes' expected degrees. The angular distances between nodes A and B are highlighted ($\Delta\theta_{A_1B_1}$ and $\Delta\theta_{A_2B_2}$). Light grey lines represent the edges in the bipartite network generated by Eq. (1).

Validation

To validate our method, we generated single realizations of synthetic bipartite networks from the bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model with different dimensions and parameters. We then used B-Mercator for each such synthetic network to infer both the model parameters and node coordinates ($\kappa^{\text{inf}}, \theta^{\text{inf}}$), and compared them with the true values ($\kappa^{\text{true}}, \theta^{\text{true}}$). Figure 2 shows comparisons between the true and inferred coordinates for type-A and -B nodes. One can observe a high Spearman correlation coefficient for $D = 1$ and $D = 2$, corroborating the effectiveness of our embedding technique. For more examples in $D = 1$ and $D = 2$ as well as $D = 3$, see Supplementary Figs. 2–4. B-Mercator can also infer the inverse temperature β_b as shown in Supplementary Fig. 7.

As a further check, we tested the reproducibility of the topological properties of the original network embedded by B-Mercator. Using the model parameters and coordinates inferred by B-Mercator, we generated an ensemble of synthetic networks using the bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model (see Supplementary Figs. 8–12). We then evaluated ensemble averages and standard deviations for various topological properties and compared them with those of the original network. The degree distributions and clustering spectra of type-A and type-B nodes were very well reproduced. We also observed good agreement between the empirical and theoretical connection probabilities. These findings confirm that B-Mercator accurately reconstructs the nodes' coordinates in synthetic networks and reliably determines other model parameters, such as hidden degrees and inverse temperature, regardless of the network dimensionality.

Bipartite greedy routing

In order to establish a meaningful geometric representation of a bipartite network at the global scale, we introduce bipartite greedy routing (BGR) as a practical tool to infer the network's effective dimension³⁶. The idea is to first embed the bipartite network into a latent geometric space with B-Mercator and then test how well nodes can route information by simply forwarding messages to their neighbors closest to the destination in that space. By systematically evaluating the success of these greedy routes –measured, for instance, by the probability that messages reach their targets without getting stuck– we gain insights into the dimensional structure underlying the network. In this way, the performance of greedy routing serves as an indicator of how well the bipartite graph can be embedded in a space of a given dimension, effectively allowing us to determine the dimension that best captures its structure.

We implemented a BGR protocol in which both the origin and destination can be either type-A or type-B nodes thus defining four variants. In Fig. 3a, we depict a schematic picture of the BGR for the A–B variant, i.e.,

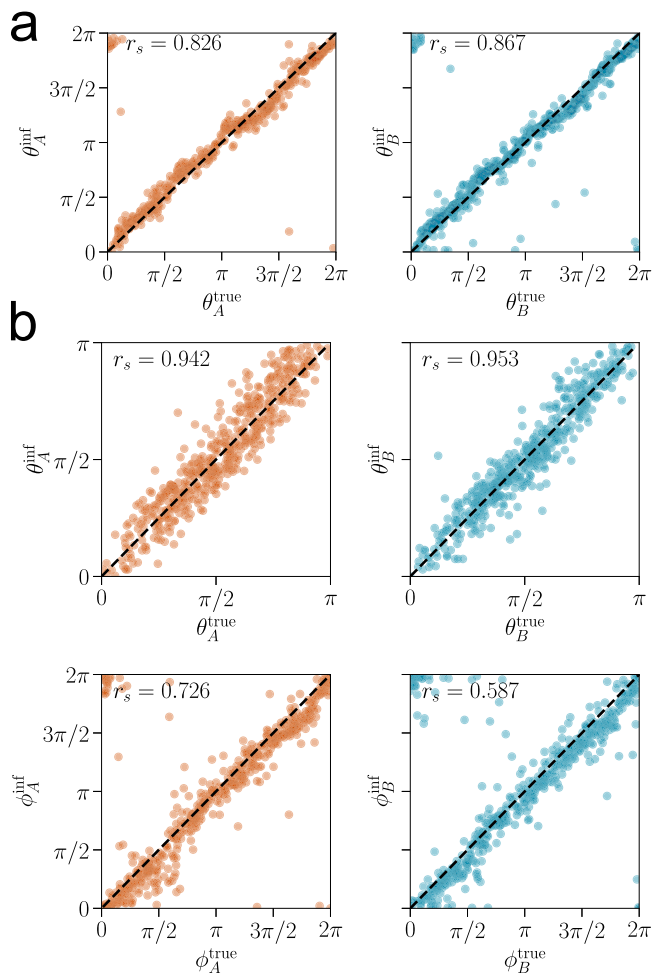


Fig. 2 | Validation of B-Mercator on synthetic bipartite networks. Relationship between the original and the inferred coordinates of the (a) bipartite- S^1 and (b) bipartite- S^2 models. In the top left corner of each figure, we report the value of the Spearman correlation coefficient between the inferred and original coordinates. Since the inferred coordinates might be rotated, we transform them to minimize the average angular distance between the original and inferred coordinates (Supplementary Section IV in ref. 32). With parameters: $N_A = 500$ (number of type-A nodes), $N_B = 1000$ (number of type-B nodes), $\gamma_A = 2.7$ (exponent of the powerlaw degree distribution of type-A nodes), $\gamma_B = 2.1$ (exponent of the powerlaw degree distribution of type-B nodes), $\langle k_A \rangle = 10$ (average degree of type-A nodes), $\beta_b = 1.5D$ (inverse temperature), with dimension $D = 1$ for (a) and $D = 2$ for (b).

when the source node is a type-A node (S_A), and the target node is a type-B node (T_B). The message is forwarded from S_A to the type-B node that is hyperbolically closest to the target. Since the node B_1 is not the destination, the message is forwarded again to a type-A node. The process is repeated until the destination is reached or the message becomes stuck. Then, the BGR protocol is executed for a large number of randomly chosen node pairs to assess the global network's geometric properties.

