Science Advances
AAAS

Supplementary Materials for

Flow stability for dynamic community detection

Alexandre Bovet*, Jean-Charles Delvenne, Renaud Lambiotte

*Corresponding author. Email: alexandre.bovet@maths.ox.ac.uk

Published 11 May 2022, *Sci. Adv.* **8**, eabj3063 (2022)
DOI: 10.1126/sciadv.abj3063

This PDF file includes:

Supplementary Text

Relations with community detection in static networks

The expression of the clustered-covariance eq. (12) encompasses several well-known heuristics for the clustering of static network as special cases. The simplest example is the case of an undirected static network with M edges described by the adjacency matrix $\mathbf{A}$ (see Supplementary Text Relations with co-clustering and Tab. S4 for more examples). Considering a discrete time random walk on this network, the transition matrix after n steps is given by $\mathbf{T}(n) = (\mathbf{D}^{-1}\mathbf{A})^n$

where $\mathbf{D}$ is the diagonal matrix with diagonal element (i,i) equal the the degree k_i of vertex i. The stationary distribution of the random walk is given by the vector $\boldsymbol{\pi}$, with elements $\pi_i = k_i/2M$. In this case, the element (i,j) of the clustered covariance (eq. 12), computed after one step and evaluated at stationarity, reduces to

$$R_{ij}(n=1;H) = \left(\pi_i \frac{A_{ij}}{k_i} - \pi_i\pi_j\right)\delta\left(c_i, c_j\right) = \frac{1}{2m}\left(A_{ij} - \frac{k_ik_j}{2m}\right)\delta\left(c_i, c_j\right). \tag{S1}$$

Or in matrix notation

$$\mathbf{R}(n=1;H) = \mathbf{H}^T\left[\mathbf{\Pi}\mathbf{D}^{-1}\mathbf{A} - \boldsymbol{\pi}^T\boldsymbol{\pi}\right]\mathbf{H} = \frac{1}{2m}\mathbf{H}^T\mathbf{B}\mathbf{H}, \tag{S2}$$

where $\mathbf{B}$ is the modularity matrix (*55,56*). We recognize the classical Newman-Girvan modularity (*57*) by taking the trace of the clustered covariance: $Q = \text{trace}\left[\mathbf{R}(n=1;H)\right]$ (*32*). Finding a partition that maximizes the modularity, i.e. the number of observed edges inside each clusters minus the number expected from a random null model, can then be seen as finding a partition that the maximizes the elements on the diagonal of $\mathbf{R}(n=1;H)$, namely the probability that the walkers stay in the same clusters after one step minus the same probability for two independent walkers, evaluated at stationarity. This analogy allows us to see that the random null model of the modularity corresponds to the outer product of the stationary distribution of the random walks. As a matter of fact, the stationary distributions of different models of random walks correspond to different generative network null models (*35*). This random walk framework has been shown to be a very fruitful way to generalize modularity optimization and unify different clustering heuristics (*32, 35, 42*). By allowing the walkers to make multiple steps (*32*), or by considering a continuous time random walk (*35*), one can use the elapsed time of the random walk as a resolution parameter allowing to recover the multiscale community structure of networks (*42*) and overcome the resolution limit of the Newman-Girvan modularity (*41*).

A particularly interesting special case of application of eq. (12) is the case of a static directed network with M edges and with adjacency matrix $\mathbf{A}$. The in-degrees of node i is $k_i^{\text{in}} = \sum_j A_{ji}$ and its out-degree is $k_i^{\text{out}} = \sum_j A_{ij}$. The transition matrix after one step is given by $T_{ij} = A_{ij}/k_i^{\text{out}}$ if $k_i^{\text{out}} \neq 0$ and $T_{ij} = 0$ if $k_i^{\text{out}} = 0$. Considering an initial distribution of walkers given by $p_i(0) = k_i^{\text{out}}/M$, the distribution after one step is given by $p_i(1) = \sum_j p_j(0)T_{ij} = k_i^{\text{in}}/M$. Replacing these expressions in eq. (12) and taking the trace of the clustered covariance matrix, we find

$$\text{trace}\left[\mathbf{R}(n=1;H)\right] = \frac{1}{m}\sum_{ij}\left(A_{ij} - \frac{k_i^{\text{out}}k_j^{\text{in}}}{m}\right)\delta\left(c_i, c_j\right) = Q^{\text{d.}}, \tag{S3}$$

which is a classical generalization of modularity to directed networks (*26, 58*).

Relations with co-clustering

It is interesting to note that the clustering of symmetric covariance matrices with the Markov stability framework is linked to the spectral approaches of graph clustering (*32*). Indeed, as the time parameter increases, the contribution of the eigenvectors of the transition matrix, which are similar to the ones of the random walk graph Laplacian, $\mathbf{L}$ (*3*), to the stability are re-weighted according to their eigenvalues to give more weight to larger and larger scales in the network. In the static undirected case, the random walk has a stationary distribution, $\boldsymbol{\pi}$, i.e. $\boldsymbol{\pi}$ is a left-eigenvector of $\mathbf{T}$ with eigenvalue 1. The covariance is given by $\mathbf{S}_{static}(\tau) = \boldsymbol{\Pi}e^{-\tau\mathbf{L}} - \boldsymbol{\pi}^T\boldsymbol{\pi}$.

In the case of asymmetric matrices, spectral clustering approaches usually rely on the singular vectors rather than on the eigenvectors to capture the structural asymmetries of a system (*59*). Similarily, the forward and backward clustering of our framework can be related to the clustering of the singular vectors of the transition matrix. In the temporal case, the existence of a stationary distribution is not guaranteed, however, we have $\mathbf{p}(t_1)\mathbf{T}(t_1,t_2)\mathbf{T}^{\text{inv}}(t_2,t_1) = \mathbf{p}(t_1)$ and $\mathbf{p}(t_2)\mathbf{T}_{\text{inv}}(t_2,t_1)\mathbf{T}(t_1,t_2) = \mathbf{p}(t_2)$, i.e. $\mathbf{p}(t_1)$ and $\mathbf{p}(t_2)$ are left-eigenvectors of $\mathbf{T}(t_1,t_2)\mathbf{T}^{\text{inv}}(t_2,t_1)$ and $\mathbf{T}^{\text{inv}}(t_2,t_1)\mathbf{T}(t_1,t_2)$, respectively, with eigenvalue 1. If the processes defined by $\mathbf{T}(t_1,t_2)\mathbf{T}^{\text{inv}}(t_2,t_1)$ and $\mathbf{T}^{\text{inv}}(t_2,t_1)\mathbf{T}(t_1,t_2)$ are irreducibles, $\mathbf{p}(t_1)$ and $\mathbf{p}(t_2)$ are their respective stationary distributions. Using the covariances of the forward and inverse backward flows (eqs. 2 & S4) is therefore a natural generalization of the Markov stability in the stationary case to the non-stationary case. Moreover, the inverse transition matrix, $\mathbf{T}^{\text{inv}}(t_2,t_1) = \mathbf{P}(t_2)^{-1}\mathbf{T}(t_1,t_2)^T\mathbf{P}(t_1)$, can be seen as the adjoint operator of $\mathbf{T}(t_1,t_2)$ with respect to the inner product $\langle x,y\rangle_t = \sum_i x_iy_i/p_i(t)$, for which we have $\langle\mathbf{p}(t_1)\mathbf{T}(t_1,t_2),\mathbf{p}(t_2)\rangle_{t_2} = \langle\mathbf{p}(t_1),\mathbf{p}(t_2)\mathbf{T}^{\text{inv}}(t_2,t_1)\rangle_{t_1}$. The vectors $\mathbf{p}(t_1)$ and $\mathbf{p}(t_2)$ are therefore singular vectors of the transition matrix with respect to this inner product.

Special cases of the random walk covariances in static networks

Table S4 shows how Modularity (*57*), directed-Modularity (*58,60*) and Markov Stability (*32,35*) can be constructed from special cases of the non-stationary clustered covariance from eq. (12). Similarly, Tab. S4 shows that the clustering of static directed networks using the bibliographic coupling and co-citation matrices (*61*) are special cases of the clustering with the forward and backward non-stationary covariances from eqs. (2) and (3), respectively.

Covariances of inverse processes

An alternative backward process than the one defined in eq. (3) can be constructed by considering the inverse of the process that started at t_1 instead of the reversed evolution of the network. In this case, the corresponding covariance is given by

$$\begin{aligned} \mathbf{S}^{\text{inv}}_{\text{back}}(t_1, t) &= \mathbf{P}(t)\mathbf{T}^{\text{inv}}(t, t_1)\mathbf{T}(t_1, t) - \mathbf{p}(t)^T\mathbf{p}(t) \\ &= \mathbf{T}(t_1, t)^T\mathbf{P}(t_1)\mathbf{T}(t_1, t) - \mathbf{p}(t)^T\mathbf{p}(t). \end{aligned} \tag{S4}$$

Similarily, the covariance of the inverse backward process of eq. 3 is given by

$$\begin{aligned} \mathbf{S}^{\text{inv}}_{\text{forw}}(t_2, t) &= \mathbf{P}(t)\mathbf{T}^{\text{inv}}_{\text{rev}}(t, t_2)\mathbf{T}_{\text{rev}}(t_2, t) - \mathbf{p}(t)^T\mathbf{p}(t) \\ &= \mathbf{T}(t_2, t)^T_{\text{rev}}\mathbf{P}(t_1)\mathbf{T}_{\text{rev}}(t_2, t) - \mathbf{p}(t)^T\mathbf{p}(t). \end{aligned} \tag{S5}$$

A difference between these two definitions is the choice of the initial condition, which is at t_2 for eq. 3 and at t_1 for eq. S4. The matrices $\mathbf{S}_{\text{forw}}(t_1, t)$ and $\mathbf{S}^{\text{inv}}_{\text{back}}(t_1, t)$ are both covariances of the same diffusion process that start at t_1 and evolve until $t > t_1$ while the matrices $\mathbf{S}_{\text{forw}}(t_1, t)$ and $\mathbf{S}_{\text{back}}(t_2, t)$ are the covariances of two different processes, the first starting at t_1 and evolving in the direction of time and the second starting at t_2 and evolving backward in time. Here, we prefer to use $\mathbf{S}_{\text{forw}}(t_1, t)$ and $\mathbf{S}_{\text{back}}(t_2, t)$ for the general clustering of temporal networks between t_1 and t_2 using $\mathbf{p}(t_1)$ and $\mathbf{p}(t_2)$ as two uniform distributions. Using the inverse covariances (S4) and (S5) may, for example, be preferred when studying a specific diffusion process.

