## ARTICLE

Check for updates

Multi pathways temporal distance unravels the hidden geometry of network-driven processes

Sebastiano Bontorin [1,2] & Manlio De Domenico [3,4,5]

Network-based interactions allow one to model many technological and natural systems, where understanding information flow between nodes is important to predict their functioning. The complex interplay between network connectivity and dynamics can be captured by scaling laws overcoming the paradigm of information spread being solely dependent on network structure. Here, we capitalize on this paradigm to identify the relevant paths for perturbation propagation. We introduce a multi-pathways temporal distance between nodes that overcomes the limitation of focussing only on the shortest path. This metric predicts the latent geometry induced by the dynamics in which the signal propagation resembles the traveling wave solution of reaction-diffusion systems. We validate the framework on a set of synthetic dynamical models, showing that it outperforms existing approaches in predicting arrival times. On a set of empirical contact-based social systems, we show that it can be reliably used also for models of infectious diseases spread - such as the Susceptible-Infected-Susceptible - with remarkable accuracy in predicting the observed timing of infections. Our framework naturally encodes the concerted behavior of the ensemble of paths connecting two nodes in conveying perturbations, with applications ranging from regulatory dynamics within cells to epidemic spreading in social networks.

[1] Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo, TN, Italy. [2] Department of Physics, University of Trento, Via Sommarive 14, 38123 Povo, TN, Italy. [3] Department of Physics and Astronomy "Galileo Galilei", University of Padua, Via F. Marzolo 8, 315126 Padova, Italy. [4] Padua Center for Network Medicine, University of Padua, Via F. Marzolo 8, 315126 Padova, Italy. [5] Istituto Nazionale di Fisica Nucleare, Sez. Padova, Padova, Italy. ✉email: sbontorin@fbk.eu; manlio.dedomenico@unipd.it

Dynamical processes on top of a network are often adopted to model the behavior of complex interdependent systems[1]. Advances in the study of spreading and reaction processes[2], synchronization phenomena[3,4] and diffusion dynamics[5] have started to uncover and to emphasize how the coupling of the system dynamics to its connectivity properties determines emergent behaviors, such as synchronizability being dependent on both the Laplacian spectrum and the oscillators dynamics equations. In these networked systems links can encode a pairwise, possibly nonlinear, interaction: from mass action kinetics[6] to switch-like processes[7], the specific set of functions adopted to mimic the real empirical mechanics give rise to profoundly different, and apparently random, nodes' responses and patterns on top of the same network structure[8]. As a consequence, a complete understanding of these systems and the functional role played by each node can not be satisfied by the sole knowledge of the topology of connections. Inter alia, the challenge of predicting the arrival time of a signal[9–12] is fundamental to expose temporal dependencies between system units, how the interplay between topology and dynamics affects the transient temporal response of the system, and ultimately to reveal and exploit the hidden geometries induced by the dynamics[2,13,14].

Recent works have shed light over this interplay[8,15], where perturbations from an equilibrium basin are used to understand how correlations between units of a complex networked dynamical systems emerge via information propagation[16].

Analytical treatment under linear response approximation uncovers the scaling laws that govern the system response to perturbations, effectively decoupling the role of the network topology (a node's degree k[17]) from the dynamics (the universal exponents[8] obtained from the model's ODEs) and synthesizes the complex interplay that determines the stability[18] and dynamical properties. This paradigm of scaling laws helps identifying different models into universality classes of dynamical regimes[19] in which units respond to and propagate information in a similar manner. These metrics can in turn be used to quantify how strongly a node can impact on the local topology[8] or the efficiency of a node as a pipeline in information flow $\mathcal{F}(k) \sim k^{\omega}$[20]. Ultimately, they estimate a node's characteristic time response from a perturbation in a neighboring node $\tau_i \sim k_i^{\theta}$[9], condensing in a universal scaling law the information of the temporal behavior of nodes in a networked system near steady state.

One of the remarkable applications of these scaling laws, is the definition of predictive observables. For instance, knowledge of τ_i allows the introduction of a temporal distance[9] between two nodes $\mathcal{L}(j \rightarrow i)$, which reflects the effective arrival time of a signal by selecting the shortest path with the minimum cumulative lag time as candidate temporal distance.

This metric was proposed under the rationale that a single, fastest, path can be considered the main artery for information flow/response[8,9]. It represents a dramatic improvement over the standard topological distance: in fact, it also embeds information of the non-linear interactions between nodes, accounting for processes in which hubs can act as fast responders ($\theta < 0$) or bottlenecks ($\theta > 0$), and explain why in certain processes the shortest-path paradigm results inadequate. Even so, it is limited by not considering the possible contribution of other relevant paths in conveying the perturbation, a key ingredient which was also recently introduced in the definition of an effective distance for diffusive dynamics on networks[14]. An advance is therefore necessary to take into account these contributions and entirely predict propagation times for systems where interactions are modeled via non-linear dynamics.