We tested the BGR protocol in synthetic networks generated from the bipartite- S^D model. First, we generated networks with specific dimensionality and topological properties, and we obtained their hyperbolic maps by embedding them using B-Mercator with different embedding dimensions. The performance of BGR was assessed based on two key measures: the proportion of messages that successfully reach their destination p_s , i.e., the success rate, and the mean stretch, where stretch is defined for each path connecting a source and target node as the ratio of the hop count of a successful greedy path to the shortest path. In Fig. 3b, we show the success rate as a function of the embedding dimension for networks generated using the bipartite- S^D model with dimensions ranging from $D = 1$ to $D = 4$. The

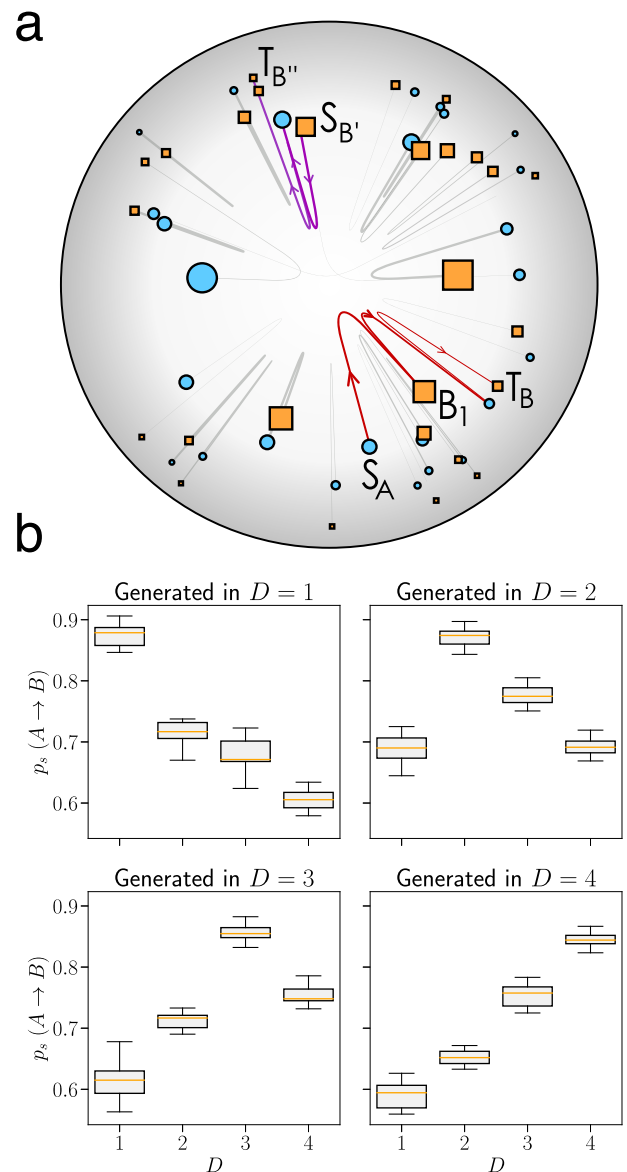


Fig. 3 | Bipartite greedy routing in synthetic networks. a Schematic view of the greedy routing protocol. We select a type-A node as an origin (S_A) and a type-B (T_B) as a destination. The red arrows show how the message is forwarded towards the destination. In the second example, we select two type-B nodes as source (S_B) and destination (T_B') and outline the greedy path with purple color. The line width is proportional to the connection probability (Eq. (1)). b Success rate as a function of embedding dimension for networks generated in $D = \{1, 2, 3, 4\}$. We consider here a navigation protocol where a source is a type-A node and a destination a type-B one. Results are obtained by averaging over 10 realizations with parameters: $N_A = 500$ (number of type-A nodes), $N_B = 500$ (number of type-B nodes), $\gamma_A = 2.5$ (exponent of the powerlaw degree distribution of type-A nodes), $\gamma_B = 2.5$ (exponent of the powerlaw degree distribution of type-B nodes), $\langle k_A \rangle = 10$ (average degree of type-A nodes), $\beta_b = 1.5D$ (inverse temperature). The box ranges from the first quartile to the third quartile. A horizontal line goes through the box at the median. The whiskers go from each quartile to the minimum or maximum.

performance of BGR is optimal (in terms of p_s and mean stretch, see Supplementary Fig. 14) when B-Mercator is used with the same dimension that was used to generate the network, thus justifying BGR as an alternative method to infer the effective dimension of real networks, different from the topological-based method introduced in ref. 33. These results are consistent across all variants of BGR (see Supplementary Fig. 13) and corroborate the results obtained for unipartite networks³².

Embedding of real bipartite networks

The significance of B-Mercator lies not only in its ability to embed synthetic networks generated by the bipartite- $S^D/\mathbb{H}^{D+1}$ model, but rather in its capacity to uncover geometric insights from real bipartite networks. Moreover, embeddings produced by B-Mercator can be applied to tasks such as node classification and link prediction on graph-structured data. As case studies, we analyze the Unicodelang dataset, which captures relationships between countries and the languages spoken within them, the human metabolic network connecting metabolites with the reactions in which they participate, and the flavor network linking ingredients to the chemical compounds they contain. These examples demonstrate the practical applicability of B-Mercator in extracting meaningful structural patterns from real-world bipartite networks.

Using B-Mercator, we embedded the Unicodelang dataset in various dimensions. The inferred embeddings are able to reproduce the topological properties of the network (see Supplementary Fig. 17). In Fig. 4, we show a dual embedding representation of this dataset in dimension $D = 1$. First, we focus on the countries in which a given language is used. For instance, in Fig. 4a, we plot all countries in which English is spoken, i.e., the neighbors of the English language in the bipartite network. One can notice that English is located closer to India or the Philippines than to the United States of America (USA). This can be explained by the fact that many different languages are spoken in the USA, which influence its position in the bipartite map. Indeed, in Fig. 4b, where we plot all neighbors of the Spanish language, the USA is located close to the Spanish language. Lastly, in Fig. 4c one can observe that countries from the French colonialism are concentrated in the similarity space and lie close to the French language, in contrast to France and other European countries. We can also shift the perspective: instead of examining the neighbors of each language, we can analyze the neighbors of a given country. In the bottom row of Fig. 4, we plot the language neighbors of India, China, and Brazil. Countries are often located in the embedding space close to the most widely spoken language.

Additional examples are provided in Supplementary Fig. 15, where we explore the neighbors of languages such as Hindi, Swahili, Catalan, Persian, Korean, Dutch, Russian, and Arabic. Similarly, in Supplementary Fig. 16, we depict the neighbors of countries including Canada, Turkey, Indonesia, Cameroon, Tanzania, Greece, the Philippines, and Bolivia.

