


<https://doi.org/10.1038/s42005-026-02725-1>

A Bayesian approach to out-of-sample network reconstruction



Mattia Marzi^{1,2} & Tiziano Squartini^{1,2,3}

Networks underpin systems that range from finance to biology, yet their structure is often only partially observed. Current reconstruction methods typically fit the parameters of a model anew to each snapshot, thus offering no guidance to predict future configurations. Here, we develop a Bayesian approach that uses the information about past network snapshots to inform a prior and predict the subsequent ones, while quantifying uncertainty. Instantiated with a single-parameter fitness model, our method infers link probabilities from node strengths and carries information forward in time. When applied to the Electronic Market for Interbank Deposit across the years 1999–2012, our method accurately recovers the number of connections per bank at subsequent times, outperforming probabilistic benchmarks designed for analogous, link prediction tasks. Notably, each predicted snapshot serves as a reliable prior for the next one, thus enabling self-sustained, out-of-sample reconstruction of evolving networks with a minimal amount of additional data.

Network theory is employed to address problems of scientific and societal relevance, from the prediction of epidemic spreading to the identification of early-warning signals of upcoming financial crises^{1–8}.

Any dynamical process is strongly affected by the topology of the underlying network; still, data restrictions may prevent it from being fully accessible - in many empirical settings, only aggregate node-level quantities are available. The inference problem we address here consists in turning the available information into a probability distribution over the unknown network portion, so as to quantify the plausibility of each link.

Such a problem can be approached by constructing a statistical benchmark, i.e. a set of configurations that preserve only certain properties while maintaining controlled randomness elsewhere: these properties, named *constraints* and indicated as **C**, represent the only information that can be employed to infer any other (inaccessible) property.

A class of models whose popularity has steadily increased over the years is that of Exponential Random Graphs (ERGs)^{9–15}. ERGs come from constrained Shannon entropy maximization^{9,11,16,17} and belong to the category of approaches enforcing constraints (only) on average.

So far, the parameters defining ERGs have been determined by employing the traditional maximum-of-the-likelihood estimation procedure^{18–22}; in the case of ERGs, in fact, it leads to the particularly simple request that the expected value of each constraint equals the corresponding empirical value: in symbols, $\langle \mathbf{C} \rangle = \mathbf{C}^*$, with $\langle \mathbf{C} \rangle$ indicating the expected value of the vector of constraints and $\mathbf{C}^*$ the empirical one.

Coming to the purpose of reconstructing a graph topology, one of the most effective models is the Undirected Binary Configuration Model

(UBCM)^{9,11–13,15}, induced by the degree sequence $\{k_i\}_{i=1}^N$, with $k_i = \sum_{j(\neq i)} a_{ij}$ indicating the number of neighbors of node i , a_{ij} indicating the entry (i, j) of the adjacency matrix **A** of a binary undirected network and N indicating the total number of nodes. More formally, the UBCM assigns **A** the factorized probability distribution reading

$$P(\mathbf{A}|\mathbf{x}) = \prod_{i=1}^N \prod_{j(i>i)} p_{ij}^{a_{ij}} (1 - p_{ij})^{1-a_{ij}}, \quad (1)$$

where

$$p_{ij} = \frac{x_i x_j}{1 + x_i x_j} \quad (2)$$

and x_i indicates the Lagrange multiplier associated to k_i ; maximizing the log-likelihood of such a model implies solving the system of equations

$$\langle k_i \rangle = \sum_{j(\neq i)} p_{ij} = k_i^*, \quad \forall i \quad (3)$$

that, in turn, induces the vector of point-wise estimates hereby indicated with $\{x_i^*\}_{i=1}^N$.

A similar perspective can be adopted to approach *link prediction*, an active research field within the broader one of network science. Close in scope to network reconstruction²³, link prediction targets specific

¹IMT School for Advanced Studies, Lucca, Italy. ²INdAM-GNAMPA Istituto Nazionale di Alta Matematica ‘Francesco Severi’, Rome, Italy. ³Scuola Normale Superiore, Pisa, Italy. e-mail: mattia.marzi@imtlucca.it

connections, aiming to uncover missing ones and predict those most likely to emerge in the future^{24–26}. Such an issue is relevant in many research areas, such as those concerning socio-economic and financial networks: knowing the structure of the commercial partnerships between firms or of the financial exchanges between banks is, in fact, relevant for a number of reasons, such as quantifying the risk associated to the propagation of a shock^{27–29}. Link prediction algorithms rank unconnected node pairs on the basis of a *score*: while some methods rely on purely structural information, others admit external information as well.

The simplest framework to carry out link prediction includes the so-called *similarity-based* algorithms: scores, here, are induced by some measure of similarity between nodes; to this aim, *local*, *quasi-local* or *global* information - such as the degree, the degree of common neighbors or the length of paths connecting any two nodes - has been employed^{24–26}.

A more refined framework includes machine learning approaches^{24–26}: specifying a model, here, amounts to learning a function that determines the probability for nodes i and j to be connected while taking as inputs two vectors $\mathbf{f}_i$ and $\mathbf{f}_j$ of (structural and/or external) node-specific features and a vector $\mathbf{g}_{ij}$ of (structural and/or external) edge-specific features.

A third alternative is represented by the so-called *likelihood-based* algorithms, defined by a likelihood function whose maximization provides the probability that any two nodes are connected, to be interpreted as a score for the link existence: this is achieved precisely by assuming that a certain amount of information is accessible, hence treating it as a constraint to account for refs. 24–26.

The third perspective justifies the interpretation of the link prediction problem as an instance of the network reconstruction one, although network reconstruction usually deals with less information to predict more aggregate properties; in this sense, the key distinction between these two approaches is informational: while link prediction exploits dyadic features to achieve an edge-level classification that is as accurate as possible, entropy-based reconstruction makes the best possible use of the available information to reproduce the largest possible portion of a network.

Let us explicitly notice that the schemes sketched above have been designed to carry out inference in an *in-sample* fashion: in other words, the probability coefficients $\{p_{ij}\}_{i,j=1}^N$ are estimated from the available information concerning $\mathbf{A}^*$ to make statements about $\mathbf{A}^*$ itself. To the best of our knowledge, a genuinely *out-of-sample* formulation of the entropy-based framework for network reconstruction is still missing: existing methods for temporal link prediction, in fact, address such a task under highly controlled circumstances such as being in presence of the same set of nodes, having full access to all sets of dyadic features, as well as to a stable representation of the latter, etc.

Here, instead, the aforementioned issue is addressed by developing a Bayesian framework in which past snapshots inform an empirical prior only through the sufficient statistics needed to calibrate the model - i.e. without requiring additional regularities that are not supported by data. When instantiated with a single-parameter model, our framework *i)* allows the number of nodes to vary over time; *ii)* allows one to reconstruct future snapshots from node-level information and propagated uncertainty (let us, in fact, notice that the derivation of ERGs offers no prescription to propagate the uncertainty affecting a parameter estimation from one snapshot to another, nor to turn past observations into an informative prior about future ones); *iii)* in its self-sustained version, evolves without requiring any other topological information than the one needed for initialization; *iv)* remains competitive when compared with an in-sample reconstruction model calibrated on the target snapshot, thus showing that an accurate out-of-sample reconstruction can be achieved with a limited amount of accessible, interpretable information.

