Network geometry inference using common neighbors

Fragkiskos Papadopoulos,[1,*] Rodrigo Aldecoa,[2] and Dmitri Krioukov[3]

[1]*Department of Electrical Engineering, Computer Engineering and Informatics, Cyprus University of Technology, Saripolou 33, Limassol 3036, Cyprus*

[2]*Northeastern University, Department of Physics, Boston, Massachusetts 02115, USA*

[3]*Northeastern University, Department of Physics, Department of Mathematics, Department of Electrical & Computer Engineering, Boston, Massachusetts 02115, USA*

(Received 24 February 2015; published 12 August 2015)

We introduce and explore a method for inferring hidden geometric coordinates of nodes in complex networks based on the number of common neighbors between the nodes. We compare this approach to the HyperMap method, which is based only on the connections (and disconnections) between the nodes, i.e., on the links that the nodes have (or do not have). We find that for high degree nodes, the common-neighbors approach yields a more accurate inference than the link-based method, unless heuristic periodic adjustments (or "correction steps") are used in the latter. The common-neighbors approach is computationally intensive, requiring $O(t^4)$ running time to map a network of t nodes, versus $O(t^3)$ in the link-based method. But we also develop a hybrid method with $O(t^3)$ running time, which combines the common-neighbors and link-based approaches, and we explore a heuristic that reduces its running time further to $O(t^2)$, without significant reduction in the mapping accuracy. We apply this method to the autonomous systems (ASs) Internet, and we reveal how soft communities of ASs evolve over time in the similarity space. We further demonstrate the method's predictive power by forecasting future links between ASs. Taken altogether, our results advance our understanding of how to efficiently and accurately map real networks to their latent geometric spaces, which is an important necessary step toward understanding the laws that govern the dynamics of nodes in these spaces, and the fine-grained dynamics of network connections.

DOI: 10.1103/PhysRevE.92.022807 PACS number(s): 89.75.Fb, 02.40.−k, 02.50.Tt

I. INTRODUCTION

The main premise of preferential attachment [1] is that popularity is attractive [2], but similarity is also attractive [3]. When combined, these two attractive forces, i.e., popularity and similarity, have been shown to form hidden hyperbolic geometries that drive the evolution of networks [4]. Since these geometries are hidden, effective, or latent, they must be inferred from the network structure. Specifically, what must be inferred are node coordinates in these underlying hyperbolic spaces. Existing approaches [5,6] to such an inference are based on the connections (and disconnections) between the nodes, i.e., on the links that the nodes have (or do not have). Connected nodes are attracted to each other, while disconnected nodes repel, and these approaches are placing nodes into a hyperbolic space based on these attraction and repulsion forces. Both approaches in [5,6] are based on *maximum likelihood estimation*. The approach in [6] embeds a given network topology into the hyperbolic plane by maximizing the likelihood that the topology is produced by the equilibrium hyperbolic network model [7], while the approach in [5] embeds the network by maximizing the likelihood that the topology is produced by the hyperbolic model of growing networks [4]. Both approaches produce similar results, even though there are fundamental differences between them. In this paper, we build on the latter approach [5], which is more recent and simpler to implement.

The work in [4] shows that tradeoffs between popularity and similarity shape the structure and dynamics of growing complex networks, and that these tradeoffs in network dynamics give rise to hyperbolic geometry. The growing network model in [4] is essentially a model of random geometric graphs growing in hyperbolic spaces. Synthetic graphs grown according to this simple model simultaneously exhibit many common structural and dynamical characteristics of real networks. We call the model in [4] the popularity × similarity optimization (PSO) model.

Given the ability of the PSO model to construct synthetic growing networks that resemble real networks across a wide range of structural and dynamical characteristics, the work in [5] showed how to reverse this synthesis, and given a real network, how to map (embed) the network into the hyperbolic plane in a way congruent with the PSO model. Specifically, the mapping method of [5], called *HyperMap*, replays the network's geometric growth, estimating at each time step the hyperbolic coordinates of new nodes by maximizing the likelihood of the network snapshot in the model. In the inferred polar coordinates of nodes, the radial coordinate r can be associated with node popularity, while the angular coordinate θ is the node coordinate in the similarity space abstracted by a circle. HyperMap has been applied to the autonomous systems (ASs) topology of the real Internet in [5], where it was shown that (i) the method can identify soft communities of ASs belonging to the same geographic region, even though the method is completely geography-agnostic; (ii) the method can predict *missing* links between ASs with high precision, outperforming popular existing methods; and (iii) the method can construct a highly navigable Internet map—greedy forwarding in the map can reach destinations with more than 90% success probability and low stretch.

Here we introduce and explore a method for inferring the node similarity coordinates, and we release its implementation to the public [8]. This method differs from the one in [5] in that it is not based on the links that the nodes have or

[*] Corresponding author: fragkiskosp@gmail.com

do not have. Instead, it is based on the *number of common neighbors* between the nodes. The method is inspired by the observation that the number of common neighbors between two nodes is a measure of similarity between the nodes; in general, the higher the number of common neighbors between two nodes is, the more similar the two nodes are, i.e., the smaller is their similarity distance [9,10]. We call the approach in [5] the *link-based approach*, and the approach considered here is the *common-neighbors approach*. We compare the two approaches and find that for high degree nodes, the common-neighbors approach yields a more accurate inference than the link-based method, unless heuristic periodic adjustments (or "correction steps" [5]) are used in the latter. On the other hand, the common-neighbors approach is computationally intensive, requiring $O(t^4)$ running time to map a network of t nodes, versus $O(t^3)$ in the link-based method.

Based on these observations, we introduce a hybrid method with $O(t^3)$ running time, which combines the common-neighbors and link-based approaches, and we explore a heuristic that can reduce its running time further to $O(t^2)$ without significantly sacrificing the embedding accuracy. We apply this method to snapshots of the real Internet to reveal how soft communities of ASs evolve over time in similarity space. We also demonstrate the method's predictive power by forecasting *future* links between ASs. Taken altogether, our results advance our understanding of how to efficiently and accurately map real networks to their latent hyperbolic spaces, which is an important necessary step toward understanding the laws that govern the dynamics of nodes in these spaces, and the fine-grained dynamics of network connections.

The rest of the paper is organized as follows. In Sec. II we review the extended PSO (E-PSO) model from [5] and the details of the HyperMap method that we need in this paper. In Sec. III we show how the angular coordinates of nodes can be inferred using the common-neighbors approach, and we describe the hybrid method. In Sec. IV we describe how to speed up the method, and in Sec. V we validate our results in synthetic networks. In Sec. VI we apply the hybrid method to the AS Internet. In Sec. VII we discuss other relevant work, and in Sec. VIII we conclude with a discussion of open problems and future work.

II. PRELIMINARIES

The E-PSO model of growing networks was introduced in [5] for HyperMap development purposes. As its name suggests, this model is a modification of the PSO model in [4]. The E-PSO model constructs growing networks using external links only, while being equivalent to the generalized PSO model in [4] that uses both external and internal links [5]. External links connect new nodes to existing nodes, while internal links appear between existing nodes only. Given a single snapshot of the topology of a real network, there is no way to distinguish external links from internal links. The E-PSO model sidesteps this obstacle, and it helps to map a given real network topology by replaying its geometric growth, treating all links in the topology *as if* they were external [5]. Below, we first review the E-PSO model and then proceed to HyperMap, which is based on this model. We limit the exposition only to the basic details that we need in the rest of the paper.

A. The E-PSO model

The E-PSO model has five input parameters: $m \geqslant 0$, $L \geqslant 0$, $\beta \in (0,1]$, $T \in [0,1)$, and $\zeta > 0$. Parameters m and L are the rates at which external and internal links appear. (We will explain shortly how we compute them in a real network.) These two parameters appear inside Eq. (4) below, and they define the average node degree in the network, $\bar{k} \approx 2(m+L)$. Parameter β defines the exponent $\gamma = 1 + 1/\beta \geqslant 2$ of the power-law degree distribution $P(k) \sim k^{-\gamma}$ in the network. Temperature T controls the average clustering $\bar{c}$ [11] in the network, which is maximized at $T = 0$ and decreases to zero nearly linearly with $T \in [0,1)$. Parameter $\zeta = \sqrt{-K}$, where K is the curvature of the hyperbolic plane. As will be explained, changing ζ rescales the node radial coordinates. This rescaling parameter does not affect any topological properties of networks generated by the model. Therefore, it can be set to any value in the model, e.g., $\zeta = 1$, without loss of generality. Having these parameters and the final size of the network $t > 0$ specified, the E-PSO model constructs a growing scale-free network up to t nodes according to the following E-PSO model definition:

(i) Initially the network is empty.

(ii) Coordinate assignment and update as follows:

(a) At time $i = 1,2,\ldots,t$, new node i is added to the hyperbolic plane at polar coordinates (r_i,θ_i), where radial coordinate $r_i = \frac{2}{\zeta}\ln i$, while the angular coordinate θ_i is sampled uniformly at random from $[0,2\pi]$.

(b) Each existing node $j = 1,2,\ldots,i-1$ moves, increasing its radial coordinate according to $r_j(i) = \beta r_j + (1-\beta) r_i$.

(iii) Creation of edges: node i connects to each existing node $j = 1,2,\ldots,i-1$ with probability $p_{ij} \equiv p(x_{ij})$ given by

$$p(x_{ij}) = \frac{1}{1 + e^{\frac{\zeta}{2T}(x_{ij}-R_i)}}. \tag{1}$$

In the last expression, x_{ij} is the hyperbolic distance between nodes i and j [12],

$$\begin{aligned} \cosh \zeta x_{ij} &= \cosh \zeta r_i \cosh \zeta r_j(i) \\ &\quad - \sinh \zeta r_i \sinh \zeta r_j(i) \cos \theta_{ij}, \\ \text{where} \quad \theta_{ij} &= \pi - |\pi - |\theta_i - \theta_j||, \end{aligned} \tag{2}$$

while R_i is given by

$$R_i = r_i - \frac{2}{\zeta}\ln\left[\frac{2T}{\sin T\pi}\,\frac{I_i}{\bar{m}_i(t)}\right], \tag{3}$$

with $I_i = \frac{1}{1-\beta}(1 - i^{-(1-\beta)})$. Equation (3) is derived from the condition that the expected number of old nodes $j < i$ that i connects to, denoted by $\bar{m}_i(t)$, is

$$\begin{aligned} \bar{m}_i(t) =& m + \frac{2L(1-\beta)}{(1-t^{-(1-\beta)})^2(2\beta-1)} \\ &\times \left[\left(\frac{t}{i}\right)^{2\beta-1} - 1\right][1 - i^{-(1-\beta)}]. \end{aligned} \tag{4}$$

The radial coordinate of a node abstracts its popularity. The smaller the radial coordinate of a node, the more popular the node is, and the more likely it attracts new connections. The angular distance between two nodes abstracts their similarity. The

```
1: Sort node degrees in decreasing order $k_1 > k_2 > \ldots > k_t$
   with ties broken arbitrarily.
2: Call node $i$, $i = 1, 2, \ldots, t$, the node with degree $k_i$.
3: Node $i = 1$ is born, assign to it initial radial coordinate
   $r_1 = 0$ and random angular coordinate $\theta_1 \in [0, 2\pi]$.
4: for $i = 2$ to $t$ do
5:     Node $i$ is born, assign to it initial radial coordinate
       $r_i = \frac{2}{\zeta} \ln i$.
6:     Increase the radial coordinate of every existing node
       $j < i$ according to $r_j(i) = \beta r_j + (1 - \beta) r_i$.
7:     Assign to node $i$ angular coordinate $\theta_i$ maximizing $\mathcal{L}_{\mathrm{L}}^i$
       given by Equation (5).
8: end for
```

FIG. 1. The HyperMap embedding algorithm.

smaller this distance, the more similar the two nodes are, and the more likely they are connected. The hyperbolic distance x_{ij} is then a single-metric representation of a combination of the two attractiveness attributes, namely radial popularity and angular similarity. The connection probability $p(x_{ij})$ is a decreasing function of x_{ij}, meaning that new connections take place by optimizing tradeoffs between popularity and similarity [4]. It has been shown [5] that the E-PSO model can reproduce not only the degree distribution and clustering of real networks such as the AS Internet, but also several other important properties. Given the ability of the model to construct growing synthetic networks that resemble real networks, Ref. [5] then showed how to reverse this synthesis, and given a real network, how to map (embed) the network into the hyperbolic plane in a way congruent with the E-PSO model. The mapping method, HyperMap, is described next.

B. HyperMap

HyperMap is based on maximum likelihood estimation (MLE) and is fully specified in Fig. 1. On its input it takes the network adjacency matrix α_{ij} ($\alpha_{ij} = \alpha_{ji} = 1$ if there is a link between nodes i and j, and $\alpha_{ij} = \alpha_{ji} = 0$ otherwise), and the network parameters m, L, γ, T, ζ. It then computes radial and angular coordinates $r_i(t), \theta_i$ for all nodes $i \leqslant t$ in the network.