Importance of early and late times on the optimal partitions

We consider a simple example of temporal network with eight nodes ($N = 8$) that initially forms two communities of four nodes each and after a time $t^\star$ split to form four communities of two nodes. The question we want to answer is how does the partition maximizing the forward flow stability (eq. 4) changes as a function of $t^\star$ when the integration goes from $t_1 = 0 < t^\star$ until $t > t^\star$. The Laplacian matrix from $t = 0$ until $t = t^\star$ is a matrix with four 4×4 blocks. The off diagonal blocks are zero matrices and the diagonal blocks are two similar matrices given by

$$\mathbf{L}_A = \begin{pmatrix} 1 & -1/3 & -1/3 & -1/3 \\ -1/3 & 1 & -1/3 & -1/3 \\ -1/3 & -1/3 & 1 & -1/3 \\ -1/3 & -1/3 & -1/3 & 1 \end{pmatrix}. \tag{S6}$$

For times $t > t^\star$, the Laplacian has the same block structure with diagonal blocks given by

$$\mathbf{L}_B = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}. \tag{S7}$$

The two Laplacians are symmetric and commute. They are therefore simultaneously diagonalisable, i.e $\mathbf{L}_A = \mathbf{U}\mathbf{\Lambda}_A\mathbf{U}^T$ and $\mathbf{L}_B = \mathbf{U}\mathbf{\Lambda}_B\mathbf{U}^T$ with $\mathbf{\Lambda}_A = \mathrm{diag}((4/3, 4/3, 4/3, 0))$, $\mathbf{\Lambda}_B = \mathrm{diag}((2, 2, 0, 0))$ and where

$$\mathbf{U} = \begin{pmatrix} 0 & -1/\sqrt{2} & -1/2 & 1/2 \\ 0 & 1/\sqrt{2} & -1/2 & 1/2 \\ -1/\sqrt{2} & 0 & 1/2 & 1/2 \\ 1/\sqrt{2} & 0 & 1/2 & 1/2 \end{pmatrix} \tag{S8}$$

is a unitary matrix. The transition matrix is given by

$$\mathbf{T}(0,t) = \begin{cases} \mathbf{U}e^{-\lambda t\mathbf{\Lambda}_A}\mathbf{U}^T = \mathbf{U}\mathbf{\Sigma}_A(t)\mathbf{U}^T & \text{if } 0 \le t \le t^\star \\ \mathbf{U}e^{-\lambda t^\star\mathbf{\Lambda}_A}e^{-\lambda(t-t^\star)\mathbf{\Lambda}_B}\mathbf{U}^T = \mathbf{U}\mathbf{\Sigma}_A(t^\star)\mathbf{\Sigma}_B(t-t^\star)\mathbf{U}^T & \text{if } t > t^\star \end{cases}, \tag{S9}$$

where λ is the random walk rate. In the rest of this section, we use the notation $\mathbf{T}(t)$ as meaning $\mathbf{T}_1(0,t)$. We can now calculate the forward covariance (eq. 2) taking $\mathbf{p} = \frac{1}{N}(1111)$ as initial condition. Note that $\mathbf{p}$ is a stationary state of the system, i.e. $\mathbf{p} = \mathbf{p}\mathbf{T}(t)\ \forall t$ such that $t \geq 0$. The forward covariance is given by $\mathbf{S}_{\text{forw}}(t) = \mathbf{P}\mathbf{T}(t)\mathbf{P}^{-1}\mathbf{T}(t)^T\mathbf{P} - \mathbf{p}^T\mathbf{p} = \frac{1}{N}\mathbf{T}(t)\mathbf{T}(t)^T - \frac{1}{N^2}\overleftrightarrow{\mathbf{1}}$. Or, using eq. (S9)

$$\mathbf{S}_{\text{forw}}(t) = \begin{cases} \frac{1}{N}\mathbf{U}\mathbf{\Sigma}_A^2(t)\mathbf{U}^T - \frac{1}{N^2}\overleftrightarrow{\mathbf{1}} & \text{if } 0 \leq t \leq t^\star \\ \frac{1}{N}\mathbf{U}\mathbf{\Sigma}_A^2(t^\star)\mathbf{\Sigma}_B^2(t-t^\star)\mathbf{U}^T - \frac{1}{N^2}\overleftrightarrow{\mathbf{1}} & \text{if } t \geq t^\star \end{cases}. \tag{S10}$$

We find the forward flow stability by integrating $\mathbf{S}_{\text{forw}}(t')$ from 0 to t and dividing by t. This yields the matrix

$$\mathbf{F}_{\text{forw}}(t) = \begin{cases} \frac{1}{tN}\mathbf{U}\int_0^t \mathbf{\Sigma}_A^2(t')dt'\mathbf{U}^T - \frac{1}{N^2}\overleftrightarrow{\mathbf{1}} & \text{if } 0 \leq t \leq t^\star \\ \frac{t^\star}{t}\mathbf{F}_{\text{forw}}(t^\star) + \frac{1}{tN}\mathbf{U}\mathbf{\Sigma}_A^2(t^\star)\int_{t^\star}^t \mathbf{\Sigma}_B^2(t'-t^\star)dt'\mathbf{U}^T - \frac{t-t^\star}{tN^2}\overleftrightarrow{\mathbf{1}} & \text{if } t \geq t^\star \end{cases}, \tag{S11}$$

where $\int_0^t \mathbf{\Sigma}_A^2(t')dt' = \text{diag}\big((\frac{3}{8\lambda}(1-e^{-\frac{8}{3}\lambda t}), \frac{3}{8\lambda}(1-e^{-\frac{8}{3}\lambda t}), \frac{3}{8\lambda}(1-e^{-\frac{8}{3}\lambda t}), t)\big)$ and $\int_{t^\star}^t \mathbf{\Sigma}_B^2(t'-t^\star)dt' = \text{diag}\big((\frac{1}{4\lambda}(1-e^{-4\lambda(t-t^\star)}), \frac{1}{4\lambda}(1-e^{-4\lambda(t-t^\star)}, t-t^\star, t-t^\star)\big)$. The resulting matrix $\mathbf{F}_{\text{forw}}(t)$ is formed by four 4×4 blocks. The off diagonal blocks are equal to $-\frac{1}{N^2}\overleftrightarrow{\mathbf{1}}$ and the diagonal blocks have the following form

$$\begin{pmatrix} F_{11}^{\text{forw}}(t) & F_{12}^{\text{forw}}(t) & F_{13}^{\text{forw}}(t) & F_{13}^{\text{forw}}(t) \\ F_{12}^{\text{forw}}(t) & F_{11}^{\text{forw}}(t) & F_{13}^{\text{forw}}(t) & F_{13}^{\text{forw}}(t) \\ F_{13}^{\text{forw}}(t) & F_{13}^{\text{forw}}(t) & F_{11}^{\text{forw}}(t) & F_{12}^{\text{forw}}(t) \\ F_{13}^{\text{forw}}(t) & F_{13}^{\text{forw}}(t) & F_{12}^{\text{forw}}(t) & F_{11}^{\text{forw}}(t) \end{pmatrix}. \tag{S12}$$

Moreover, for $0 \leq t \leq t^\star$, $F_{12}^{\text{forw}}(t) = F_{13}^{\text{forw}}(t)$.

The partition maximizing the forward flow stability groups together positive elements of $\mathbf{F}_{\text{forw}}(t)$ and avoids its negative elements. As the off diagonal blocks are always negative, regardless of t, the optimal partition depends on the signs of $F_{11}^{\text{forw}}(t)$, $F_{12}^{\text{forw}}(t)$ and $F_{13}^{\text{forw}}(t)$ as a function of t.

We have

$$F_{11}^{\text{forw}}(t) = \begin{cases} \frac{9}{32N\lambda t}\left(1 - e^{-\frac{8}{3}\lambda t}\right) + \frac{N-4}{4N^2} & \text{if } 0 \le t \le t^\star \\ \frac{1}{32N\lambda t}\left(e^{-\frac{8}{3}\lambda t^\star}\left(8\lambda(t-t^\star) - 4e^{-4\lambda(t-t^\star)} - 5\right) + 9 + 8\lambda t\right) - \frac{1}{N^2} & \text{if } t \ge t^\star \end{cases}, \tag{S13}$$

which is always positive, with $F_{11}^{\text{forw}}(0) = \frac{N-1}{N^2}$ and $F_{11}^{\text{forw}}(t) \to \frac{1}{4N^2}\left(N\left(e^{-\frac{8}{3}\lambda t^\star} + 1\right) - 4\right)$ as $t \to \infty$.

