

<https://doi.org/10.1038/s42005-025-02403-8>

Indirect influence on network diffusion

Check for updates

Lluís Torres-Hugas , Jordi Duch & Sergio Gómez

Models of network diffusion typically rely on the Laplacian matrix, capturing interactions via direct connections. Beyond direct interactions, information in many systems can also flow via indirect pathways, where influence typically diminishes over distance. In this work, we analyze diffusion dynamics incorporating such indirect connections using the d -path Laplacian framework. We introduce a parameter, the indirect influence, based on the change in the second smallest eigenvalue of the generalized path Laplacian, to quantify the impact of these pathways on diffusion timescales relative to direct-only models. Using perturbation theory and mean-field approximations, we derive analytical expressions for the indirect influence in terms of structural properties of random networks. Theoretical predictions align well with numerical simulations, providing a phase diagram for when indirect influence becomes significant. We also identify a structural phase transition governed by the emergence of d -paths and derive the critical connection probability above which they dramatically alter diffusion. This study provides a quantitative understanding of how indirect pathways shape network dynamics and reveals their collective structural onset.

Diffusion processes on networks are fundamental in understanding a wide range of phenomena in complex networks¹; these include information spreading in social systems², disease propagation³, and transport dynamics in physical and biological systems^{4,5}. These processes are often modeled using the Laplacian matrix of the network⁶, which captures the dynamics of local interactions based on the structure of the direct connections between nodes.

In some systems, considering only the direct connections within a network fails to capture all diffusion pathways, and underlying mechanisms may alter how the spreading occurs. To address the limitations of simplifying complex networks as basic graphs, various extensions within the framework of higher-order networks have been proposed. These include multiplex networks, multilayer networks, hypergraphs, simplicial complexes, and others^{7–9}. These efforts to go beyond pairwise interactions often focus on incorporating group interactions. However, they still overlook an important feature of network dynamics where the strength of interaction gradually decreases as the separation between nodes increases in the underlying graph structure.

In particular, this drawback has been emphasized in the diffusion of information in social systems¹⁰, where the decaying influence between nodes is necessary to capture the underlying dynamics, making the introduction of indirect connections, or “weak ties”, particularly relevant. An indirect connection allows a weaker spread of information between two nodes, i and k , that are not directly connected but are connected through a third node j , in a path of length two. While j is implicitly involved in the process, it does so passively, without being influenced by the transmitted information. The importance of indirect connections has also been pointed out in

other contexts. For example, second-order neighbors in online networks can significantly influence political mobilization and social influence¹¹, and in ecological systems, indirect connections among species can shape coevolutionary dynamics and network structure¹².

One proposed approach to incorporate these indirect connections is through the use of the d -path Laplacian and the d -path adjacency matrices¹³, a natural generalization of the Laplacian and the adjacency matrices to include higher-order paths and establish the theoretical possibility of non-local diffusion, even demonstrating super-diffusive behavior under certain conditions for single-particle dynamics^{14,15}. Other options to incorporate indirect influence in networks include the multi-hop random walks model¹⁶, and the fractional Laplacian^{17,18}, which show important qualitative and quantitative differences with respect to the d -path Laplacian approach. Beyond diffusion, indirect influence can also affect other dynamics, such as synchronization¹⁹ and consensus models²⁰.

Indirect diffusion has proven effective in modeling real-world scenarios, e.g., capturing the long-range dispersal of pathogens via vectors in epidemics²¹, showing how peer pressure can shape consensus and leadership in social groups²², and accounting for the accelerated diffusion of innovations through socially-close peers in experiments²³. Despite its practical importance, the interplay between network structure and diffusion dynamics remains poorly understood, although it has proven central in other frameworks^{24,25}.

In this work, we extend standard network diffusion models to quantitatively assess the impact of influence propagating indirectly along network paths. Employing the d -path Laplacian framework, we introduce the indirect influence parameter ζ to precisely measure the change in diffusion

timescales caused by these non-local interactions, linking it directly to the spectral properties of the generalized Laplacian. We provide a theoretical foundation for this parameter, deriving analytical expressions for ζ in random networks using perturbation theory and mean-field approximations, leading to a predictive phase diagram validated by numerical simulations. Furthermore, we uncover and characterize a structural phase transition associated with the emergence of d -paths, deriving its critical threshold $p_c(N)$ and revealing its fundamental role in altering diffusion dynamics. This framework offers a robust method for analyzing the contribution of indirect pathways in complex network processes.

Results

Indirect influence

To analyze diffusion dynamics incorporating non-local interactions, we employ the d -path Laplacian framework¹³. This approach extends the standard Laplacian model by considering pathways of length $d > 1$ between nodes. The framework uses d -path adjacency matrices, $A^{(d)}$, where $(A^{(d)})_{ij} = a_{ij}^{(d)} = 1$ if the shortest path distance between nodes i and j is exactly d , and 0 otherwise. From these, the corresponding d -path Laplacian, $L^{(d)} = K^{(d)} - A^{(d)}$, is constructed, where $K^{(d)}$ is the diagonal matrix of d -path degrees $k_i^{(d)} = \sum_j a_{ij}^{(d)}$ (see Methods).

To capture diffusion across paths of all lengths up to the network diameter D , while allowing for influence to decay with distance, we consider the path Laplacian as:

$$\hat{L} = \sum_{d=1}^D d^{-\alpha} L^{(d)}, \quad (1)$$

where $\alpha \geq 0$ is a parameter controlling the decay rate of influence with path length d . Setting $\alpha = 0$ implies uniform influence regardless of distance, while a large α recovers the standard local diffusion dynamics dominated by $L^{(1)} = L$. The path Laplacian \hat{L} retains the essential properties of the standard Laplacian, being symmetric and positive semi-definite. Notice that this path Laplacian can be interpreted as the Laplacian matrix of the path network, see Fig. 1, where each link is weighted according to its diffusion rate $d^{-\alpha}$.