HyperMap first estimates the MLE appearance (or birth) times of nodes $i = 1, 2, \ldots, t$. As shown in [5], the higher the degree of a node in the E-PSO model, the earlier is its MLE appearance time. Therefore, HyperMap uses the following procedure for finding the MLE of the node appearance times in a given network with t nodes. It sorts all nodes in decreasing order of their degrees $k_1 > k_2 > \cdots > k_t$, with ties broken arbitrarily, and sets their MLE appearance times $i = 1, 2, \ldots, t$ in the same order. That is, the node with the largest degree k_1 is expected to appear first, $i = 1$, the second largest degree node k_2 appeared second, $i = 2$, and so on. The node born at time i is called node i.

Having a sequence of MLE node birth times, HyperMap replays the hyperbolic growth of the network in accordance with the E-PSO model as follows. When a node is born at time $1 \leqslant i \leqslant t$, it is assigned an initial radial coordinate $r_i = \frac{2}{\zeta} \ln i$, and every existing node $j < i$ moves increasing its radial coordinate according to $r_j(i) = \beta r_j + (1 - \beta) r_i$. The method assigns to a new node $i > 1$ the angular coordinate θ_i that maximizes its local likelihood,

$$\mathcal{L}_{\mathrm{L}}^i = \prod_{1 \leqslant j < i} p(x_{ij})^{\alpha_{ij}} [1 - p(x_{ij})]^{1 - \alpha_{ij}}. \quad (5)$$

This likelihood is a function of θ_i, since x_{ij} depends on θ_i [see Eq. (2)], $p(x_{ij})$ depends on x_{ij} [see Eq. (1)], and $\mathcal{L}_{\mathrm{L}}^i$ depends on $p(x_{ij})$. The product in Eq. (5) goes over all the old nodes $j < i$. The likelihood $\mathcal{L}_{\mathrm{L}}^i$ is called *local* as it depends only on the connections (and disconnections) between new node i and existing nodes $j < i$. For example, if new node $i = 4$ is connected to nodes 1,2 but not to node 3, i.e., $\alpha_{41} = 1, \alpha_{42} = 1, \alpha_{43} = 0$, then $\mathcal{L}_{\mathrm{L}}^4$ would be $\mathcal{L}_{\mathrm{L}}^4 = p(x_{41}) p(x_{42}) [1 - p(x_{43})]$. We use the subscript "L" to emphasize that $\mathcal{L}_{\mathrm{L}}^i$ depends on the links between new node i and existing nodes $j < i$, i.e., it is a *link-based* approach. In the next section, we will derive an alternative local likelihood, $\mathcal{L}_{\mathrm{CN}}^i$, which depends on the number of common neighbors between new node i and existing nodes $j < i$.

The maximization of $\mathcal{L}_{\mathrm{L}}^i$ is performed numerically by sampling the likelihood $\mathcal{L}_{\mathrm{L}}^i$ at different values of θ in $[0, 2\pi]$ separated by intervals $\Delta\theta = \frac{1}{i}$ and then setting θ_i to the value of θ that yields the largest value of $\mathcal{L}_{\mathrm{L}}^i$. Since to compute $\mathcal{L}_{\mathrm{L}}^i$ for a given θ we need to compute the connection probability between node i and all existing nodes $j < i$, we need a total of $O(i^2)$ steps to perform the maximization. If there are t nodes in total, HyperMap needs $O(t^3)$ running time to map the full network.

Specifying input parameters. Parameter $\zeta > 0$ can be set to any value, e.g., $\zeta = 1$. As mentioned, changing the value of this parameter corresponds to radial coordinate rescaling. Specifically, the radial coordinates of nodes will be rescaled by the factor ζ, since, as can be seen by steps 5 and 6 in Fig. 1, at the final time $i = t$, $r_j(t) = \beta r_j + (1 - \beta) r_t = \frac{2\beta}{\zeta} \ln j + \frac{2(1-\beta)}{\zeta} \ln t, j \leqslant t$. Furthermore, the likelihood $\mathcal{L}_{\mathrm{L}}^i$ in Eq. (5) does not depend on ζ, as it cancels out in the connection probability $p(x_{ij})$ in Eq. (1). That is, different values of ζ will yield exactly the same angular coordinates. Parameter m can be obtained from historical data of the evolution of the network. If such data are available, then m is the average number of connections that nodes have once they first appear in the data. If no historical data are available, m can be set, as an approximation, to the minimum observed node degree in the network. Given the average node degree $\bar{k}$ in the network, and knowing m and $\bar{k}$, we get $L = \frac{\bar{k} - 2m}{2}$. The power-law exponent γ can be obtained from the degree distribution of the network, while parameter T is found experimentally [5]. We emphasize that the parameters for HyperMap come directly from the observation of the real network. With these five parameters (m, L, γ, T, ζ), and the network adjacency matrix α_{ij}, HyperMap infers $2t$ hyperbolic node coordinates in a network of t nodes (a radial and angular coordinate for each node), and consequently $O(t^2)$ hyperbolic distances between nodes.

III. INFERRING NODE SIMILARITY COORDINATES USING THE NUMBER OF COMMON NEIGHBORS

We now show how the angular (similarity) coordinates of nodes can be inferred using the number of common neighbors between new and old nodes instead of the connections and disconnections between them. Specifically, we first derive an alternative local likelihood, $\mathcal{L}_{\mathrm{CN}}^{i}$, which uses the observed number of common neighbors between each new node i and each existing node $j < i$ at final time t. Then, we use this likelihood in place of $\mathcal{L}_{\mathrm{L}}^{i}$ in Eq. (5) in order to infer the angular coordinate of each node.

In Sec. V, we show that for small i's, i.e., for nodes that appear at early MLE times, which are the high degree nodes, $\mathcal{L}_{\mathrm{CN}}^{i}$ yields a more accurate angular coordinate inference than $\mathcal{L}_{\mathrm{L}}^{i}$. This is because, for all node pairs i,j, $j < i$, $\mathcal{L}_{\mathrm{CN}}^{i}$ utilizes more information, since it uses the final number of common neighbors between the pairs. That is, it considers the *full* network adjacency matrix, i.e., the network adjacency matrix at the final time t, and it uses the number of common neighbors between the node pairs at that time. In contrast, $\mathcal{L}_{\mathrm{L}}^{i}$ in Eq. (5) uses less information, since at each time $i \leqslant t$ it considers only the connections and disconnections between node i and old nodes $j < i$. To derive $\mathcal{L}_{\mathrm{CN}}^{i}$, we first need to compute the distribution of the number of common neighbors between node pairs in the E-PSO model, which is the task we perform next.

A. Distribution of the number of common neighbors

Consider a network that has grown up to t nodes according to E-PSO (Sec. II A), where nodes are numbered according to the order in which they appear. Consider two nodes i,j with $j < i$ and a third node k. The initial radial coordinates of these nodes are $r_i = \frac{2}{\zeta}\ln i$, $r_j = \frac{2}{\zeta}\ln j$, and $r_k = \frac{2}{\zeta}\ln k$. We first need to find $p(i,j,\theta_i,\theta_j;k)$, which is the probability that i and j are both connected to k given their angular coordinates θ_i,θ_j. Below, we distinguish three cases and compute corresponding probabilities $p_1(i,j,\theta_i,\theta_j;k)$, $p_2(i,j,\theta_i,\theta_j;k)$, and $p_3(i,j,\theta_i,\theta_j;k)$.

Case 1: $i > j > k$. In this case, the connections to k happen when j and i first appear, i.e., at times j and i respectively. Therefore,

$$\begin{aligned}
p_1(i,j,\theta_i,\theta_j;k) &= \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{1+e^{\frac{\zeta}{2T}(x_{jk}-R_j)}}\frac{1}{1+e^{\frac{\zeta}{2T}(x_{ik}-R_i)}}d\theta_k,\\
\text{where}\quad x_{jk} &= \frac{1}{\zeta}\operatorname{arccosh}[\cosh\zeta r_j\cosh\zeta r_k(j)-\sinh\zeta r_j\sinh\zeta r_k(j)\cos\theta_{jk}],\\
x_{ik} &= \frac{1}{\zeta}\operatorname{arccosh}[\cosh\zeta r_i\cosh\zeta r_k(i)-\sinh\zeta r_i\sinh\zeta r_k(i)\cos\theta_{ik}],\\
R_j &= r_j-\frac{2}{\zeta}\ln\left[\frac{2T}{\sin T\pi}\frac{I_j}{\bar{m}_j(t)}\right],\quad R_i = r_i-\frac{2}{\zeta}\ln\left[\frac{2T}{\sin T\pi}\frac{I_i}{\bar{m}_i(t)}\right],\\
r_k(j) &= \beta r_k+(1-\beta)r_j,\quad r_k(i)=\beta r_k+(1-\beta)r_i.
\end{aligned}\tag{6}$$

Case 2: $i > k > j$. Here the connection between i and k happens when i first appears, i.e., at time i, and the connection between j and k happens when k first appears, i.e., at time k. Thus,

$$\begin{aligned}
p_2(i,j,\theta_i,\theta_j;k) &= \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{1+e^{\frac{\zeta}{2T}(x_{kj}-R_k)}}\frac{1}{1+e^{\frac{\zeta}{2T}(x_{ik}-R_i)}}d\theta_k,\\
\text{where}\quad x_{kj} &= \frac{1}{\zeta}\operatorname{arccosh}[\cosh\zeta r_k\cosh\zeta r_j(k)-\sinh\zeta r_k\sinh\zeta r_j(k)\cos\theta_{jk}],\\
x_{ik} &= \frac{1}{\zeta}\operatorname{arccosh}[\cosh\zeta r_i\cosh\zeta r_k(i)-\sinh\zeta r_i\sinh\zeta r_k(i)\cos\theta_{ik}],\\
R_k &= r_k-\frac{2}{\zeta}\ln\left[\frac{2T}{\sin T\pi}\frac{I_k}{\bar{m}_k(t)}\right],\quad R_i = r_i-\frac{2}{\zeta}\ln\left[\frac{2T}{\sin T\pi}\frac{I_i}{\bar{m}_i(t)}\right],\\
r_j(k) &= \beta r_j+(1-\beta)r_k,\quad r_k(i)=\beta r_k+(1-\beta)r_i.
\end{aligned}\tag{7}$$

Case 3: $k > i > j$. In this final case, both connections with k happen when k appears, i.e., at time k. Therefore,

$$\begin{aligned}
p_3(i,j,\theta_i,\theta_j;k) &= \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{1+e^{\frac{\zeta}{2T}(x_{kj}-R_k)}}\frac{1}{1+e^{\frac{\zeta}{2T}(x_{ki}-R_k)}}d\theta_k,\\
\text{where}\quad x_{kj} &= \frac{1}{\zeta}\operatorname{arccosh}[\cosh\zeta r_k\cosh\zeta r_j(k)-\sinh\zeta r_k\sinh\zeta r_j(k)\cos\theta_{jk}],\\
x_{ki} &= \frac{1}{\zeta}\operatorname{arccosh}[\cosh\zeta r_k\cosh\zeta r_i(k)-\sinh\zeta r_k\sinh\zeta r_i(k)\cos\theta_{ik}],\\
R_k &= r_k-\frac{2}{\zeta}\ln\left[\frac{2T}{\sin T\pi}\frac{I_k}{\bar{m}_k(t)}\right],\quad r_j(k)=\beta r_j+(1-\beta)r_k,\quad r_i(k)=\beta r_i+(1-\beta)r_k.
\end{aligned}\tag{8}$$

The integrals in Eqs. (6)–(8) can only be computed numerically. Since the connection events are statistically independent, the number of common neighbors between nodes i and j, $j < i$, given their angles θ_i,θ_j, is a sum of independent Bernoulli trials with different success probabilities, given by Eqs. (6)–(8). Therefore, by the central limit theorem [13], for sufficiently large network sizes t, the distribution of the number of common neighbors n_{ij} between i and j is approximately normally distributed, i.e., its probability density is approximately

$$f(n_{ij}|\theta_i,\theta_j) = \frac{1}{\sigma(i,j,\theta_i,\theta_j)\sqrt{2\pi}} e^{-\frac{[n_{ij}-\mu(i,j,\theta_i,\theta_j)]^2}{2\sigma^2(i,j,\theta_i,\theta_j)}}, \tag{9}$$

where its mean $\mu(i,j,\theta_i,\theta_j)$ and variance $\sigma^2(i,j,\theta_i,\theta_j)$ are

$$\mu(i,j,\theta_i,\theta_j) = \sum_{k=1}^{j-1} p_1(i,j,\theta_i,\theta_j;k) + \sum_{k=j+1}^{i-1} p_2(i,j,\theta_i,\theta_j;k) + \sum_{k=i+1}^{t} p_3(i,j,\theta_i,\theta_j;k), \tag{10}$$

$$\sigma^2(i,j,\theta_i,\theta_j) = \sum_{k=1}^{j-1} p_1(i,j,\theta_i,\theta_j;k)(1 - p_1(i,j,\theta_i,\theta_j;k)) + \sum_{k=j+1}^{i-1} p_2(i,j,\theta_i,\theta_j;k)(1 - p_2(i,j,\theta_i,\theta_j;k))$$
$$+ \sum_{k=i+1}^{t} p_3(i,j,\theta_i,\theta_j;k)(1 - p_3(i,j,\theta_i,\theta_j;k)). \tag{11}$$

To compute $\mu(i,j,\theta_i,\theta_j)$ and $\sigma(i,j,\theta_i,\theta_j)$ we use the fact that the mean of a Bernoulli random variable with success probability p is p, and its variance is $p(1-p)$. The computation of $\mu(i,j,\theta_i,\theta_j)$ and $\sigma(i,j,\theta_i,\theta_j)$ for each i,j pair requires $O(t)$ steps.