We have

$$F_{12}^{\text{forw}}(t) = \begin{cases} \frac{3}{32N\lambda t}\left(e^{-\frac{8}{3}\lambda t} - 1\right) + \frac{N-4}{4N^2} & \text{if } 0 \le t \le t^\star \\ \frac{1}{32N\lambda t}\left(e^{-\frac{8}{3}\lambda t^\star}\left(8\lambda(t-t^\star) + 4e^{-4\lambda(t-t^\star)} - 1\right) - 3 + 8\lambda t\right) - \frac{1}{N^2} & \text{if } t \ge t^\star \end{cases}, \tag{S14}$$

which is negative at $t = 0$ with $F_{12}^{\text{forw}}(0) = -\frac{1}{N^2}$ and then increases monotonically. At $t^\star$, $F_{12}^{\text{forw}}(t^\star)$ is positive only if $t^\star > \hat{t}_{\text{f}}$, where $\hat{t}_{\text{f}}$ is the time at which $F_{12}^{\text{forw}}(t)$ crosses the x-axis if this happens before $t^\star$. Its value is given by

$$\hat{t}_{\text{f}} = \frac{3}{8\lambda}\left(\frac{N}{N-4} + W_0\left(-\frac{N}{N-4}e^{-\frac{N}{N-4}}\right)\right), \tag{S15}$$

where W_0 is the principal branch of the Lambert W function. We see that $\hat{t}_{\text{f}}$ is made of two terms, a linear coefficient $\frac{1}{\lambda}$ and a constant term depending only on N (which is fixed at $N = 8$ here). By varying λ one can therefore adjust $\hat{t}_{\text{f}}$ in order to make $F_{12}^{\text{forw}}(t)$ positive or negative for any t such that $0 < t \le t^\star$. As $t \to \infty$, $F_{12}^{\text{forw}}(t) \to \frac{1}{4N^2}\left(N\left(e^{-\frac{8}{3}\lambda t^\star} + 1\right) - 4\right) > 0$ indicating that even if $F_{12}^{\text{forw}}(t^\star)$ is negative, $F_{12}^{\text{forw}}(t)$ eventually becomes positive.

For $F_{13}^{\text{forw}}(t)$, we have

$$F_{13}^{\text{forw}}(t) = \begin{cases} F_{12}^{\text{forw}}(t) & \text{if } 0 \le t \le t^\star \\ \frac{1}{32N\lambda t}\left(e^{-\frac{8}{3}\lambda t^\star}\left(3 - 8\lambda(t-t^\star)\right) - 3 + 8\lambda t\right) - \frac{1}{N^2} & \text{if } t \ge t^\star \end{cases}. \tag{S16}$$

Similarly to $F_{12}^{\text{forw}}(t)$, $F_{13}^{\text{forw}}(t^\star) > 0$ only if $t^\star > \hat{t}_{\text{f}}$. As $t \to \infty$, $F_{13}^{\text{forw}}(t) \to \frac{1}{4N^2}\left(N\left(1 - e^{-\frac{8}{3}\lambda t^\star}\right) - 4\right)$ which is positive only if $t^\star > t_{13} = \frac{3}{8\lambda}\ln\frac{N}{N-4}$ whose value can again be controlled by varying λ.

we remark that $F^{\text{forw}}_{11}(t) \geq F^{\text{forw}}_{12}(t) \geq F^{\text{forw}}_{13}(t)$, we have three possible configurations: A) $F^{\text{forw}}_{11}(t) > 0$, $F^{\text{forw}}_{12}(t) > 0$ and $F^{\text{forw}}_{13}(t) > 0$: the optimal partition is composed of two communities of size 4; B) $F^{\text{forw}}_{11}(t) > 0$, $F^{\text{forw}}_{12}(t) > 0$ and $F^{\text{forw}}_{13}(t) < 0$: the optimal partition is composed of four communities of size 2; C) $F^{\text{forw}}_{11}(t) > 0$, $F^{\text{forw}}_{12}(t) < 0$ and $F^{\text{forw}}_{13}(t) < 0$: the optimal partition is composed of 8 singleton communities.

Noticing that $t_{13} < \hat{t}_{\text{f}}$, we therefore have three scenarios:

1) $t^\star < t_{13} < \hat{t}_{\text{f}}$: $F^{\text{forw}}_{12}(t)$ and $F^{\text{forw}}_{13}(t)$ are negative at $t = t^\star$. The switch to four communities happens too fast for the walkers to have time to explore the two community structure. After $t^\star$, $F^{\text{forw}}_{13}(t)$ remains negative while $F^{\text{forw}}_{12}(t)$ eventually becomes positive. For short intervals, the configuration C is optimal. For longer intervals, the configuration B becomes optimal. This change can be controlled by varying λ.

2) $t_{13} < t^\star < \hat{t}_{\text{f}}$: As before, $F^{\text{forw}}_{12}(t)$ and $F^{\text{forw}}_{13}(t)$ are negative at $t = t^\star$. While the walkers have not fully explored the two communities structure at $t^\star$, they have sufficiently done so that after some time $F^{\text{forw}}_{12}(t)$ and $F^{\text{forw}}_{13}(t)$ both become positive and the optimal forward partition is given by configuration A.

3) $t_{13} < \hat{t}_{\text{f}} < t^\star$: $F^{\text{forw}}_{12}(t^\star >$ and $F^{\text{forw}}_{13}(t^\star)$ are already positive at $t = t^\star$ and remain positive afterward. The walkers have already fully explored the two community structure before $t^\star$ and the optimal forward partition remains given by configuration A.

Figure S1 shows a graph of $F^{\text{forw}}_{12}(t)$ and $F^{\text{forw}}_{13}(t)$ for different values of λ. We see that the importance of early or late times on the forward partition can be controlled by varying the value of the random walk rate. Indeed, the three conditions corresponding to the three scenarios are expressed as inequalities between $\lambda t^\star$ and constants that depends only on the structure of the network.

Considering the backward evolution by reversing time, the system starts in the configuration with four communities at $t = 0$ until $t^\star$ and then forms the structure in two communities. For the backward case, we have

$$\mathbf{T}(t)_{\text{rev}} = \begin{cases} \mathbf{U}\mathbf{\Sigma}_B(t)\mathbf{U}^T & \text{if } 0 \leq t \leq t^\star \\ \mathbf{U}\mathbf{\Sigma}_B(t^\star)\mathbf{\Sigma}_A(t - t^\star)\mathbf{U}^T & \text{if } t > t^\star \end{cases}, \tag{S17}$$

$$\mathbf{S}_{\text{back}}(t) = \begin{cases} \frac{1}{N}\mathbf{U}\mathbf{\Sigma}_B^2(t)\mathbf{U}^T - \frac{1}{N^2}\overleftrightarrow{\mathbf{1}} & \text{if } 0 \le t \le t^\star \\ \frac{1}{N}\mathbf{U}\mathbf{\Sigma}_B^2(t^\star)\mathbf{\Sigma}_A^2(t-t^\star)\mathbf{U}^T - \frac{1}{N^2}\overleftrightarrow{\mathbf{1}} & \text{if } t \ge t^\star \end{cases}, \tag{S18}$$

$$\mathbf{F}_{\text{back}}(t) = \begin{cases} \frac{1}{tN}\mathbf{U}\int_0^t \mathbf{\Sigma}_B^2(t')dt'\mathbf{U}^T - \frac{1}{N^2}\overleftrightarrow{\mathbf{1}} & \text{if } 0 \le t \le t^\star \\ \frac{t^\star}{t}\mathbf{F}_{\text{back}}(t^\star) + \frac{1}{tN}\mathbf{U}\mathbf{\Sigma}_B^2(t^\star)\int_{t^\star}^t \mathbf{\Sigma}_A^2(t'-t^\star)dt'\mathbf{U}^T - \frac{t-t^\star}{tN^2}\overleftrightarrow{\mathbf{1}} & \text{if } t \ge t^\star \end{cases}. \tag{S19}$$

The backward flow stability matrix has the same structure than in the forward case (eq. S12), but with

$$F_{11}^{\text{back}}(t) = \begin{cases} \frac{1}{8N\lambda t}\left(1 - e^{-4\lambda t}\right) + \frac{N-2}{2N^2} & \text{if } 0 \le t \le t^\star \\ \frac{1}{32N\lambda t}(e^{-4\lambda t^\star}(2 - 6e^{-\frac{8}{3}\lambda(t-t^\star)} - 3e^{-\frac{4}{3}\lambda(2t-5t^\star)}) + \\ 7 + 8\lambda(t+t^\star)) - \frac{1}{N^2} & \text{if } t \ge t^\star \end{cases}, \tag{S20}$$

which is always positive with a value of $\frac{N-1}{N^2}$ at $t = 0$ and $\frac{N-4}{4N^2}$ as $t \to \infty$.