In this article we further expand this theoretical framework and develop a simple perspective of the multiple-paths nature of spatiotemporal propagation. We first analyze paths' contribution in spreading a perturbation between two nodes, taking advantage of the exponential decay of correlations with a path's topological length. We then define a temporal distance which synthesizes the paths' concerted behavior in dispersing the perturbation and completely predicts the propagation time required for a perturbation generated in a source node to affect another. Embedding targets nodes in the vector space induced by this metric reveals the intuitive, hidden geometry of perturbation propagation (see Fig. 1)[2]. Finally, we apply this metric on an empirical network of
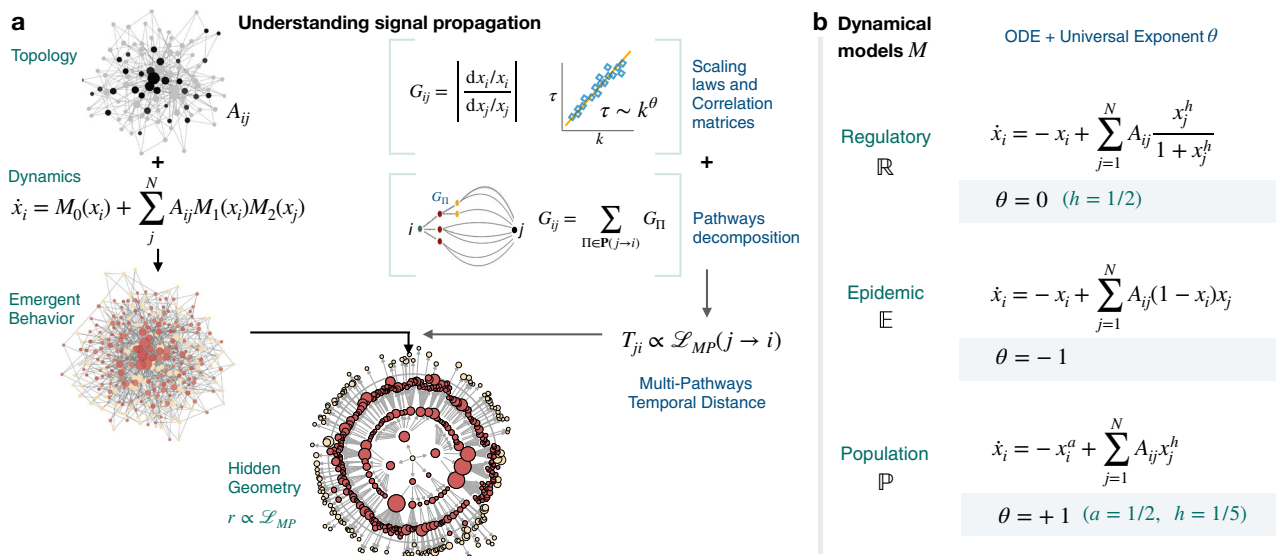

Fig. 1 Revealing the hidden geometry. a The interplay between structure and non-linear interactions results in an emergent, complex propagation pattern. We combine analytical knowledge of scaling laws[9] and correlation matrices[8] with a network's paths decomposition of propagation, to predict a temporal multi-pathways distance and unravel the signal propagation. **b** Dynamical models' ordinary differential equations (ODE) employed representing three widely adopted and diverse processes. $\mathbb{R}$ Regulatory dynamics via Michaelis–Menten model[7], Susceptible-Infected-Susceptible compartmental model[26] $\mathbb{E}$ as a case of Epidemic dynamics, and finally $\mathbb{P}$ Population processes via birth-death dynamics. The specific choice of exponents (a,h) is necessary in order to reproduce the three classes of time response regimes encoded in the exponent θ of the scaling law $\tau_i(k) \sim k_i^{\theta}$[9].

social interactions, modeling an epidemic dynamics. We show in this last section how the predictive nature near the steady state can be employed to better understand the infection times which occur far from the linear regime.

Results

Dynamical processes on top of networks. We now introduce the analytical framework for a dynamical model on top of a network[8,15], under which these measures have been proposed. For consistency and continuity, we employ the same notation and rely on the same dynamical processes as testing ground. In this framework each node is characterized by a time-dependent variable $x_i(t)$, $i = 1, \ldots, N$, which can represent expression of genes in a regulatory network[7], the fraction of infected individuals in an airport network[2], or any other scalar value of interest. In this general framework the system dynamics is described by a set of non-linear functions $\mathbf{M} = \{M_0(x), M_1(x), M_2(x)\}$, and the time evolution of $x_i(t)$ is driven by a non-linear dynamical equation:

$$\dot{x}_i(t) = M_0\big(x_i(t)\big) + \sum_{j=1}^{N} A_{ij} M_1\big(x_i(t)\big) M_2(x_j(t)). \quad (1)$$

Here A_{ij} represents the adjacency matrix element, M_0 accounts for element i's self dynamics, while the second term in the r.h.s. describes the factorized interaction between i and its direct neighbors j; e.g.: for SIS compartmental model ($M_0 = -\beta x, M_1 = 1 - x, M_2 = \alpha x$)[9]. Expansions of $\mathbf{M}$ functions as power series of the degrees define the set of aforementioned scaling laws and universal exponents[8], which are necessary to construct the following analyses. The set of steady-state dynamics that can be casted in this framework are shown to be a reasonable approximation to reproduce the behavior for many biological, social and technological systems. The models and framework we work with are presented in Fig. 1. Small constant perturbations in the system's equilibrium state ($x_i \rightarrow x_i + dx_i$) are used as an analytical probe to understand the relation between the connectivity structure and dynamics in signal propagation. As a permanent perturbation brings the system to a new steady state, the notion of a network Global Correlation Function G has been introduced to encode the full system response to a perturbation from node j[15]:

$$G_{ij} = \left| \frac{dx_i/x_i}{dx_j/x_j} \right|. \quad (2)$$

Refer to Supplementary Notes 1 and 2 for more details on this framework.