Results

Towards a Bayesian approach

To carry out an out-of-sample network reconstruction, let us, then, move towards a different design, trying to turn the UBCM into a Bayesian model. Bayes' theorem provides the key relation $P(\mathbf{x}|\mathbf{A})P(\mathbf{A}) = P(\mathbf{A}, \mathbf{x}) = P(\mathbf{A}|\mathbf{x})\pi(\mathbf{x})$

which, in turn, induces the following definition of *posterior distribution*:

$$P(\mathbf{x}|\mathbf{A}) = \frac{P(\mathbf{A}|\mathbf{x})\pi(\mathbf{x})}{P(\mathbf{A})} = \frac{P(\mathbf{A}|\mathbf{x})\pi(\mathbf{x})}{\int P(\mathbf{A}|\mathbf{x})\pi(\mathbf{x})d\mathbf{x}}; \quad (4)$$

adapting the parameter estimation procedure accordingly would amount to calculating either the *mean* or the *mode* of the posterior distribution, respectively defined as

$$\langle \mathbf{x}|\mathbf{A} \rangle = \int \mathbf{x}P(\mathbf{x}|\mathbf{A})d\mathbf{x} \quad (5)$$

and

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x}} \{P(\mathbf{x}|\mathbf{A})\}, \quad (6)$$

the acronym standing for *maximum-a-posteriori*. Both procedures, however, provide a point-wise estimate solely depending on the observation made at time t . A better suited instrument seems to be the so-called *posterior predictive distribution*³⁰. Upon calling $\mathbf{A}_t$ the adjacency matrix representing our graph at time t and $\mathbf{A}_{t+1}$ the adjacency matrix representing our graph at time $t + 1$, it reads

$$\begin{aligned} P(\mathbf{A}_{t+1}|\mathbf{A}_t) &= \int P(\mathbf{A}_{t+1}, \mathbf{x}|\mathbf{A}_t)d\mathbf{x} \\ &= \int P(\mathbf{A}_{t+1}|\mathbf{x}, \mathbf{A}_t)P(\mathbf{x}|\mathbf{A}_t)d\mathbf{x} \\ &= \int P(\mathbf{A}_{t+1}|\mathbf{x})P(\mathbf{x}|\mathbf{A}_t)d\mathbf{x} \\ &= \int \frac{P(\mathbf{A}_{t+1}|\mathbf{x})P(\mathbf{A}_t|\mathbf{x})\pi(\mathbf{x})}{P(\mathbf{A}_t)}d\mathbf{x}, \end{aligned} \quad (7)$$

where the chain rule, $f(x, y|z) = f(x|y, z)f(y|z)$ has been employed at the second passage and conditional independence has been assumed at the third passage.

The multivariate nature of the formula above, however, makes its resolution cumbersome; moreover, the UBCM is not a viable reconstruction model, as its calibration requires the knowledge of the entire degree sequence per snapshot - a requirement that is hardly satisfied in a realistic scenario. For such a reason, let us look for a simpler, single-parameter model. Our choice, then, turns Eq. (7) into

$$P(\mathbf{A}_{t+1}|\mathbf{A}_t) = \int_0^{+\infty} \frac{P(\mathbf{A}_{t+1}|z)P(\mathbf{A}_t|z)\pi(z)}{P(\mathbf{A}_t)}dz, \quad (8)$$

where $P(\mathbf{A}_t) = \int_0^{+\infty} P(\mathbf{A}_t|z)\pi(z)dz$. Let us stress that assuming conditional independence simplifies the computation of the formula above, as the dependence of the future snapshot on the past snapshot is established only via z ; moreover, nothing prevents us from considering a varying number of nodes, i.e. N_t and N_{t+1} can differ.

Since z is integrated out, we are left with a probability depending on $\mathbf{A}_t$, that we know, and on $\mathbf{A}_{t+1}$, that we want to infer; the problem can be simplified even more upon considering that the marginal (i.e. edge-specific) probability reads

$$q_{ij}^{t+1} = P(a_{ij}^{t+1} = 1|\mathbf{A}_t) = \int_0^{+\infty} p_{ij}^{t+1}(z) \frac{P(\mathbf{A}_t|z)\pi(z)}{P(\mathbf{A}_t)}dz, \quad (9)$$

where p_{ij}^{t+1} is the probability that the corresponding entry of $\mathbf{A}_{t+1}$, i.e. a_{ij}^{t+1} , is 1 (see also Supplementary Note 1).

Such an expression offers us a viable way to carry out our inference exercise, since it allows the following quantities to be evaluated analytically,

i.e. the (expected) total number of links, reading

$$\langle L_{t+1} \rangle = \sum_{i=1}^N \sum_{j(>i)}^N q_{ij}^{t+1}, \quad (10)$$

the (expected) degree of each node, reading

$$\langle k_i^{t+1} \rangle = \sum_{j(>i)} q_{ij}^{t+1}, \quad \forall i \quad (11)$$

and the entries of the so-called *confusion matrix*, i.e.

$$\langle TP_{t+1} \rangle = \sum_{i=1}^N \sum_{j(>i)} a_{ij}^{t+1} q_{ij}^{t+1}, \quad (12)$$

$$\langle FP_{t+1} \rangle = \sum_{i=1}^N \sum_{j(>i)} (1 - a_{ij}^{t+1}) q_{ij}^{t+1}, \quad (13)$$

$$\langle TN_{t+1} \rangle = \sum_{i=1}^N \sum_{j(>i)} (1 - a_{ij}^{t+1})(1 - q_{ij}^{t+1}), \quad (14)$$

$$\langle FN_{t+1} \rangle = \sum_{i=1}^N \sum_{j(>i)} a_{ij}^{t+1}(1 - q_{ij}^{t+1}); \quad (15)$$

in words, $\langle TP_{t+1} \rangle$ represents the (expected) number of true positives, i.e. the number of correctly recovered connections; $\langle FP_{t+1} \rangle$ represents the (expected) number of false positives, i.e. the number of incorrectly recovered connections; $\langle TN_{t+1} \rangle$ represents the (expected) number of true negatives, i.e. the number of correctly recovered missing connections; $\langle FN_{t+1} \rangle$ represents the (expected) number of false negatives, i.e. the number of incorrectly recovered missing connections.

The Bayesian Erdős–Rényi Model (BERM)

Let us, now, instantiate our framework with the Erdős–Rényi Model. Its classical formulation reads

$$P(\mathbf{A}|p) = p^{L(\mathbf{A})}(1-p)^{V-L(\mathbf{A})}, \quad (16)$$

where $V = N(N-1)/2$ indicates the total number of pairs of nodes. Fully determining it, however, requires the *prior distribution* to be specified: a popular choice is that of considering the *conjugate prior*, introduced to maintain the functional form of the likelihood function. Upon doing so, we are led to expression

$$\begin{aligned} P(p|\mathbf{A}) &= \frac{P(\mathbf{A}|p)\pi(p)}{\int_0^1 P(\mathbf{A}|p)\pi(p)dp} \\ &= \frac{p^{L(\mathbf{A})+\alpha-1}(1-p)^{V-L(\mathbf{A})+\beta-1}}{B(L(\mathbf{A})+\alpha, V-L(\mathbf{A})+\beta)}, \end{aligned} \quad (17)$$

coming from posing

$$\pi(p) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)}, \quad (18)$$

where $B(\alpha, \beta) = \int_0^1 p^{\alpha-1}(1-p)^{\beta-1}dp$ is the Beta function. As a consequence, we can write

$$\begin{aligned} P(L_{t+1} = k|\mathbf{A}_t) &= \int_0^1 P(L_{t+1} = k|p)P(p|\mathbf{A}_t)dp \\ &= \int_0^1 \binom{V_{t+1}}{k} p^k (1-p)^{V_{t+1}-k} \frac{P(\mathbf{A}_t|p)\pi(p)}{P(\mathbf{A}_t)} dp \\ &= \text{BetaBin}(V_{t+1}, L_t + \alpha, V_t - L_t + \beta), \end{aligned} \quad (19)$$

where the explicit dependence of the third expression on $\mathbf{A}_t$ has been dropped; in words, the total number of links at time $t+1$, conditional to the observation of $\mathbf{A}_t$, obeys a beta-binomial distribution. Since

$$\langle L_{t+1}|\mathbf{A}_t \rangle = \sum_{k=0}^{V_{t+1}} k P(L_{t+1} = k|\mathbf{A}_t), \quad (20)$$

swapping the operations of sum and integration leads to

$$\begin{aligned} \langle L_{t+1}|\mathbf{A}_t \rangle &= \int_0^1 \sum_{k=0}^{V_{t+1}} k P(L_{t+1} = k|p) P(p|\mathbf{A}_t) dp \\ &= V_{t+1} \int_0^1 p \frac{P(\mathbf{A}_t|p)\pi(p)}{P(\mathbf{A}_t)} dp \end{aligned} \quad (21)$$

and comparing it with Eq. (9) further leads us to recognize that

$$\begin{aligned} q^{t+1} &= \int_0^1 p \frac{P(\mathbf{A}_t|p)\pi(p)}{P(\mathbf{A}_t)} dp \\ &= \frac{B(1+L_t+\alpha, V_t-L_t+\beta)}{B(L_t+\alpha, V_t-L_t+\beta)} \\ &= \frac{L_t+\alpha}{V_t+\alpha+\beta} \end{aligned} \quad (22)$$

(see also Supplementary Note 2).