The diffusion dynamics on a network governed by the path Laplacian are described by:

$$\dot{\mathbf{x}}(t) = -\hat{L}\mathbf{x}(t), \quad (2)$$

where $\mathbf{x}(t)$ is the vector of node states at time t . Assuming the network is connected, the general solution is given by:

$$\mathbf{x}(t) = \sum_{i=1}^N C_i e^{-\hat{\lambda}_i t} \hat{\mathbf{v}}_i \quad (3)$$

where $0 = \hat{\lambda}_1 < \hat{\lambda}_2 \leq \dots \leq \hat{\lambda}_N$ are the eigenvalues and $\hat{\mathbf{v}}_i$ are the corresponding eigenvectors of \hat{L} , and C_i are constants determined by initial

conditions. The convergence to the equilibrium state, where $x_i = x_j$ for all i, j , is governed by the second smallest eigenvalue, $\hat{\lambda}_2$, often called the algebraic connectivity or Fiedler value^{26,27}. The characteristic diffusion time is $\tau = 1/\hat{\lambda}_2$; a larger $\hat{\lambda}_2$ means faster convergence.

In contrast, standard diffusion models only consider direct connections, described by $\dot{\mathbf{x}} = -L\mathbf{x}$, where $L = L^{(1)}$, with a diffusion time $\tau = 1/\lambda_2$, where λ_2 is the second smallest eigenvalue of the standard Laplacian L . To quantify the impact of including indirect pathways on the diffusion timescale, we introduce the *indirect influence*, defined as the normalized change in the algebraic connectivity:

$$\zeta = \frac{\hat{\lambda}_2 - \lambda_2}{N}. \quad (4)$$

This parameter ranges from 0 to 1, since N is the maximum value that any eigenvalue can take in a Laplacian matrix²⁶, and $\hat{\lambda}_2 \geq \lambda_2$ since the original 1-path network is a (weighted) spanning graph of the path network²⁸. A value of $\zeta = 0$ indicates that indirect connections do not alter the diffusion timescale compared to the direct-only model, while $\zeta > 0$ means that indirect pathways accelerate the diffusion process, with larger values indicating a greater impact. This parameter provides a quantitative measure to address when and how significantly indirect connections reshape network diffusion dynamics.

Theoretical approximation of indirect influence

To understand how network structure influences the indirect influence, we derive an analytical expression for it, particularly focusing on general random networks characterized by their size N and expected connection probability $p \equiv \langle a_{ij} \rangle$. We begin by considering the regime where indirect connections act as a perturbation to the direct diffusion process, which corresponds to a strong decay of influence with distance, $\alpha \gg 1$.

In the large α limit, the path Laplacian \hat{L} is dominated by the $d = 1$ term. We start by including the most significant indirect contribution, the $d = 2$ term, scaled by a small parameter $\epsilon = 2^{-\alpha} \ll 1$:

$$\hat{L} = L + 2^{-\alpha} L^{(2)} = L + \epsilon L^{(2)}. \quad (5)$$

We can now apply standard perturbation theory to find the change in the second smallest eigenvalue λ_2 of the standard Laplacian L . Let λ_2 and \mathbf{v}_2 be the eigenvalue and corresponding normalized eigenvector of L ($L\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, $\mathbf{v}_2^T\mathbf{v}_2 = 1$). The perturbed eigenvalue $\hat{\lambda}_2$ of \hat{L} is, to first order in ϵ (see Methods):

$$\hat{\lambda}_2 = \lambda_2 + \epsilon \mathbf{v}_2^T L^{(2)} \mathbf{v}_2 + O(\epsilon^2). \quad (6)$$

Substituting this into the definition of the indirect influence, we obtain:

$$\zeta \approx \frac{2^{-\alpha} \mathbf{v}_2^T L^{(2)} \mathbf{v}_2}{N}, \quad (7)$$

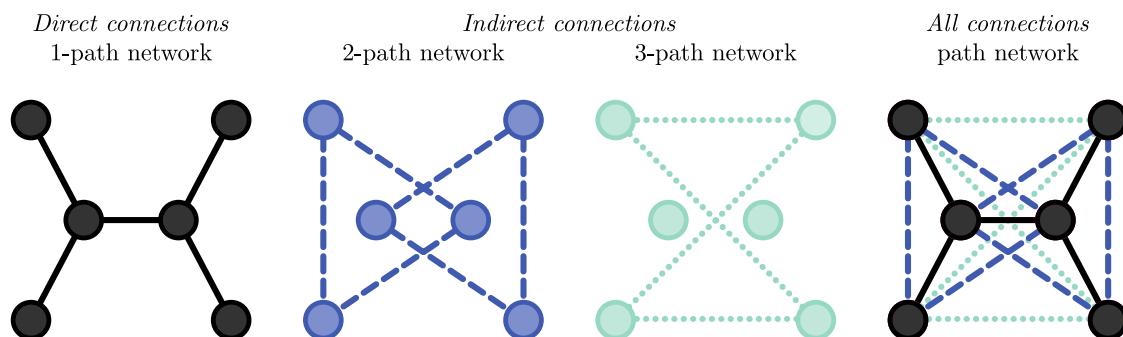


Fig. 1 | Diagrams of d -path networks and the path network. The figure shows the d -path networks derived from a base network and the corresponding path network that combines all d -paths.

to first order in ϵ . This expression quantifies the indirect influence arising from 2-paths in the perturbative regime, but it still depends on the specific eigenvector \mathbf{v}_2 and the 2-path Laplacian $L^{(2)}$ of a given network.

To obtain an expression dependent only on average network properties, we employ a mean-field approach to approximate the expectation value of the quadratic form $\mathbf{v}_2^T L^{(2)} \mathbf{v}_2$ over an ensemble of general random networks. Assuming the entries of the eigenvector \mathbf{v}_2 are uncorrelated with the entries of $L^{(2)}$ and are independent identically distributed random variables satisfying $\sum_{i=1}^N (\mathbf{v}_2)_i^2 = 1$ and $\sum_{i=1}^N (\mathbf{v}_2)_i = 0^{26}$, we can approximate the expectation:

$$\begin{aligned} \langle \mathbf{v}_2^T L^{(2)} \mathbf{v}_2 \rangle &= \left\langle \sum_{i=1}^N k_i^{(2)} (\mathbf{v}_2)_i^2 \right\rangle - \left\langle \sum_{i=1}^N \sum_{j=1}^N a_{ij}^{(2)} (\mathbf{v}_2)_i (\mathbf{v}_2)_j \right\rangle \\ &\approx N \langle k_i^{(2)} \rangle \langle (\mathbf{v}_2)_i^2 \rangle - N(N-1) \langle a_{ij}^{(2)} \rangle \langle (\mathbf{v}_2)_i (\mathbf{v}_2)_j \rangle. \end{aligned} \quad (8)$$