Path-driven analysis of perturbation propagation. In order to provide new insight on paths' contribution in signal propagation, we now propose a path-driven alternative description of G_{ij} elements, recasting them as a sum over the ensemble of paths that connect nodes i and j. Furthermore, this formulation will pave the way to the definition of a multi-pathways temporal distance. We introduce $\mathbf{P}(j \rightarrow i)$ as the set of all paths Π connecting source node j to ending node i (in the definition, we include paths that traverse vertices multiple times, also defined as walks in[21]) and we define the amount of correlation carried along each path G_Π as the product of local correlation elements $R_{ij} = \left| \frac{\partial x_i/x_i}{\partial x_j/x_j} \right|$[8] associated to each edge in the path. R_{ij} quantifies the dependence of node i's state on small changes in its neighbor j's state, via partial derivative in the steady state of Eq. (1).

In this notation a path Π of length L defines a set of L nodes, sorted by visiting order: $\Pi(1 \rightarrow L) = \{n_1, \ldots, n_L\}$. Hence the path's

correlation is defined as $G_\Pi(L) = R_{L,L-1} \times \ldots \times R_{2,1}$:

$$G_\Pi(L) = \prod_{i=1}^{L-1} \left| \frac{\partial x_{i+1}/x_{i+1}}{\partial x_i/x_i} \right| = \prod_{i=1}^{L-1} R_{i+1,i} \quad (3)$$

Therefore G_{ij} can be rewritten as an ensemble sum of contributions G_Π of all the possible unique paths that connect j to i:

$$G_{ij} = \sum_{\Pi \in \mathbf{P}(j \rightarrow i)} G_\Pi \quad (4)$$

In Fig. 2, we make use of this formulation to reconstruct G_{ij} with a limited subset of paths, under the rationale that most paths' contributions will be negligible. This assumption is driven by the known result that the average individual node's correlation at distance l is governed by an exponential decay $G(l) \approx \exp(-l/\lambda)$[8]. In this particular case, we are interested in the exponential decay of single path's correlation G_Π. In similar fashion to $G(l)$, we derive an equation for for $G_\Pi(l)$ that confirms the exponential decay also for unique paths: $G_\Pi(l) \approx C_M^l \cdot e^{-\alpha\beta l} \cdot e^{-\alpha l}$. Where the known universal exponent β[8] defines whether a dynamic process is conservative ($\beta = 0$) or dissipative ($\beta > 0$), i.e., characterized by fast decay of correlations. The derivation of this result is discussed in Supplementary Note 3. Consequently, if $L_{ij} = L$ is the topological distance between nodes i and j, due to the exponential decay of $G_\Pi(l)$, it can be expected that only paths of order (length) L in $\mathbf{P}(j \rightarrow i)$ add relevant contributions to G_{ij}.

As the ensemble $\mathbf{P}(j \rightarrow i)$ can be decomposed as a collection of subsets of paths having the same topological distance l: $\mathbf{P}(j \rightarrow i) = \bigcup_{l=1}^{\infty} \mathbf{P}_l(j \rightarrow i)$, we therefore seek to reconstruct the propagation from j to i by using possibly only paths in $\mathbf{P}_L(j \rightarrow i)$. Formally, this translates to limiting the sum (4) only to the subset of shortest paths of degenerate lengths L_{ij}, simplifying the equation:

$$G_{ij} \approx \bar{G}_{ij} = \sum_{\Pi \in \mathbf{P}_L(j \rightarrow i)} G_\Pi \quad (5)$$

We aim to verify numerically this assumption, comparing directly exact values for G_{ij} against approximate values $\bar{G}_{ij}$. $\bar{G}_{ij}$ are compared with numerical G_{ij} by estimating the absolute relative errors $\delta G = |G_{ij} - \bar{G}_{ij}|/G_{ij}$. We evaluate the reconstruction of G_{ij} by first only using single terms G_{Π^*} associated to the shortest path carrying the maximum correlation ($\Pi^* = \max_\Pi(G_\Pi)$), then considering also all shortest paths in the subset $\mathbf{P}_L(j \rightarrow i)$ and finally by integrating paths of successive orders $L_{ij} + N$.

Analytical predictions of Eq. (5) are tested against simulations in Fig. 2, highlighting how the introduction of successive orders reduces the error δG and improves the estimation of G_{ij}. These results support the idea that when analyzing perturbation propagation between two nodes, focusing only on the set of shortest paths (and eventually extending to a successive order $L + 1$) is enough to understand which pathways retain the relevant information in the spatiotemporal propagation of perturbations.