The Bayesian Fitness Model (BFM)

The main limitation of the BERM lies in its homogeneity: in other words, all nodes are treated equally. Since different nodes are known to play different roles, according to their structural importance, let us look for a single-parameter model, *heterogeneous in nature*.

A very natural choice is that of considering the variant of the UBCM named *density-corrected Gravity Model* (dcGM)^{31–36}, induced by node-specific fitnesses typically identified with the strengths $\{s_i\}_{i=1}^N$, where $s_i = \sum_{j \neq i} w_{ij}$ and $\mathbf{W}$ is the weighted adjacency matrix associated to our graph, while solely enforcing (a proxy of) the link density, defined as $c = 2L/N(N-1) = 2\sum_{i=1}^N \sum_{j(>i)} a_{ij}/N(N-1)$. For consistency, $\mathbf{A} = \Theta(\mathbf{W})$, where $\Theta(x)$ denotes the Heaviside step function, here defined as $\Theta(x) = 1$ if $x > 0$ and $\Theta(x) = 0$ otherwise, applied element-wise. More quantitatively, the dcGM is defined by the fitness ansatz $x_i = \sqrt{z}s_i$, which exploits the empirical correlation between the Lagrange multiplier x_i and the strength s_i , thus turning Eq. (2) into

$$p_{ij} = \frac{zs_i s_j}{1 + zs_i s_j}; \quad (23)$$

the only parameter z is, then, determined by maximizing the log-likelihood $\ln P(\mathbf{A}|z) = \sum_{i=1}^N \sum_{j(>i)} [a_{ij} \ln p_{ij} + (1-a_{ij}) \ln(1-p_{ij})]$, a recipe implying that the only equation to be solved reads

$$\langle L \rangle = \sum_{i=1}^N \sum_{j(>i)} p_{ij} = \sum_{i=1}^N \sum_{j(>i)} \frac{zs_i s_j}{1 + zs_i s_j} = L^*, \quad (24)$$

where L^* and z_{ML}^* respectively indicate the empirical value of the total number of links and the related parameter estimation.

Instantiating the expression in Eq. (9) with the one defining the dcGM leads to

$$q_{ij}^{t+1} = \int_0^{+\infty} \left(\frac{zs_i^{t+1} s_j^{t+1}}{1 + zs_i^{t+1} s_j^{t+1}} \right) \frac{P(\mathbf{A}_t|z)\pi(z)}{P(\mathbf{A}_t)} dz, \quad (25)$$

where

$$P(\mathbf{A}_t|z) = \prod_{i=1}^{N_t} \prod_{j(>i)} (p_{ij}^t)^{a_{ij}^t} (1 - p_{ij}^t)^{1-a_{ij}^t} \quad (26)$$

and

$$p_{ij}^t = \frac{zs_{ij}^t s_j^t}{1 + zs_{ij}^t s_j^t}; \quad (27)$$

let us stress that the dependence on the observed snapshot does not imply that the adjacency matrix at time t must be entirely known: once the strength sequence is given, the binary information entering the likelihood is, in fact, summarized by the sufficient statistic L_t , i.e. the total number of links at time t (see also Supplementary Note 3).

The rationale that led us to the BFM motivates us to determine $\pi(z)$ by adopting the recipe called *empirical prior*: in words, one *i*) estimates z on each of the snapshots preceding the one under consideration; *ii*) deduces the functional form of $\pi(z)$; *iii*) plugs it into the expression of the snapshot to be reconstructed (see also Supplementary Note 3).

At this point, one needs to compute q_{ij}^{t+1} . A possible choice is that of evaluating the integrand in correspondence of z_{MAP}^* , i.e. the value that maximizes the log-posterior $\ln P(z|\mathbf{A}_t) = \ln P(\mathbf{A}_t|z) + \ln \pi(z)$, a recipe implying that the only equation to be solved reads

$$\frac{\partial \ln P(z|\mathbf{A}_t)}{\partial z} = \frac{\partial \ln P(\mathbf{A}_t|z)}{\partial z} + \frac{\partial \ln \pi(z)}{\partial z} = 0 \quad (28)$$

or, more explicitly,

$$\frac{1}{z} \left[L_t^* - \sum_{i=1}^{N_t} \sum_{j(>i)} \frac{zs_{ij}^t s_j^t}{1 + zs_{ij}^t s_j^t} \right] + \frac{\partial \ln \pi(z)}{\partial z} = 0; \quad (29)$$

although viable, this choice rests upon the assumption that the integrand is peaked around z_{MAP}^* ; besides, it privileges the likelihood term over the prior - i.e. a system current realization over its history.

To avoid relying on this assumption, and fully account for uncertainty, one is forced to proceed numerically. To this aim, we move to the logarithmic coordinate $u = \ln z$ and employ the integration scheme named *Gauss-Hermite quadrature* (see also the Methods section and Supplementary Note 4).

Out-of-sample network reconstruction of the eMID dataset

Let us, now, test our recipes on the transaction-level data constituting the overnight segment of the Electronic Market for Interbank Deposits (eMID), a screen-based market for unsecured deposits^{8,37-39}. Starting from directed, weighted trades (lender, borrower, notional amount), we aggregate transactions within ISO weeks, symmetrize exposures, and then binarize them to obtain weekly undirected snapshots (see also Supplementary Note 5).

The performance of our Bayesian ERGs has been tested on the following properties: the *total number of links*, by comparing

$$L_{t+1} = \sum_{i=1}^N \sum_{j(>i)} a_{ij}^{t+1} \quad (30)$$

with $\langle L_{t+1} \rangle$; the *degree of each node*, by comparing

$$k_i^{t+1} = \sum_{j(\neq i)} a_{ij}^{t+1}, \quad \forall i \quad (31)$$

with $\langle k_i^{t+1} \rangle$, $\forall i$; the *true positive rate* (also known as *recall* or *sensitivity*)

$$\langle \text{TPR}_{t+1} \rangle = \frac{\langle \text{TP}_{t+1} \rangle}{L_{t+1}}, \quad (32)$$

the *positive predictive value* (also known as *precision*)

$$\langle \text{PPV}_{t+1} \rangle = \frac{\langle \text{TP}_{t+1} \rangle}{\langle L_{t+1} \rangle}, \quad (33)$$

the *true negative rate* (also known as *specificity*)

$$\langle \text{TN}_{t+1} \rangle = \frac{\langle \text{TN}_{t+1} \rangle}{V_{t+1} - L_{t+1}} \quad (34)$$

and the *accuracy*

$$\langle \text{ACC}_{t+1} \rangle = \frac{\langle \text{TP}_{t+1} \rangle + \langle \text{TN}_{t+1} \rangle}{V_{t+1}}, \quad (35)$$

quantifying the ability of an algorithm in capturing both the number of true positives and the number of true negatives.