Using the properties $\langle (\mathbf{v}_2)_i^2 \rangle = \frac{1}{N}$ and $\langle (\mathbf{v}_2)_i (\mathbf{v}_2)_j \rangle = -\frac{1}{N(N-1)}$ (see Methods), this simplifies to:

$$\langle \mathbf{v}_2^T L^{(2)} \mathbf{v}_2 \rangle \approx \langle k_i^{(2)} \rangle + \langle a_{ij}^{(2)} \rangle. \quad (9)$$

Defining the expected connection probability of the 2-path network as $p^{(2)} \equiv \langle a_{ij}^{(2)} \rangle$, and noting that the expected 2-path degree is $\langle k^{(2)} \rangle = (N-1)p^{(2)}$, we arrive at:

$$\langle \mathbf{v}_2^T L^{(2)} \mathbf{v}_2 \rangle \approx N p^{(2)} \quad (10)$$

Substituting this mean-field result back into Eq. (7), the expected indirect influence due to 2-paths is approximately:

$$\zeta \approx 2^{-\alpha} p^{(2)}. \quad (11)$$

The expected 2-path connection probability $p^{(2)}$ can be expressed in terms of the direct expected connection probability p of the underlying network. A 2-path exists between nodes i and j if they are not directly connected ($a_{ij} = 0$) but there exists at least one intermediate node k such that $a_{ik} = 1$ and $a_{kj} = 1$. Using a mean-field argument for the existence probability over the $N-2$ potential intermediate nodes:

$$\begin{aligned} p^{(2)} &= P\{a_{ij}^{(2)} = 1\} = P\{a_{ij} = 0\} P\{\exists k \in G \mid a_{ik} a_{kj} = 1\} \\ &= P\{a_{ij} = 0\} \left(1 - P\{\forall k \in G \mid a_{ik} a_{kj} = 0\}\right) \\ &\approx P\{a_{ij} = 0\} \left(1 - (1 - P\{k \in G \mid a_{ik} a_{kj} = 1\})^{N-2}\right) \\ &= (1-p) \left(1 - (1-p^2)^{N-2}\right) \end{aligned} \quad (12)$$

This provides an explicit expression for $p^{(2)}$ based on the network parameters N and p .

The perturbation and mean-field approach can be extended recursively to account for contributions from d -paths with $d > 2$. Each d -path term contributes approximately $d^{-\alpha} p^{(d)}$ to the indirect influence ζ . The d -path expected connection probability $p^{(d)}$ can be approximated similarly to $p^{(2)}$, considering the condition that no shorter path exists between the nodes and that at least one path of length d exists (see Methods):

$$p^{(d)} \approx \left(1 - \sum_{i=1}^{d-1} p^{(i)}\right) \left(1 - (1-p^d)^{\binom{N-2}{d-1}}\right). \quad (13)$$

Therefore, the total indirect influence ζ can be approximated by summing the contributions from all relevant path lengths ($d = 2, \dots, D$):

$$\zeta(N, p, \alpha) \approx \sum_{d=2}^D d^{-\alpha} \left(1 - \sum_{i=1}^{d-1} p^{(i)}\right) \left(1 - (1-p^d)^{\binom{N-2}{d-1}}\right). \quad (14)$$

Figure 2 shows the behavior of these calculated expected d -path probabilities $p^{(d)}$ as a function of the direct connection probability p for both Erdős-Rényi (ER) and Random Regular (RR) networks. The theoretical predictions are compared with results obtained from direct numerical simulations. The plots demonstrate that the recursive mean-field calculation accurately captures the dependence of $p^{(d)}$, especially for smaller values of d . As d increases, the accuracy of the mean-field approximation slightly decreases, likely due to increasing correlations between paths that are ignored in the calculation.

Therefore, Eq. (14), combined with the recursive calculation of $p^{(d)}$ validated in Fig. 2, provides a theoretical framework to predict the total indirect influence ζ based on the structural network properties and the influence decay parameter α .

Phase diagram for indirect influence

Having derived the theoretical approximation for the indirect influence, we now test its validity by comparing it against direct numerical simulations across a wide range of network parameters. We computed ζ numerically for ensembles of Erdős-Rényi (ER) and Random Regular (RR) networks by calculating the second smallest eigenvalues $\hat{\lambda}_2$ and λ_2 of the full path Laplacian \hat{L} and the standard Laplacian L , respectively.

Figure 3 presents the phase diagram of the average indirect influence ζ in the parameter space defined by the connection probability p and the decay exponent α , for both ER and RR networks with $N = 1000$. The agreement between the theoretical predictions and the simulation results is remarkably

Fig. 2 | Expected connection probabilities of d -path networks relative to the 1-path network. Expected connection probabilities for different d -path networks $p^{(d)}$ as a function of the expected 1-path connection probability p , with $N = 1000$ nodes. **a** Erdős-Rényi (ER) networks and **b** Random Regular (RR) networks. Each point corresponds to the average over 50 different networks with error bars of one standard deviation, and the solid lines correspond to their respective analytical expression. The vertical gray dashed line marks the connectivity threshold $p < \log N/N$, below which the network is almost surely disconnected³² and therefore not considered in the analysis.

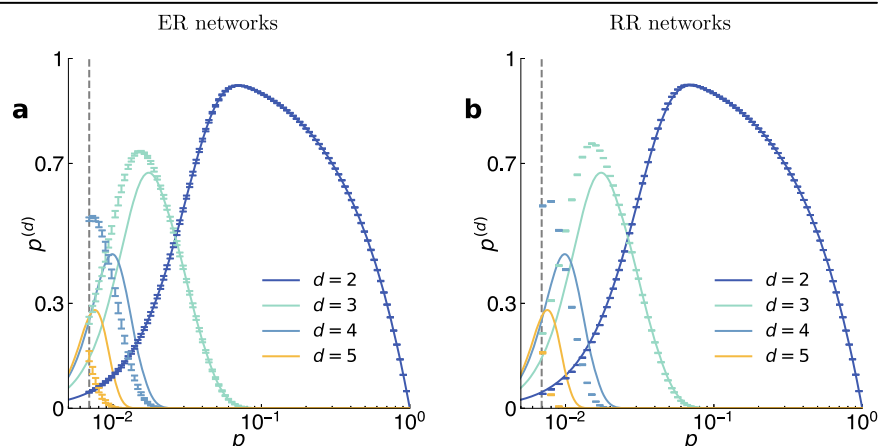
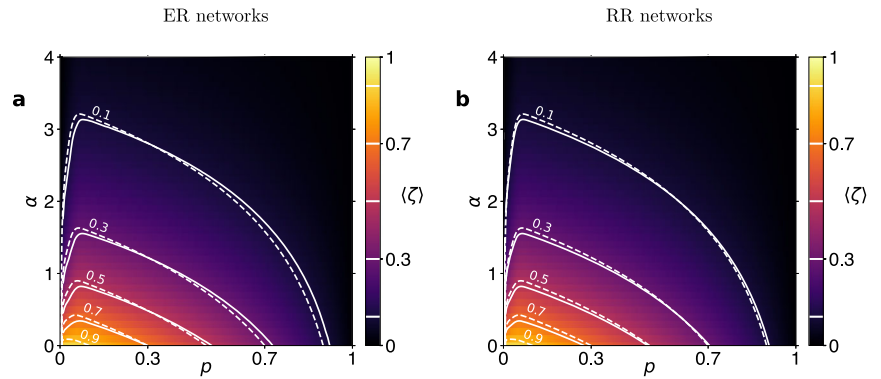