$$F_{12}^{\text{back}}(t) = \begin{cases} \frac{1}{8N\lambda t}\left(e^{-4\lambda t} - 1\right) + \frac{N-2}{2N^2} & \text{if } 0 \le t \le t^\star \\ \frac{1}{32N\lambda t}(e^{-4\lambda t^\star}(6e^{-\frac{8}{3}\lambda(t-t^\star)} - 2 - 3e^{-\frac{4}{3}\lambda(2t-5t^\star)}) - \\ 1 + 8\lambda(t+t^\star)) - \frac{1}{N^2} & \text{if } t \ge t^\star \end{cases}, \tag{S21}$$

which starts with a negative value of $-\frac{1}{N^2}$ at $t = 0$ and increases until $t = t^\star$. $F_{12}^{\text{back}}(t^\star)$ is positive only if $t^\star > \hat{t}_{\text{b}}$ where

$$\hat{t}_{\text{b}} = \frac{1}{4\lambda}\left(\frac{N}{N-2} + W_0\left(-\frac{N}{N-2}e^{-\frac{N}{N-2}}\right)\right). \tag{S22}$$

As $t \to \infty$, $F_{12}^{\text{back}}(t) \to \frac{N-4}{4N^2}$, i.e. $F_{12}^{\text{back}}(t)$ eventually becomes positive if it was not the case at $t^\star$ or stays positive otherwise.

$$F_{13}^{\text{back}}(t) = \begin{cases} -\frac{1}{N^2} & \text{if } 0 \le t \le t^\star \\ \frac{1}{32N\lambda t}\left(3e^{-\frac{8}{3}\lambda(t-t^\star)} + 8\lambda(t-t^\star)\right) - \frac{1}{N^2} & \text{if } t \ge t^\star \end{cases}, \tag{S23}$$

which is negative until $t = t^\star$ and then increases monotonically, eventually becomes positive and reaches a value of $\frac{N-4}{4N^2}$ as $t \to \infty$. The time at which $F_{13}^{\text{back}}(t)$ becomes positive is given by

$$\hat{t}_{\text{b}}' = \frac{1}{8\lambda(N-4)}\left(3N + 8\lambda N t^{\star} + (3N-12)W_0\left(-\frac{N}{N-4}e^{-\frac{3N+32\lambda t^{\star}}{3N-12}}\right)\right). \tag{S24}$$

Contrary to $\hat{t}_{\text{f}}$ (eq. S15) and $\hat{t}_{\text{b}}$ (eq. S22) that both tend to zero as the speed of the walkers is increased, $\hat{t}_{\text{b}}'$ tends to $2t^{\star}$ as $\lambda \to \infty$. This indicates that $F_{13}^{\text{back}}(t)$ can become positive only for times larger than $2t^{\star}$. The importance of early and late time can therefore also be controlled with λ, however there is a limit on the possibility of detecting the first structure (two communities) if it lasts for a shorter time than the second structure (four communities). This is due to the fact that the second structure (configuration B) is composed of smaller communities. When looking at the network evolution from the point of view of the backward partition, i.e. from the end of the interval backward in time, the vision of the first structure can be obstructed by the second smaller structures. In this case, the first structure can be captured by the forward partition.

As for the forward partition, for the backward case we have $F_{11}^{\text{back}}(t) \geq F_{12}^{\text{back}}(t) \geq F_{13}^{\text{back}}(t)$ for $t > 0$. We therefore have the following scenarios: 1) $t < \hat{t}_{\text{b}}'$, $F_{11}^{\text{back}}(t)$ is positive and $F_{13}^{\text{back}}(t)$ is negative. The sign of $F_{12}^{\text{back}}(t)$ is controlled by the RW rate λ. For slow RWs, configuration C is optimal (singletons communities). For fast RWs, configuration B is optimal (four communities). 2) $t > \hat{t}_{\text{b}}'$, $F_{11}^{\text{back}}(t)$ is positive and the signs of $F_{12}^{\text{back}}(t)$ and $F_{13}^{\text{back}}(t)$ is controlled by λ. As the RW rate increases, the optimal backward partition changes from configurations C to B and finally A.

Figure S1 shows a graph of $F_{12}^{\text{back}}(t)$ and $F_{13}^{\text{back}}(t)$ for different values of λ showing that the evolution of the network can be captured by combining the solutions of the optimal forward and backward partitions for different values of λ.

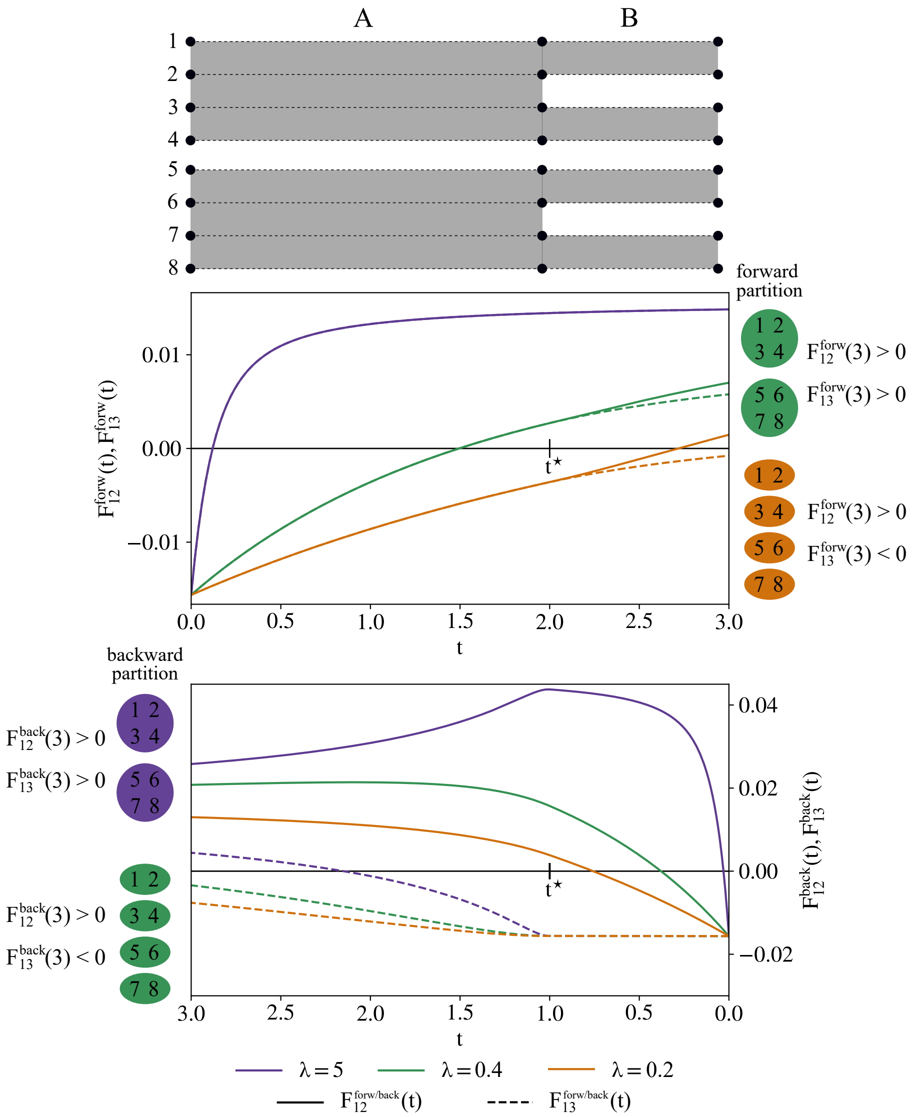

Figure S1: Graph of the functions $F_{12}^{\mathbf{forw}}(t)$ & $F_{13}^{\mathbf{forw}}(t)$ (top) and $F_{12}^{\mathbf{back}}(t)$ & $F_{13}^{\mathbf{back}}(t)$ (bottom) for different values of the random walk rate λ. The sign of these functions control whether the forward, respectively backward, optimal partitions take the form of the early times or later times. Here, the network splits from a structure, A, in 2 communities to a structure, B, in four communities at $t^\star = 2$. By varying the value of λ, we can give give more importance to the structure at early times (fast diffusion) or to the structure at later times (slow diffusion). By considering a time interval that lasts 3 time units, the backward process starts at $t_2 = 3$ with a reverse time evolution (bottom). $F_{13}^{\text{back}}(t)$ can be positive only for values of the reverse time larger than 2. We show three scenarios depending on the value of λ: 1) $\lambda = 5$ (purple), the optimal forward and backward partitions have the form of A; 2) $\lambda = 0.4$ (green), the optimal forward partition has the form of A and the optimal backward partition has the form of B; 3) $\lambda = 0.2$ (yellow), the optimal forward and backward partitions have the form of B.