B. Likelihood and likelihood maximization

We are now ready to derive the likelihood $\mathcal{L}^i_{\rm CN}$ that we can use in place of $\mathcal{L}^i_L$ in Eq. (5) in order to infer the node angular coordinates. Consider new node $i \leqslant t$ in a network that grows according to E-PSO up to time t. We denote by $\mathcal{L}^i_1 \equiv \mathcal{L}(\theta_i|r_i,\{r_j(i),\theta_j\},\{n^t_{ij}\},m,L,\gamma,T,\zeta)_{j<i}$ the likelihood that i's angular coordinate takes value θ_i, given its r_i, the coordinates of the old nodes $\{r_j(i),\theta_j\} \equiv \{r_1(i),\theta_1,r_2(i),\theta_2,\ldots,r_{i-1}(i),\theta_{i-1}\}$, the number of common neighbors between i and each old node $j < i$ at the final time t, $\{n^t_{ij}\} \equiv \{n^t_{i1},n^t_{i2},\ldots,n^t_{ii-1}\}$, and the network parameters m,L,γ,T,ζ. Since the distribution of the angular coordinates is uniform on $[0,2\pi]$, we can rewrite $\mathcal{L}^i_1$ using Bayes' rule as

$$\mathcal{L}^i_1 = \frac{1}{2\pi}\frac{\mathcal{L}^i_2}{\mathcal{L}^i_3}, \tag{12}$$

where $\mathcal{L}^i_2 \equiv \mathcal{L}(\{n^t_{ij}\}|r_i,\theta_i,\{r_j(i),\theta_j\},m,L,\gamma,T,\zeta)_{j<i}$ is the likelihood to have the numbers of common neighbors $\{n^t_{ij}\}$, if the angular coordinate of node i has value θ_i, conditioned on its radial coordinate, the coordinates of the old nodes, and the network parameters. Likelihood $\mathcal{L}^i_3 \equiv \mathcal{L}(\{n^t_{ij}\}|r_i,\{r_j(i),\theta_j\},m,L,\gamma,T,\zeta)_{j<i}$, independent of θ_i, is the probability that i has the numbers of common neighbors with old nodes specified by $\{n^t_{ij}\}$, conditioned as shown by notation.

We are looking for the angle θ^*_i that maximizes the likelihood $\mathcal{L}^i_1$ in Eq. (12), or equivalently, $\mathcal{L}^i_2$. We can compute $\mathcal{L}^i_2$ using Eq. (9),

$$\mathcal{L}^i_2 = \prod_{1\leqslant j<i} f\left(n^t_{ij}|\theta_i,\theta_j\right) \equiv \mathcal{L}^i_{\rm CN}. \tag{13}$$

The product goes over all the old nodes $j < i$. Equation (13) gives the likelihood $\mathcal{L}^i_{\rm CN}$ that we can use in HyperMap in place of $\mathcal{L}^i_L$ in Eq. (5), where n^t_{ij} is the observed number of common neighbors between nodes appearing at MLE times i,j, computed from the given network adjacency matrix α_{ij}. Note that maximizing $\mathcal{L}^i_{\rm CN}$ is equivalent to maximizing its logarithm $\ln \mathcal{L}^i_{\rm CN}$,

$$\ln \mathcal{L}^i_{\rm CN} = C - \sum_{j=1}^{i-1} \ln \sigma(i,j,\theta_i,\theta_j)$$
$$- \sum_{j=1}^{i-1} \frac{\left[n^t_{ij} - \mu(i,j,\theta_i,\theta_j)\right]^2}{2\sigma^2(i,j,\theta_i,\theta_j)}, \tag{14}$$

where $C = (i-1)\ln\frac{1}{\sqrt{2\pi}}$, independent of θ_i.

C. $\mathcal{L}^i_{\rm CN}$ versus $\mathcal{L}^i_{\rm L}$, and the hybrid method

As with $\mathcal{L}^i_{\rm L}$, the maximization of $\mathcal{L}^i_{\rm CN}$ can only be performed numerically. This can be done in the same way as with $\mathcal{L}^i_{\rm L}$ (Sec. II B), i.e., by sampling the likelihood $\mathcal{L}^i_{\rm CN}$ at different values of θ in $[0,2\pi]$ separated by intervals $\Delta\theta = \frac{1}{i}$, and then setting θ^*_i to the value of θ that yields the largest value of $\mathcal{L}^i_{\rm CN}$. (To be more precise, we will be using sampling intervals $\Delta\theta = \min\{0.01,\frac{1}{i}\}$). Since, to compute $\mathcal{L}^i_{\rm CN}$ for a given θ we need to compute $\mu(i,j,\theta_i,\theta_j)$ and $\sigma(i,j,\theta_i,\theta_j)$ between node i and every existing node $j < i$, we need a total of $O(i^2t)$ steps to perform the maximization of $\mathcal{L}^i_{\rm CN}$. Therefore, if there are t nodes in total, HyperMap with $\mathcal{L}^i_{\rm CN}$ requires $O(t^4)$ running time to map the full network, versus $O(t^3)$ with $\mathcal{L}^i_{\rm L}$.

Likelihoods $\mathcal{L}^i_{\rm CN}$ and $\mathcal{L}^i_{\rm L}$ yield different results for the first few nodes appearing at early MLE times. Specifically, all nodes i for which their average number of connections to previous nodes in Eq. (4) is $\bar{m}_i(t) \geqslant i-1$ are expected to be connected to all previous nodes $j \leqslant i-1$ with a high probability. This condition holds for high degree nodes appearing at early MLE times, rendering their exact angular coordinate inference with $\mathcal{L}^i_{\rm L}$ infeasible. This is because $\mathcal{L}^i_{\rm L}$ uses the connections and disconnections between new and old nodes in order to place the nodes at the right angles; if new node i is connected to all previous nodes $j < i$ with high probability, then large zones of different angular coordinates are all quite likely with $\mathcal{L}_{\mathrm{L}}^{i}$. This effect was noted in [5]. In contrast, $\mathcal{L}_{\mathrm{CN}}^{i}$ can accurately infer the angular coordinates of nodes appearing early because it effectively utilizes "future" connectivity information as well, i.e., the number of common neighbors between the nodes at the *final* time t. This important difference between $\mathcal{L}_{\mathrm{CN}}^{i}$ and $\mathcal{L}_{\mathrm{L}}^{i}$ is illustrated in Sec. V. We note that since the inference of the angular coordinates of new nodes appearing at later MLE times depends on the inferred angles of high degree nodes appearing early, then if the latter are not accurately inferred, the former will not be accurately inferred either.

Given the angular coordinates of high degree nodes appearing at early MLE times, the inference of the angular coordinates of nodes appearing at later MLE times, e.g., of nodes i for which $\bar{m}_i(t) < i - 1$, using either $\mathcal{L}_{\mathrm{CN}}^{i}$ or $\mathcal{L}_{\mathrm{L}}^{i}$, yields similar results, i.e., the two likelihoods infer approximately the same angular coordinates for later nodes. This effect is also illustrated in Sec. V, and it means that one can use the following *hybrid approach*: use $\mathcal{L}_{\mathrm{CN}}^{i}$ for the first nodes i for which $\bar{m}_i(t) \geqslant i - 1$, and then use $\mathcal{L}_{\mathrm{L}}^{i}$ for the rest of the nodes for which $\bar{m}_i(t) < i - 1$. The benefit of this approach is running time, as the number of nodes for which $\bar{m}_i(t) \geqslant i - 1$ is usually quite small, e.g., on the order of a few tens of nodes. Therefore, HyperMap with this hybrid approach will still have $O(t^3)$ running time. In the next section, we describe a simple heuristic to reduce this running time to $O(t^2)$.

D. Hybrid method versus $\mathcal{L}_{\mathrm{L}}^{i}$ with correction steps

It was shown in [5] that the accuracy of HyperMap can be improved by occasionally running "correction steps" right after step 7 in Fig. 1. Specifically, at some predefined set of times i, we visit each existing node $j \leqslant i$, and having the coordinates of the rest of the nodes $l \leqslant i, l \neq j$, we update j's angle to the value θ_j' that maximizes

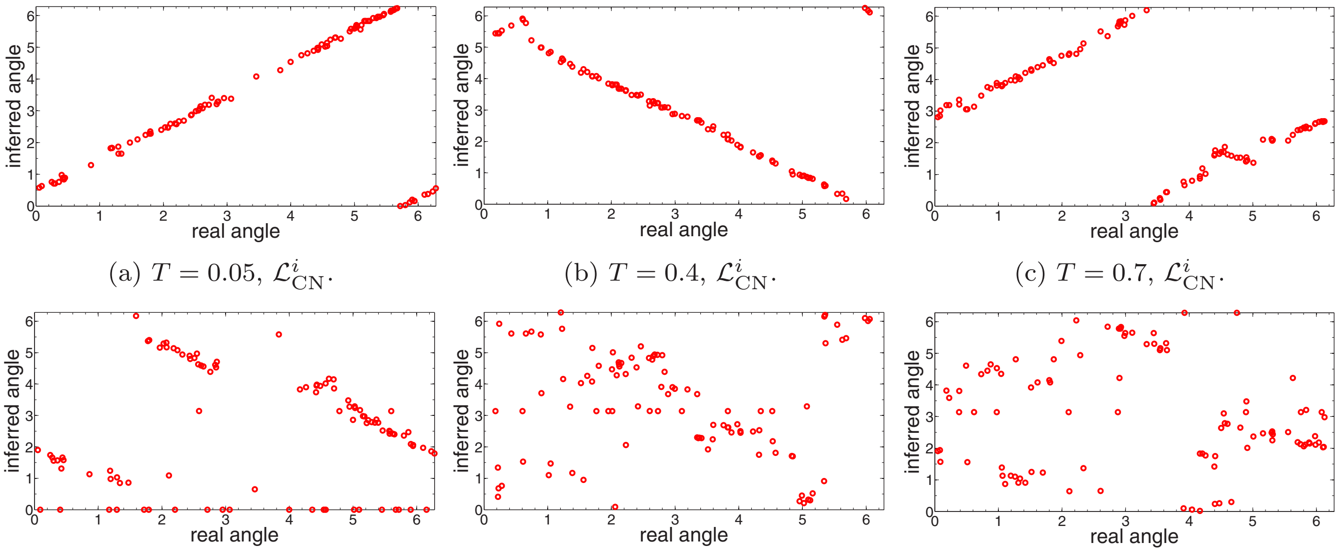
$$\widetilde{\mathcal{L}_{\mathrm{L}}^{j}} = \prod_{1 \leqslant l \leqslant i} p(x_{jl})^{\alpha_{jl}} [1 - p(x_{jl})]^{1-\alpha_{jl}}, \quad l \neq j, \tag{15}$$

where x_{jl} is the hyperbolic distance between j and l when the youngest of the two nodes appeared, and $p(x_{jl})$ is given by Eq. (1), using in it R_j if $j > l$ or R_l if $j < l$. It has been observed in [5] that these correction steps are beneficial when run at relatively small times i. This observation fully agrees with our results in this paper.

These correction steps are a heuristic that tries to effectively recompute improved angles for the first (high degree) nodes by considering not only the connections to their previous nodes, but also connections to nodes that appear later, i.e., future connectivity information, as in the common-neighbors approach. In Sec. V, we show that HyperMap with $\mathcal{L}_{\mathrm{L}}^{i}$ and correction steps yields similar results to the hybrid method that does not use correction steps.