Fig. 3 | Average indirect influence across parameter space.

Average indirect influence (ζ) computed over all possible $d = 1, \dots, D$ paths for networks with $N = 1000$ nodes, as a function of the connection probability p and path decay rate α . **a** Erdős-Rényi (ER) networks. **b** Random Regular (RR) networks. Solid white lines correspond to the contour levels of the color map that corresponds to the simulations, and the dashed white lines correspond to the theoretical contour levels. We compute 100 different networks for each combination of α ranging from 0 to 4 in steps of 0.1 and p ranging from 0 to 1 in steps of 0.001. The colorbar represents the average indirect influence (ζ).



good across the entire phase space for both network types. This alignment holds even for small values of α , which are technically outside the $\alpha \gg 1$ regime where the perturbation theory was initially applied. This indicates that the mean-field approximation for the expected d -path probabilities $p^{(d)}$ and the additive contribution assumption in Eq. (14) effectively capture the collective impact of indirect pathways on diffusion dynamics.

The phase diagram provides valuable insights into when indirect influence becomes significant. As expected, ζ decreases as α increases, reflecting the diminished contribution of longer paths when influence decays rapidly. For a fixed α , ζ generally decreases with increasing p , as denser networks offer less potential indirect pathways. This phase diagram serves as a practical tool to determine, based on network structure and the assumed decay mechanism, whether incorporating indirect connections is likely to substantially alter the system's diffusion timescale compared to a standard direct-only model. Furthermore, the framework provides a quantitative measure, ζ , of the magnitude of this effect.

Analysis of the uniform weighting limit ($\alpha = 0$)

The theoretical framework based on perturbation theory assumes $\alpha \gg 1$. We now investigate the opposite limit, $\alpha = 0$, where influence does not decay with distance, i.e., $d^{-\alpha} = 1$ for all d . In this limit, the path Laplacian becomes $\hat{L} = \sum_{d=1}^D L^{(d)}$ and the perturbation approach is no longer valid, requiring an alternative analysis.

If we consider only direct and 2-path connections, the path Laplacian is $\hat{L} = L^{(1)} + L^{(2)}$. This structure resembles the Laplacian used in studies of diffusion on duplex networks^{24,29}. Analogously, we can approximate the combined 1-path and 2-path network as a single effective network. Since the edge sets of the 1-path and 2-path networks are disjoint, by definition of $A^{(2)}$, we approximate this effective network as an Erdős-Rényi network with N nodes and an effective connection probability $p_{\text{eff}} = p + p^{(2)}$, where p and $p^{(2)}$ are given by Eq. (12). The algebraic connectivity λ_2 for this truncated $\alpha = 0$ case can then be approximated by the known results for the algebraic connectivity of ER and RR networks^{30,31}. The indirect influence can be written as:

$$\zeta_{\alpha=0}^{\text{ER/RR}} = \frac{\lambda_2^{\text{ER}}(N, p + p^{(2)}) - \lambda_2^{\text{ER/RR}}(N, p)}{N}. \quad (15)$$

The approximation for large ($N \gg 1$) unweighted RR networks of degree $k = p(N - 1)$ is:

$$\lambda_2^{\text{RR}}(k) \approx k - 2\sqrt{k - 1}. \quad (16)$$

And for large unweighted ER networks is:

$$\begin{aligned} \lambda_2^{\text{ER}}(N, p) \approx & p(N - 1) - \sqrt{2p(1 - p)(N - 1) \log N} \\ & + \sqrt{\frac{(N - 1)p(1 - p)}{2 \log N} \log \sqrt{2\pi \log \left(\frac{N^2}{2\pi} \right)}} \\ & - \sqrt{\frac{(N - 1)p(1 - p)}{2 \log N} \gamma}, \end{aligned} \quad (17)$$

where p is the connection probability of the nodes and γ is the Euler-Mascheroni constant.

To assess the validity of this ER approximation for the combined 1-path and 2-path network, we examine its degree distribution in Fig. 4. While the average degree matches the expectation, the shape of the distribution differs from the theoretical binomial distribution predicted for an equivalent ER graph. For ER base networks, Fig. 4a, the combined network exhibits a broader tail, whereas for RR base networks, Fig. 4b, the distribution is narrower. These differences indicate that the combined 1-path and 2-path network structure deviates from a pure ER graph, which may introduce inaccuracies in the $\hat{\lambda}_2$ approximation based on λ_2^{ER} .

If we consider the full path Laplacian at $\alpha = 0$, assuming the underlying network is connected, the path network effectively becomes the complete graph K_N . This is because for any two nodes i, j in a connected graph, there exists a shortest path of some length $d \leq D$. The algebraic connectivity of the complete unweighted network is $\lambda_2 = N^2$. Therefore, in this limit, the indirect influence reaches its maximum possible value:

$$\zeta_{\alpha=0}^{\text{ER/RR}} = \frac{N - \lambda_2^{\text{ER/RR}}(N, p)}{N}. \quad (18)$$

Figure 5 compares the theoretical approximations for $\alpha = 0$ with simulation results. Overall, the results are in strong agreement. However, for ER networks, there are slight deviations that align with the differences in the degree distributions shown in Fig. 4. This analysis highlights that while the $\alpha = 0$ case represents maximal indirect influence, accurately predicting it requires either considering the full path network or carefully accounting for the structural specifics of truncated path networks beyond simple ER approximations.