The hyperbolic bipartite embedding enables us to examine the concentration of countries sharing the same language. We select the top 15 languages with the highest degree and compute the angular distance from each language to its neighboring countries. A small average angular distance may indicate that these countries are highly similar, whereas a broad distribution suggests a more international language. Figure 5a displays the *diversity* for these high-degree languages. As expected, the angular distance distribution for English is broad, indicating connections with countries distributed throughout the similarity space. Instead, the angular distances for Fula are relatively small. Fula, a Senegambian language spoken primarily in West and Central Africa, is concentrated in a specific region of our embedding space. Similarly, we can examine the linguistic diversity within each country. Following the previous approach, we selected the top 15 countries with the highest degree and computed the angular distances to each neighboring language. In a given country, if the languages are more similar, their angular positions in the similarity space should be more concentrated. Figure 5b shows that India and Canada exhibit a broader distribution of angular distances, reflecting the presence of a diverse array of languages. In contrast, Russia and Cameroon display a narrower distribution, indicating a more homogeneous set of languages. Notably, all these countries are linguistic hubs, with approximately 30 or more languages spoken in each. These results suggest that our embedding can serve as an indicator of a language's international reach and linguistic diversity, which is not solely reflected by its degree.

In addition, we investigated the human metabolic network, defined as metabolites connected to the reactions they participate in Ref.³⁴ and the network of food ingredients based on the flavor compounds they share³⁵ (see

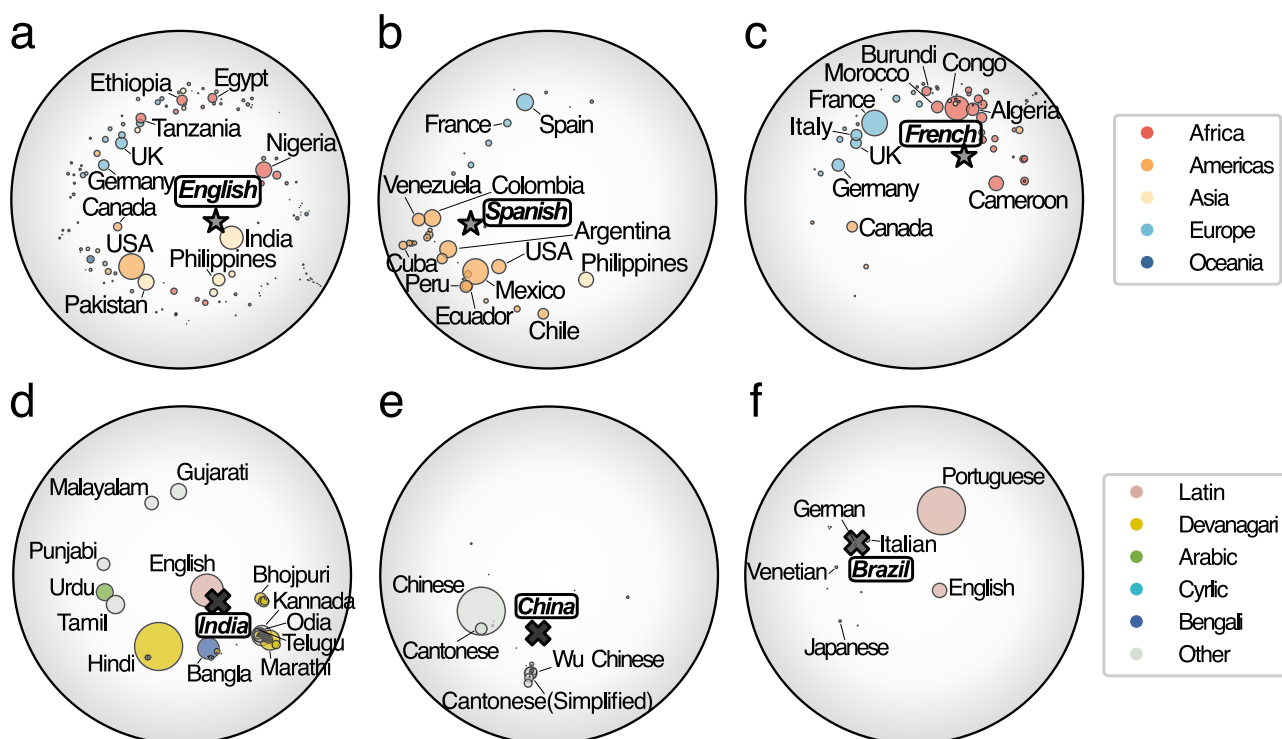


Fig. 4 | Visualization of the bipartite- S^1 embedding of the Unicodelang dataset per country or language. a–c Countries where a given language is spoken, i.e., the neighbors of the language node. The size of the nodes is proportional to the number of language speakers in that country. The color corresponds to the geographical region in which the country is located. A star marker indicates the position of a given

language. In (d–f), we depict all languages spoken in a given country, i.e., the neighbors of the country node. The size of the nodes is proportional to the fraction of speakers of a given language. The color represents that language's script. A cross marker indicates the position of a given country.