IV. SPEEDING UP THE METHOD

As explained in Sec. III C, the running time of HyperMap with either the hybrid or link-based approaches is $O(t^3)$. Here we introduce a simple heuristic that reduces this running time to $O(t^2)$ without significantly sacrificing embedding accuracy, as we verify in the next section. We first observe that connected nodes are attracted to each other, and they are expected to be placed close to each other in the angular space [6]. This means that for each node i, we can get an *initial estimate* for its angular coordinate, θ_i^{init}, by considering only the previous nodes $j < i$ in $\mathcal{L}_{\mathrm{L}}^{i}$ [Eq. (5)] that are its neighbors. This requires only $O(k_i)$ steps, where k_i is i's degree, and $k_i = O(\bar{k})$ for sufficiently large i. That is, we can estimate θ_i^{init} by maximizing



(a) $T = 0.05$, $\mathcal{L}_{\mathrm{CN}}^{i}$.
(b) $T = 0.4$, $\mathcal{L}_{\mathrm{CN}}^{i}$.
(c) $T = 0.7$, $\mathcal{L}_{\mathrm{CN}}^{i}$.
(d) $T = 0.05$, $\mathcal{L}_{\mathrm{L}}^{i}$.
(e) $T = 0.4$, $\mathcal{L}_{\mathrm{L}}^{i}$.
(f) $T = 0.7$, $\mathcal{L}_{\mathrm{L}}^{i}$.

FIG. 2. (Color online) Inferred vs real angles (in radians) for synthetic networks with $t = 5000$ nodes and parameters $m = 1.5$, $L = 2.5$, $\gamma = 2.1$, and T as shown in the captions. The plots juxtapose the inferred against the real angles for the first 100 nodes, i.e., the nodes that appear at MLE times $1 \leqslant i \leqslant 100$. In (a)–(c) the common-neighbors method is used, while in (d)–(f) the link-based method is used.

the likelihood

$$\mathcal{L}^{i}_{\text{L-init}} = \prod_{1\leqslant j<i,\alpha_{ij}=1} p(x_{ij}), \tag{16}$$

where the product goes over all previous nodes $j < i$ that are i's neighbors. The maximization of Eq. (16) can be performed numerically by sampling the likelihood at intervals $\Delta\theta = \frac{1}{i}$ as before, yielding a total running time of $O(\bar{k}i) = O(i)$ to find $\theta_i^{\rm init}$.

Once we estimate $\theta_i^{\rm init}$, we can consider a region around it, $[\theta_i^{\rm init} - \frac{C}{i}, \theta_i^{\rm init} + \frac{C}{i}]$, where $0 < C \ll t$ is a constant, and we set the angular coordinate of node i, θ_i, to the value of θ that yields the largest value of $\mathcal{L}^i_{\rm L}$ [Eq. (5)] in this region. Since we sample the likelihood at intervals $\Delta\theta = \frac{1}{i}$, we need $O(C)$ steps to perform this maximization. Taken altogether, at sufficiently large times $i \gg C$ we need $O(iC) = O(i)$ steps to find θ_i. Therefore, if we have $t \gg C$ nodes in total, the total running time to find their angles following this procedure is $O(t^2)$. The larger the value of C, the better the results are expected to be in general, as we are searching for the optimal value of θ_i over a larger region, but the procedure will also be slower. We validate this speedup heuristic in the next section, where we set $C = 200$, and we show that it produces good results.

V. VALIDATION

In this section, we validate our mapping method and its variations. To do so, we first grow synthetic networks according to E-PSO up to $t = 5000$ nodes, with $m = 1.5$, $L = 2.5$, $\gamma = 2.1$, $\zeta = 1$, and $T = 0.05, 0.4, 0.7$. Similar results hold for other parameter values. Then, we pass these synthetic networks to HyperMap, using their corresponding m, L, γ, T, ζ values, and we compute radial and angular coordinates for the

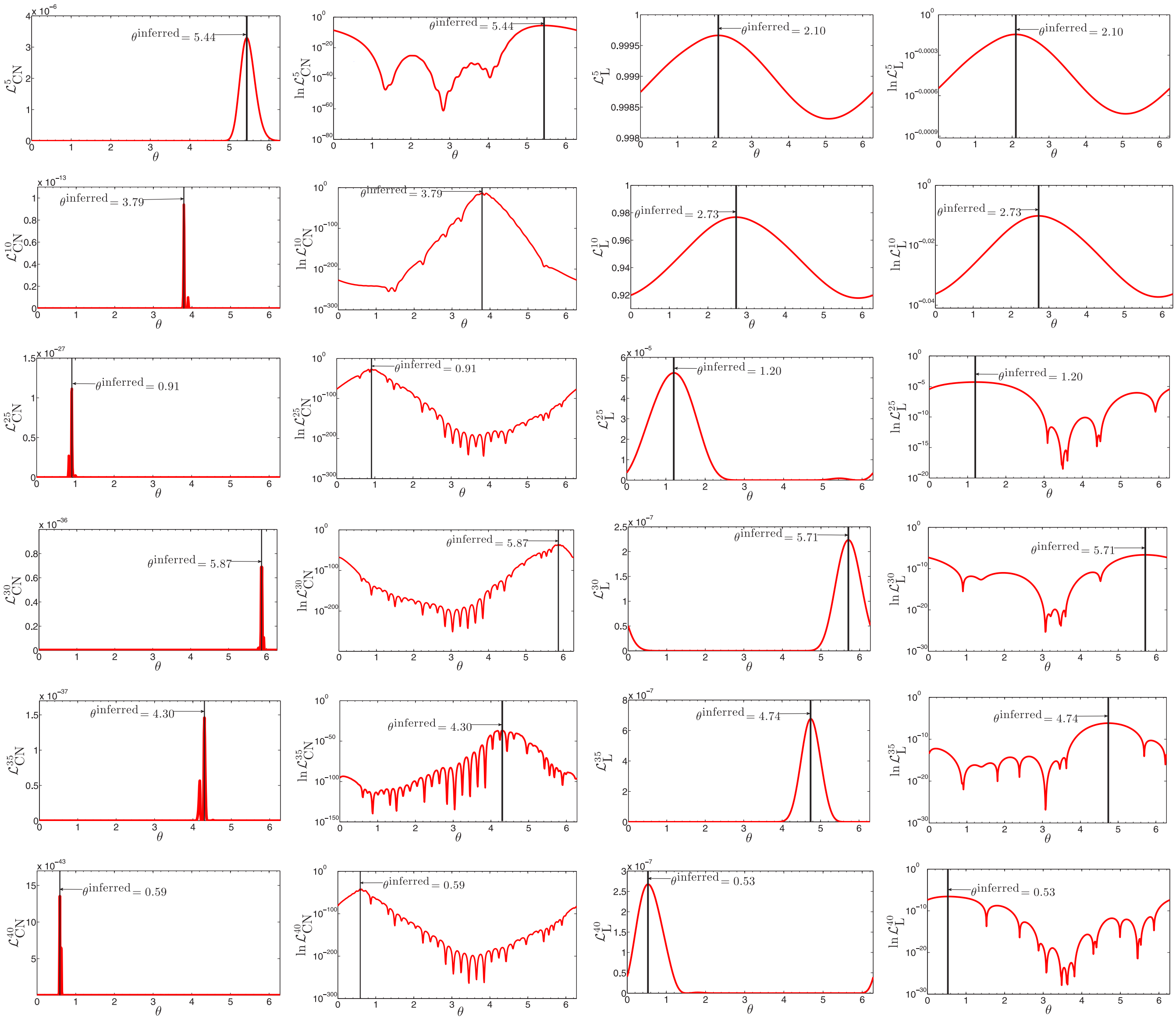
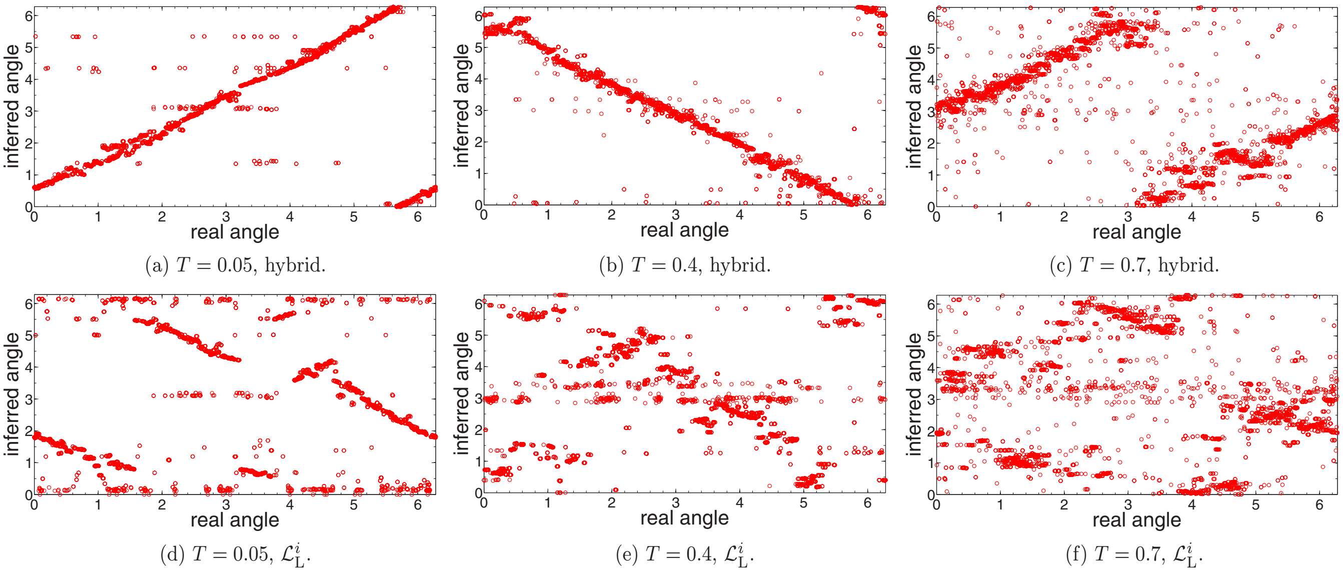


FIG. 3. (Color online) Likelihood landscapes for different nodes in a synthetic network with $t = 5000$ nodes and parameters $m = 1.5$, $L = 2.5$, $\gamma = 2.1$, and $T = 0.4$. The plots show the likelihoods $\mathcal{L}^i_{\rm CN}$, $\mathcal{L}^i_{\rm L}$ [Eqs. (13) and (5)] and the log-likelihoods $\ln\mathcal{L}^i_{\rm CN}$, $\ln\mathcal{L}^i_{\rm L}$, for nodes appearing at MLE times $i = 5$, 10, 25, 30, 35, and 40, as a function of the angular coordinate θ (in radians). The vertical line in each plot shows the inferred angle $\theta^{\rm inferred}$ (in radians), which always corresponds to the global maximum of the likelihood.



(a) $T = 0.05$, hybrid.

(b) $T = 0.4$, hybrid.

(c) $T = 0.7$, hybrid.

(d) $T = 0.05$, $\mathcal{L}^i_{\mathrm{L}}$.

(e) $T = 0.4$, $\mathcal{L}^i_{\mathrm{L}}$.

(f) $T = 0.7$, $\mathcal{L}^i_{\mathrm{L}}$.

FIG. 4. (Color online) Inferred vs real angles (in radians) for all the nodes in the synthetic networks of Fig. 2. In (a)–(c) the hybrid method is used, while in (d)–(f) the link-based method is used.

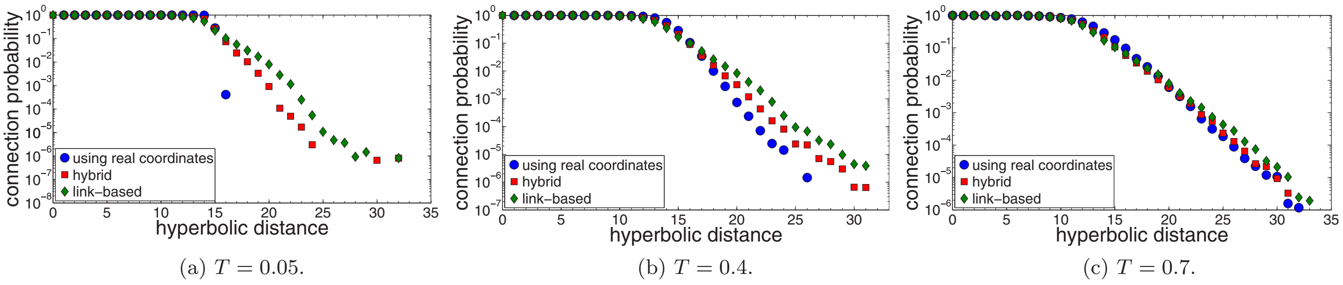
nodes, using either $\mathcal{L}^i_{\mathrm{CN}}$ (the common-neighbors method), $\mathcal{L}^i_{\mathrm{L}}$ (the link-based method), or the hybrid method. We consider the real Internet in the next section.