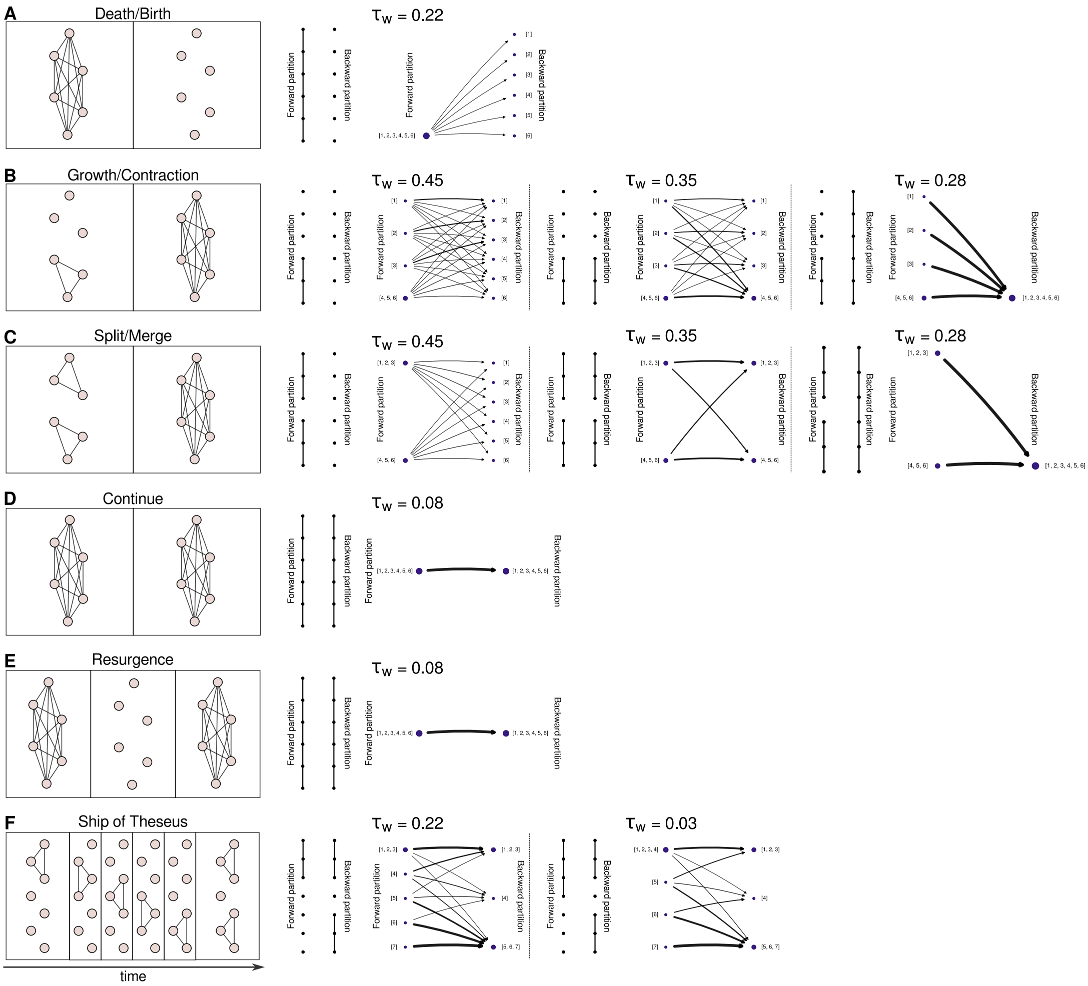

Figure S2: Results of the flow stability community detection applied to toy examples of dynamic community events. We reproduce the community events from Ref. (*27*) with the addition of the Ship of Theseus (*62*). For each event, we show the schematic evolution on the left and the flow stability results on the right. Note that our framework does not distinguish between a node that is absent or a node that is inactive. Such nodes are usually clustered in singleton communities. We show the results for several values of the resolution (waiting time τ_w) when several non trivial solutions exists. For each result, we represent the partitions in two manners: 1) the nodes as dots and the forward and backward communities as lines joining the dots, 2) as a bipartite graph where the nodes represents the communities and the edges represents the probability transitions from forward to backward communities of the random walk. The death/birth (**A**), growth/contraction (**B**), split/merge (**C**) and continue (**D**) events are well detected by our method. However, the flow stability is unable to distinguish the resurgence event (**E**) from the continue event as the absence of all connections in the middle results in an unchanged diffusion. To distinguish such situations, one has to split the time window in two resulting in a sequence of 'death' and 'birth' (**A**). In the ship of Theseus (**F**), the initial and final states are captures by the forward and backward partitions. To capture the dynamics between those states one needs to look at the transitions probabilities: there are non-zero probabilities to go from the initial ship to the two final ships, however there is a zero probability to go from nodes of the bottom ship at the beginning to the top ship at the end. We understand that the two final ships are linked to the initial ship.

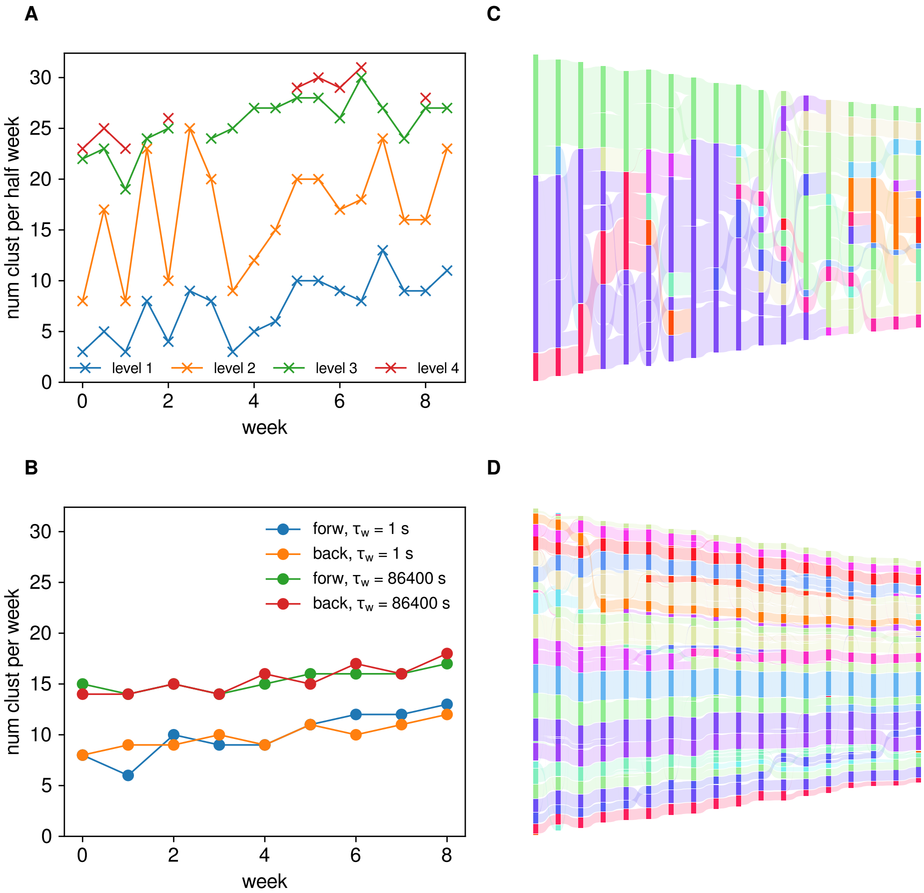

Figure S3: Community detection of the free-ranging house mice contact network using a series of static networks with an aggregation window of a half week. The hierarchical infomap algorithm (*50*) is used on each time slice and the evolution of the communities is tracked using the method developed in Ref. (*29*) with a history parameter of 8 time points. The infomap algorithm is run with the default parameters for the hierarchical case. (**A**) Number of communities per time slice found by infomap for different hierarchical levels. An issue with this approach is that the method does not necessarily find the same number of hierarchical levels at each time slice which makes the comparison from slice to slice not clear. (**B**) Number of communities per week found with the flow stability for the two resolutions shown in the main manuscript. As the resolution parameter of the flow stability can be interpreted as a physical quantity (characteristic waiting time of the random walk) the comparison between different time slices is done in a principled manner. We see that, contrary to the case in (**A**), the number of clusters per week varies more smoothly with the flow stability. (**C** & **D**) Result of the tracking of communities found with infomap for the coarsest level (i.e. level 1) and the finest level (i.e. 2, 3 or 4 depending on the time slice), respectively.