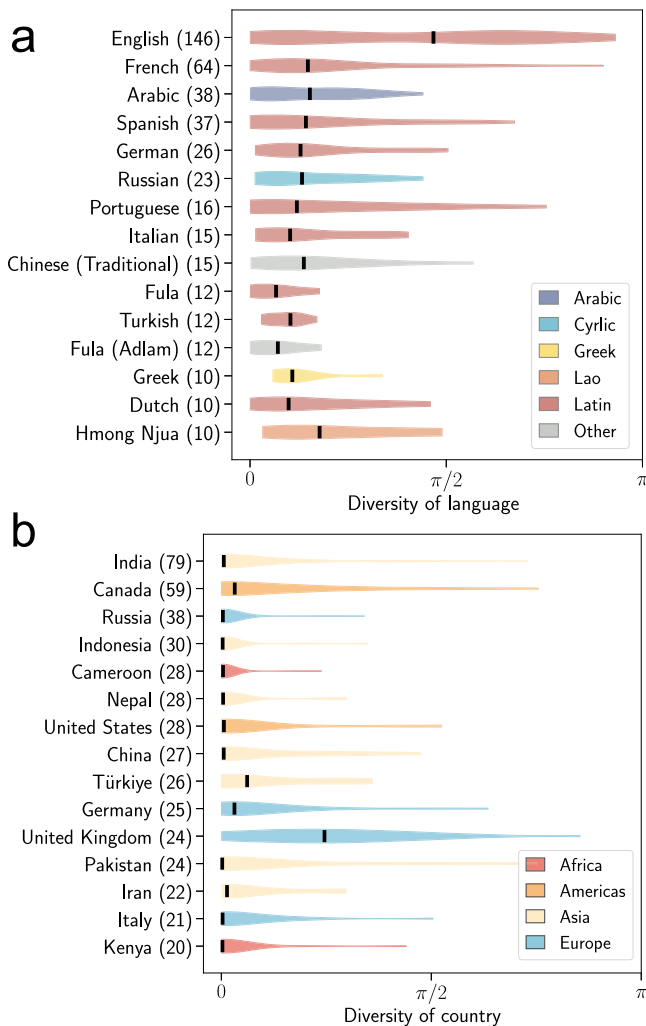


Fig. 5 | Language or country diversity for the top 15 highest-degree nodes. a Violin plots of angular distances between each language and its neighboring countries, with colors indicating script type. **b** Presents analogous plots for countries, with colors representing geographic region. In both panels, nodes are ordered in descending order of degree, with each node's degree (in brackets) indicated next to its label. Only the central 95% of the data is plotted—that is, data between the 2.5th and 97.5th quantiles are shown. A black line highlights the median value in each plot.

Section IIIC). In both cases, B-Mercator is able to reproduce topological properties of this bipartite network such as the degree distributions and clustering spectra. See Supplementary Note 6 for more details.

Finally, we applied the bipartite greedy routing protocol to the embeddings derived from these real-world networks. For the Unicodelang network, the highest success rate-based on the $A - A$ BGR variant-is observed at an embedding dimension of $D = 4$. However, variations in the success probability (p_s) across different embedding dimensions are minimal, likely due to the relatively low value of β_b . In contrast, the highest p_s for both the Metabolic and Flavor networks is achieved at $D = 1$. A summary of these findings is provided in Supplementary Tables S2–S5.

Case study on the graph machine learning tasks

Graph Machine Learning (Graph ML) focuses on extracting patterns, making predictions, and uncovering insights from graph-structured data^{37,38}. This data is typically defined as a set of entities (nodes) with complex relationships (links), defining a unipartite graph $\mathcal{G}_n$. Nodes are classified in different categories (labels) and are enriched with a set of features, defining a bipartite network of nodes and features $\mathcal{G}_{n,f}$. Within the Graph ML community, two key tasks are commonly used to evaluate and

rank network embedding models: node classification (NC) and link prediction (LP). In turn, network embeddings can be broadly categorized into supervised and unsupervised approaches. Supervised embeddings, such as those learned by graph neural networks (GNNs), use node labels in the training set to inform the learning process for classification tasks. In contrast, unsupervised methods leverage only the network structure and, optionally, node features to generate low-dimensional representations of the data. These maps can then be used for multiple downstream tasks by integrating additional classification models.

Our embedding method B-Mercator belongs to the class of unsupervised graph embeddings that leverage only the nodes' feature matrices. This is possible thanks to the findings in ref. 15, which show that nodes and their associated features define a bipartite network, $\mathcal{G}_{n,f}$, with strong geometric properties. We used B-Mercator to find an embedding of the nodes-features bipartite network in the common similarity space, which we subsequently used to perform NC and LP tasks in a supervised manner. To highlight the importance of such embeddings, we compared B-Mercator with D-Mercator³², which produces multidimensional hyperbolic maps of unipartite networks $\mathcal{G}_n$ without using information from the nodes' features. It has been shown that real complex networks can be accurately embedded in low dimensional hyperbolic spaces³⁹. Thus, in this work we selected embedding dimensions $D = 1$ and $D = 2$ to map the node features into the bipartite- $\mathbb{S}^1$ and bipartite- $\mathbb{S}^2$ models using B-Mercator, and to map the unipartite network into the $\mathbb{S}^1$ and $\mathbb{S}^2$ models using D-Mercator.

We also compared our model-driven approach with existing state-of-the-art graph embedding methods in both node classification and link prediction tasks. Prior work shows that efficient encoding can often be achievable in low-dimensional spaces⁴⁰. Therefore, we consider embedding dimensions between 2 and 16. For further details on the methods used, see Supplementary Note 7. We selected seven graph datasets commonly used in machine learning research, each with varying levels of correlations between the graph, nodes' features, and nodes' labels (see Supplementary Note 8 for details). It has been shown recently that the performance metrics of Graph ML tasks can vary significantly depending on these correlations³⁹. For instance, in the node classification tasks, adding features can be detrimental when the correlation between node features and network structure is very low. In contrast, adding features significantly enhances the results when the correlation is high. Among the analyzed networks, Cora and Citeseer exhibit a strong correlation between network structure and node features. In contrast, IMDB, Wisconsin, Texas, and Cornell show relatively low correlation, while Film demonstrates almost no correlation (see Supplementary Table S6 for more details).

To compute the accuracy of the node classification task, we applied a KNeighborsClassifier from the scikit-learn library⁴¹ to each network embedding with $K = 10$ as the number of nearest neighbors. We split the data into training and testing subsets with a 20/80% ratio. In the case of B-Mercator and D-Mercator, we computed distances among pairs of nodes using their angular separation on the D -sphere. For the rest of the methods, we computed the Euclidean distance between the nodes' positions from the corresponding embedding. In Fig. 6a, we report the performance of the NC task on the Wisconsin dataset. This is a network of web pages from the Computer Science department of the University of Wisconsin. Each web page is enriched with 1613 features and manually classified into one of five categories: student, project, course, staff, and faculty. For this dataset, feature-based methods demonstrate superior performance, with our methods (B-Mercator in $D = 1$ and $D = 2$) achieving a significant margin of improvement over competitors. This is due to the strong correlation between the nodes' features and labels, as shown in ref. 39. In contrast, network-based methods tend to perform less effectively, as they rely solely on the graph topology, which is weakly correlated with the nodes' labels³⁹.