Inferred versus real angles for nodes appearing at early MLE times. Figure 2 juxtaposes the inferred angles against the real angles for the first 100 nodes, i.e., for the nodes that appear at MLE times $1 \leqslant i \leqslant 100$ for each considered network, when $\mathcal{L}^i_{\mathrm{CN}}$ or $\mathcal{L}^i_{\mathrm{L}}$ is used. We observe that the common-neighbors method is more accurate at inferring the angles of these first nodes. The reason for this was explained in Sec. III C. Specifically, we see in Figs. 2(a)–2(c) that $\mathcal{L}^i_{\mathrm{CN}}$ can infer the real angles of the nodes quite accurately, subject only to a *global* phase shift. This phase shift can take any value in $[0,2\pi]$, and it is due to the rotational symmetry of the model. The exact value of this shift is not important, and it depends on the initialization of the angle of the first node in HyperMap, which can be any random value in $[0,2\pi]$ (cf. step 3 in Fig. 1).

Likelihood landscapes. To gain a deeper understanding of the behavior of $\mathcal{L}^i_{\mathrm{CN}}$ and $\mathcal{L}^i_{\mathrm{L}}$, we show in Fig. 3 the corresponding likelihood landscapes for different nodes that appear at early MLE times, $i = 5,10,25,30,35,40$. To enable comparison between the two methods, the link-based likelihood $\mathcal{L}^i_{\mathrm{L}}$ is computed after fixing the angles of the old nodes $j < i$ to the angles inferred by the common-neighbors method. We observe that at small i, $i = 5,10$, $\mathcal{L}^i_{\mathrm{CN}}$ and $\mathcal{L}^i_{\mathrm{L}}$ behave quite differently, achieving their maximum at different values of θ. As discussed in Sec. III C, nodes appearing at early MLE times are connected to all previous nodes with high probability. Therefore, large zones of angular coordinates are nearly equally likely according to $\mathcal{L}^i_{\mathrm{L}}$, which is not the case with $\mathcal{L}^i_{\mathrm{CN}}$. This difference is evident in the first two rows of Fig. 3, showing the landscapes of $\mathcal{L}^5_{\mathrm{CN}}$, $\mathcal{L}^{10}_{\mathrm{CN}}$ and $\mathcal{L}^5_{\mathrm{L}}$, $\mathcal{L}^{10}_{\mathrm{L}}$. We also observe that $\mathcal{L}^i_{\mathrm{L}}$ of all possible angular coordinates is quite high for early nodes: $\mathcal{L}^5_{\mathrm{L}}$ of any angle is above 99%, and $\mathcal{L}^{10}_{\mathrm{L}}$ is above 92% for all angles.

At larger times i, $i \geqslant 25$, the two likelihoods achieve their maximum around the same angle, while their landscapes vary in a somewhat similar manner. This justifies the hybrid approach of Sec. III C, which uses $\mathcal{L}^i_{\mathrm{CN}}$ to infer the angles of the first i nodes for which $\bar{m}_i(t) \geqslant i - 1$, and then $\mathcal{L}^i_{\mathrm{L}}$ to infer the angles of the rest of the nodes. For the considered networks, relation $\bar{m}_i(t) \geqslant i - 1$ holds only for the first 33 nodes, while for the AS Internet snapshots in the next section it holds only for the first 36–40 nodes.

Inferred versus real angles for all the nodes. Figure 4 juxtaposes the inferred angles against the real angles for all nodes in each considered network, when the hybrid and link-based methods are used. We observe that (i) the hybrid method



(a) $T = 0.05$.

(b) $T = 0.4$.

(c) $T = 0.7$.

FIG. 5. (Color online) Connection probabilities with inferred (radial and angular) node coordinates obtained by the hybrid and link-based methods, and with real node coordinates. The results correspond to the mappings of Fig. 4.

TABLE I. Logarithmic losses in the mappings of Fig. 4.

Network	$\mathrm{LL}^{\mathrm{real}}$	$\mathrm{LL}^{\mathrm{inf}}$, hybrid	$\mathrm{LL}^{\mathrm{inf}}$, link-based	$\mathrm{LL}^{\mathrm{rand}}$	r_{LL}, hybrid	r_{LL}, link-based
$T = 0.05$	1.1×10^4	9.6×10^4	24.8×10^4	123×10^4	$e^{1134000}$	e^{982000}
$T = 0.4$	2.4×10^4	3.5×10^4	5.4×10^4	17×10^4	e^{135000}	e^{116000}
$T = 0.7$	4.1×10^4	4.4×10^4	5.2×10^4	11×10^4	e^{66000}	e^{58000}

is more accurate than the link-based method, as expected; (ii) Figs. 4(a)–4(c) are similar to Figs. 2(a)–2(c), meaning that as long as the angular coordinates of the first few nodes are accurately inferred, then the angular coordinates of the rest of the nodes will also be accurately inferred; (iii) the inference is in general better at lower temperatures T; and (iv) the inference is in general better for higher degree nodes appearing at early MLE times; cf. Figs. 2(a)–2(c) and 4(a)–4(c).

Connection probability. In Fig. 5 we report the connection probability, which is the probability that there is a link between a pair of nodes located at hyperbolic distance x, using real and inferred node coordinates. This probability is computed as the ratio of the number of connected node pairs to the total number of pairs of nodes located at distance x. From the figure, we observe that all inferred connection probabilities are close to the real ones, except for some discrepancies at their tails, which are more pronounced at lower T's. Furthermore, we see that the results with the hybrid method are only slightly better compared to the link-based method in terms of the connection probability. This suggests that the link-based method also produces relatively good mappings, even though it cannot infer as well the real angular coordinates.

We also quantify the quality of the obtained mappings using two other metrics: (i) the logarithmic loss, and (ii) the performance of greedy routing.

Logarithmic loss. The logarithmic loss is a quality metric for statistical inference defined as $\mathrm{LL} = -\ln \mathcal{L}$, where $\mathcal{L}$ in our case is the global likelihood

$$\mathcal{L} = \prod_{1\leqslant j<i\leqslant t} p[x_{ij}(t)]^{\alpha_{ij}}\{1 - p[x_{ij}(t)]\}^{1-\alpha_{ij}}. \qquad (17)$$

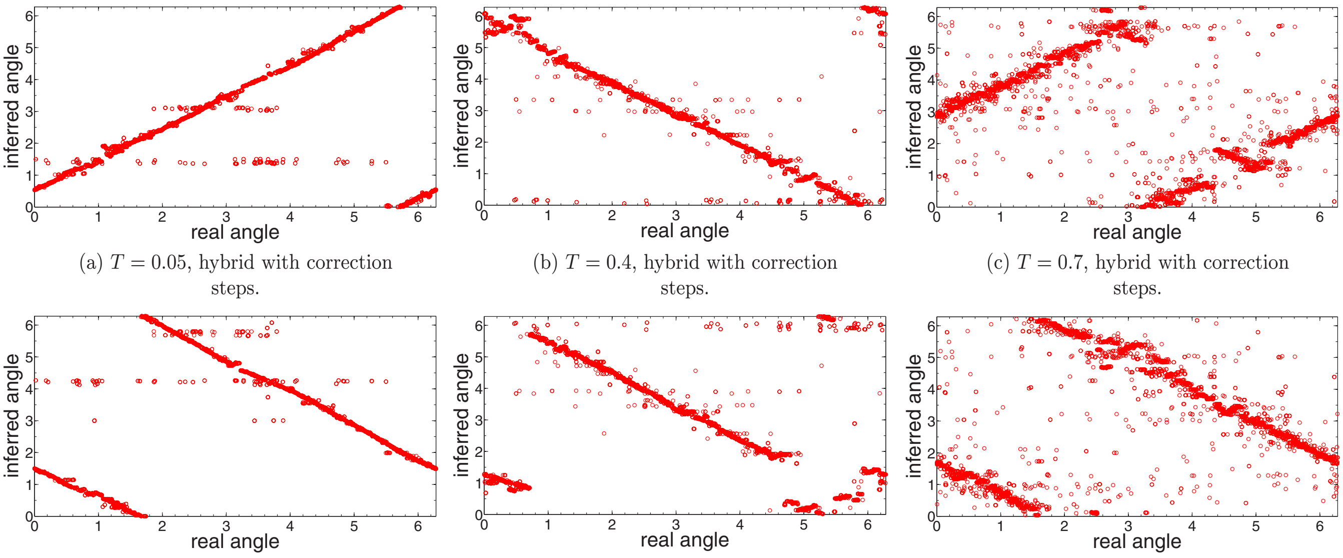
The product goes over all node pairs i,j in the network, $x_{ij}(t)$ is the hyperbolic distance between pair i,j, and $p[x_{ij}(t)] = 1/(1 + e^{\frac{\zeta}{2T}[x_{ij}(t)-R_t]})$ is the connection probability. We use LL to quantify the quality of the inference of the node angular coordinates. Specifically, we first compute LL using the inferred node coordinates $\{r_i(t),\theta_i\}$, and then we compare the result to the case in which LL is computed using the inferred $r_i(t)$'s and *random* θ_i's drawn uniformly from $[0,2\pi]$. We denote the former by $\mathrm{LL}^{\mathrm{inf}}$ and the latter by $\mathrm{LL}^{\mathrm{rand}}$. The smaller the $\mathrm{LL}^{\mathrm{inf}}$ compared to $\mathrm{LL}^{\mathrm{rand}}$, the better the quality of the mapping. In particular, the ratio $r_{\mathrm{LL}} = e^{-\mathrm{LL}^{\mathrm{inf}}}/e^{-\mathrm{LL}^{\mathrm{rand}}} = e^{(\mathrm{LL}^{\mathrm{rand}}-\mathrm{LL}^{\mathrm{inf}})}$ is the ratio of the likelihood with the inferred angular coordinates to the likelihood with random angular coordinates. The higher this ratio, the better the mapping quality. Table I reports the logarithmic losses $\mathrm{LL}^{\mathrm{inf}}$, $\mathrm{LL}^{\mathrm{rand}}$, the ratio r_{LL}, as well as $\mathrm{LL}^{\mathrm{real}}$, which is the logarithmic loss if we use the real radial and angular coordinates of nodes. We observe that (i) the hybrid method yields lower logarithmic losses compared to the link-based method, which is expected since it infers the node angular coordinates more accurately; and (ii) the logarithmic losses for both methods are significantly lower than those obtained with random angular coordinates, and closer to the logarithmic losses obtained with the real coordinates. These results suggest that the link-based method yields relatively good results, but the hybrid approach is better, as expected.

Performance of greedy routing. One specific class of network functions that are impossible without underlying geometry are efficient targeted transport processes without global knowledge of the network structure. Many real networks have this routing or navigation function in common; in some networks, including the Internet, this function is their primary function [14]. Therefore, navigability can be used as an alternative indirect metric of embedding quality. Navigability of an embedding is also of independent interest for some applications, such as Internet routing [6]. A network embedded in a geometric space is said to be *navigable* if *greedy routing (GR)* is efficient according to the metrics considered below. In GR, a node's address is its coordinates in the space, and each node knows only the addresses of its neighbors and the destination node address of a "packet." Upon receipt of such a packet, the GR node, if it is not a destination, forwards the packet to its neighbor closest to the destination in the geometric space, and it drops the packet if a local minimum loop is detected, i.e., if this neighbor is the same as the previous node visited by the packet.

We evaluate the efficiency of GR in the synthetic networks of Fig. 4 using both the HyperMap-inferred (hybrid, link-based) and the real node coordinates. We consider the following two GR efficiency metrics [14]: (i) the percentage of successful paths, p_s, which is the proportion of paths that do not get looped and reach their destinations; and (ii) the average hop-length $\bar{h}$ of the successful paths. The results are shown in Table II, where we see that (i) both the hybrid and link-based methods yield mappings where GR is quite efficient, yielding high p_s's and low path lengths $\bar{h}$, as is

TABLE II. Success ratio p_s and average hop length $\bar{h}$ of greedy paths in the mappings of Fig. 4.

Network	Using real coordinates	Using inferred coordinates, hybrid	Using inferred coordinates, link-based
$T = 0.05$	$p_s = 0.99, \bar{h} = 3.0$	$p_s = 0.96, \bar{h} = 3.1$	$p_s = 0.82, \bar{h} = 3.3$
$T = 0.4$	$p_s = 0.94, \bar{h} = 3.2$	$p_s = 0.95, \bar{h} = 3.4$	$p_s = 0.87, \bar{h} = 3.5$
$T = 0.7$	$p_s = 0.77, \bar{h} = 3.5$	$p_s = 0.92, \bar{h} = 3.8$	$p_s = 0.89, \bar{h} = 3.9$



(a) $T = 0.05$, hybrid with correction steps.

(b) $T = 0.4$, hybrid with correction steps.

(c) $T = 0.7$, hybrid with correction steps.

(d) $T = 0.05$, link-based with correction steps.