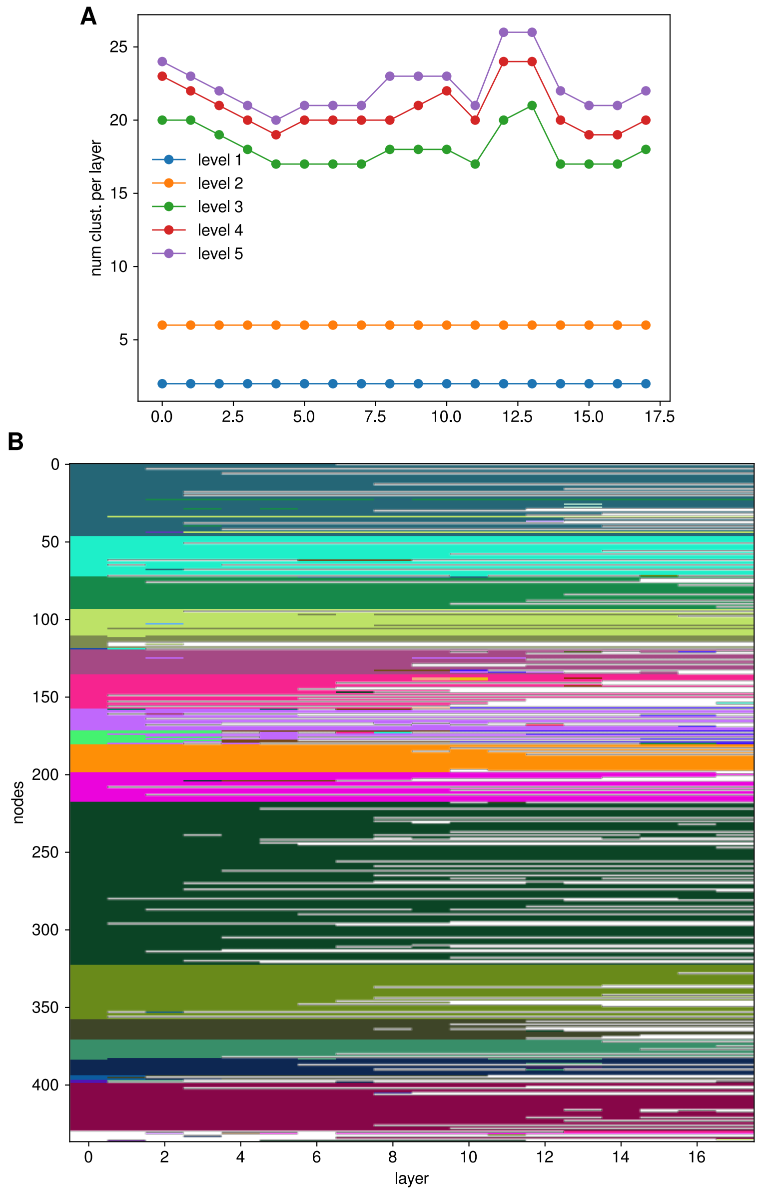

Figure S4: Hierarchical multilayer community detection with the Infomap algorithm applied to the free-ranging house mice contact network. A multilayer representation of this dataset is constructed by aggregating the activity in static networks over half-week time windows. The Infomap algorithm (*30*) is run with the parameters `flow_model='undirected'`, `multilayer_relax_by_jsd` for neighborhood flow coupling for temporal networks (*21*), `multilayer_relax_limit=1` limiting the RW to jump only to neighboring layers in order to encode the temporal ordering, and with a value `multilayer_relax_rate` of 0.001. (**A**) Number of communities per layer (half-week) found by the multilayer Infomap algorithm at different hierarchical levels. (**B**) Multilayer partition at hierarchical level 3. Here, infomap find 5 scales of communities considering all time slices simultaneously, however the communities found are all elongated in time and the dynamics of communities splitting found by the flow stability is not recovered.

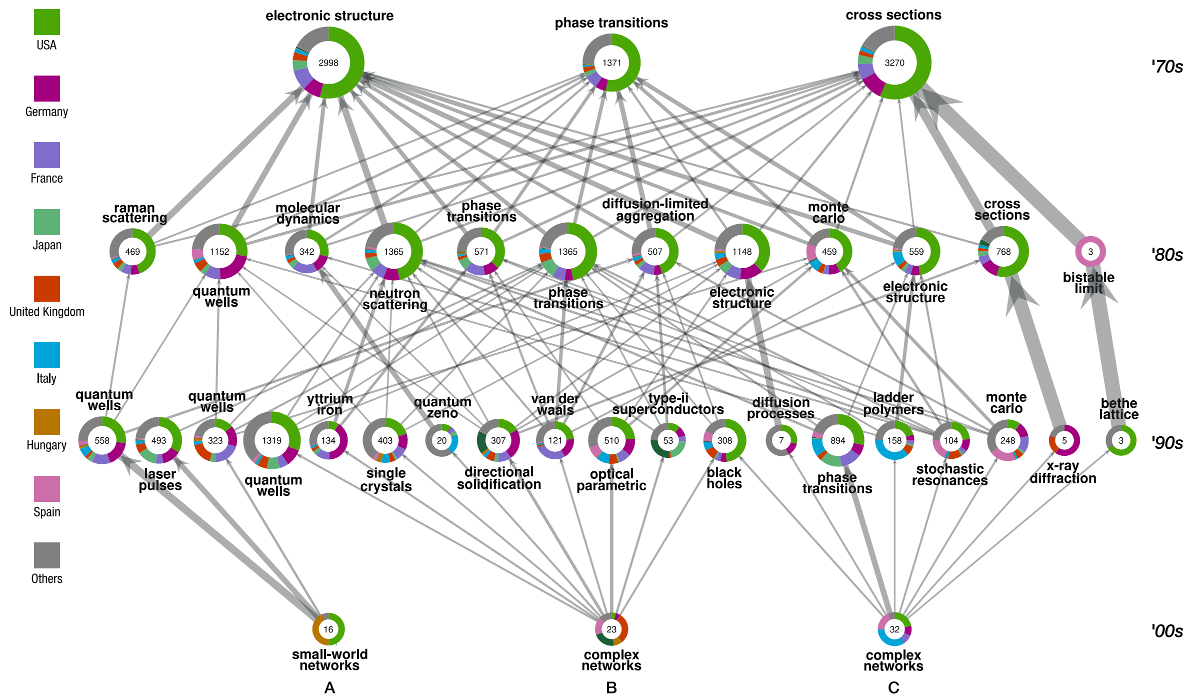

Figure S5: Influential communities of authors of articles published in the APS journals for three communities of network scientists in the 2000s. Each node represent a community and its size is indicated in the center. The colors represent the distribution of countries inside each community where each author is associated to the country most often associated with their affiliations. The pair of words next to each node indicate one of the most frequent pair of words of all the titles of the articles belonging to the community. Arrows between the communities represents probability transitions ($> 5\%$) from community to community of the diffusive process starting in 2010 and finishing in 1970.

slice length	static aggregation from start	from end	NMI with fast flow stability forward	backward	NMI with slow flow stability forward	backward
0.1	{a, b, c}, {d}	{a}, {b, c, d}	1.00	1.00	0.54	0.31
0.2	{a, b, c}, {d}	{a}, {b, c, d}	1.00	1.00	0.54	0.31
0.3	{a, b, c}, {d}	{a, d}, {b, c}	1.00	0.31	0.54	1.00
0.4	{a, b, c}, {d}	{a, d}, {b, c}	1.00	0.31	0.54	1.00
0.5	{a, d}, {b, c}	{a, d}, {b, c}	0.31	0.31	0.67	1.00
0.6	{a, d}, {b, c}	{a, d}, {b, c}	0.31	0.31	0.67	1.00
0.7	{a, d}, {b, c}	{a, d}, {b, c}	0.31	0.31	0.67	1.00
0.8	{a, d}, {b, c}	{a, d}, {b, c}	0.31	0.31	0.67	1.00
0.9	{a, d}, {b, c}	{a, d}, {b, c}	0.31	0.31	0.67	1.00
1.0	{a, d}, {b, c}	{a, d}, {b, c}	0.31	0.31	0.67	1.00

Table S1: Comparison between static partitions with different aggregation length and the flow stability partitions from the example in Fig. 1. The first column shows the slice lengths expressed as a ratio of the total length. The second and third columns show the partitions found by optimizing modularity on the networks found by aggregating from the beginning and end of the network with increasing slice lengths. The remaining columns show the value of the Normalized Mutual Information computed between the initial and final static partitions and the forward and backward flow stability partitions, respectively, for the case of the fast and slow diffusion of Fig. 1. Similar results are obtained when self-loops with weight corresponding to the inactivity time of nodes are added during the aggregation. We see that the NMI with the slow forward partition is never equal to one, indicating that the static aggregations cannot fully reproduce the results of the flow stability.

Forward	1	2	3	4	5	6	7	8	9	10
size	49	50	46	22	24	47	1	1	1	1
$\bar{T}_{\text{first}}$	Day 1 09:00	Day 1 09:32	Day 1 08:56	Day 1 09:04	Day 1 09:56	Day 1 09:04	Day 2 08:43	Day 2 08:41	Day 2 08:42	Day 2 08:42

Backward	1	2	3	4	5	6	7	8	9	10
size	24	26	51	46	24	22	46	1	1	1
$\bar{T}_{\text{last}}$	Day 2 17:09	Day 2 17:00	Day 2 17:03	Day 2 17:03	Day 2 13:21	Day 2 11:58	Day 2 17:00	Day 1 11:59	Day 1 17:08	Day 1 17:04

Table S2: Size, average time of the first ($\bar{T}_{\text{first}}$) and last ($\bar{T}_{\text{last}}$) contact for each cluster of the forward and backward flow stability partitions at scale $\tau_w = 1\,\text{h}$.

Forward	1	2	3	4	5	6	7	8	9	10	11
size	114	52	67	1	1	1	1	1	2	1	1
$\bar{T}_{\text{first}}$	Day 1 08:55	Day 1 09:04	Day 1 08:55	Day 1 13:21	Day 2 09:40	Day 2 08:42	Day 2 08:42	Day 2 08:40	Day 1 14:17	Day 2 08:41	Day 2 08:43

Backward	1	2	3	4	5	6	7	8	9	10	11	12
size	141	26	10	21	23	1	1	1	1	15	1	1
$\bar{T}_{\text{last}}$	Day 2 17:07	Day 2 17:08	Day 2 11:58	Day 2 12:52	Day 2 17:02	Day 1 17:04	Day 2 12:18	Day 2 13:44	Day 1 17:08	Day 2 14:18	Day 1 11:59	Day 1 17:10

Table S3: Size, average time of the first ($\bar{T}_{\text{first}}$) and last ($\bar{T}_{\text{last}}$) contact for each cluster of the forward and backward flow stability partitions at scale $\tau_w = 63\,\text{s}$.