For the distance-based link prediction task, we randomly selected a small fraction $q = 0.1$ of the existing links from the network structure as positive links. Similarly, an equal number of non-existing edges was selected as negative links to form a balanced test set. The remaining positive links were considered as the training network, which was then embedded using

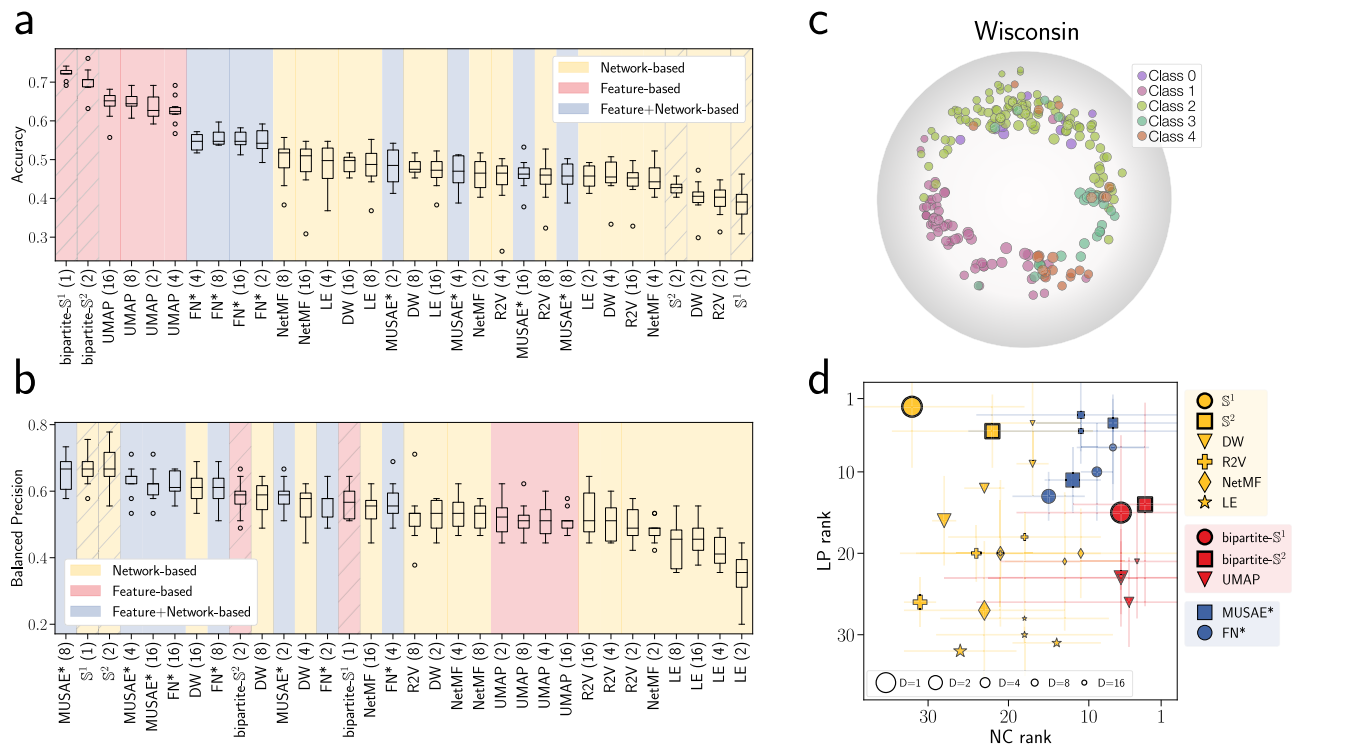


Fig. 6 | Case study on machine learning tasks. **a** Accuracy of the node classification (NC) task and **(b)** balanced precision of the distance-based link prediction (LP) task for the Wisconsin dataset. The train/test split is 20/80 for nodes in NC and 90/10 for links in LP, where the test set is balanced by randomly adding an equal number of negative links selected from non-existing links. The results are averaged over 10 different splits. Our methods are highlighted with diagonal hatches. The abbreviations of the algorithms are as follows: DW - DeepWalk, R2V - Role2Vec, LE - Laplacian Eigenmaps, FN - FeatherNode. With a star, we highlighted methods that use both network topology and node features for embedding. The numeric value in brackets indicates the embedding dimension. All other parameters are set to their default values. The methods are sorted by the median accuracy in NC and balanced precision in LP tasks. In addition, the methods are grouped and colored by input data

type, i.e., Network-based methods use only network topology, Feature-based methods use only nodes' feature matrix, and Feature + Network-based methods merge two things to construct the network embedding. **c** Visualization of the bipartite embedding with B-Mercator in $D=1$ for the Wisconsin dataset. We plot the positions of the nodes and color them based on the metadata. **d** Average rank in link prediction and node classification tasks across seven datasets. We plot the median rank value with the error bars depicting the interquartile range (IQR), i.e., the interval between the first and third quartiles that captures the middle 50% of the rank distribution. The size of the markers is inversely proportional to the embedding dimension and their shapes correspond to the input data type. Our methods are outlined with a thicker marker border.

various network-based methods. In the case of B-Mercator, the bipartite network between nodes and features was embedded, and it was not affected by splitting the links in the network structure into training and test sets. For B-Mercator and D-Mercator, we used the inverse of the hyperbolic distance between node pairs as the similarity measure, while for the other methods, we used the inverse of the Euclidean distance between node positions. As a result, links in the test set with the smallest distance were ranked highest when sorted in ascending order of their similarity scores.

In Fig. 6b, we show the precision of the different methods. In this case, methods using only the network topology, or a combination of network structure and node features, deliver the highest precisions. For this task, our embeddings with D-Mercator ($D=1$ and $D=2$), using only the network topology, are competitive and achieve a performance similar to MUSAE in dimension $D=8$, demonstrating the versatility of low-dimensional embeddings derived from our approach. Moreover, among feature-based methods, B-Mercator in $D=1$ and $D=2$ outperforms all other approaches. This highlights the adaptability of our methods to different tasks depending on the underlying data representation and task requirements. Detailed results for additional datasets can be found in Supplementary Figs. S25–S50.

Figure 6d provides a summary of the results across multiple datasets. For each dataset, the rank of each method is calculated for both NC and LP tasks, and the median rank is plotted along with the interquartile range (IQR). B-Mercator in low dimensions ($D=1$ and $D=2$) consistently ranks among the top approaches for the NC task and outperforms all feature-based methods for the LP task, for which D-Mercator is the best method in

dimension $D=1$. These results demonstrate the reliability and adaptability of our methods across datasets. The marker size in the plot, which inversely represents the embedding dimension, further underscores the efficiency of our low-dimensional embeddings compared to high-dimensional alternatives. In Supplementary Figs. S32 and S42, we provide a more detailed view of the rankings for the NC and LP tasks.