(e) $T = 0.4$, link-based with correction steps.

(f) $T = 0.7$, link-based with correction steps.

FIG. 6. (Color online) Inferred vs real angles (in radians) for all nodes in the networks of Fig. 4. In (a)–(c) the hybrid method is used, while in (d)–(f) the link-based method is used. In both methods, the correction steps are run as described in the text.

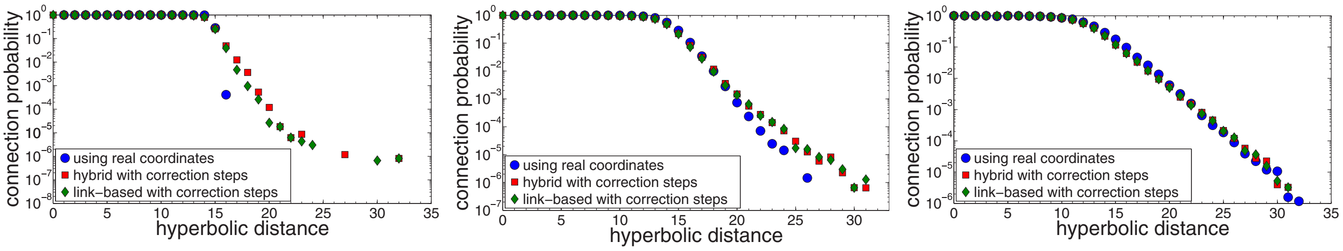
the case with the real node coordinates; and (ii) the hybrid method performs better, as expected, especially at lower T's.

Correction steps. We now repeat the same experiments applying the link-based and hybrid methods *with* correction steps in order to investigate the differences. Specifically, for each method we run four correction steps as described in Sec. III D, right after all nodes with degrees $k \geqslant 60$, 40, 20, and 10 appear in the network. Each of these correction steps is repeated eight times, which equals the average degree $\bar{k}$ in each network. In the hybrid method, the node angular coordinates that were inferred using the common-neighbors approach are not altered by the correction steps.

Figures 6(d)–6(f) show that the node angular coordinates are now inferred quite accurately with the link-based method. Figures 6(a)–6(c) show the results for the hybrid method, which look similar to Figs. 4(a)–4(c); this means that the effect of correction steps in this case is not as significant. All results are in agreement with Secs. III C and III D. In all cases, the inference is better at lower T's, as in Fig. 4.

The corresponding connection probabilities are shown in Fig. 7. Compared to Fig. 5, we observe that correction steps can help to better capture the connection probability tail, in both the hybrid and link-based methods. Finally, the logarithmic losses and the performance of GR are reported in Tables III and IV. In Table III, we observe that all logarithmic losses are smaller compared to those in Table I, and even closer to the logarithmic losses obtained with the real coordinates. This means that correction steps improve the quality of the obtained mappings in all cases. The improvement is quite significant for the link-based method, as expected, which at lower temperatures yields even lower logarithmic losses than the hybrid method. From Table IV, we see that the efficiency of GR is better compared to the results in Table II, especially for the link-based method. We also note from Tables IV and II that in some high-temperature cases, GR with inferred node coordinates performs even better than GR with real node coordinates. A possible explanation for this effect is given in Sec. VIII of [5].

Fast methods. Finally, we present results with the speedup heuristic described in Sec. IV, where we set constant $C = 200$ (Sec. IV). We consider the hybrid and link-based methods with correction steps as before, and we run the speedup heuristic for all nodes with degrees $k < k_{\text{speedup}} = 10$. We call these versions of the methods *fast versions*. Figure 8 shows likelihood landscapes sampled by the link-based method, and



(a) $T = 0.05$.

(b) $T = 0.4$.

(c) $T = 0.7$.

FIG. 7. (Color online) Connection probabilities with inferred (radial and angular) node coordinates obtained by the hybrid and link-based methods, and with real node coordinates. The results correspond to the mappings of Fig. 6.

TABLE III. Logarithmic losses in the mappings of Fig. 6.

Network	$\mathrm{LL}^{\mathrm{real}}$	$\mathrm{LL}^{\mathrm{inf}}$, hybrid	$\mathrm{LL}^{\mathrm{inf}}$, link-based	$\mathrm{LL}^{\mathrm{rand}}$	r_{LL}, hybrid	r_{LL}, link-based
$T = 0.05$	1.1×10^4	5.5×10^4	3.9×10^4	123×10^4	$e^{1175000}$	$e^{1191000}$
$T = 0.4$	2.4×10^4	2.9×10^4	2.8×10^4	17×10^4	e^{141000}	e^{142000}
$T = 0.7$	4.1×10^4	4.1×10^4	4.1×10^4	11×10^4	e^{69000}	e^{69000}

the corresponding regions of the likelihoods sampled by its fast version. We observe that the fast version infers the same angle as the original version, which always corresponds to the maximum of the likelihood. We also observe that the initial estimate of the angle is very close to the final inferred angle, as expected. Figure 9 juxtaposes the inferred angles with the original and fast version of the hybrid method for all the network nodes. Similar results hold for the link-based method. From the figure, we observe a very good match for almost all the node angles, especially at lower temperatures. Tables V and VI show the logarithmic losses and the performance of GR, where the results are very similar to those in Tables III and IV.

Summary of the results. To summarize, in this section we have validated that (i) the common-neighbors method is more accurate than the link-based method for nodes appearing at early MLE times; (ii) at larger MLE times, the two methods yield approximately the same results; (iii) the hybrid method performs significantly better from the link-based method if correction steps are not used; (iv) if correction steps are used, then hybrid and link-based methods perform similarly; (v) correction steps can help to improve the quality of the obtained mappings in all cases, but their effect on the hybrid method is not as significant as in the link-based method; and (vi) the fast and original versions of the methods perform almost the same. Our results indicate that the best options are the fast versions of either the hybrid or link-based methods with correction steps. However, we note that the correction steps are an ad hoc and computationally intensive heuristic, requiring $O(i^3)$ computations if run at time i. We have observed that these steps are beneficial when run at relatively small times i, not exceeding a few hundred nodes [5]. But being a heuristic, there are no universal guidelines of when exactly they should be invoked on a given real network to be embedded with the best results. Since correction steps do not have a significant effect on the hybrid approach, *the fast hybrid method without correction steps* might be the best option in general in terms of accuracy and computational complexity tradeoffs.

VI. APPLICATION TO THE INTERNET

We now consider the autonomous systems (ASs) Internet topology extracted from the data collected by the Archipelago active measurement infrastructure (ARK) developed by CAIDA [15], which is available at [16]. The connections in the topology are not physical but logical, representing AS relationships [16]. Specifically, an AS is a part of the Internet infrastructure administrated by a single company or organization. Pairs of ASs peer to exchange traffic. These peering relationships in the AS graph are represented as links between AS nodes. CAIDA's IPv4 Routed /24 AS Links Dataset [16] provides regular snapshots of AS links derived from ongoing traceroute-based IP-level topology measurements. A detailed description of the measurement process is given in [16]. The AS topology has a stable power-law degree distribution with exponent $\gamma = 2.1$, average node (AS) degree $\bar{k} \approx 5$, and average clustering $\bar{c} \approx 0.6$. We consider six snapshots of the topology spaced by three-month intervals from September 2009 to December 2010. These snapshots consist, respectively, of $t = 24\,091$, 25 910, 26 307, 26 756, 28 353, and 29 333 ASs.

Logarithmic loss and greedy routing efficiency. In Fig. 10 we mapped the Sept. 2009 snapshot using the fast hybrid and link-based methods with and without correction steps. The correction steps were applied as described in the previous section. In all cases, we used the estimated $m = 1.5$, $L = \frac{\bar{k}-2m}{2} = 1$, $\gamma = 2.1$, $\zeta = 1$, and different values of T in [0.1,0.9]. The speedup heuristic was applied for all nodes with degrees $k < k_{\mathrm{speedup}} = 3$. Figures 10(a) and 10(b) show the obtained logarithmic losses and the efficiency of greedy routing (GR) in all cases. We observe that correction steps do not have a significant effect on the hybrid method, whose lowest logarithmic loss is obtained at $T = 0.6$. This value is close to the value $T = 0.45$–0.5 required to construct synthetic networks with the same clustering $\bar{c}$ as in the Internet [5]. The link-based method without correction steps yields significantly higher logarithmic losses than the hybrid method, for almost all temperature values T. These losses decrease when correction steps are used and become similar to the ones in the hybrid method. These results agree with our observations in the previous section on synthetic networks, which indicated that the link-based method without correction steps is not as accurate at inferring the angular coordinates of nodes, while correction steps are not as important for the hybrid method; cf. Figs. 4 and 6 and Tables I and III.

GR is also very efficient. In the hybrid method, with or without correction steps, the success ratios are close to 90% for a wide range of T in [0.3,0.6]. In the link-based method

TABLE IV. Success ratio p_s and average hop length $\bar{h}$ of greedy paths in the mappings of Fig. 6.

Network	Using real coordinates	Using inferred coordinates, hybrid	Using inferred coordinates, link-based
$T = 0.05$	$p_s = 0.99, \bar{h} = 3.0$	$p_s = 0.97, \bar{h} = 3.1$	$p_s = 0.98, \bar{h} = 3.1$
$T = 0.4$	$p_s = 0.94, \bar{h} = 3.2$	$p_s = 0.96, \bar{h} = 3.3$	$p_s = 0.97, \bar{h} = 3.3$
$T = 0.7$	$p_s = 0.77, \bar{h} = 3.5$	$p_s = 0.93, \bar{h} = 3.7$	$p_s = 0.93, \bar{h} = 3.7$

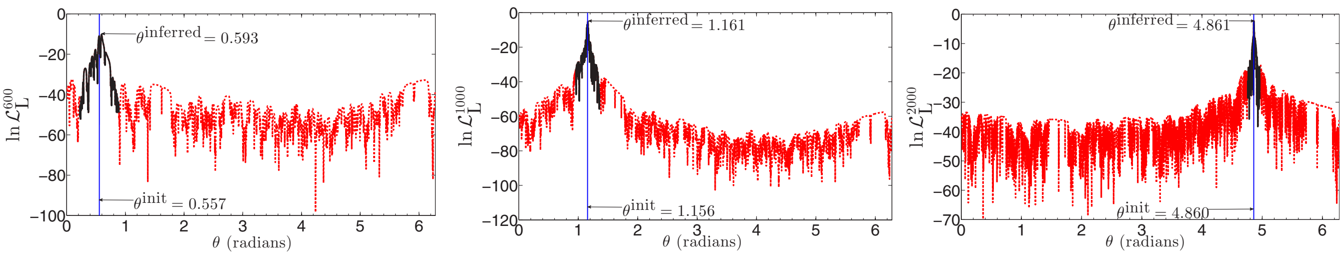


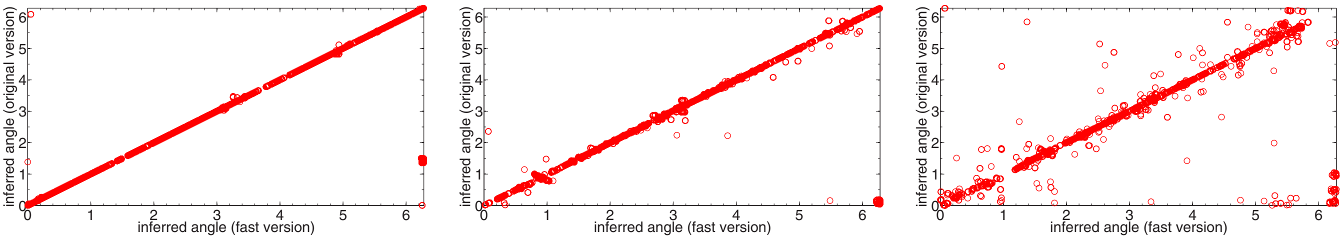
FIG. 8. (Color online) Likelihood landscapes for different nodes in a synthetic network with $t = 5000$ nodes and parameters $m = 1.5$, $L = 2.5$, $\gamma = 2.1$, and $T = 0.4$. The plots show the log-likelihoods $\ln \mathcal{L}_{\text{L}}^{i}$, $i = 600$, 1000, and 2000, with the original version of the method that samples the likelihood over the whole $[0,2\pi]$ domain (dashed red line), and with its fast version that samples the likelihood only over the θ region shown by the solid black line. The vertical line in each plot shows the initial estimate for the angle, θ^{init}, while θ^{inferred} is the final inferred angle.

without correction steps, the success ratios are smaller, and they become similar to the ratios of the hybrid method only if correction steps are used. These results agree again with our previous observations on synthetic networks; cf. Tables II and IV.