Structural phase transition of indirect connections

The preceding analyses show that the significance of indirect influence ζ is closely related to the prevalence of d -paths, particularly those corresponding to non-negligible weights $d^{-\alpha}$. The sharp changes observed in the $p^{(d)}$ curves and the behavior of ζ at low p suggest the existence of an underlying structural transition related to the emergence of these indirect connections. We now characterize this phenomenon more formally.

Inspired by the concept of connectivity transitions in random graphs³², we focus on the completeness of the network formed by considering both direct and indirect paths. For simplicity, let's first consider the network formed by only 1-paths and 2-paths, ignoring the weighting $d^{-\alpha}$ and focusing solely on the presence or absence of connections. The density of connections in this combined network can be quantified by its completeness v , defined as the probability that a randomly chosen pair of nodes i and j is connected by either a 1-path or a 2-path:

$$v = p + p^{(2)} \quad (19)$$

Fig. 4 | Degree distribution of the path network. Degree distribution of the path network $P(k)$, with average degree $k = k^{(1)} + k^{(2)}$, combining the average degrees of 1-path $k^{(1)}$ and 2-path $k^{(2)}$ links, averaged over 50 realizations of (a) Erdős-Rényi (ER) networks and (b) Random Regular (RR) networks. In both cases, $N = 1000$ nodes and 1-path average degree $k^{(1)} = 30$. Solid lines represent the theoretical binomial distribution $B(N, p + p^{(2)})$ expected for an ER network with N nodes and connection probability of $p + p^{(2)}$, which accounts for both 1-path p and 2-path $p^{(2)}$ link probabilities.

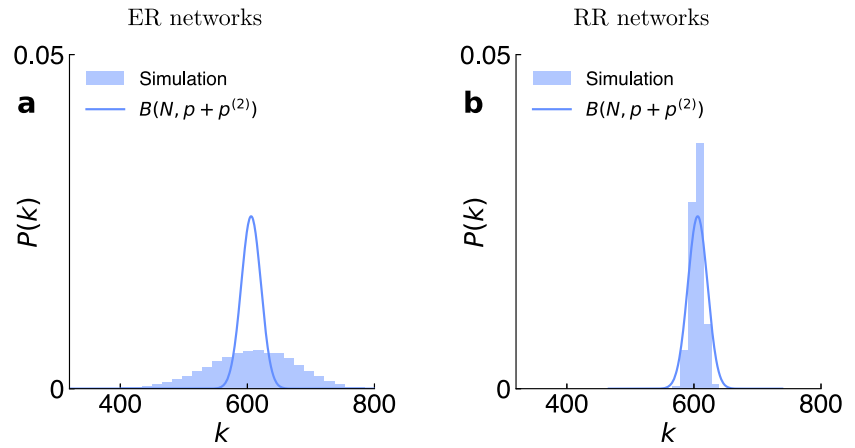
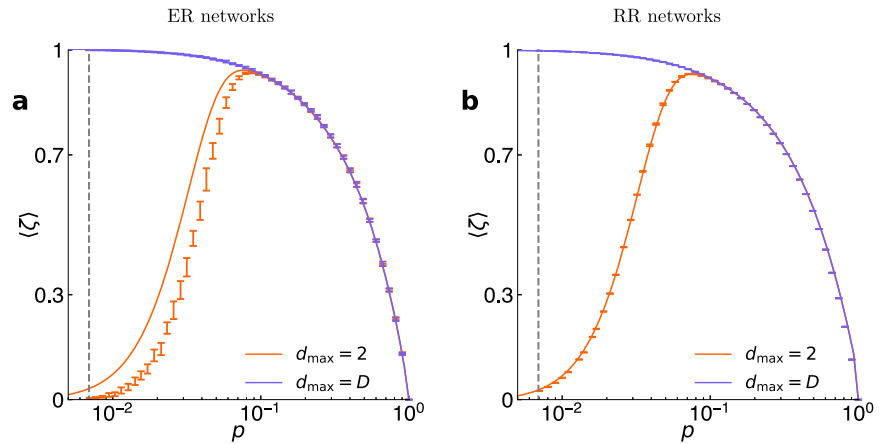


Fig. 5 | Average indirect influence of the uniform weight limit. Average indirect influence $\langle \zeta \rangle$ for (a) Erdős-Rényi networks and (b) Random Regular networks as a function of the expected connection probability p for the path Laplacian truncated at $d_{\max} = 2$ and $d_{\max} = D$. Each point corresponds to the average over 50 different networks with error bars of one standard deviation, and the solid lines correspond to the theoretical curves for the respective truncation. The vertical gray dashed line marks the connectivity threshold $p < \log N/N$, below which the network is almost surely disconnected³² and therefore not considered in the analysis.



Substituting the expression of $p^{(2)}$ from Eq. (12), we get:

$$v = p + (1 - p) \left(1 - (1 - p^2)^{N-2} \right). \quad (20)$$

We consider this completeness v as an order parameter, with the direct expected connection probability p acting as the control parameter. We look for a phase transition, marked by a critical value $p_c(N)$, where the probability of the combined network being fully connected sharply transitions from 0 to 1 as $N \rightarrow \infty$, i.e.,

$$\begin{aligned} p < p_c &\Rightarrow P\{v = 1 | N \rightarrow \infty\} = 0, \\ p > p_c &\Rightarrow P\{v = 1 | N \rightarrow \infty\} = 1. \end{aligned}$$

The probability that the combined 1-path and 2-path network is a complete network is equivalent to the probability that no pair i, j lacks both a 1-path and 2-path connection. Assuming independence between pairs for large N :

$$\begin{aligned} P\{v = 1\} &= 1 - P\{\exists i, j | a_{ij} = a_{ij}^{(2)} = 0\} \\ &= 1 - (1 - (p + p^{(2)}))^N = p + p^{(2)} \\ &= p + (1 - p) \left(1 - (1 - p^2)^{N-2} \right) \\ &= 1 - (1 - p^2)^N \\ &= 1 - e^{-p^2 N} \end{aligned} \quad (21)$$

In the last step, we use the property that $\lim_{N \rightarrow \infty} (1 - \frac{a}{N})^N = e^{-a}$. Therefore, from this development it is immediate to see that $p_c(N) = N^{-1/2}$ is a threshold function. This threshold marks a structural phase transition. Below $p_c(N)$, the network is sparse in terms of 2-paths, while above it, the network rapidly becomes dense with 2-paths, causing the completeness v to approach 1 and significantly boosting the indirect influence.