Multi pathways temporal distance. The previous section has shed light on multiple paths contributions to perturbation transmission, which will prove to be foundational in introducing a new candidate metric. A node i will respond from a constant perturbation starting at node j by shifting its steady state activity to a new steady state $x_i^{new}(t \rightarrow \infty) = x_i + \Delta x_i(t \rightarrow \infty)$. The operative definition of temporal distance proposed in ref. [9] defines $T(j \rightarrow i)$ as the time in which node i reaches a fraction η of its final response: $\Delta x_i(t = T(j \rightarrow i)) = \eta \Delta x_i(t \rightarrow \infty)$. As the final response encoded in G_{ij} is the result of the concerted behavior of the ensemble of paths, if the shortest path[17] is degenerate and more paths introduce a contribution G_Π which is comparable in magnitude, then it is necessary to introduce their time delay in transferring the perturbation. We propose here a more precise
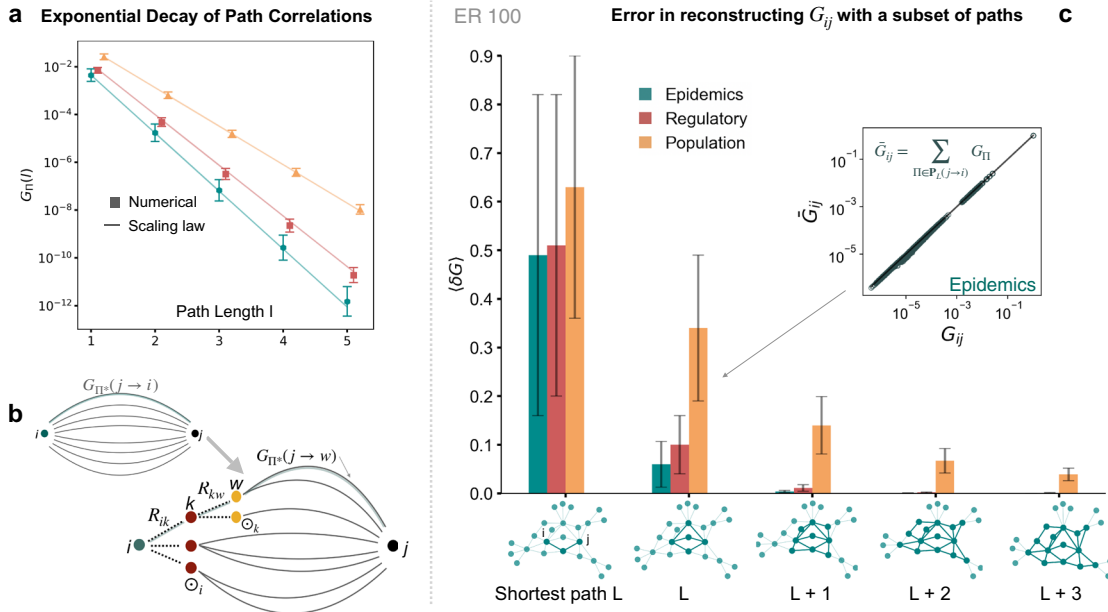
Fig. 2 Path driven analysis of G_{ij}. a Exponential decay of $G_{\Pi}(l)$: Log-Linear plot of numerical results versus predictions, tested on an Erdős-Rényi (ER)[27] synthetic network ($N = 100$, $\rho = 0.15$) for each model. The exponential decay predicted fits the average path correlation at a given length. **b** Ensemble of paths connecting two nodes i and j. Regrouping the paths via local correlation matrix elements allows for a path decomposition of G_{ij} as in Eq. (4). **c** Average absolute relative errors $\langle \delta G \rangle$ are plotted with their standard deviation as a function of the number of path orders included in Eq. (4). Relying on one shortest path on average justifies only 50% of final correlation. Using the entire set $\mathbf{P}_L(j \to i)$ of degenerate shortest paths increases this to above ~90% for processes with fast decay of correlations ($\beta > 0$), effectively retaining all necessary information to understand the amount of perturbation exchanged. Extending the sum to other orders further improves the accuracy in reproduction of G_{ij}. Inset: Exact values for each matrix element G_{ij} against approximate values $\bar{G}_{ij}$ for Epidemic $\mathbb{E}$ dynamics on top of the ER network, considering only the subset $\mathbf{P}_L$.

version of temporal distance: this new metric $\mathcal{L}_{MP}(j \to i)$ takes into account the fact that information follows multiple paths (MP), not just shortest paths.

The simplest ansatz to account for the different path times $\tau_{\Pi} = \sum_{p \in \Pi, p \neq j} \tau_p$, is to perform a linear combination. We weight a path's relaxation time τ_{Π} by the fraction of perturbation it conveys G_{Π}/G_{ij}, which effectively acts as a weight as $\sum_{\Pi} G_{\Pi} / G_{ij} = 1$. Therefore we define a new temporal distance $\mathcal{L}_{MP}(j \to i)$ as a weighted sum of the accumulated lag time along each path:

$$\mathcal{L}_{MP}(j \to i) = \sum_{\Pi \in \mathbf{P}(j \to i)} \tau_{\Pi} \cdot W_{\Pi} \qquad (6)$$

where:

$$W_{\Pi} = \frac{G_{\Pi}}{G_{ij}}. \qquad (7)$$

We provide a more extensive understanding of this metric in Supplementary Note 4. This measure integrates effectively the two relevant descriptors of signal propagation: local perturbations R_{ij} (which enter in the definition of G_{Π}) and relaxation times τ_i. As a blueprint for propagation patterns, these measures can be analytically derived and do not require numerical simulations of perturbation propagation in order to predict $T(j \to i)$. In our calculations R_{ij} values have been obtained exactly from analytical derivation: by isolating the steady state condition from $\dot{x}_i = f(x_1...x_N) = 0$, and taking its derivative $\partial x_i/\partial x_j$ at steady state. Steady state activities can be either obtained numerically or approximated analytically, avoiding simulations, using a known scaling law: $x_i(k) \sim R^{-1}(1/k_i)$[8,19].

Numerical validation on synthetic models. In this section we evaluate $\mathcal{L}_{MP}(j \to i)$ predictions against previous temporal metric $\mathcal{L}(j \to i)$ and traditional shortest path length L_{ij}[17]. As testing

Table 1 Accuracy of temporal metrics.