We have also approached the problem of reconstructing a network at time $t+1$, on the basis of the information at time t , from the perspective of link prediction (although the related problem of spurious links detection can be approached within the same framework as well, we have ignored it, here): to this aim, we have considered the top $\langle L_{t+1} \rangle$ links, i.e. the links characterized by the $\langle L_{t+1} \rangle$ largest probability coefficients, and checked their position via the *Jaccard Index*, reading

$$J_{t+1} = \frac{|E_{t+1}^{\text{pred}} \cap E_{t+1}^{\text{true}}|}{|E_{t+1}^{\text{pred}} \cup E_{t+1}^{\text{true}}|}, \quad (36)$$

where E_{t+1}^{true} is the ground-truth edge set, with $|E_{t+1}^{\text{true}}| = L_{t+1}$. We have also considered the Area Under the Receiver Operating Curve (AUROC), defined as the area under the curve obtained upon scattering the index

$$\text{TPR}_{t+1} = \frac{|E_{t+1}^{\text{pred}} \cap E_{t+1}^{\text{true}}|}{|E_{t+1}^{\text{true}}|} \quad (37)$$

(notice that such an index is nothing but the deterministic version of the one defined in Eq. (32)) versus the index

$$\text{FPR}_{t+1} = \frac{|E_{t+1}^{\text{pred}} \cap \overline{E_{t+1}^{\text{true}}}|}{|\overline{E_{t+1}^{\text{true}}}|} \quad (38)$$

(where $\overline{E_{t+1}^{\text{true}}}$ is the complementary set of E_{t+1}^{true} , with $|\overline{E_{t+1}^{\text{true}}}| = V_{t+1} - L_{t+1}$), as the list of links ranked in decreasing order of the chosen score is gone through: the AUROC quantifies the extent to which a given link prediction algorithm performs better than a random one - i.e. one that flips a coin to classify each non-observed link as either non-existent or missing.

Let us stress that the evaluation protocol above addresses two different aspects of the inference task, reflecting (the difference between) the reconstruction and the prediction tasks described in the Introduction: although the explicit sampling of individual network realizations is avoided in both cases, the quantities evaluated in a probabilistic fashion address the problem of reproducing the either global or local *number of links*, while the ones evaluated in a deterministic fashion address the problem of reproducing their *position*.

As a first analysis, we have considered the entire dataset, spanning the years 1999–2012, at the weekly time scale: the first three years have been used for the initial calibration of the prior; from 2002 on, we have performed the analysis over the remaining years, letting the prior be updated over a rolling window of three years. As all banks observed at each date are retained, here we have worked with an ‘unbalanced’ panel of nodes - i.e. their number is allowed to vary from snapshot to snapshot. For what concerns the BERM, the empirical prior on p^* , across the initial calibration window, is well described by a beta

distribution whose parameters read $\alpha \simeq 58.261$ and $\beta \simeq 541.543$ (see Supplementary Fig. 2). For what concerns the BFM, instead, the distribution of the z values across the same calibration window is well described by the Gamma distribution $\pi(z) = z^{\kappa-1} e^{-z/\theta} / \Gamma(\kappa) \theta^\kappa$ whose parameters read $\kappa \simeq 27.538$ and $\theta \simeq 0.007$ - and the Kolmogorov-Smirnov test⁴⁰ confirms it to be the distribution with the smallest number of parameters (only two) not rejected as a plausible parent distribution (see Supplementary Fig. 4). In both cases, the prior is updated by adding each, new, weekly value to it: from this perspective, we register no substantial difference between enriching it with ML or MAP estimates (see also the Methods section).

As Fig. 1 shows, both the BERM and the BFM are capable of recovering the total number of links. When coming to the degrees, instead, the BFM recovers their heterogeneity to a substantially better extent: more quantitatively, *i*) the relative error $RE_L = |L - \langle L \rangle|/L$ amounts, on average, to $\simeq 0.1$, under both the BFM and the BERM; *ii*) the average relative error $ARE_k = \sum_{i=1}^N (|k_i - \langle k_i \rangle|/k_i)/N$ is smaller under the BFM than under the BERM. Let us explicitly notice that, under the BFM, a deviation from the identity is visible for a limited set of nodes with very small degree. Such a mismatch is likely imputable to the underlying fitness ansatz, proxying the degree-controlling Lagrange multipliers with the node strengths: as observed in similar circumstances, this approximation works better for intermediate and large degrees than for small ones^{31,32,36}.

Although both models achieve a large ACC, amounting, on average, to $\simeq 0.80$, the BFM doubles the other scores, intended to quantify the ability of a model in recovering the position of connections - and not just that of the missing ones; the explanation of such a result lies in the evidence that the BERM cannot predict too dense configurations, although it is not capable of treating different pairs of nodes in a different way. Figure 2 further refines the picture above, by plotting the TPR, the JI and the AUROC for the BFM: while the TPR amounts, on average, to $\simeq 0.40$, the JI amounts, on average, to $\simeq 0.20$ - but it should be noticed that the TPR of the Bayesian Fitness Model is twice the one achieved in ref. 41 by the Directed Binary Configuration Model, implemented to carry out the in-sample version of the same exercise, on the same dataset. We explicitly stress that these ranking-based diagnostics are meaningful only for the BFM, whose heterogeneous set of probability coefficients induces a non-trivial ordering of the candidate links; the BERM, instead, would assign the same score to each pair, therefore inducing a 'random' classification.

A simpler version of our analysis (carried out by considering a 'balanced' panel of 73 banks across the weeks of 2002 - i.e. retaining only the nodes that appear in all snapshots of both the calibration and test period) is described in Supplementary Notes 2 and 3.

Self-sustained inference of the eMID dataset

The previous exercise was carried out by employing A_0 (where $t = 0$ indicates the last week of 2001) to predict A_1 , A_1 to predict A_2 and so on. Let us, now, test the capability of our inference procedure to *self-sustain* itself, i.e. to operate recursively without accessing any other adjacency matrix beyond the one employed for initialization. In this regime, both the prior on the parameter and the network representation are updated in a fully Bayesian fashion: the empirical prior keeps rolling over a three-year window but each new parameter value is estimated from inferred quantities and the matrix Q_t replaces A_t in the predictive step.

More specifically, we still consider the entire dataset by letting the prior roll over a time window of three years but *i*) determine each, new, weekly value of the parameter by solving Eq. (29) where L_t^* is replaced by $\langle L_t \rangle$, i.e. an estimate itself; *ii*) substitute the generic coefficient q_{ij}^{t+1} with the expression

$$r_{ij}^{t+1} = \int_0^{+\infty} \left(\frac{z s_i^{t+1} s_j^{t+1}}{1 + z s_i^{t+1} s_j^{t+1}} \right) \frac{P(Q_t|z)\pi(z)}{P(Q_t)} dz, \quad (39)$$

where A_t is replaced by Q_t , again an estimate itself: stated otherwise, we employ A_0 to predict Q_1 , Q_1 to predict Q_2 and so on.

As Fig. 3 shows, such a 'self-sustained' inference highlights patterns that are very similar to those in Fig. 1, an evidence revealing the accuracy of our algorithm to carry out an out-of-sample network reconstruction in presence of very little information (see also Supplementary Note 6). To provide a more quantitative assessment of such an agreement, let us scatter the values of the Kullback-Leibler divergence between the adjacency matrix $A = \{a_{ij}\}$ and its ensemble average $Q = \{q_{ij}\}$ at time t , i.e.

$$D_{KL}(A_t||Q_t) = \sum_{i=1}^N \sum_{j(>i)} a_{ij}^t \ln \left(\frac{a_{ij}^t}{q_{ij}^t} \right), \quad (40)$$

versus the values of the Kullback-Leibler divergence between the adjacency matrix $A = \{a_{ij}\}$ and its 'self-sustained', inferred version $R = \{r_{ij}\}$ at time t , i.e.,

$$D_{KL}(A_t||R_t) = \sum_{i=1}^N \sum_{j(>i)} a_{ij}^t \ln \left(\frac{a_{ij}^t}{r_{ij}^t} \right); \quad (41)$$

let us stress that the quantities above are not intended to represent Kullback-Leibler divergences between joint distributions: they are, in fact, *defined* in an element-wise fashion to compare all empirical entries with their reconstructed counterparts at once - in the same spirit of the discrepancy measure employed in ref. 42.