Figure 6 visually confirms this transition by plotting the completeness against the expected connection probability rescaled by the critical threshold.

This analysis can be generalized to consider the emergence of d -paths for $d > 2$. The probability that a d -path exists between nodes i, j given that no shorter path exists, is controlled by terms approximating $(1 - p^d)^{N^{d-1}}$, therefore $P\{v = 1\} \propto e^{-p^d N^{d-1}}$. Since each successive exponential term decays faster, the general threshold function is

$$p_c(N) = N^{-\frac{d-1}{d}}. \quad (22)$$

This threshold function provides the critical value of the expected connection probability for the onset of this new structural phase transition involving up to d -path indirect connections. Notably, as N increases, and with it the network diameter D , the maximum possible d term in the path Laplacian also grows. For sufficiently large d , this critical value can fall below the connectivity threshold $\log(N)/N$, causing this structural phase transition to disappear.

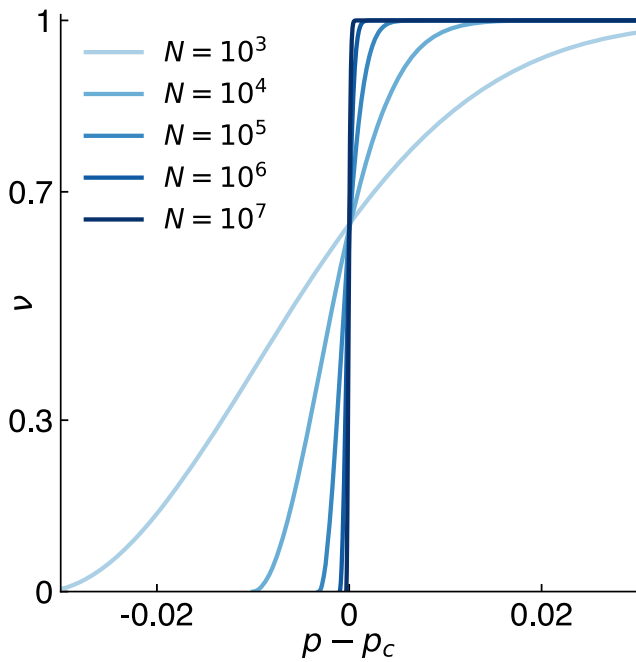


Fig. 6 | Completeness of the path network. Completeness ν of the path network, accounting for 1-path and 2-path links, as a function of the expected connection probability p , rescaled by the corresponding critical threshold p_c , for increasing numbers of nodes N .

Discussion

This work provides a quantitative framework for analyzing the impact of indirect pathways on network diffusion. The central results, including the indirect influence parameter ζ , its predictive phase diagram, and the critical threshold for the onset of d -paths, $p_c(N)$, offer a practical method for assessing the significance of non-local interactions. The theory, derived for analytically tractable random networks, is validated by numerical simulations, confirming that indirect pathways emerge as a collective structural phenomenon rather than a gradual effect.

The theoretical approach relies on mean-field approximations that are well-suited for random graphs. This scope invites future work on networks with more complex topologies, such as those with high clustering or community structure, where structural correlations would likely alter the quantitative predictions. A key implication of our framework is the potential for empirical measurement of ζ . Observing the global diffusion timescale of a process on a network with a known structure would allow for the calculation of an empirical ζ . This value could then be used to test the model and estimate system-specific parameters, such as the influence decay rate α , bridging the gap between theory and real-world data.

The model employs a power-law decay of influence $d^{-\alpha}$, a flexible choice corresponding to a Mellin transform. While other functional forms, like exponential decay, are possible, the qualitative features of the phase transition are expected to be robust. This study's findings provide a rigorous guide for model selection in network science, with direct applications in fields such as epidemiology and the study of social influence.

Methods

Indirect diffusion model

We model diffusion dynamics on a network represented as an unweighted, connected graph $G = (V, E)$, where V is the set of N nodes (indexed $i = 1, \dots, N$) and E is the set of edges representing direct connections between nodes. To incorporate the influence of interactions beyond direct neighbors, we employ the d -path Laplacian framework¹³.

This framework makes use of the d -path adjacency matrices, $A^{(d)} \in \{0, 1\}^{N \times N}$, that account for all shortest paths of length $d = 1, \dots, D$ connecting nodes in the network, where D is the diameter of the network.

They are defined as

$$(A^{(d)})_{ij} = a_{ij}^{(d)} = \begin{cases} 1 & \text{if } \text{dist}(i, j) = d \\ 0 & \text{otherwise} \end{cases}, \quad (23)$$

where $\text{dist}(i, j)$ denotes the shortest path distance between nodes i and j . Note that $A^{(1)}$ corresponds to the standard adjacency matrix A .

By extending the direct diffusion formalism¹, the diffusion equations for a network with indirect connections are

$$\frac{dx_i}{dt} = \gamma^{(1)} \sum_{j=1}^N a_{ij}^{(1)} (x_j - x_i) + \dots + \gamma^{(D)} \sum_{j=1}^N a_{ij}^{(D)} (x_j - x_i), \quad (24)$$

where x_i is the state of node i , and $\gamma^{(d)} > 0$ is the diffusion rate along d -path connections. It is natural to assume that the diffusion rate decays as a function of the distance d between nodes. This choice can be conveniently tuned depending on the problem at hand. Moreover, it is convenient to set the time scale of the process with $\gamma^{(1)} = 1$. Without loss of generality, we consider an exponential decay of the form $\gamma^{(d)} = d^{-\alpha}$ where α is the new general control parameter. This formulation corresponds to a Mellin transformation¹³.