	Probability density at t_1: $\mathbf{p}(t_1)$	Transition matrix: $\mathbf{T}(t_1,t_2)$	Probability density at t_2: $\mathbf{p}(t_2)$	Covariance matrix: $\mathbf{S}$	Partition quality function:
1) **General** non-stationary	$\mathbf{p}_1$	Eq. (7)	$\mathbf{p}_2 = \mathbf{p}_1\mathbf{T}(t_1,t_2)$	$\mathbf{P}_1\mathbf{T}(t_1,t_2) - \mathbf{p}_1^T\mathbf{p}_2$	
2)**Markov Stability**: static undirected	$\pi_i = \frac{k_i}{2M}, \forall i$	$\tau = t_2 - t_1$ $e^{-\tau\mathbf{L}}$	$\pi_i = \frac{k_i}{2M}, \forall i$	$\mathbf{\Pi} e^{-\tau\mathbf{L}} - \boldsymbol{\pi}^T\boldsymbol{\pi}$	trace $\left[\mathbf{H}^T\mathbf{S}(\tau)\mathbf{H}\right]$
3) **Modularity**: static undirected	$\pi_i = \frac{k_i}{2M}, \forall i$	discrete-time one step $\mathbf{D}^{-1}\mathbf{A}$	$\pi_i = \frac{k_i}{2M}, \forall i$	$\mathbf{\Pi}\mathbf{D}^{-1}\mathbf{A} - \boldsymbol{\pi}^T\boldsymbol{\pi}$	$\frac{1}{2M}\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2M}\right)\delta(c_i,c_j)$
4) **Dir. Modularity**: Static directed	$p_{1,i} = \frac{k_i^{\mathrm{o}}}{M}, \forall i$	discrete-time one step $\mathbf{D}_{\mathrm{o}}^{-1}\mathbf{A}$	$p_{2,i} = \frac{k_i^{\mathrm{i}}}{M}, \forall i$	$\frac{1}{M}\mathbf{D}_{\mathrm{o}}\mathbf{D}_{\mathrm{o}}^{-1}\mathbf{A} - \mathbf{p}_1^T\mathbf{p}_2$	$\frac{1}{M}\sum_{ij}\left(A_{ij} - \frac{k_i^{\mathrm{o}}k_j^{\mathrm{i}}}{M}\right)\delta(c_i,c_j)$
5) **Forward**: general non-stationary	$\mathbf{p_1}$	Eq. (7)	$\mathbf{p}_2 = \mathbf{p}_1\mathbf{T}(t_1,t_2)$	$\mathbf{P}_1\mathbf{T}(t_1,t_2)\mathbf{P}_2^{-1}\mathbf{T}(t_1,t_2)^T\mathbf{P}_1$ $-\mathbf{p}_1^T\mathbf{p}_1$	
6) **Backward**: general non-stationary	$\mathbf{p}_1 = \mathbf{p}_2\mathbf{T}_{\text{rev}}(t_2,t_1)$	Eq. (7) time-reversed $\mathbf{T}_{\text{rev}}(t_2,t_1)$	$\mathbf{p}_2$	$\mathbf{P}_2\mathbf{T}_{\text{rev}}(t_2,t_1)\mathbf{P}_1^{-1}\mathbf{T}_{\text{rev}}(t_2,t_1)^T\mathbf{P}_2$ $-\mathbf{p}_2^T\mathbf{p}_2$	
7) **Bib. coupling**: Static directed	$p_{1,i} = \frac{k_i^{\mathrm{o}}}{M}, \forall i$	discrete-time one step $\mathbf{D}_{\mathrm{o}}^{-1}\mathbf{A}$	$p_{2,i} = \frac{k_i^{\mathrm{i}}}{M}, \forall i$	$\frac{1}{M}\mathbf{A}\mathbf{D}_{\mathrm{i}}^{-1}\mathbf{A}^T - \mathbf{p}_1^T\mathbf{p}_1$	$\frac{1}{M}\sum_{ij}\left(\sum_k \frac{A_{ik}A_{kj}}{k_k^{\mathrm{i}}} - \frac{k_i^{\mathrm{o}}k_j^{\mathrm{o}}}{M}\right)\delta(c_i,c_j)$
8) **Co-citation**: Static directed	$p_{1,i} = \frac{k_i^{\mathrm{o}}}{M}, \forall i$	reversed one step $\mathbf{D}_{\mathrm{i}}^{-1}\mathbf{A}^T$	$p_{2,i} = \frac{k_i^{\mathrm{i}}}{M}, \forall i$	$\frac{1}{M}\mathbf{A}^T\mathbf{D}_{\mathrm{o}}^{-1}\mathbf{A} - \mathbf{p}_2^T\mathbf{p}_2$	$\frac{1}{M}\sum_{ij}\left(\sum_k \frac{A_{ki}A_{jk}}{k_k^{\mathrm{o}}} - \frac{k_i^{\mathrm{i}}k_j^{\mathrm{i}}}{M}\right)\delta(c_i,c_j)$

Table S4: Relation between non-stationary random walk covariance matrices and known partition quality functions.

complex network
scale-free network
small-world network
weighted network
directed network
growing network
evolving network

Table S5: Keywords used to find authors of articles about complex networks in the American Physical Society journals. The titles and abstracts of articles published between 2000 and 2010 were searched.

Community A			Community B			Community C		
A. L. Barabási	USA	PhysRevE	Chiu Fan Lee	United Kingdom	PhysRevE	Abolfazl Ramezanpour	Italy	PhysRevE
András Lukács	Hungary	PhysRevE	David D. Smith	USA	PhysRevA	Alain Barrat	France	PhysRevE
Balázs Rácz	Hungary	PhysRevE	Douglas J. Ashton	United Kingdom	PhysRevE	Alessandro Vespignani	Italy	PhysRevE
Brad R. Trees	USA	PhysRevB	J. P. Saramäki	Finland	PhysRevE	Alexei Vázquez	USA	PhysRevE
David G. Stroud	USA	PhysRevB	Jukka Pekka Onnela	United Kingdom	PhysRevE	Andrea Baronchelli	Italy	PhysRevE
E. Almaas	USA	PhysRevE	Jussi M. Kumpula	Finland	PhysRevE	Andrea Lancichinetti	Italy	PhysRevE
Erzsébet Ravasz	Romania	PhysRevE	János Kertész	Hungary	PhysRevE	Bruno Gonçalves	Brazil	PhysRevD
Gergely Palla	Hungary	PhysRevE	K. Tucci	Venezuela	PhysRevE	C. L. Zhang	USA	PhysRevB
I. Szakadát	Hungary	PhysRevE	Kimmo K. Kaski	Finland	PhysRevB	Claudio Castellano	Italy	PhysRevE
Illés J. Farkas	Hungary	PhysRevE	Konstantin Klemm	Germany	PhysRevE	Daniele Vilone	Germany	PhysRevE
Imre Derényi	Hungary	PhysRevE	L. Kullmann	Hungary	PhysRevE	Dmitri Krioukov	USA	PhysRevE
M. Argollo De Menezes	USA	Other	Mario G. Cosenza	Venezuela	PhysRevE	Eric D. Kolaczyk	USA	PhysRevE
R. V. Kulkarni	USA	PhysRevB	Mark Fricker	United Kingdom	PhysRevE	Fabien Viger	France	PhysRevE
Tamás Vicsek	Hungary	PhysRevE	Martín G. Zimmermann	Spain	PhysRevE	Filippo Radicchi	Italy	PhysRevE
V. Saranathan	USA	PhysRevE	Maxi San Miguel	Spain	PhysRevA	Francesca Colaiori	Italy	PhysRevE
Zoltán Dezsö	USA	PhysRevE	Mikko Kivelä	Finland	PhysRevE	Gary G. Yen	USA	PhysRevE
			Neil F. Johnson	United Kingdom	PhysRevB	José J. Ramasco	Spain	PhysRevE
			R. Toivonen	Finland	PhysRevE	Luca Dall'Asta	France	PhysRevE
			Renaud Lambiotte	Belgium	PhysRevE	Ma Ángeles Serrano	Spain	PhysRevE
			T. S. Evans	United Kingdom	PhysRevD	Marián Boguñá	Spain	PhysRevE
			Timothy C. Jarrett	United Kingdom	PhysRevE	Michel L. Goldstein	USA	PhysRevE
			Víctor M. Eguíluz	Spain	PhysRevE	Michele Catanzaro	Spain	PhysRevE
			X. Castelló	Spain	PhysRevE	Miguel A. Muñoz	Spain	PhysRevE
						Nicola Perra	Italy	PhysRevE
						Philippe Blanchard	Germany	PhysRevE
						Romualdo Pastor-Satorras	Spain	PhysRevE
						S. Mehdi Vaez Allaei	Iran	PhysRevE
						Santo Fortunato	Italy	PhysRevE
						Steven A. Morris	USA	PhysRevE
						Tyll Krüger	Germany	PhysRevE
						Vittoria Colizza	Italy	PhysRevE
						Vittorio Loreto	Italy	PhysRevE

Table S6: Authors in the three selected initial communities of the 2000s.

REFERENCES AND NOTES

1. Y. Bar-Yam, *Dynamics of Complex Systems* (CRC Press, 2019).

2. P. Holme, J. Saramäki, *Temporal Networks* (Understanding Complex Systems, Springer Berlin Heidelberg, 2013).

3. N. Masuda, M. A. Porter, R. Lambiotte, Random walks and diffusion on networks. *Phys. Rep.* **716-717**, 1–58 (2017).

4. P. Holme, J. Saramäki, *Temporal Network Theory* (Computational Social Sciences, Springer International Publishing, 2019).