	$\mathbb{E}$	$\mathbb{R}$	$\mathbb{P}$
L_{ij}	0.83 ± 0.01	0.88 ± 0.01	0.40 ± 0.02
$\mathcal{L}$	0.56 ± 0.01	0.88 ± 0.01	0.85 ± 0.01
$\mathcal{L}_{MP}$	0.99 ± 4e−4	0.99 ± 2e−5	0.99 ± 4e−4

To estimate the goodness of these measures in reproducing arrival times we average the Spearman's ρ_s over an ensemble of several perturbation realizations for $\mathbb{E}$ Epidemic, $\mathbb{R}$ Regulatory and $\mathbb{P}$ Population dynamics. 10 synthetic networks of $N = 300$ nodes are generated via B-A algorithm[22]. For each network, we select randomly 10 nodes as sources of a perturbation, which is obtained numerically. ρ_s is then computed for the set of predictions versus empirical arrival times $T(s \to i)$ for each source. Average ρ_s and 95% C.I. for each dynamics is presented in this table. We find that, despite the approximation introduced by scaling laws in estimating $\tau_i \sim k_i^\theta$ time responses, ρ_s for our proposed metric is close to 1, meaning that the multi-pathways (MP) metric $\mathcal{L}_{MP}$ fully predicts temporal distances.