As evident from the figure, the two series of weekly values closely resemble each other. An even more explicit representation of such an agreement is provided in Fig. 4, depicting the empirical adjacency matrix of the week #20 of the year 2007, its ensemble average Q and its 'self-sustained' inferred version R .

Out-of-sample versus in-sample reconstruction

A particularly stringent test of our predictive scheme is obtained by contrasting it with an in-sample reconstruction that is allowed to use the quantity that our method must, instead, infer. More specifically, we compare the set of r_{ij} s with the set of probability coefficients returned by the dcGM, calibrated by employing the information about the total number of links per (weekly) snapshot. Despite the 'informational advantage' of the dcGM, the predictive performances of the two models remain remarkably close: as Fig. 5 shows, upon averaging over 561 weekly snapshots we obtain $\overline{TPR}_{\text{Bayes}} = 0.2498$ versus $\overline{TPR}_{\text{dcGM}} = 0.2494$, with the Bayesian predictor outperforming the dcGM on the $\simeq 54\%$ of snapshots. For what concerns the remaining scores, we have $\overline{PPV}_{\text{Bayes}} = 0.2495$ versus $\overline{PPV}_{\text{dcGM}} = 0.2494$, $\overline{TNR}_{\text{Bayes}} = 0.9124$ versus $\overline{TNR}_{\text{dcGM}} = 0.9122$ and $\overline{ACC}_{\text{Bayes}} = 0.8429$ versus $\overline{ACC}_{\text{dcGM}} = 0.8432$; degree-level errors, instead, remain slightly smaller for the dcGM, with $\overline{ARE}_k^{\text{Bayes}} = 0.8999$ versus $\overline{ARE}_k^{\text{dcGM}} = 0.8624$ and $\overline{MRE}_k^{\text{Bayes}} = 14.6836$ versus $\overline{MRE}_k^{\text{dcGM}} = 14.6267$ (with $\overline{MRE}_k = \max_{i=1}^N \{|k_i - \langle k_i \rangle|/k_i\}$), although the Bayesian predictor is still better for a non-negligible fraction of weeks, i.e. on the $\simeq 46\%$ of the snapshots.

Discussion

In the economic and financial domain, reconstruction methods based on Shannon entropy maximization have been widely studied from a frequentist perspective^{15,23,36}. Within such a framework, fitness-based models provide a powerful approach to reconstructing the structure of complex networks from partial information: among them, the dcGM offers a minimal, yet effective, parametrization for estimating link probabilities that combines node-level attributes with a global parameter.

The dcGM accurately captures many structural features of real-world networks but relies on a point-wise, maximum-of-the-likelihood estimation

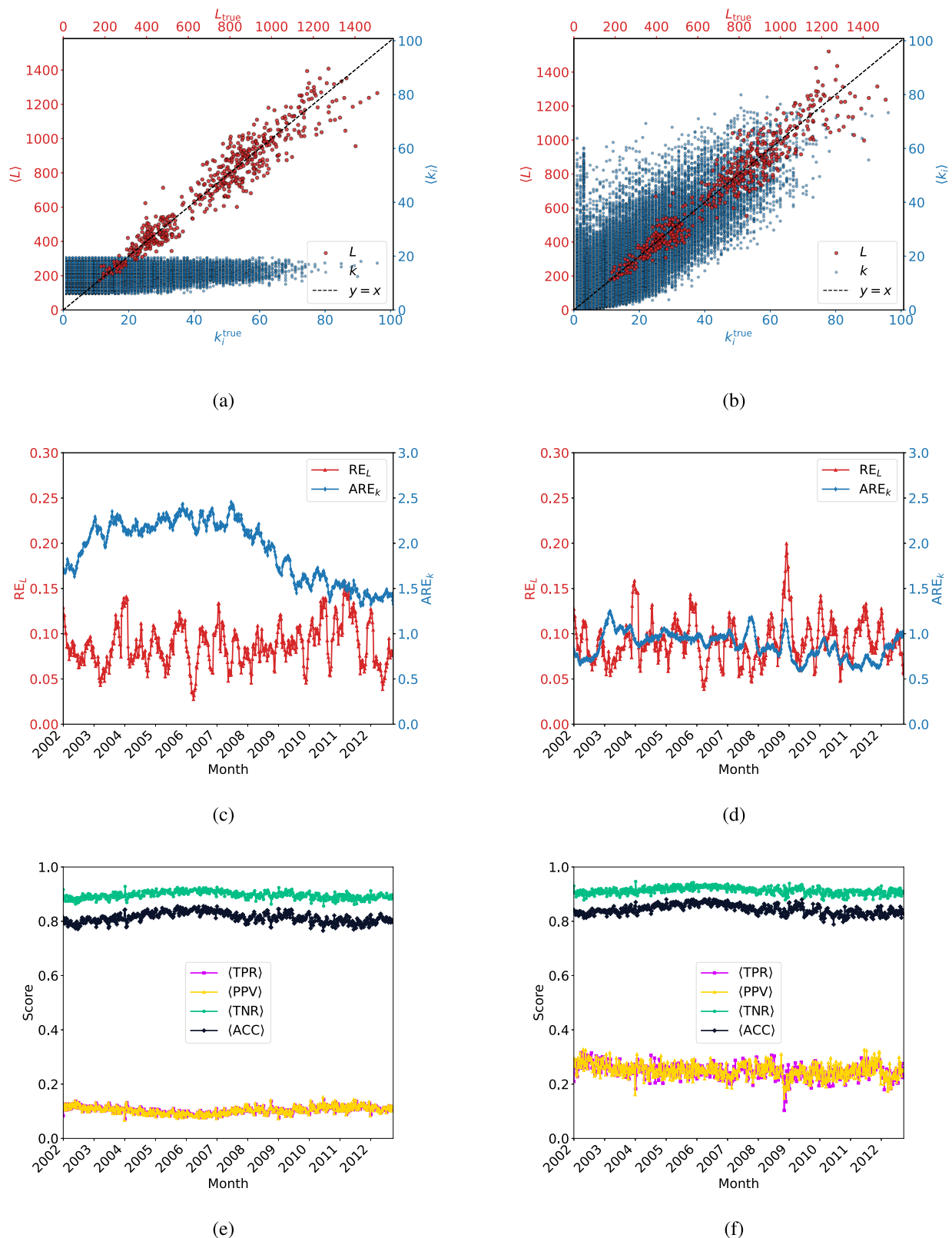


Fig. 1 | Out-of-sample reconstruction performance of the BERM and BFM.

Panels **a** and **b**: empirical values of the total number of links (red) and node degrees (blue) scattered versus the predicted ones, pooled across the weeks constituting our dataset; the dashed line marks the identity. Panels **c** and **d**: evolution of the relative error on the total number of links (red) and the average relative error on the nodes degrees (blue), across the weeks constituting our dataset. Panels **e** and **f**: evolution of

the $\langle \text{TPR} \rangle$, the $\langle \text{PPV} \rangle$, the $\langle \text{TNR} \rangle$ and the $\langle \text{ACC} \rangle$ across the weeks constituting our dataset. Panels **a**, **c** and **e** refer to the BERM, while panels **b**, **d** and **f** refer to the BFM: while both models recover the total number of links and achieve a large $\langle \text{ACC} \rangle$ score, driven by the large value of the $\langle \text{TNR} \rangle$, only the BFM is capable of recovering the degree sequence to an acceptable degree of accuracy - as well as more than doubling the other scores.