5. M. A. Porter, Nonlinearity + networks: A 2020 vision, in *Emerging Frontiers in Nonlinear Science* , P. G. Kevrekidis, J. Cuevas-Maraver, A. Saxena, Eds. (Springer International Publishing, 2020), pp. 131–159.

6. P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, J.-P. Onnela, Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).

7. P. Holme, J. Saramäki, Temporal networks. *Phys. Rep.* **519**, 97–125 (2012).

8. J. Sun, C. Faloutsos, S. Papadimitriou, P. S. Yu, *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '07* (ACM Press, 2007), pp. 687.

9. M. Latapy, T. Viard, C. Magnien, Stream graphs and link streams for the modeling of interactions over time. *Soc. Netw. Anal. Min.* **8**, 61 (2018).

10. C. Stadtfeld, P. Block, Interactions, actors, and time: Dynamic network actor models for relational events. *Sociol. Sci.* **4**, 318–352 (2017).

11. C. T. Butts, 4. A relational event framework for social action. *Sociol. Methodol.* **38**, 155–200 (2008).

12. J. Moody, D. McFarland, S. Bender-deMoll, Dynamic network visualization. *Am. J. Sociol.* **110**, 1206–1241 (2005).

13. M. Takaffoli, F. Sangi, J. Fagnan, O. R. Zaiane, *Fifth international AAAI conference on weblogs and social media* (2011), pp. 626–629.

14. G. Rossetti, L. Pappalardo, D. Pedreschi, F. Giannotti, Tiles: An online algorithm for community discovery in dynamic social networks. *Mach. Learn.* **106**, 1213–1241 (2017).

15. F. Folino, C. Pizzuti, An evolutionary multiobjective approach for community discovery in dynamic networks. *IEEE Trans. Knowl. Data Eng.* **26**, 1838–1852 (2014).

16. T. Aynaud, J.-L. Guillaume, *Proceedings of the 5th SNA-KDD workshop* (2011), vol. 11.

17. T. Viard, M. Latapy, C. Magnien, Computing maximal cliques in link streams. *Theor. Comput. Sci.* **609**, 245–252 (2016).

18. P. Holme, F. Liljeros, Birth and death of links control disease spreading in empirical contact networks. *Sci. Rep.* **4**, 4999 (2014).

19. E. Valdano, L. Ferreri, C. Poletto, V. Colizza, Analytical computation of the epidemic threshold on temporal networks. *Phys. Rev. X* **5**, 021005 (2015).

20. T. P. Peixoto, M. Rosvall, Modelling sequences and temporal networks with dynamic community structures. *Nat. Commun.* **8**, 582 (2017).

21. U. Aslak, M. Rosvall, S. Lehmann, Constrained information flows in temporal networks reveal intermittent communities. *Phys. Rev. E* **97**, 062312 (2018).

22. G. Petri, P. Expert, Temporal stability of network partitions. *Phys. Rev. E* **90**, 022813 (2014).

23. G. Palla, A.-L. Barabási, T. Vicsek, Quantifying social group evolution. *Nature* **446**, 664–667 (2007).

24. C. Matias, V. Miele, Statistical clustering of temporal networks through a dynamic stochastic block model. *J. R. Stat. Soc. Series B Stat. Methodology* **79**, 1119–1141 (2017).

25. A. Ghasemian, P. Zhang, A. Clauset, C. Moore, L. Peel, Detectability thresholds and optimal algorithms for community structure in dynamic networks. *Phys. Rev. X* **6**, 031005 (2016).

26. S. Fortunato, Community detection in graphs. *Phys. Rep.* **486**, 75–174 (2010).

27. G. Rossetti, R. Cazabet, Community discovery in dynamic networks. *ACM Comput. Surv.* **51**, 1–37 (2018).

28. P. Holme, Modern temporal network theory: A colloquium. *Eur. Phys. J. B.* **88**, 234 (2015).

29. J. I. Liechti, S. Bonhoeffer, A time resolved clustering method revealing longterm structures and their short-term internal dynamics. arXiv:1912.04261 [stat.ML] (9 December 2019).

30. M. De Domenico, A. Lancichinetti, A. Arenas, M. Rosvall, Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Phys. Rev. X* **5**, 011027 (2015).

31. M. Rosvall, C. T. Bergstrom, Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1118–1123 (2008).

32. J. C. Delvenne, S. N. Yaliraki, M. Barahona, Stability of graph communities across time scales. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12755–12760 (2010).

33. T. Aynaud, J. L. Guillaume, Static community detection algorithms for evolving networks, in *Proceedings of the 8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)* (IEEE, 2010), pp. 513–519.

34. C. Guo, J. Wang, Z. Zhang, Evolutionary community structure discovery in dynamic weighted networks. *Physica A* **413**, 565–576 (2014).

35. R. Lambiotte, J.-C. Delvenne, M. Barahona, Random walks, Markov processes and the multiscale modular organization of complex networks. *IEEE Trans. Netw. Sci. Eng.* **1**, 76–90 (2014).

36. M. T. Schaub, J.-C. Delvenne, R. Lambiotte, M. Barahona, Multiscale dynamical embeddings of complex networks. *Phys. Rev. E* **99**, 062308 (2019).

37. I. Scholtes, N. Wider, R. Pfitzner, A. Garas, C. J. Tessone, F. Schweitzer, Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks. *Nat. Commun.* **5**, 5024 (2014).

38. M. M. Pérez-Nimo, J. A. Camúñez-ruiz, Matrix form of the Bayes theorem and diagnostic tests. *IOSR J. Math.* **14**, 1–6 (2018).

39. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech.*, P10008 (2008).

40. R. C. Lupton, J. M. Allwood, Hybrid Sankey diagrams: Visual analysis of multidimensional data for understanding resource use. *Resour. Conserv. Recycl.* **124**, 141–151 (2017).

41. S. Fortunato, M. Barthelemy, Resolution limit in community detection. *Proc. Natl. Acad. Sci.* **104**, 36–41 (2007).

42. M. T. Schaub, J.-C. Delvenne, S. N. Yaliraki, M. Barahona, Markov dynamics as a zooming lens for multiscale community detection: Non clique-like communities and the field-of-view limit. *PLOS ONE* **7**, e32210 (2012).

43. E. Seneta, *Non-negative Matrices and Markov Chains* (Springer Series in Statistics, Springer New York, 1981).

44. V. A. Traag, L. Waltman, N. J. van Eck, From Louvain to Leiden: Guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).

45. V. Traag, F. Zanini, R. Gibson, O. Ben-Kiki, D. van Kuppevelt, vtraag/leidenalg 0.8.2 (2020).

46. J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, P. Vanhems, High-resolution measurements of face-to-face contact patterns in a primary school. *PLOS ONE* **6**, e23176 (2011).

47. B. König, A. K. Lindholm, P. C. Lopes, A. Dobay, S. Steinert, F. J.-U. Buschmann, A system for automatic recording of social behavior in a free-living wild house mouse population. *Anim. Biotelemetry* **3**, 39 (2015).

48. B. Konig, A. K. Lindholm, The complex social environment of female house mice (*Mus domesticus*), in *Evolution of the House Mouse*, M. Macholán, S. J. E. Baird, P. Munclinger, J. Piálek, Eds. (Cambridge Univ. Press, 2012), pp. 114–130.

49. J. I. Liechti, B. Qian, B. König, S. Bonhoeffer, Contact patterns reveal a stable dynamic community structure with fission-fusion dynamics in wild house mice. bioRxiv 963512 [**Preprint**]. 26 February 2020. https://doi.org/10.1101/2020.02.24.963512.

50. M. Rosvall, C. T. Bergstrom, Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLOS ONE* **6**, e18209 (2011).

51. N. Tremblay, P. Borgnat, Graph wavelets for multiscale community mining. *IEEE Trans. Signal Process.* **62**, 5227–5239 (2014).

52. L. Gutiérrez-Gómez, A. Bovet, J.-C. Delvenne, Multi-scale anomaly detection on attributed networks, in *Proceedings of the 34th AAAI Conference on Artificial Intelligence* (AAAI Press, 2020), vol. 34, pp. 678–685.

53. E. W. Montroll, G. H. Weiss, Random Walks on Lattices. II. *J. Math. Phys.* **6**, 167–181 (1965).

54. R. Sinatra, D. Wang, P. Deville, C. Song, A.-L. Barabasi, Quantifying the evolution of individual scientific impact. *Science* **354**, eaaf5239(2016).

55. M. E. J. Newman, Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8577–8582 (2006).

56. M. E. J. Newman, Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* **74**, 036104 (2006).

57. M. E. J. Newman, M. Girvan, Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).

58. A. Arenas, J. Duch, A. Fernández, S. Gómez, Size reduction of complex networks preserving modularity. *New J. Phys.* **9**, 176–176 (2007).

59. K. Rohe, T. Qin, B. Yu, Co-clustering directed graphs to discover asymmetries and directional communities. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12679–12684 (2016).

60. Y. Kim, S.-W. Son, H. Jeong, Finding communities in directed networks. *Phys. Rev. E* **81**, 016103 (2010).

61. V. Satuluri, S. Parthasarathy, *Proceedings of the 14th International Conference on Extending Database Technology - EDBT/ICDT '11*, no. i (ACM Press, 2011), pp. 343.

62. R. Cazabet, S. Boudebza, G. Rossetti, Evaluating community detection algorithms for progressively evolving graphs. *J. Complex Netw.* **8**, cnaa027 (2021).