Discussion

The ability to embed real-world systems into a geometric space is a pivotal step toward understanding their intrinsic structure, function, and underlying organization. While numerous network embedding techniques have been proposed, most are not derived from a model-based perspective, limiting their interpretability and capacity to reconstruct empirical data. Model-based approaches offer a principled way of capturing the generative mechanisms that shape network topologies, thereby enabling researchers to interpret embeddings in a manner closely aligned with the data's underlying structure.

Several model-based embedding methods have already proved effective for unipartite networks, yet comparable solutions for bipartite networks have remained underexplored. In this work, we addressed this gap by introducing B-Mercator, a geometric model-based embedding algorithm specifically designed for bipartite networks. By mapping bipartite structures into hyperbolic space, B-Mercator offers a powerful and interpretable way to capture community structure, hierarchical organization, and topological relationships.

To demonstrate the versatility of B-Mercator, we applied it to embed real-world bipartite systems. The analysis of the language network (countries–spoken languages), the metabolic network (metabolites–reactions), and the Flavor network (ingredients–chemical compounds) show that the embeddings not only correlate well with metadata but also retain the essential characteristics of each dataset. Furthermore, we evaluated B-Mercator on several datasets on the node classification and link prediction tasks. It consistently outperformed all unsupervised methods for node classification and emerged as the best performer among feature-based embeddings for link prediction.

We stress an important advantage of our model-based embeddings. The hyperbolic maps can be used to generate synthetic networks with node features that closely resemble their real-world counterparts (see Supplementary Figs. 51–53). By doing so, we safeguard any sensitive information derived from real complex networks, such as personal connections, transactional data, or proprietary interactions. Thus, we enable the secure sharing of structural data without compromising the integrity of the original network or revealing sensitive information.

These findings underscore the value of geometric model-based embeddings for both theoretical analyses and practical applications, ranging from community detection and studying network hierarchies to advanced machine learning tasks. B-Mercator's robust performance highlights its capacity to reveal meaningful insights into bipartite systems – a domain often overlooked in the current embedding literature – while providing substantially more accurate analyses than those based on corresponding one-mode projections. Overall, B-Mercator represents a significant advancement in bipartite network analysis, paving the way for more accurate, interpretable, and generative representations of complex real-world systems.

Methods

Bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model

In the bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ model – an extension of the bipartite- $\mathbb{S}^1/\mathbb{H}^{2,15}$ – we assign to each node (of type-A or type-B) a hidden degree (κ_A or κ_B) and the position in the D -dimensional similarity space chosen uniformly at random, and represented as a point on a D -dimensional sphere. Each node of type-A and -B is assigned a vector $\mathbf{x}_i \in \mathbb{R}^{D+1}$ with $||\mathbf{x}_i|| = R$. For instance, when $D = 1$ the similarity space is represented as a circle, whereas for $D = 2$ it is a sphere (see Fig. 1).

The connection probability between type-A node u and type-B node v takes the form of gravity law:

$$p_{uv} = \frac{1}{1 + \chi_{uv}^{\beta_b}}, \quad \text{with } \chi_{uv} = \frac{R\Delta\theta_{uv}}{(\mu_b \kappa_u \kappa_v)^{1/D}}. \quad (1)$$

The number of type-A (type-B) nodes is $N_A(N_B)$, for convenience and without loss of generality, we set the density of type-A nodes in the D -sphere to one so that

$$R = \left[\frac{N_A}{2\pi^{\frac{D+1}{2}}} \Gamma\left(\frac{D+1}{2}\right) \right]^{\frac{1}{D}}. \quad (2)$$

The separation $\Delta\theta_{uv} = \arccos\left(\frac{\mathbf{x}_u \cdot \mathbf{x}_v}{R^2}\right)$ represents the angular distance between nodes u and v in the D -dimensional similarity space. The parameter β_b (with $\beta_b > D$) controls the coupling between the resulting topology and the underlying metric space. Lastly, the parameter μ_b controls the average degree of type-A nodes and is defined as

$$\mu_b = \frac{\beta_b \Gamma\left(\frac{D}{2}\right) \sin\frac{D\pi}{\beta_b}}{2\pi^{1+\frac{D}{2}} \langle \kappa_A \rangle}, \quad (3)$$

whereas the average degree of type-B nodes is set by $\langle \kappa_B \rangle = \frac{N_A}{N_B} \langle \kappa_A \rangle$. By choosing distributions for the hidden degrees κ_A and κ_B and inverse temperature β_b , we can generate bipartite networks with any desired degree distributions and varying levels of geometric properties.

The bipartite- $\mathbb{S}^D$ model can be represented in purely geometric terms as the bipartite- $\mathbb{H}^{D+1}$ model. This is achieved by mapping the hidden degrees of each type-A and -B nodes to radial coordinates while preserving their positions on the D -sphere. Specifically, the transformation for type-A nodes has the form (similarly for type-B nodes)

$$r_u = \hat{R} - \frac{2}{D} \ln \frac{\kappa_u}{\kappa_{u,0}}, \quad \text{with } \hat{R} = 2 \ln \left(\frac{2R}{(\mu_b \kappa_{u,0} \kappa_{v,0})^{1/D}} \right). \quad (4)$$

where $\kappa_{u,0}(\kappa_{v,0})$ is the smallest hidden degree for type-A (type-B) nodes. We can rewrite Eq. (1) as

$$p_{uv} = \frac{1}{1 + e^{\frac{\beta_b}{2}(x_{uv} - \hat{R})}}, \quad \text{with } x_{uv} = r_u + r_v + 2 \ln \frac{\Delta\theta_{uv}}{2}. \quad (5)$$

With this transformation, the space represented by the radial position of each node, along with its angular position on the sphere, becomes the native representation of the hyperbolic space of dimension $D + 1$. Consequently, the connection probability between nodes u and v becomes a function of x_{uv} , which is a good approximation of the hyperbolic distance between them⁴².

B-Mercator in detail

We adopt the code of D -Mercator³² for embedding bipartite networks. Here, we provide an overview of the differences between the embeddings for the unipartite and bipartite networks. See also Supplementary Note 12 for more technical details.