Fig. 2 | Ranking-based performance of the BFM. Panel a: evolution of the TPR, the JI and AUROC across the weeks constituting our dataset. The markers report pointwise weekly values; no error bars are shown. Panel b: ROC curves for all snapshots. The purple one represents the average ROC, obtained by interpolating each snapshot-specific ROC on a common grid of FPR values and averaging the corresponding TPR values. The dashed diagonal line represents the random-classifier benchmark. These ranking-based diagnostics are meaningful only for the BFM, that induces a non-trivial ordering of candidate links.

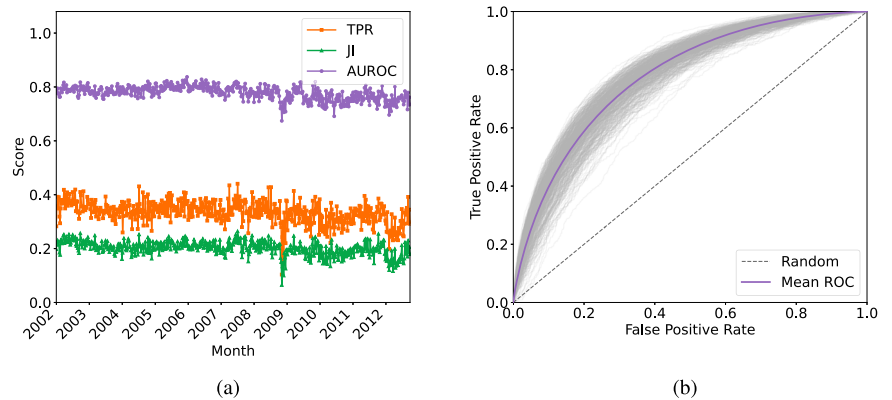


Fig. 3 | Comparison between direct and self-sustained reconstruction. Panel a: values of the Kullback–Leibler divergence between $\mathbf{A}$ and its ensemble average $\mathbf{Q}$ scattered versus the values of the Kullback–Leibler divergence between $\mathbf{A}$ and its 'self-sustained' inferred version $\mathbf{R}$, pooled across the weeks constituting our dataset. Panel b: values of the total number of links (red) and node degrees (blue) predicted by employing $\mathbf{Q}$ scattered versus the values predicted by employing $\mathbf{R}$, pooled across the weeks constituting our dataset. Both plots confirm that $\mathbf{Q}$ represents a reliable surrogate of $\mathbf{A}$ – in fact, so accurate to constitute a valid prior for subsequent inference.

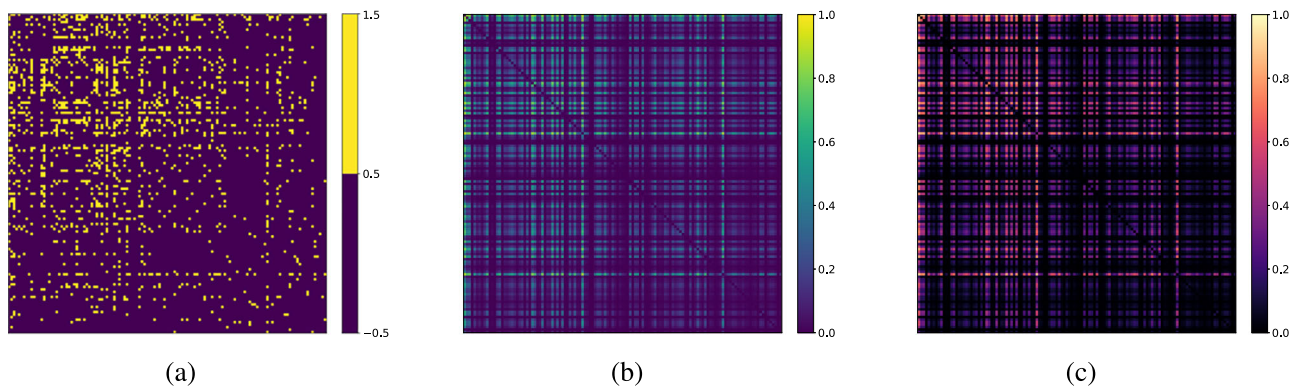
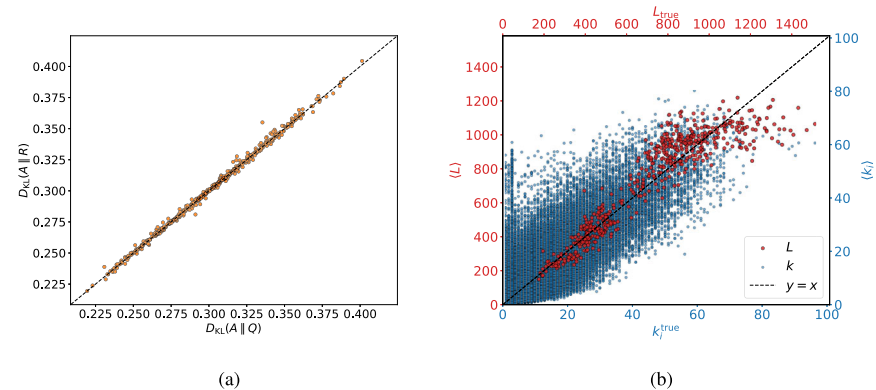


Fig. 4 | Observed adjacency and Bayesian reconstructions for a representative snapshot. Panel a: empirical adjacency matrix $\mathbf{A}_{t+1}$ corresponding to the week #20 of the year 2007. Panel b: ensemble average of $\mathbf{A}_{t+1}$, i.e. $\mathbf{Q}_{t+1}$. Panel c: 'self-sustained', inferred version of $\mathbf{A}_{t+1}$, i.e. $\mathbf{R}_{t+1}$. While $\mathbf{Q}_{t+1}$ needs the information

provided by $\mathbf{A}_t$, $\mathbf{R}_{t+1}$ 'only' needs the information provided by $\mathbf{Q}_t$, i.e. an estimate of $\mathbf{A}_t$. More quantitatively, $2\sum_{i=1}^N \sum_{j(i>j)} |q_{ij} - r_{ij}|/N(N-1) \simeq 0.006$ and $2\sum_{i=1}^N \sum_{j(i>j)} |a_{ij} - q_{ij}|/N(N-1) \simeq 0.152 \simeq 2\sum_{i=1}^N \sum_{j(i>j)} |a_{ij} - r_{ij}|/N(N-1)$.

procedure that offers no principled way to incorporate temporal dependencies: in other words, the dcGM can be employed to carry out an *in-sample* network reconstruction.

Related predictive tasks have been studied within the literature on temporal networks: a first example is provided by Temporal Exponential Random Graphs that, however, can be employed to predict future edges only in case a sequence of empirical networks with the same set of nodes is available for parameter estimation^{43,44}; other examples are provided by machine learning and embedding-based approaches that, however, operate under a richer informational regime than the one considered here, requiring fully observable past snapshots, stable representations of features, ad-hoc

mechanisms to handle nodes entering or leaving the system, validation procedures for tuning hyperparameters^{24–26}.