The d -path Laplacian $L^{(d)} \in \mathbb{R}^{N \times N}$ is defined as

$$(L^{(d)})_{ij} = \begin{cases} -1 & \text{if } \text{dist}(i, j) = d \\ k_i^{(d)} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

where $k_i^{(d)} = (\mathbf{1}^T A^{(d)})_i$ is the degree of node i in the d -path adjacency matrix. This definition corresponds to the matrix form

$$L^{(d)} = K^{(d)} - A^{(d)}, \quad (26)$$

where $K^{(d)}$ is the diagonal matrix of node degrees for $A^{(d)}$. To capture diffusion across paths of all lengths, we define the path Laplacian as

$$\hat{L} = \sum_{d=1}^D d^{-\alpha} L^{(d)}. \quad (27)$$

Using the path Laplacian definition, Eq. (24) can be compactly written as

$$\dot{\mathbf{x}} = -\hat{L}\mathbf{x}. \quad (28)$$

Perturbation theory derivation

To derive the first-order perturbation term of the d -path Laplacian, we recall the perturbation framework

$$\hat{L}\hat{\mathbf{v}}_2 = \hat{\lambda}_2 \hat{\mathbf{v}}_2 \rightarrow \begin{cases} \hat{\lambda}_2 = \lambda_2 + \epsilon \lambda' + o(\epsilon^2) \\ \hat{\mathbf{v}}_2 = \mathbf{v}_2 + \epsilon \mathbf{v}' + o(\epsilon^2) \end{cases}. \quad (29)$$

By introducing the perturbations into the eigenvalue equation, we obtain:

$$(L + \epsilon L^{(2)})(\mathbf{v}_2 + \epsilon \mathbf{v}') = (\lambda_2 + \epsilon \lambda')(\mathbf{v}_2 + \epsilon \mathbf{v}'). \quad (30)$$

After expanding the equality and ignoring second-order terms, we get

$$L\mathbf{v}' + L^{(2)}\mathbf{v}_2 = \lambda_2 \mathbf{v}' + \lambda' \mathbf{v}_2. \quad (31)$$

We then multiply both sides of the equation by \mathbf{v}_2^T and simplify

$$\mathbf{v}_2^T L^{(2)} \mathbf{v}_2 = \mathbf{v}_2^T \lambda' \mathbf{v}_2. \quad (32)$$

Lastly, since the eigenvectors are orthonormal, we end up with

$$\lambda' = \mathbf{v}_2^T L^{(2)} \mathbf{v}_2. \quad (33)$$

Mean-field approximation details

To derive an analytical expression for the indirect influence, we used the first and second moments of the entries of the eigenvector \mathbf{v}_2 . These moments are obtained from the properties $\sum_{i=1}^N (\mathbf{v}_2)_i^2 = 1$ and $\sum_{i=1}^N (\mathbf{v}_2)_i = 0$, which correspond to the normalization of \mathbf{v}_2 , and orthogonality of \mathbf{v}_2 and $\mathbf{v}_1 = \mathbf{1}$, respectively. To compute the first moment, we begin by taking the expectation of both sides of the first property, yielding:

$$\sum_{i=1}^N \langle (\mathbf{v}_2)_i^2 \rangle = 1. \quad (34)$$

If we assume that the eigenvector entries are identically independent, we get

$$\langle (\mathbf{v}_2)_i^2 \rangle = \frac{1}{N}. \quad (35)$$

For the second moment, we start from

$$\sum_{i=1}^N (\mathbf{v}_2)_i = 0 \Rightarrow 0 = \left(\sum_{i=1}^N (\mathbf{v}_2)_i \right) \left(\sum_{j=1}^N (\mathbf{v}_2)_j \right). \quad (36)$$

Then, by isolating the diagonal terms, we obtain:

$$\begin{aligned} 0 &= \sum_{i=1}^N \sum_{j=1}^N (\mathbf{v}_2)_i (\mathbf{v}_2)_j = \sum_{i=1}^N (\mathbf{v}_2)_i^2 + \sum_{i=1}^N \sum_{j \neq i}^N (\mathbf{v}_2)_i (\mathbf{v}_2)_j \\ &= 1 + \sum_{i=1}^N \sum_{j \neq i}^N (\mathbf{v}_2)_i (\mathbf{v}_2)_j. \end{aligned} \quad (37)$$

We used the first eigenvector property to evaluate the first term after splitting the summations. Finally, after taking the expected value on both sides, we get:

$$\langle (\mathbf{v}_2)_i (\mathbf{v}_2)_j \rangle = -\frac{1}{N(N-1)}. \quad (38)$$

d-path expected connection probability

We derive the general d -path expected connection probability in terms of the expected connection probability of the base network. The existence of a d -path connection requires two conditions: the absence of all shorter d' -path connections for $d' < d$, and the existence of at least one d -path between the nodes. Therefore:

$$\begin{aligned} p^{(d)} &= P\{a_{ij}^{(d)} = 1\} = P\{a_{ij} = 0, \dots, a_{ij}^{(d-1)} = 0\} \\ &= P\{\exists k_1, \dots, k_d \in G \mid a_{ik_1} \dots a_{k_{d-1}j} = 1\}. \end{aligned} \quad (39)$$

The second term is equivalent to one minus the probability of the absence of all possible d -path connections. Using the mean-field approximation, which assumes independence between the absence of these paths, we obtain:

$$\begin{aligned} p^{(d)} &= P\{a_{ij} = 0, \dots, a_{ij}^{(d-1)} = 0\} \\ &= (1 - P\{\forall k_1, \dots, k_{d-1} \in G \mid a_{ik_1} \dots a_{k_{d-1}j} = 0\}) \\ &\approx P\{a_{ij} = 0, \dots, a_{ij}^{(d-1)} = 0\} \\ &= \left(1 - (1 - P\{k_1, \dots, k_{d-1} \in G \mid a_{ik_1} \dots a_{k_{d-1}j} = 1\}) \right)^{\binom{N-2}{d-1}}. \end{aligned} \quad (40)$$

Finally, we can express the first probability term in terms of the previously derived expected connection probabilities, while the second term simplifies to p^d , since the existence of a d -path connection requires the presence of each direct link along the path.

$$p^{(d)} \approx \left(1 - \sum_{i=1}^{d-1} p^{(i)} \right) \left(1 - (1 - p^d)^{\binom{N-2}{d-1}} \right). \quad (41)$$

Since the $p^{(i)}$ probabilities can be recursively expressed in terms of p , each $p^{(d)}$ ultimately depends only on the expected connection probability of the direct links.

Numerical simulations

Numerical results for comparison with analytical predictions were obtained by generating ensembles of Erdős-Rényi (ER) and Random Regular (RR) graphs using standard algorithms. For each set of parameters (N, p), results were typically averaged over multiple independent network realizations, as indicated in the figure captions. The shortest path distances required for constructing d -path matrices were computed using the Bellman-Ford algorithm, and the eigenvalues of both the standard Laplacian L and the path Laplacian \hat{L} were calculated using the Arpack numerical solver.

Data availability

All generated data is available at the persistent repository [10.5281/zenodo.17395848](https://doi.org/10.5281/zenodo.17395848).

Code availability

The code for replicating the results and figures is available at github.com/llu2/indirect-diffusion and [10.5281/zenodo.17395848](https://doi.org/10.5281/zenodo.17395848).