Inferring the hidden degrees and parameter β_b . The inference of hidden degrees for type-A and -B nodes, and the inverse temperature β_b is implemented as an iterative process. We begin with the initial guess for the parameter $\beta_b \in (D, 2D)$ where D is the embedding dimension, and initialize the hidden degrees as the observed degrees in the original network for type-A ($k_A = \{k_{A,i}\}_{i=1}^{N_A}$) and type-B ($k_B = \{k_{B,j}\}_{j=1}^{N_B}$) nodes. The aim of the estimation is to modify the hidden degrees in order to ensure that the expected degree of each node within the model aligns with the degree observed in the original network. After the hidden degrees for both nodes A and B are computed, the synthetic graph from the bipartite- $\mathbb{S}^D/\mathbb{H}^{D+1}$ is constructed and the bipartite clustering coefficient is calculated. If the computed bipartite clustering coefficient deviates from that of the original network, $\bar{c}_b$, the value of β_b is adjusted. Then, the process is repeated using the current estimation of hidden degrees until a pre-determined precision is reached.

Bipartite- $\mathbb{S}^D$ model corrected Laplacian Eigenmaps. Since the biadjacency matrix of a bipartite graph $\mathbf{A}$ is not symmetric, to apply the Laplacian Eigenmaps⁴³, we transform it to the adjacency matrix as

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{bmatrix} \quad (6)$$

Similarly to³², the expected angular distance between nodes u (of type-A) and v (of type-B) in the bipartite- $\mathbb{S}^D$ model, conditioned to the fact that they are connected, can be computed as

$$\langle \Delta\theta_{uv} \rangle = \int_0^\pi \Delta\theta_{uv} \rho(\Delta\theta_{uv} | a_{uv} = 1) d\Delta\theta_{uv} \quad (7)$$

Additionally, for $D = 1$, we keep the ordering inferred by LE and distribute type-A and type-B nodes evenly on the circle.

Likelihood maximization. The nodes' coordinates in the similarity space inferred using LE are adjusted by Maximum Likelihood Estimation (MLE) to optimize the probability that the bipartite- $\mathbb{S}^D$ model generates

the observed network. We define an order of nodes sorted by their degree for A and B type nodes separately. Fixing the positions of a subset of B type nodes, we find new optimal coordinates for a subset of type-A nodes. First, we compute the mean coordinates of type-A node u 's neighbors.

$$\mathbf{x}_u = \sum_v \frac{1}{\kappa_v^2} \mathbf{x}_v \quad (8)$$

where the sum goes over all neighbors of node u , i.e., type-B nodes. Later, the new positions around $\mathbf{x}_u$ are proposed using a multivariate normal distribution. Finally, we select the most likely candidate position based on the local log-likelihood

$$\ln \mathcal{L}_u = \sum_{v=0}^{N_B} a_{uv} \ln p_{uv} + (1 - a_{uv}) \ln (1 - p_{uv}) \quad (9)$$

After iterating over a subset of type-A nodes, we apply a similar approach to type-B nodes, and repeat the process until all node positions are adjusted.

Final adjustment of hidden degrees. Lastly, we adjust the hidden degrees to compensate deviations from $\bar{k}_A(\kappa_u) = \kappa_u$ and $\bar{k}_B(\kappa_v) = \kappa_v$, which might have been introduced during the estimation of the nodes' coordinates in the similarity space.

Real bipartite datasets

The Unicodelang dataset was downloaded from the Unicode CLDR Project GitHub repository⁹. The dataset⁴⁴ contains information about the languages spoken in a given territory. We preprocessed this dataset and matched the country codes to their geographical regions⁴⁵. In total, the bipartite graph contains 246 countries, 717 languages, and 1487 edges.

The metabolic network was extracted from the BiGG webpage³⁴. We focus on the RECON1 model, which corresponds to *Homo sapiens*. The data was preprocessed and cleaned, resulting in 1497 metabolites and 2212 chemical reactions.

The flavor network is a network of food ingredients based on the flavor compounds they share³⁵. After preprocessing, the total number of ingredients is 602, and the number of compounds is 1138. The bipartite graph contains 15,382 edges. Additional network properties and inferred values of β_b for different embedding dimensions are shown in Supplementary Table 1.

Data availability

The network datasets used in this study are available from the sources referenced in the manuscript and the Supplementary Information.

Code availability

The open-source code for B-Mercator, along with the code to reproduce the figures, is available on GitHub⁴⁶.

Received: 16 July 2025; Accepted: 12 December 2025;

Published online: 07 January 2026

References

- Newman, M. E. J. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci.* **98**, 404–409 (2001).
- Newman, M. E. J. Scientific collaboration networks. i. network construction and fundamental results. *Phys. Rev. E* **64**, 016131 (2001).
- Newman, M. E. J. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E* **64**, 016132 (2001).
- Newman, M. E. J. Coauthorship networks and patterns of scientific collaboration. *Proc. Natl. Acad. Sci.* **101**, 5200–5205 (2004).
- Koren, Y., Bell, R. & Volinsky, C. Matrix factorization techniques for recommender systems. *Computer* **42**, 30–37 (2009).
- Konstan, J. A. et al. Grouplens: Applying collaborative filtering to usenet news. *Commun. ACM* **40**, 77–87 (1997).
- Linden, G., Smith, B. & York, J. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet Comput.* **7**, 76–80 (2003).
- Davis, A., Gardner, B. B., and Gardner, M. R. Deep South: A social anthropological study of caste and class (Univ of South Carolina Press, 2009).
- Unicode Consortium, Unicode Common Locale Data Repository (CLDR), (2024), available at: <https://github.com/unicode-org/cldr> (Accessed: 2024-11-15).
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabási, A.-L. The large-scale organization of metabolic networks. *Nature* **407**, 651–654 (2000).
- Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barabási, A.-L. Hierarchical organization of modularity in metabolic networks. *Science* **297**, 1551–1555 (2002).
- Serrano, M. Ángeles, Boguñá, M. arián & Sagués, F. Uncovering the hidden geometry behind metabolic networks. *Mol. Biosyst.* **8**, 843 (2012).
- Bascompte, J., Jordano, P., Melián, C. J. & Olesen, J. M. The nested assembly of plant–animal mutualistic networks. *Proc. Natl. Acad. Sci.* **100**, 9383–9387 (2003).
- Olesen, J. M., Bascompte, J., Dupont, Y. L. & Jordano, P. The modularity of pollination networks. *Proc. Natl. Acad. Sci.* **104**, 19891–19896 (2007).
- Aliakbarisani, R., Serrano, M. Ángeles & Boguñá, M. arián Feature-enriched hyperbolic network geometry. *Phys. Rev. Res.* **7**, 033036 (2025).
- Zhou, T., Ren, J., Medo, M. atuš & Zhang, Y. i-C. heng Bipartite network projection and personal recommendation. *Phys. Rev. E—Stat., Nonlinear, Soft Matter Phys.* **76**, 046115 (2007).
- Neal, Z. The backbone of bipartite projections: Inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors. *Soc. Netw.* **39**, 84–97 (2014).
- Hoff, P. D., Raftery, A. E. & Handcock, M. S. Latent Space Approaches to Social Network Analysis. *J. Am. Stat. Assoc.* **97**, 1090–1098 (2002).
- Friel, N., Rastelli, R., Wyse, J. & Raftery, A. E. Interlocking directorates in irish companies using a latent space model for bipartite networks. *Proc. Natl. Acad. Sci.* **113**, 6629–6634 (2016).
- Filho, D. V. and O'Neale, D. R. J. Latent space generative model for bipartite networks. in *Proceedings of NetSci-X 2020: Sixth International Winter School and Conference on Network Science*, edited by Masuda, N., Goh, K.-I., Jia, T., Yamanoi, J., and Hiroki, J. Sayama (Springer International Publishing, Cham, 2020) pp. 3–16.
- Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A. & Boguñá, M. arián Hyperbolic geometry of complex networks. *Phys. Rev. E* **82**, 036106 (2010).
- Ortiz, E., García-Pérez, G. & Serrano, M. Ángeles Geometric detection of hierarchical backbones in real networks. *Phys. Rev. Res.* **2**, 033519 (2020).
- Boguñá, M. arián et al. Network geometry. *Nat. Rev. Phys.* **3**, 114–135 (2021).
- Serrano, M. Á, Krioukov, D. & Boguñá, M. Self-Similarity of Complex Networks and Hidden Metric Spaces. *Phys. Rev. Lett.* **100**, 078701 (2008).
- Allard, A., Serrano, M. Ángeles & Boguñá, M. arián Geometric description of clustering in directed networks. *Nat. Phys.* **20**, 150–156 (2024).
- Allard, A., Serrano, M. Ángeles, García-Pérez, G. & Boguñá, M. arián The geometric nature of weights in real complex networks. *Nat. Commun.* **8**, 14103 (2017).
- Kleineberg, K. aj-K. olja, Boguñá, M. arián, Serrano, M. Ángeles & Papadopoulos, F. Hidden geometric correlations in real multiplex networks. *Nat. Phys.* **12**, 1076–1081 (2016).