ground, we opt for a structure with heterogeneous connectivity by generating a Scale Free network A_{ij} via B-A preferential attachment model[22]; we then obtain numerically the steady state activity x_i of nodes for Epidemic, Regulatory and Population dynamics (see Fig. 1). Afterwards, we introduce a small *constant* perturbation from a source node and solve numerically the set of coupled equations. As a result, other network units' states will be eventually shifted and their response times $T(j \to i)$ are obtained numerically following the operative definition[9]. This set of empirical perturbation times $T(j \to i)$ is used to test the goodness of candidate metrics. We measure the accuracy in predicting observed times using Spearman's correlation coefficient ρ_s. Results for average ρ_s are in Table 1. Empirical times versus predictions are presented in Fig. 3, we can appreciate how $\mathcal{L}_{MP}$ improves over the existing metric in predicting the arrival time for farther nodes, where multiple paths propagation is dominant. Moreover, in Supplementary Note 5 we validate the metric against network's sparseness and the degeneracy of shortest paths. In the limit of large and
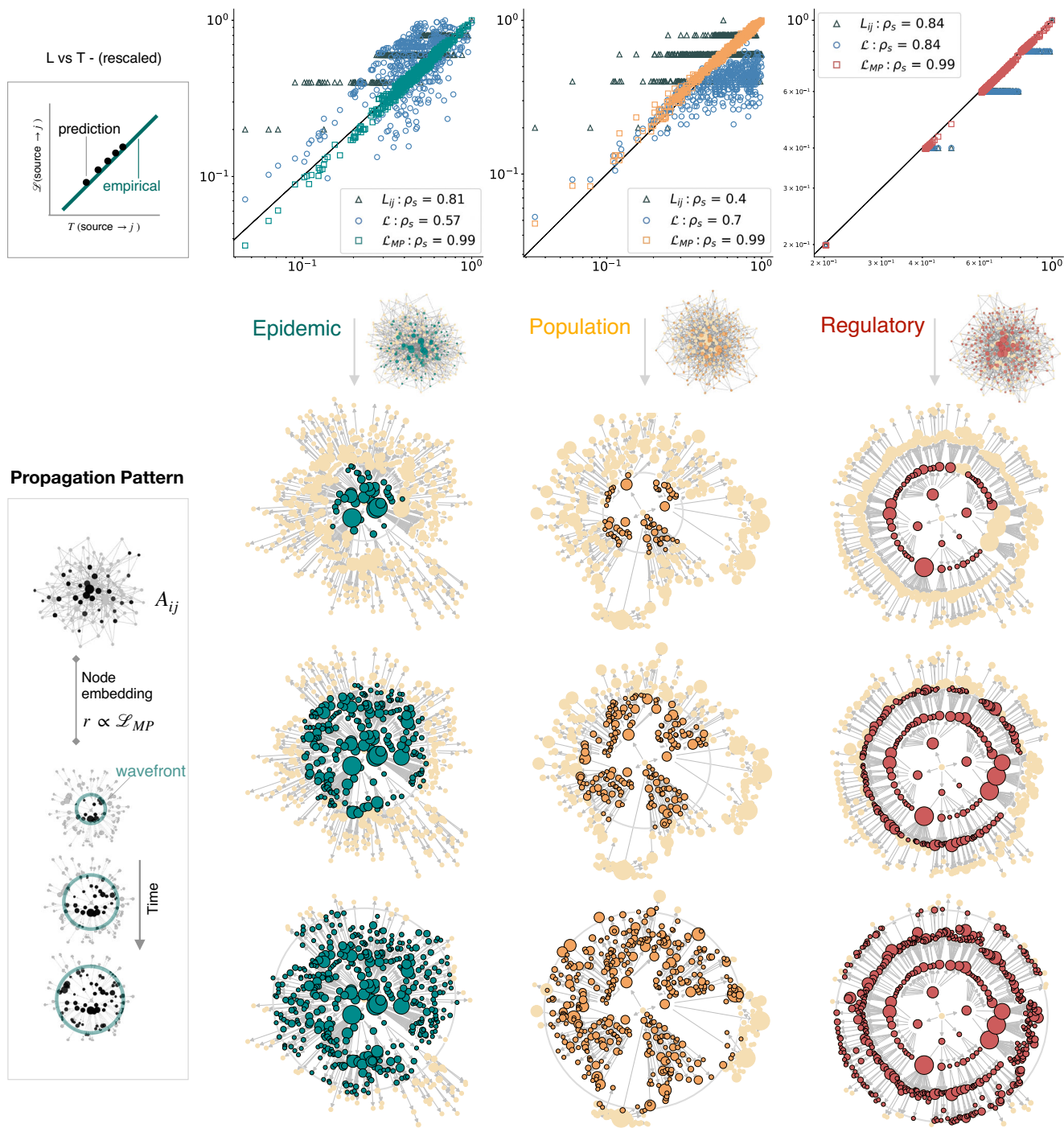
Fig. 3 Metrics predictions vs Empirical $T(j \to i)$. Comparison for a single propagation instance (with associated Spearman's scores) for Population $\mathbb{P}$, Epidemics $\mathbb{E}$ and Regulatory $\mathbb{R}$ models on a scale free network ($N = 500$, $\gamma = 3$). Topological shortest path length L_{ij}, single path $\mathcal{L}(j \to i)$ and multi paths (MP) $\mathcal{L}_{MP}(j \to i)$ distances are compared against empirical times $T(j \to i)$: integrating multiple paths in $\mathcal{L}_{MP}$ results in an exact linear relationship with T, empirical and predicted values have been rescaled for a better representation. Note that $\mathcal{L}(j \to i)$'s lower Spearman's score with respect to L_{ij} in the Epidemics case is due to the single-path distance being affected by the broad distribution of degrees of the Scale Free model, where accounting for multiple paths becomes fundamental. We discuss this point in detail in Supplementary Note 5. Wavefronts: the geometry predicted by $\mathcal{L}_{MP}$ allows to transform the complex propagation in a concentric pattern[9] where the propagation of the perturbation resembles the traveling wave solution of reaction–diffusion systems[2,23]. Interestingly, different perturbations patterns arise from the same topology A_{ij}. Node size is proportional to its degree and nodes are located with radial distance from source $r \propto \mathcal{L}_{MP}$. Edges are drawn following the path Π^* with highest correlation G_{Π^*}. We can appreciate qualitatively the Fisher-like[23] wavefront expansion.
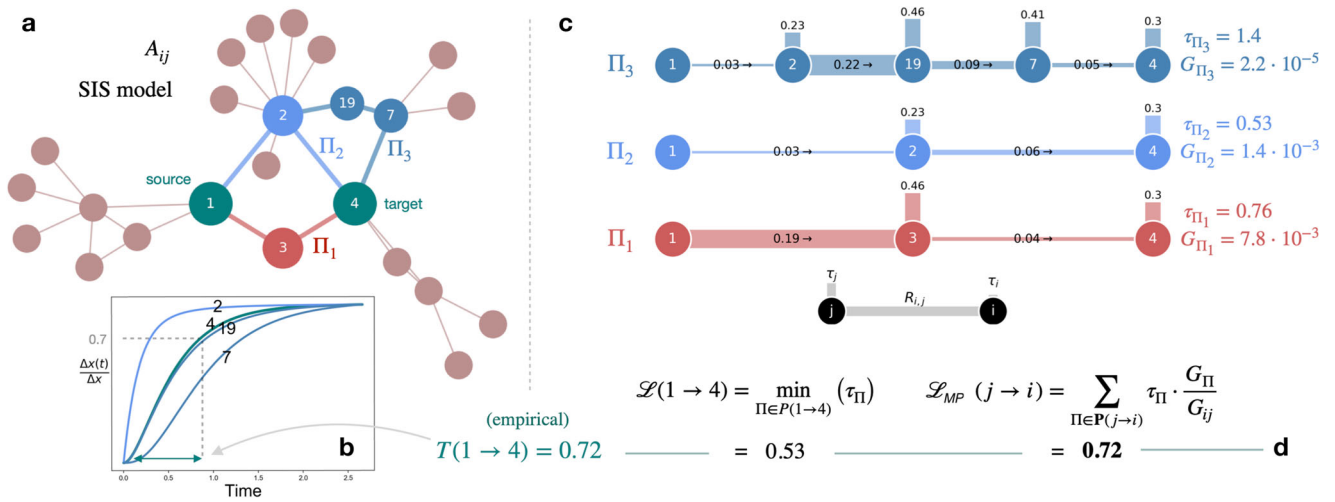
Fig. 4 Detailed view of spatiotemporal perturbation propagation. a Descriptive network A_{ij}, in which correlations R_{ij} and time delays τ_i (numerically estimated, see Supplementary Note 4) can be explicitly shown for an $\mathbb{E}$ Epidemic process. Perturbation from source node 1 impacts target node 4 via the ensemble of paths $P(1 \to 4)$. **b** Selected nodes' normalized responses to the perturbation in 1. **c** We highlight with different colors the 3 principal paths that connect the two nodes: it is interesting to notice the strong heterogeneity found in pair-wise $R_{i+1,i}$ factors that constitute each path G_Π. Path correlations G_Π and path's lag times τ_Π are the building blocks of time distances. **d** $\mathcal{L}(1 \to 4)$ selects τ_{Π_2}[9] being the smallest term, however, since Π_1 is the main actor in determining the linear response of 4 (G_{Π_1}) its delay time has to be considered.

sparse networks, single paths prevail, and one could expect single and multi-paths metrics to converge. Nevertheless, we highlight the importance of accounting for the existence of multiple paths even in this regime, especially for scale free topologies (see Fig. 4 for a detailed view of propagation on multiple paths on a simple model).