Prediction of future links. Figure 10(c) shows the empirical probability that a future link appears between two disconnected ASs as a function of their hyperbolic distance in Sept. 2009. To compute this probability, we consider all disconnected AS pairs in Sept. 2009 and all *future links* that appear between these pairs in the period Sept. 2009–Dec. 2010 (48 119 new links). We then bin the range of hyperbolic distances between these pairs from zero to the maximum distance into small bins. For each bin, we find all the disconnected pairs located at the hyperbolic distances falling within the bin. The percentage of pairs in this set of pairs that get connected with a future link is the value of the empirical future-link probability at the bin. From Fig. 10(c), we observe that this probability decreases with the hyperbolic distance between disconnected ASs, as expected. Furthermore, this decrease is exponential at large distances. We note that the shape of this probability is similar to the connection probability in our model, cf. Fig. 7, but it has a slope that does not depend on T; in fact, different values of $T \leqslant 0.7$ yield very similar results.

To provide a deeper insight into the ability of the fast hybrid and link-based methods to predict future links, we also compute the area under the receiver operating characteristic curve (AUC) [17]. The AUC here is defined as the probability that a randomly selected link from the set of our future links is given a better score (i.e., a higher existence likelihood) than a randomly selected nonexistent link, where the "nonexistent links" are the disconnected AS pairs in Sept. 2009 that never get connected in Sept. 2009–Dec. 2010. The score s_{ij} between two disconnected ASs (i,j) is the hyperbolic distance x_{ij} between them. The smaller this score, i.e., the smaller the hyperbolic distance between two disconnected ASs, the more likely it is that these two ASs will get connected; cf. Fig. 10(c). The degree to which the AUC exceeds 0.5 indicates how much better the method performs than pure chance, while AUC = 1 is the best possible AUC.

The results are shown in Fig. 11(a) for different values of T, and they are juxtaposed to the results obtained with the preferential attachment (PA) and common-neighbors (CN) heuristics [17]. In PA, the score between two disconnected ASs (i,j) is $s_{ij} = k_i \times k_j$, where k_i,k_j are the degrees of the ASs, while in CN $s_{ij} = n_{ij}$, where n_{ij} is the number of common neighbors between the ASs. The higher these scores, the higher the chance of a future link between the disconnected ASs. From Fig. 11(a), we observe that the fast hybrid and link-based methods yield very high AUC values, around 0.97 for almost all T, outperforming the PA and CN heuristics. Note that hybrid and link-based methods perform similarly with respect to this performance metric. This is not surprising since, as we have seen in the previous section, the resulting connection probabilities in the two methods are quite similar; cf. Figs. 5 and 7. In particular, even though the link-based method without correction steps is not as accurate at inferring the real angular coordinates of nodes [cf. Figs. 4(d)–4(f)], its resulting connection probabilities are close to the ones obtained by the hybrid method; cf. Fig. 5. That is, these results also agree with our previous observations on synthetic networks. In Fig. 11(b), we compute the AUC by considering



(a) $T = 0.05$, hybrid with correction steps.

(b) $T = 0.4$, hybrid with correction steps.

(c) $T = 0.7$, hybrid with correction steps.

FIG. 9. (Color online) Inferred angles (in radians) with the original and fast versions of the hybrid method for synthetic networks with $t = 5000$ nodes, $m = 1.5$, $L = 2.5$, $\gamma = 2.1$, and T as shown in the captions.

TABLE V. Logarithmic losses obtained by the fast version of the methods (with correction steps).

Network	$\mathrm{LL}^{\mathrm{inf}}$, fast hybrid	$\mathrm{LL}^{\mathrm{inf}}$, fast link-based
$T = 0.05$	6.2×10^4	4.0×10^4
$T = 0.4$	3.0×10^4	2.9×10^4
$T = 0.7$	4.2×10^4	4.1×10^4

TABLE VI. Success ratio p_s and average hop length $\bar{h}$ of greedy paths obtained by the fast version of the methods (with correction steps).

Network	Fast hybrid	Fast link-based
$T = 0.05$	$p_s = 0.97, \bar{h} = 3.1$	$p_s = 0.98, \bar{h} = 3.1$
$T = 0.4$	$p_s = 0.96, \bar{h} = 3.3$	$p_s = 0.97, \bar{h} = 3.3$
$T = 0.7$	$p_s = 0.91, \bar{h} = 3.7$	$p_s = 0.92, \bar{h} = 3.7$

only disconnected AS pairs with no common neighbors and the future links among these pairs. In this case, CN performs as well as pure chance since it assigns a zero score to all the pairs, while the fast hybrid and link-based methods still perform remarkably well, with AUC values between 0.89 and 0.92. Finally, in Fig. 11(c), we compute the AUC by considering only disconnected AS pairs with low degrees, less than the average degree $\bar{k} = 5$, and the future links among these pairs. The figure shows that the methods still perform very well, with AUC values between 0.79 and 0.85 for $T \leqslant 0.8$, significantly outperforming the PA and CN heuristics.

To summarize, our results indicate that our methods have a very strong predictive power. Specifically, they perform remarkably well not only in predicting the "easy-to-predict" future links, i.e., the links that appear among nodes with high degrees or many common neighbors, but also in predicting the "hard-to-predict" future links, i.e., the links that appear among nodes with low degrees or no common neighbors. In that sense, one can say that the measure of proximity (hyperbolic distances) between nodes in our approach reflects reality more accurately than the PA and CN approaches do, and that our methods can infer these distances in the real Internet quite accurately. The predictive power of our methods is not very sensitive to the exact value of T, with the best results obtained for $T \leqslant 0.8$; cf. Fig. 11.

Evolution of soft AS communities. In Fig. 12 we map our six AS snapshots, using the fast hybrid method with correction steps as before, with $T = 0.6$, which yielded the lowest logarithmic loss, and $k_{\mathrm{speedup}} = 3$. In all cases, the angle θ_1 of node $i = 1$ (see step 3 of Fig. 1) is fixed to $\theta_1 = \pi$. Figures 12(a)–12(f) show that the method produces meaningful mappings, in the sense that the method infers soft communities of ASs belonging to the same country, where by soft communities we mean groups of nodes located close to each other in the space. For each mapped snapshot, Fig. 12 shows the angular distribution of ASs belonging to the same country for 20 different countries. For comparison among the distributions, for each snapshot after Sept. 2009 we consider only the ASs that were also present in Sept. 2009. The x axis in Figs. 12(a)–12(f) (angular coordinate) uses bins of size 3.6°. The AS-to-country mapping is taken from the CAIDA AS ranking project [18]. We observe that the fast hybrid method places ASs belonging to the same country close to each other in the angular space. The reason for this is that ASs belonging to the same country tend to connect more densely to each other than to the rest of the world. Connected ASs are attracted to each other, while disconnected ASs repel, and the fast hybrid method feels these attraction/repulsion forces, placing groups of densely connected ASs in narrow regions, close to each other. As expected, due to significant geographic spread in ASs belonging to the United States, these ASs are more widespread. We note that other reasons besides geographic proximity may affect the connectivity between ASs, such as economical, political, and performance-related reasons. The mapping method does not favor any specific reason but relies only on the connectivity between ASs in order to place the ASs at the right angular (and consequently hyperbolic) distances.

Figures 12(g)–12(i) also show how the angular center of masses of the considered AS communities evolves in the similarity space during the period Sept. 2009–Dec. 2010. We observe that some communities, e.g., the United States and several European countries, have a more stable position in this space than others, e.g., Argentina and Brazil. The observed dynamics in the similarity space is likely due to a combination of two classes of factors: (i) stochastic fluctuations and noise coming from the data (our mapping does not introduce any additional randomness since the algorithm is deterministic), and (ii) real dynamics of nodes and communities in the similarity space, caused by new connections and disconnections

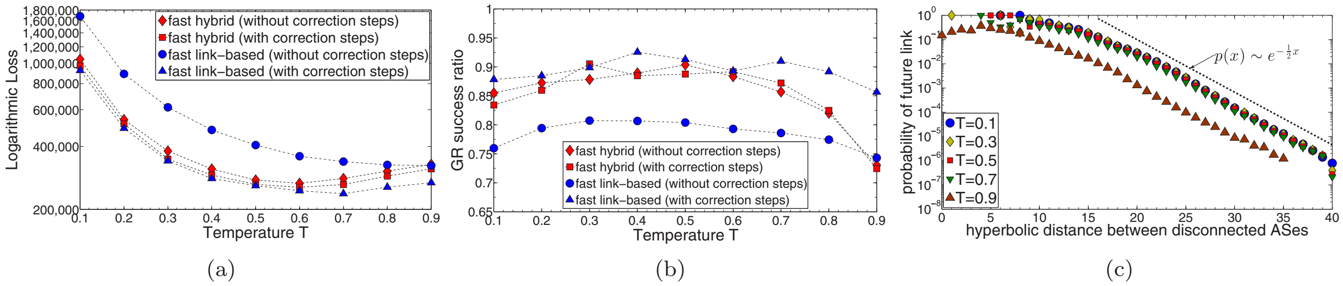
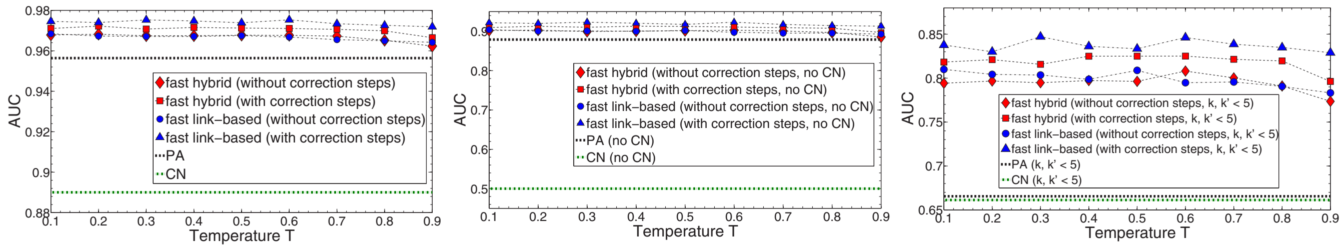


FIG. 10. (Color online) Logarithmic loss ($\mathrm{LL}^{\mathrm{inf}}$), GR success ratio ($p_s$), and future-link probability in a mapped snapshot of the AS Internet (Sept. 2009 snapshot). In (a) and (b), the results are obtained by the fast hybrid and link-based methods, with and without correction steps. The results in (c) are obtained by the fast hybrid method with correction steps. In all cases $k_{\mathrm{speedup}} = 3$, and the results are shown for different values of the temperature parameter T.



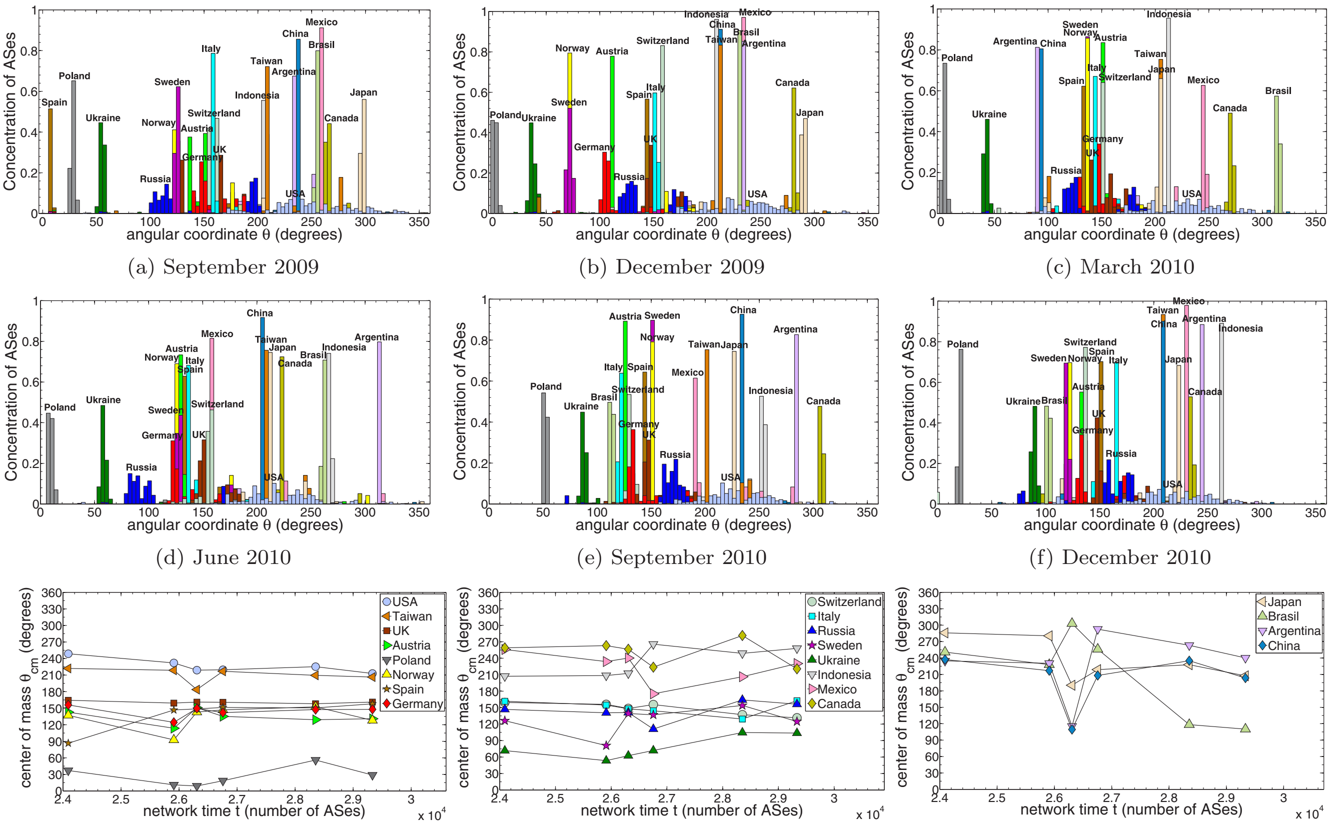
(a) (b) (c)