Against this background, building on the approach developed in refs. 42,45,46, our contribution addresses the problem of carrying out an *out-of-sample* network reconstruction, by presenting a fully Bayesian framework to infer a network structure from a previous observation. As our tests reveal, our reconstruction procedure is capable of 'self-sustaining' itself, each predicted configuration becoming a reliable prior for inference at the subsequent step. In practice, after the lastly observed snapshot is used, no additional topological information is fed into the algorithm and the reconstruction is propagated forward in time solely

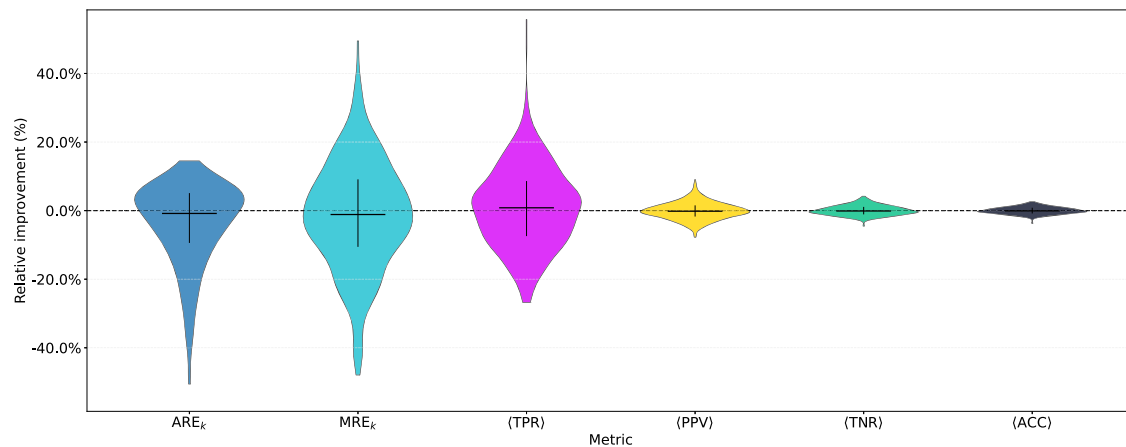


Fig. 5 | Metric-specific improvement of the ‘self-sustained’ BFM over the in-sample dcGM. For each snapshot and metric m , we define the score I_m in two different ways, i.e. as $I_m = (m_{\text{dcGM}} - m_{\text{Bayes}})/m_{\text{dcGM}}$ for the ARE_k and the MRE_k and as $I_m = (m_{\text{Bayes}} - m_{\text{dcGM}})/m_{\text{dcGM}}$ for the (TPR), the (PPV), the (TNR) and the (ACC): in both cases, values above the 0% dashed line indicate that the Bayesian

predictor performs better than the (in-sample) dcGM. Each violin plot summarizes the distribution of the improvement, showing that our fully predictive procedure frequently matches, and sometimes exceeds, the in-sample reconstruction calibrated by taking L as input at each time step.

through the posterior predictive distribution: such an experiment probes the information-retention capability of the model, as prediction errors are allowed to accumulate while our procedure attempts to reconstruct future configurations at increasingly large temporal distances from the lastly observed topology.

Methods

Data preprocessing

We analyze the binary undirected representation of eMID across the years 1999–2012: weekly snapshots are built from all transactions settled within ISO weeks; self-loops and multi-edges are discarded. For the unbalanced specification, the empirical prior is calibrated on the years 1999–2001 while the out-of-sample analysis is performed across the years 2002–2012, retaining all the banks that are active in each week. For the balanced specification, discussed in Supplementary Note 2 (BERM) and Supplementary Note 3 (BFM), the empirical prior is calibrated on the years 1999–2001 and the node set is restricted to the banks that are active in all weekly snapshots from 1999 to 2002; the out-of-sample analysis is, then, performed only over the year 2002. A complete description of the raw records, weekly aggregation, symmetrization and binarization steps is provided in Supplementary Note 5.

Prior calibration

Within each rolling window of three years, we compute a snapshot-specific point estimate for each parameter and, then, fit the empirical prior on the resulting set of values. Such a set is augmented via the jack-knife procedure, i.e. by removing one trading day at a time from each weekly aggregation and repeating the point-estimation step.

For what concerns the BERM, the ML estimate reads $p^* = L^*/V^*$; for what concerns the BFM, instead, the ML estimate is obtained by solving Eq. (24): in practice, we solve the corresponding one-dimensional problem via Newton iterations with backtracking, using a tolerance of 10^{-8} and a maximum number of 80 iterations. The corresponding fits on the initial calibration window are shown in Supplementary Note 2 for the BERM and in Supplementary Note 3 for the BFM.

Evaluation of the posterior predictive distribution

Adopting the Beta distribution as a (conjugate) prior allows us to treat the posterior predictive distribution of the BERM analytically.

For what concerns the BFM, instead, the posterior predictive distribution must be evaluated numerically. To this aim, we pose $u = \ln z$ and perform the calculation in such a logarithmic coordinate, which maps the

original domain $(0, +\infty)$ onto $\mathbb{R}$ and prevents numerical overflow when z becomes large.

A first integration scheme is the one named *Gauss-Hermite quadrature*⁴⁷: after having ‘gaussianized’ the integrand and evaluated it in correspondence of the posterior mode, the predictive probability reduces to a weighted sum over K quadrature nodes; for our experiments, we have used $K = 25$.

An alternative integration scheme is the one named *slice sampling*⁴⁸; it samples the log-posterior M times without requiring the computation of the normalization constant and providing an estimator that converges to the true distribution as M increases; for our experiments, $M = 3000$ values are sampled after a burn-in phase of 600 steps, thus yielding the estimator

$$q_{ij}^{t+1} \simeq \frac{1}{M} \sum_{m=1}^M \frac{e^{u^{(m)} s_i^{t+1} s_j^{t+1}}}{1 + e^{u^{(m)} s_i^{t+1} s_j^{t+1}}} \quad (42)$$

While the snapshot-specific computational cost of the first scheme scales as $O(KN^2)$, the one of the second scheme scales as $O(MN^2)$: as the Gauss-Hermite quadrature one is computationally cheaper, it is employed as the default method (see also Supplementary Note 4).

Data availability

Raw eMID data are subject to restrictions and are not publicly available, as they contain confidential interbank transaction records. Access to the raw data can be requested from the corresponding author; requests will be forwarded to the data provider and considered subject to its approval and to the applicable data-use restrictions. Source data are provided with this paper. Source data for Figs. 1–5 are provided as Supplementary Data 1–5, respectively. Source data for Supplementary Figs. 1–5, 7 and 8 are provided as Supplementary Data S1–S5, Supplementary Data S7 and Supplementary Data S8, respectively. Supplementary Fig. 6 is illustrative and has no associated source-data file.

Code availability

The Python package named OR4CLE (*Out-of-sample Bayesian Reconstruction 4 Complex Networks*), implementing the algorithms described in the main text, is available on PyPI, on GitHub at <https://github.com/mattiamarzi/OR4CLE>, and on Zenodo⁴⁹.

Received: 8 December 2025; Accepted: 8 June 2026;