Testing different processes shows the adaptive nature of the new metric: starting from the same underlying network A_{ij}, a scale-free (B-A[22]) synthetic graph of $N = 500$ nodes, $\mathcal{L}_{MP}(j \to i)$ uncovers the hidden geometry induced by the set of interactions **M**. For instance, in the case of Population dynamics $\mathbb{P}$ hubs are slow responders and placed far from the source, whereas for Regulatory $\mathbb{R}$ the degree-independence of $\tau(k)$ results in a discretized propagation pattern driven by the topological distance L_{ij}[9]. In Supplementary Note 6 we provide a complete comparison of networks' embeddings using the different metrics. Embedding nodes in this new effective distance[2,9] with a radius from source $r \propto \mathcal{L}_{MP}(j \to i)$, the complex propagation of the perturbation front resembles the traveling wave solution for a reaction-diffusion system modeled by a Fisher-like equation[23].

Predicting an infection outbreak on real social networks. The framework discussed is employed to predict arrival times of an epidemic outbreak on two real social network of contacts. To account for the social ties that drive the infection, networks reconstructed from real datasets of social contact patterns and population structure offer valuable information and provide the ideal environment to perform epidemiological simulations and analyses. The BBC Four Pandemic[24] dataset was built with the aim of simulating an infectious disease outbreak on a real network of human interactions and was obtained by tracking 469 individuals in Haslemere, London. Additionally, we also make use of the network of contacts[25] of 327 students of a high school in Marseille, France. To test our metric, we adopt an SIS (Susceptible-Infectious-Susceptible) epidemic model: $\dot{x}_i = -\mu x_i + \beta \sum_j A_{ij}(1 - x_i)x_j$, with parameters $\beta = 0.015$ and $\mu = 0.01$, yielding a reproduction ratio $R_0 = \beta/\mu = 1.5$, and we select the node with maximum degree as source of the infection outbreak. Albeit simple, this compartmental model can be employed to model infectious processes in which recovering from an infection does not yield a long-lasting immunity, such as influenza. The
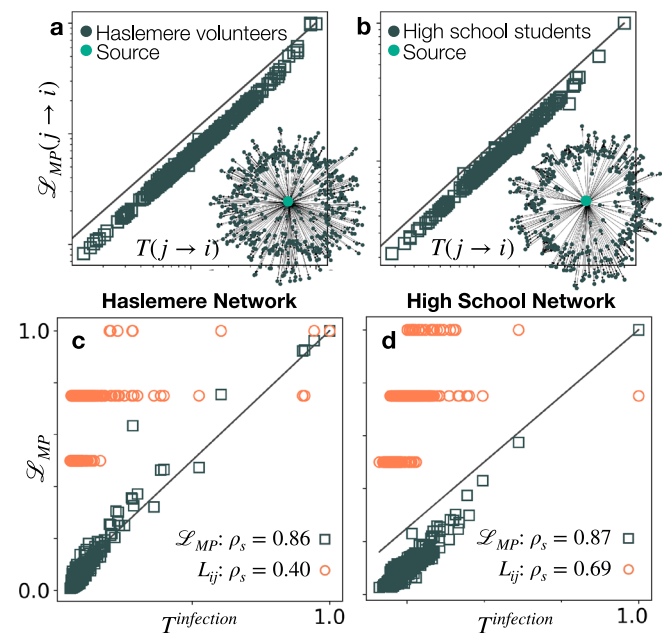

Fig. 5 Epidemic spreading. We model a Susceptible-Infected-Susceptible dynamics on real networks of contacts, and predict the arrival times of a perturbation in the epidemic state. The BBC Four Pandemic[24] dataset (**a**) and the High school[25] network of contacts between students (**b**) are paired with the latent geometry of the spreading phenomenon with susceptible nodes at a radial distance $r \propto \mathcal{L}_{MP}$ from the source. In **c**, **d**, the infection times $T^{infection}(s \to i)$ are compared with the temporal multi-pathways (MP) distance $\mathcal{L}_{MP}$.

measures described so far have been devised to estimate the temporal propagation of a signal in the proximity of a steady state, exploiting the linearity conditions; the purpose of this section is to test this measure in a non-linearity regime. In Supplementary Note 7 we also discuss an application of $\mathcal{L}_{MP}$ on diffusion dynamics. We first validate the distance $\mathcal{L}_{MP}$ by comparing it with perturbation propagation times $T(j \to i)$, panels A and B in Fig. 5. We then simulate an infection outbreak from the

source node s, with the initial condition for the network state: $x_i = 0, \forall\, i, i \neq s$, and source node state $x_s(t=0) = 1.0$. We then solve the set of coupled equations up to the steady state, and calculate the empirical time required for a target node i to get infected $T^{\text{infection}}(s \rightarrow i)$ as the time in which its probability overcomes a 0.5 threshold: $x_i(T^{\text{infection}}) = 0.5$. Results of infection times versus $\mathcal{L}_{MP}$ are presented for the two empirical networks in panels C and D of Fig. 5. We find that, despite losing its accuracy found in the linearity regime, the measure improves over L_{ij} shortest path distance and allows for a better understanding of infection times even in absence of more detailed contact information (possibly encoded in edges weights).