FIG. 11. (Color online) Performance of future-link prediction in the AS Internet with the fast hybrid and link-based methods ($k_{\text{speedup}} = 3$), and comparison to the preferential attachment (PA) and common-neighbors (CN) heuristics. In each case, the AS snapshot of Sept. 2009 is considered. In (a), the AUC is computed over all disconnected AS pairs and the new links that appear between them in Sept. 2009–Dec. 2010 (48 119 new links); in (b), the AUC is computed only over the disconnected AS pairs that have no common neighbors (95% of all disconnected pairs) and the new links between them (9279 new links); and in (c), the AUC is computed only over the disconnected AS pairs with degrees $k,k' < \bar{k} = 5$ (72% of all disconnected pairs) and the new links between them (2050 new links).

within and across the communities. Similar results hold for the link-based method with correction steps.

VII. OTHER RELATED WORK

A different mapping of the AS Internet to the hyperbolic plane was performed in [19]. The authors found that the hop lengths of the shortest AS paths in the Internet can be embedded into the hyperbolic plane with low distortion, and that the resulting embedding can be used for efficient overlay network construction and accurate path distance estimation. Our work is different from [19] in that hyperbolic distances between ASs in our case are not directly defined by their "observable" AS path lengths. Instead, they are defined by



(a) September 2009 (b) December 2009 (c) March 2010

(d) June 2010 (e) September 2010 (f) December 2010

(g) (h) (i)

FIG. 12. (Color online) Distributions of angular coordinates of ASs belonging to the same country during September 2009–December 2010 (a)–(f), and the evolution of the angular center of masses of the corresponding communities over time (g)–(i). For each snapshot in (a)–(f), the angular center of mass of each country is $\theta_{\text{c.m.}} = (1/n)\sum_b \theta(b)n(b)$, where n is the number of ASs belonging to the country, $n(b)$ is the number of such ASs falling within bin b, $\theta(b)$ is the value of θ in the bin, and the summation is over all the bins. For each country, $\theta_{\text{c.m.}}$ is shown (g)–(i) as a function of the network time t, i.e., as a function of the number of ASs in the snapshots (a)–(f), $t = 24\,091$, 25 910, 26 307, 26 756, 28 353, and 29 333, respectively.

"hidden" popularity and similarity node coordinates that manifest themselves indirectly via the nodes' connections and disconnections. As indicated by the performance of greedy routing in Sec. V, short paths follow well the underlying hyperbolic geodesics in our mappings. However, nodes at short path distances are not always hyperbolically closer than nodes separated by longer paths. For the same reason, our approach differs from multidimensional scaling (MDS) techniques, which try to compute coordinates for points in low-dimensional geometric spaces (see, e.g., [20]), such that the distances between the points in these spaces match as closely as possible some given distances between the points.

In addition to [5,6], perhaps the most relevant earlier work is [21]. In that work, the authors considered a model of social networks in which nodes reside in a latent Euclidean space [22]. Nodes that are sufficiently close in this space have higher chances of being connected. Based on this model, the authors presented a combined MDS and maximum likelihood estimation (MLE) procedure for inferring the node coordinates in the latent space. The procedure can take into consideration previously estimated node positions, e.g., estimated node positions in a previous closely spaced network snapshot, and penalize large displacements from these positions, in an attempt to yield more accurate embeddings. The authors applied this procedure to create embeddings for link prediction, and to illustrate how relationships between authors in coauthorship data change over time. The main difference between our work and that in Ref. [21] is that in our case the latent space is not Euclidean but hyperbolic, the latter providing a more accurate reflection of the geometry of real networks [4,6,7]. In contrast to earlier work on latent network geometry inference, here we have departed from the traditional link-based inference methods, and we based our inference entirely on a higher-order similarity statistics—the statistics of the number of common neighbors between nodes.

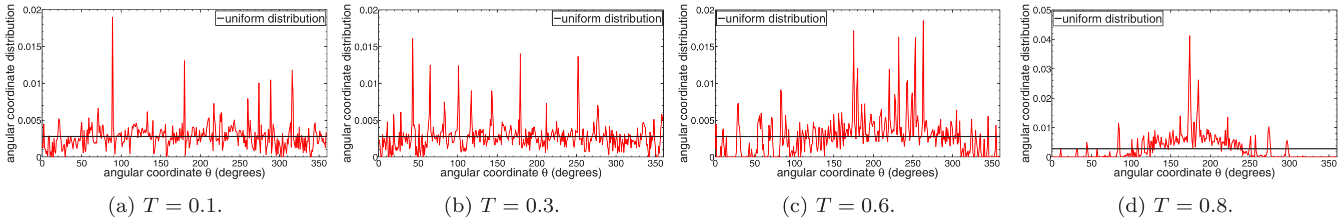
VIII. CONCLUSION

In summary, we have introduced and explored a method for inferring node similarity coordinates based on the number of common neighbors between nodes, and we have released the software package implementing this network mapping method to the public [8]. We have shown that this approach is more accurate than the link-based approach [5], unless heuristic periodic adjustments (or correction steps) are used. The common-neighbors approach is more computationally intensive, but we have devised a hybrid method that combines the common-neighbors and link-based approaches, and we showed how to reduce its running time to $O(t^2)$. The correction steps can be used in this hybrid approach as well, but their effect is not significant. Therefore, they can be entirely avoided to reduce running time. We have validated this method on synthetic model networks, and we applied it to the evolving AS Internet. Taken altogether, our results advance our understanding of how to efficiently and accurately map real networks to their underlying hyperbolic spaces.

An interesting open problem is whether more computationally efficient but also more sophisticated numerical optimization methods [23] can be applied to the latent network geometry inference problem. Such methods may expedite the maximization of the likelihoods $\mathcal{L}_{\mathrm{L}}^{i}$ and $\mathcal{L}_{\mathrm{CN}}^{i}$ in Eqs. (5) and (13), without sacrificing the embedding quality. We note that our "brute-force" approach of sampling the likelihoods at small $\Delta\theta$ intervals in order to find their global maximum appears currently to be the best option among all other methods that we have investigated. These methods [23] tend to work reliably only if the function to maximize is relatively smooth, has only one easily detectable global maximum, or has only a few local maxima. In contrast, the likelihood profiles we have to deal with, Figs. 3 and 8, are very rugged and rough, abundant with sharp local maxima, rendering unusable all the other methods with which we have experimented.

All the inference methods presented here and in [5,6] use the uniform distribution as the prior [24] for the angular distribution of nodes, Eq. (12). This means that the methods do not make any prior assumption about the node angular positions. Instead, they assume that all positions are equiprobable, and they allow the given data, i.e., the given network adjacency matrix, to determine the positions. The distributions of the inferred angular coordinates can then be nonuniform in mappings of real networks produced by these methods, since many real networks tend to have some nontrivial community structure. For example, Fig. 13 shows the distribution of the inferred AS angles in September 2009. The lowest logarithmic loss [Fig. 10(a)] is achieved at $T = 0.6$, and the corresponding distribution of angular coordinates is clearly nonuniform. In this context, an interesting open problem is to consider extensions of network geometry models that are capable of explaining the emergence of soft community structure in networks and nonuniform distribution of nodes in the similarity space (see, e.g., [25]), and to develop mapping methods for such models that would use nonuniform priors.

Finally, given an efficient and accurate method to map real complex networks into their underlying hyperbolic spaces, one of the most interesting open problems is to decipher



(a) $T = 0.1$. (b) $T = 0.3$. (c) $T = 0.6$. (d) $T = 0.8$.

FIG. 13. (Color online) Distribution of the inferred AS angles (Sept. 2009 snapshot) with the fast hybrid method ($k_{\mathrm{speedup}} = 3$) and different values of the temperature parameter T.

the laws that govern the dynamics of nodes in these spaces, Fig. 12. As real networks are characterized by a hierarchical organization and nontrivial community structure [11,26], we expect this dynamics to be also highly nontrivial, but definitely not random. This observation suggests that it might be possible to accurately predict the future positions of nodes in the underlying hyperbolic spaces. The precise knowledge of this spatial dynamics of nodes can then be used to predict *fine-grained* network dynamics, forecasting future connections and disconnections among nodes over different time scales.

ACKNOWLEDGMENTS

We thank M. Kitsak, M. Boguñá, and C. Psomas for useful discussions and suggestions. This work was supported by a Marie Curie International Reintegration Grant within the 7th European Community Framework Programme, by an AWS in Education grant award, by DARPA Grant No. HR0011-12-1-0012, NSF Grants No. CNS-1344289, No. CNS-1442999, No. CNS-0964236, No. CNS-1441828, No. CNS-1039646, and No. CNS-1345286, and by Cisco Systems.

[1] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[2] S. N. Dorogovtsev, J. Mendes, and A. Samukhin, arXiv:cond-mat/0009090 (2000).

[3] M. McPherson, L. Smith-Lovin, and J. M. Cook, Annu. Rev. Sociol. **27**, 415 (2001).

[4] F. Papadopoulos, M. Kitsak, M. A. Serrano, M. Boguñá, and D. Krioukov, Nature (London) **489**, 537 (2012).

[5] F. Papadopoulos, C. Psomas, and D. Krioukov, IEEE/ACM Trans. Netw. **23**, 198 (2015).

[6] M. Boguñá, F. Papadopoulos, and D. Krioukov, Nat. Commun. **1**, 62 (2010).

[7] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá, Phys. Rev. E **82**, 036106 (2010).

[8] HyperMap-CN Software Package, https://bitbucket.org/dk-lab/2015_code_hypermap.

[9] P. Sarkar, D. Chakrabarti, and A. W. Moore, in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence* (AAAI, Menlo Park, 2011), pp. 2722–2727.

[10] M. Kitsak and D. Krioukov, Phys. Rev. E **84**, 026114 (2011).

[11] S. N. Dorogovtsev, *Lectures on Complex Networks* (Oxford University Press, Oxford, 2010).

[12] F. Bonahon, *Low-Dimensional Geometry* (AMS, Providence, 2009).

[13] R. B. Ash and C. A. Doléans-Dade, *Probability & Measure Theory*, 2nd ed. (Academic, San Diego, 1999).

[14] M. Boguñá, D. Krioukov, and K. Claffy, Nat. Phys. **5**, 74 (2009).

[15] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov, in *CATCH* (IEEE Computer Society, Los Alamitos, 2009), http://www.caida.org/projects/ark/.

[16] IPv4 Routed /24 AS Links Dataset, http://www.caida.org/data/active/ipv4_routed_topology_aslinks_dataset.xml.

[17] L. Lu and T. Zhou, Physica A **390**, 1150 (2011).

[18] X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, K. Claffy, and G. Riley, Comput. Commun. Rev. **37**, 29 (2007).

[19] Y. Shavitt and T. Tankel, IEEE/ACM Trans. Netw. **16**, 25 (2008).

[20] E. Begelfor and M. Werman, Tech. Rep. HUJI-CSE-LTR-2006-191, School of Engineering and Computer Science, Hebrew University of Jerusalem (2005), http://www.cs.huji.ac.il/∼werman/Papers/cmds.pdf.

[21] P. Sarkar and A. W. Moore, SIGKDD Explor. Newsl. **7**, 31 (2005).

[22] P. D. Hoff, A. E. Raftery, and M. S. Handcock, J. Am. Stat. Assoc. **97**, 1090 (2002).

[23] J. Nocedal and S. Wright, *Numerical Optimization* (Springer, New York, 2000).

[24] E. Jaynes, IEEE Trans. Syst. Sci. Cybern. **4**, 227 (1968).

[25] K. Zuev, M. Boguñá, G. Bianconi, and D. Krioukov, Nat. Sci. Rep. **5**, 9421 (2015).

[26] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).