Received: 9 May 2025; Accepted: 28 October 2025;
Published online: 11 December 2025

References

- Masuda, N., Porter, M. A. & Lambiotte, R. Random walks and diffusion on networks. *Phys. Rep.* **716**, 1–58 (2017).
- Liu, C. & Zhang, Z.-K. Information spreading on dynamic social networks. *Commun. Nonlinear Sci. Numer. Simul.* **19**, 896–904 (2014).
- Newman, M. E. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).
- De Domenico, M., Solé-Ribalta, A., Gómez, S. & Arenas, A. Navigability of interconnected networks under random failures. *Proc. Natl Acad. Sci. USA* **111**, 8351–8356 (2014).
- Gentili, M., Martini, L., Sponziello, M. & Becchetti, L. Biological random walks: multi-omics integration for disease gene prioritization. *Bioinformatics* **38**, 4145–4152 (2022).
- Merris, R. Laplacian matrices of graphs: a survey. *Linear Algebra Appl.* **197**, 143–176 (1994).
- Bick, C., Gross, E., Harrington, H. A. & Schaub, M. T. What are higher-order networks? *SIAM Rev.* **65**, 686–731 (2023).
- De Domenico, M. et al. Mathematical formulation of multilayer networks. *Phys. Rev. X* **3**, 041022 (2013).
- Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 2485 (2019).
- Granovetter, M. S. The strength of weak ties. *Am. J. Sociol.* **78**, 1360–1380 (1973).
- Bond, R. M. et al. A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295–298 (2012).
- Guimaraes Jr, P. R., Pires, M. M., Jordano, P., Bascompte, J. & Thompson, J. N. Indirect effects drive coevolution in mutualistic networks. *Nature* **550**, 511–514 (2017).

13. Estrada, E. Path Laplacian matrices: introduction and application to the analysis of consensus in networks. *Linear Algebra Appl.* **436**, 3373–3391 (2012).
14. Estrada, E., Hameed, E., Hatano, N. & Langer, M. Path Laplacian operators and superdiffusive processes on graphs. I. One-dimensional case. *Linear Algebra Appl.* **523**, 307–334 (2017).
15. Estrada, E., Hameed, E., Langer, M. & Puchalska, A. Path Laplacian operators and superdiffusive processes on graphs. II. Two-dimensional lattice. *Linear Algebra Appl.* **555**, 373–397 (2018).
16. Estrada, E. et al. Random multi-hopper model: super-fast random walks on graphs. *J. Complex Netw.* **6**, 382–403 (2018).
17. Riascos, A. P. & Mateos, J. L. Fractional dynamics on networks: Emergence of anomalous diffusion and lévy flights. *Phys. Rev. E* **90**, 032809 (2014).
18. Estrada, E. Path laplacians versus fractional laplacians as nonlocal operators on networks. *N. J. Phys.* **23**, 073049 (2021).
19. Estrada, E., Gambuzza, L. V. & Frasca, M. Long-range interactions and network synchronization. *SIAM J. Appl. Dynamical Syst.* **17**, 672–693 (2018).
20. Gambuzza, L. V., Frasca, M. & Estrada, E. Second-order consensus protocols based on transformed d-path Laplacians. *Appl. Math. Comput.* **343**, 183–194 (2019).
21. Arias, J. H., Gómez-Gardeñes, J., Meloni, S. & Estrada, E. Epidemics on plants: Modeling long-range dispersal on spatially embedded networks. *J. Theor. Biol.* **453**, 1–13 (2018).
22. Estrada, E. & Vargas-Estrada, E. How peer pressure shapes consensus, leadership and innovations in social groups. *Sci. Rep.* **3**, 2905 (2013).
23. Miranda, M., Pereda, M., Sánchez, A. & Estrada, E. Indirect social influence and diffusion of innovations: An experimental approach. *PNAS Nexus* **3**, 409 (2024).
24. Gómez, S. et al. Diffusion dynamics on multiplex networks. *Phys. Rev. Lett.* **110**, 028701 (2013).
25. Solé-Ribalta, A. et al. Spectral properties of the Laplacian of multiplex networks. *Phys. Rev. E* **88**, 032807 (2013).
26. Fiedler, M. Algebraic connectivity of graphs. *Czechoslovak Math. J.* **23**, 298–305 (1973).
27. Mohar, B. Laplace eigenvalues of graphs—a survey. *Discret. Math.* **109**, 171–183 (1992).
28. Mohar, B. The Laplacian spectrum of graphs. In Alavi, Y., Chartrand, G., Oellermann, O. & Schwenk, A. (eds.) *Graph Theory, Combinatorics, and Applications*, 871–898 (Wiley, New York, NY, USA, 1991). <https://users.fmf.uni-lj.si/mohar/Papers/Spec.pdf>.
29. Torres-Hugas, L., Duch, J. & Gómez, S. Structural prediction of super-diffusion in multiplex networks. *Chaos, Solitons Fractals* **186**, 115265 (2024).
30. Jamakovic, A. & Van Mieghem, P. On the robustness of complex networks by using the algebraic connectivity. In Das, A., Pung, H. K., Lee, F. B. S. & Wong, L. W. C. (eds.) *NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*, Lecture Notes in Computer Science, 183–194 (Springer, 2008). https://doi.org/10.1007/978-3-540-79549-0_16.
31. McKay, B. D. The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.* **40**, 203–216 (1981).
32. Erdős, P. & Rényi, A. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica* **12**, 261–267 (1961).

Acknowledgements

L.T.-H. acknowledges financial support from Diputació de Tarragona and Universitat Rovira i Virgili, Spain (2023PMF-PIPF-21). This work was supported by Ministerio de Ciencia e Innovación, Spain (PID2021-128005NB-C21, RED2022-134890-T), Generalitat de Catalunya, Spain (2021SGR-633), MICIU/AEI/10.13039/501100011033 FEDER EU (PID2022-142600NB-I00), and Universitat Rovira i Virgili, Spain (2023PFR-URV-00633).

Author contributions

L.T.-H. contributed to the conceptualization, formal analysis, investigation, methodology, software development, writing the original draft, and manuscript review and editing. J.D. contributed to the conceptualization, funding acquisition, and manuscript review and editing. S.G. contributed to the conceptualization, formal analysis, funding acquisition, and manuscript review and editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Sergio Gómez.

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

© The Author(s) 2025