Discussion

The task of predicting the temporal interval for a perturbation originating from a node to reach all nodes in the network is of paramount importance in understanding spreading behaviors and predicting signaling and response properties in the context of network control[9–11]. We address this challenge working in a theoretical framework that allows simple and intuitive understanding of perturbation patterns, building on a recent advances in this direction[2,5,13,14,19]. In this article we have performed a path-driven analysis of perturbation propagation in complex networks, moving beyond the traditional ansatz that perturbations spread mostly along shortest paths among network units. Our findings highlight that multiple paths contribute to propagate perturbations and, under the rationale of the exponential decay of correlations, we redefine G_{ij} matrix elements by exploiting the vast ensemble of pathways that connect any pair of nodes. We have introduced a new metric, $\mathcal{L}_{MP}(j \rightarrow i)$, which accurately predicts the arrival times of a perturbation for any dynamical process that can be described by a set of non-linear functions, combining universal properties predicted by scaling laws that can be estimated analytically without relying on simulations. We have validated numerically these predictions using models for Epidemic spreading, Regulatory interactions and Population dynamics, finding that despite the necessary approximations introduced by scaling laws predictions, the concerted behavior of different paths in transmitting the signal is well captured by the new metric. Finally, we exploit the prediction of the latent geometry of these processes to embed the network in a space in which the spreading of a perturbation resembles the traveling wavefront of a reaction–diffusion-like process, further developing the bridge between network-driven processes and the geometry they induce[13].

Data availability

The data used in this work are publicly available from the original references.

Code availability

The code to perform the analysis will be available upon request.

References

1. Pastor-Satorras, R., Castellano, C., Mieghem, P. V. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
2. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342 (2013).
3. Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y. & Zhou, C. Synchronization in complex networks. *Phys. Rep.* **469**, 93–153 (2008).
4. Acebrón, J. A., Bonilla, L. L., Vicente, C. J. P., Ritort, F. & Spigler, R. The kuramoto model: a simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* **77**, 137–185 (2005).
5. De Domenico, M. Diffusion geometry unravels the emergence of functional clusters in collective phenomena. *Phys. Rev. Lett.* **118**, 168301 (2017).
6. Maslov, S. & Ispolatov, I. Propagation of large concentration changes in reversible protein-binding networks. *Proc. Natl Acad. Sci. USA* **104**, 13655–13660 (2007).
7. Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, 2006).
8. Barzel, B. & Barabási, A.-L. Universality in network dynamics. *Nat. Phys.* **9**, 673–681 (2013).
9. Hens, C., Harush, U., Haber, S., Cohen, R. & Barzel, B. Spatiotemporal signal propagation in complex networks. *Nat. Phys.* **15**, 403–412 (2019).
10. Zhang, X. et al. Topological determinants of perturbation spreading in networks. *Phys. Rev. Lett.* **125**, 218301 (2020).
11. Schroder, M., Zhang, X., Wolter, J. & Timme, M. Dynamic perturbation spreading in networks. *IEEE Trans. Netw. Sci. Eng.* **7**, 1019–1026 (2020).
12. Gautreau, A., Barrat, A. & Barthélemy, M. Arrival time statistics in global disease spread. *J. Stat. Mech. Theory Exp.* **2007**, L09001 (2007).
13. Boguñá, M. et al. Network geometry. *Nat. Rev. Phys.* **3**, 114–135 (2021).
14. Iannelli, F., Koher, A., Brockmann, D., Hövel, P. & Sokolov, I. M. Effective distances for epidemics spreading on complex networks. *Phys. Rev. E* **95**, 012313 (2017).
15. Barzel, B. & Biham, O. Quantifying the connectivity of a network: the network correlation function method. *Phys. Rev. E* **80**, 046104 (2009).
16. Ghavasieh, A., Bontorin, S., Artime, O., Verstraete, N. & De Domenico, M. Multiscale statistical physics of the pan-viral interactome unravels the systemic nature of SARS-COV-2 infections. *Commun. Phys.* **4**, 83 (2021).
17. Newman, M. *Networks* (Oxford University Press, 2010).
18. Meena, C. et al. Emergent stability in complex network dynamics. *Nat. Phys.* 1–10 (2023).
19. Barzel, B., Liu, Y.-Y. & Barabási, A.-L. Constructing minimal models for complex system dynamics. *Nat. Commun.* **6**, 7186 (2015).
20. Harush, U. & Barzel, B. Dynamic patterns of information flow in complex networks. *Nat. Commun.* **8**, 2181 (2017).
21. Estrada, E. & Hatano, N. Communicability in complex networks. *Phys. Rev. E* **77**, 036111 (2008).
22. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
23. Fisher, R. A. The wave of advance of advantageous genes. *Ann. Eugenics* **7**, 355–369 (1937).
24. Klepac, P., Kissler, S. & Gog, J. Contagion! The BBC Four Pandemic – The model behind the documentary. *Epidemics* **24**, 49–59 (2018).
25. Fournet, J. & Barrat, A. Contact patterns among high school students. *PLoS ONE* **9**, e107878 (2014).
26. Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge University Press, 2008).
27. Gilbert, E. N. Random graphs. *Ann. Math. Stat.* **30**, 1141–1144 (1959).

Acknowledgements

Author contributions

S.B. and M.D.D. contributed to design the study: S.B. contributed the numerical experiments; S.B. and M.D.D. contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-023-01204-1.

Correspondence and requests for materials should be addressed to Sebastiano Bontorin or Manlio De Domenico.

Peer review information *Communications Physics* thanks Hoi-To Wai and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission 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.