Published online: 19 June 2026

References

- Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc. Natl. Acad. Sci. USA*. **103**, 2015–2020 (2006).
- Barrat, A., Barthélemy, M., Vespignani, A. *Dynamical Processes on Complex Networks*. Cambridge University Press: Cambridge, 2008.
- Newman, M. E. J. *Networks: An Introduction*. Oxford University Press: Oxford, 2010.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
- Squartini, T., van Lelyveld, I. & Garlaschelli, D. Early-warning signals of topological collapse in interbank networks. *Sci. Rep.* **3**, 3357 (2013).
- Battiston, S. et al. Complexity theory and financial regulation. *Science* **351**, 818–819 (2016).
- Bardoscia, M., Battiston, S., Caccioli, F. & Caldarelli, G. Pathways towards instability in financial networks. *Nat. Commun.* **8**, 14416 (2017).
- Macchiati, V., Marchese, E., Mazzarisi, P., Garlaschelli, D. & Squartini, T. Spectral signatures of structural change in financial networks. *Chaos Solit. Fractals* **193**, 116065 (2025).
- Park, J. & Newman, M. E. J. Statistical mechanics of networks. *Phys. Rev. E* **70**, 066117 (2004).
- Bianconi, G. Entropy of network ensembles. *Phys. Rev. E* **79**, 036114 (2009).
- Squartini, T. & Garlaschelli, D. Analytical maximum-likelihood method to detect patterns in real networks. *N. J. Phys.* **13**, 083001 (2011).
- Fronczak, A. & Fronczak, P. Statistical mechanics of the international trade network: Structural correlations and modeling. *Phys. Rev. E* **85**, 056113 (2012).
- Squartini, T., Mastrandrea, R. & Garlaschelli, D. Unbiased sampling of network ensembles. *N. J. Phys.* **17**, 023052 (2015).
- Saracco, F., Di Clemente, R., Gabrielli, A. & Squartini, T. Randomizing bipartite networks: The case of the world trade web. *Sci. Rep.* **5**, 10595 (2015).
- Cimini, G. et al. The statistical physics of real-world networks. *Nat. Rev. Phys.* **1**, 58–71 (2019).
- Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630 (1957).
- Squartini, T. & Garlaschelli, D. *Maximum-Entropy Networks*. SpringerBriefs in Complexity (Springer International Publishing, Cham, 2017).
- Handcock, M. S. Statistical models for social networks: Inference and degeneracy. In *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers* (The National Academies Press, 2003).
- Garlaschelli, D. & Loffredo, M. I. Maximum likelihood: extracting unbiased information from complex networks. *Phys. Rev. E* **78**, 015101 (2008).
- van Duijn, M. A. J., Gile, K. J. & Handcock, M. S. A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models. *Soc. Netw.* **31**, 52–62 (2009).
- Vallarano, N. et al. Fast and scalable likelihood maximization for exponential random graph models with local constraints. *Sci. Rep.* **11**, 15227 (2021).
- Di Vecce, M., Garlaschelli, D. & Squartini, T. Deterministic, quenched, and annealed parameter estimation for heterogeneous network models. *Phys. Rev. E* **108**, 054301 (2023).
- Squartini, T., Caldarelli, G., Cimini, G., Gabrielli, A. & Garlaschelli, D. Reconstruction methods for networks: the case of economic and financial systems. *Phys. Rep.* **757**, 1–47 (2018).
- Zhou, T. Progresses and challenges in link prediction. *iScience* **24**, 103217 (2021).
- Lü, L. & Zhou, T. Link prediction in complex networks: A survey. *Phys. A* **390**, 1150–1170 (2011).
- Santucci, F., Cimini, G. & Squartini, T. Missing links prediction: comparing machine learning with physics-rooted approaches. Preprint at <https://arxiv.org/abs/2601.23061> (2026).
- Bardoscia, M. et al. The physics of financial networks. *Nat. Rev. Phys.* **3**, 490–507 (2021).
- Ialongo, L. N. et al. Reconstructing firm-level interactions in the dutch input–output network from production constraints. *Sci. Rep.* **12**, 11847 (2022).
- Mungo, L., Lafond, F., Astudillo-Estévez, P. & Farmer, J. D. Reconstructing production networks using machine learning. *J. Econ. Dyn. Control* **148**, 104607 (2023).
- Gelman, A. et al. *Bayesian Data Analysis*, 3rd ed. (Chapman and Hall/CRC, New York, 2013).
- Cimini, G., Squartini, T., Garlaschelli, D. & Gabrielli, A. Systemic risk analysis on reconstructed economic and financial networks. *Sci. Rep.* **5**, 15758 (2015).
- Mazzarisi, P. & Lillo, F. Methods for reconstructing interbank networks from limited information: a comparison. In *Econophysics and Sociophysics: Recent Progress and Future Directions*, New Economic Windows (eds, Abergel, F. et al.) 201–215 (Springer, Cham, 2017).
- Anand, K. et al. The missing links: a global study on uncovering financial network structures from partial data. *J. Financ. Stab.* **35**, 107–119 (2018).
- Lebacher, M., Cook, S., Klein, N. & Kauermann, G. In search of lost edges: a case study on reconstructing financial networks. *J. Netw. Theory Financ.* **5**, 29–61 (2019).
- Ramadijah, A., Caccioli, F. & Fricke, D. Reconstructing and stress testing credit networks. *J. Econ. Dyn. Control* **111**, 103817 (2020).
- Cimini, G., Mastrandrea, R. & Squartini, T. *Reconstructing Networks. Elements in the Structure and Dynamics of Complex Networks* (Cambridge Univ. Press, 2021).
- Marzi, M., Giuffrida, F., Garlaschelli, D. & Squartini, T. Reproducing the first and second moments of empirical degree distributions. *Phys. Rev. Res.* **8**, 013047 (2026).
- Iori, G., Masi, G. D., Precup, O., Gabbi, G. & Caldarelli, G. A network analysis of the italian overnight money market. *J. Econ. Dyn. Control* **32**, 259–278 (2008).
- Finger, K., Fricke, D. & Lux, T. *Network Analysis of the e-mid Overnight Money Market: The Informational Value of Different Aggregation Levels for Intrinsic Dynamic processes*. Kiel Working Paper 1782, (Kiel Institute for the World Economy (IfW Kiel), 2012).
- Massey, F. J. The kolmogorov-smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **46**, 68–78 (1951).
- Parisi, F., Caldarelli, G. & Squartini, T. Entropy-based approach to missing-links prediction. *Appl. Netw. Sci.* **3**, 17 (2018).
- Young, J.-G., Cantwell, G. T. & Newman, M. E. J. Bayesian inference of network structure from unreliable data. *J. Complex Netw.* **8**, cnaa046 (2020).
- Hanneke, S., Fu, W. & Xing, E. P. Discrete temporal models of social networks. *Electron. J. Stat.* **4**, 585–605 (2010).
- Krivitsky, P. N. & Handcock, M. S. A separable model for dynamic networks. *J. R. Stat. Soc. B*, 29–46 (2014).
- Peixoto, T. P. Reconstructing networks with unknown and heterogeneous errors. *Phys. Rev. X* **8**, 041011 (2018).
- Peixoto, T. P. Network reconstruction via the minimum description length principle. *Phys. Rev. X* **15**, 011065 (2025).
- Abramowitz, M. & Stegun, I. A. (eds.) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, vol. 55 of *Applied Mathematics Series* (National Bureau of Standards, Washington, D.C., 1964).
- Neal, R. M. Slice sampling. *Ann. Stat.* **31**, 705–767 (2003).
- Marzi, M. OR4CLE: out-of-sample Bayesian reconstruction 4 complex networks. <https://doi.org/10.5281/zenodo.20410903> (2026).

Acknowledgements

M.M. and T.S. acknowledge support from the project 'SoBigData.it - Strengthening the Italian RI for Social Mining and Big Data Analytics' - IR0000013 - CUP B53C22001760006, financed by European Union - Next Generation EU - National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR) - M4C2 I.3.1; T.S. acknowledges support from the projects 'RE-Net: Reconstructing economic networks: from physics to machine learning and back' - 2022MTBB22, Funded by the European Union Next Generation EU, PNRR Mission 4 Component 2 Investment 1.1, CUP: D53D23002330006; 'C2T - From Crises to Theory: towards a science of resilience and recovery for economic and financial systems' - P2022E93B8, Funded by the European Union Next Generation EU, PNRR Mission 4 Component 2 Investment 1.1, CUP: D53D23019330001.

Author contributions

Study conception and design: M.M., T.S. Analysis and interpretation of results: M.M., T.S. Draft manuscript preparation: M.M., T.S.

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/s42005-026-02725-1>.

Correspondence and requests for materials should be addressed to Mattia Marzi.

Peer review information *Communications Physics* thanks Tao Zhou, Giacomo Vaccario and the other, anonymous, reviewer for their contribution to the peer review of this work.

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