28. Kleineberg, K. aj-K. olja, Buzna, L., Papadopoulos, F., Boguñá, M. arián & Serrano, M. Ángeles Geometric Correlations Mitigate the Extreme Vulnerability of Multiplex Networks against Targeted Attacks. *Phys. Rev. Lett.* **118**, 218301 (2017).
29. Kitsak, M., Papadopoulos, F. & Krioukov, D. Latent geometry of bipartite networks. *Phys. Rev. E* **95**, 032309 (2017).
30. Opsahl, T. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Soc. Netw.* **35**, 159–167 (2013).
31. García-Pérez, G., Allard, A., Serrano, M. Á & Boguñá, M. Mercator: uncovering faithful hyperbolic embeddings of complex networks. *N. J. Phys.* **21**, 123033 (2019).
32. Jankowski, R., Allard, A., Boguñá, M. arián & Serrano, M. Ángeles The D-Mercator method for the multidimensional hyperbolic embedding of real networks. *Nat. Commun.* **14**, 7585 (2023).
33. Almagro, P., Boguñá, M. arián & Serrano, M. Ángeles Detecting the ultra low dimensionality of real networks. *Nat. Commun.* **13**, 6096 (2022).
34. King, Z. A. et al. Bigg models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic acids Res.* **44**, D515–D522 (2016).
35. Ahn, Y. ong-Y. eol, Ahnert, S. E., Bagrow, J. P. & Barabási, A. lbert-L. ászló Flavor network and the principles of food pairing. *Sci. Rep.* **1**, 196 (2011).
36. Boguñá, M. arián, Krioukov, D. & Claffy, K. C. Navigability of complex networks. *Nat. Phys.* **5**, 74–80 (2009).
37. Hu, W. et al. Open graph benchmark: Datasets for machine learning on graphs. *Adv. neural Inf. Process. Syst.* **33**, 22118–22133 (2020).
38. Xia, F. et al. Graph learning: A survey. *IEEE Trans. Artif. Intell.* **2**, 109–127 (2021).
39. Jankowski, R., Hozhabrierdi, P., Boguñá, M. arián & Serrano, M. Ángeles Feature-aware ultra-low dimensional reduction of real networks. *npj Complex.* **1**, 13 (2024).
40. Gu, W., Tandon, A., Ahn, Y. ong-Y. eol & Radicchi, F. Principled approach to the selection of the embedding dimension of networks. *Nat. Commun.* **12**, 1–10 (2021).
41. Pedregosa, F. et al. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
42. Serrano, M. Á. and Boguñá, M., *The Shortest Path to Network Geometry: A Practical Guide to Basic Models and Applications* (Cambridge University Press, 2022).
43. Belkin, M. & Niyogi, P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**, 1373–1396 (2003).
44. The Unicode Consortium, Unicode cldr project supplementalData.xml, <https://github.com/unicode-org/cldr/blob/main/common/supplemental/supplementalData.xml> (2025), [Accessed 28-11-2025].
45. Duncalfe, L., Iso-3166 country and dependent territories lists with un regional codes, <https://github.com/luke/ISO-3166-Countries-with-Regional-Codes> (2025), [Accessed 28-11-2025].
46. networkgeometry, b-mercator, <https://github.com/networkgeometry/b-mercator> (2025), [Accessed 28-11-2025].

Acknowledgements

We acknowledge support from: Grant TED2021-129791B-I00 funded by MCIN/AEI/10.13039/501100011033 and the “European Union NextGenerationEU/PRTR”; Grant PID2022-137505NB-C22 funded by MCIN/AEI/10.13039/501100011033 and by “ERDF/EU”. R. J. acknowledges support from the fellowship FI-SDUR funded by Generalitat de Catalunya. M. B. acknowledges the ICREA Academia award, funded by the Generalitat de Catalunya.

Author contributions

M. B., M. A. S., R. J., and R. A. designed research. R. J., and R. A. performed research. M. B., M. A. S., R. J. wrote the paper. All authors discussed results, reviewed the manuscript, and approved the final version.

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/s42005-025-02464-9>.

Correspondence and requests for materials should be addressed to Marián Boguñá.

Peer review information : *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2026