

# SUPPLEMENTARY MATERIALS

## Optimal Localization of Diffusion Sources in Complex Networks

Zhao-Long Hu, Xiao Han, Ying-Cheng Lai, Wen-Xu Wang

### I. Locating sources in continuous-time dynamical networks

A linear, time-invariant, and continuous-time dynamical network system can be described in the following state-space form

$$\begin{cases} \dot{\mathbf{x}}(t) = A\mathbf{x}(t) \\ \mathbf{y}(t) = C\mathbf{x}(t), \end{cases} \quad (\text{S1})$$

where  $\mathbf{x}(t) \in \mathbb{R}^N$  represents the complete state of the network system at time  $t$ ,  $N$  is the number of nodes,  $\mathbf{y}(t)$  is the vector of  $q$  outputs at time  $t$ ,  $A \in \mathbb{R}^{N \times N}$  is the system matrix, and  $C \in \mathbb{R}^{q \times N}$  is the output matrix. If the full initial state of the system,  $\mathbf{x}(t_0)$ , can be obtained from the outputs in the time interval  $[t_0, t]$ , the system is observable [S1].

To be concrete, we present a general method of reconstructing the initial states of an arbitrary linear time-invariant network using the diffusion model

$$\dot{x}_i(t) = \beta \sum_{j=1}^N [w_{ij}x_j(t) - w_{ji}x_i(t)], \quad (\text{S2})$$

where  $x_i(t)$  is the state of node  $i$  at the time  $t$ ,  $\beta$  is the diffusion coefficient (constant), and  $w_{ij}$  ( $w_{ji}$ ) is the weight of a directed link from node  $j$  to  $i$  ( $i$  to  $j$ ). Combining Eqs. (S2) and (S1), we have

$$\begin{cases} \dot{\mathbf{x}}(t) = \beta L\mathbf{x}(t) \\ \mathbf{y}(t) = C\mathbf{x}(t), \end{cases} \quad (\text{S3})$$

where  $L = (W - D)$ ,  $W \in \mathbb{R}^{N \times N}$  is the adjacency matrix of elements  $w_{ij}$ ,  $D \in \mathbb{R}^{N \times N}$  is the diagonal matrix with element  $d_i$  representing the total out-weight  $\sum_{j \in \Gamma_i} w_{ji}$  of node  $i$  ( $\Gamma_i$  is the set of neighbors of node  $i$ ). The output response of the system is

$$\mathbf{y}(t) = Ce^{\beta L(t-t_0)}\mathbf{x}(t_0). \quad (\text{S4})$$

For convenience, we can stack all the outputs  $\mathbf{y}(t)$  into a vector:  $\mathbf{Y} = [\mathbf{y}(t_0), \dots, \mathbf{y}(t_0+0.1), \dots, \mathbf{y}(t_0+$

$0.2), \dots, \mathbf{y}(t_0 + t)]^T$ . Intuitively,  $N$  snapshot measurements of the network state are needed to achieve a unique solution. Without loss of generality, we sample the same time interval  $T$  to obtain

$$\begin{pmatrix} \mathbf{y}(t_0) \\ \mathbf{y}(t_0 + T) \\ \vdots \\ \mathbf{y}(t_0 + (N-1)T) \end{pmatrix} = \begin{pmatrix} C \\ Ce^{\beta LT} \\ \vdots \\ Ce^{(N-1)\beta LT} \end{pmatrix} \mathbf{x}(t_0) = O \cdot \mathbf{x}(t_0), \quad (\text{S5})$$

where the matrix  $O \in \mathbb{R}^{qN \times N}$  is the so-called observability matrix in canonical control theory. A unique solution of Eq. (S5) exists and the state vector  $\mathbf{x}(t_0)$  at initial time is observable if and only if the rank condition  $\text{rank}(O) = N$  is satisfied [S2]. Our goal is to identify the minimum set of messenger nodes to satisfy the observability full rank condition.

To achieve our goal, we exploit the recently developed exact controllability theory [S1] and the duality between controllability and observability [S2], which enables us to find  $N_m$ , the minimum number of messengers in an efficient manner. In particular, for an arbitrary network,  $N_m$  is determined by the maximum geometric multiplicity  $\max_i \{\mu(\lambda_i^L)\}$  of the eigenvalues  $\lambda_i^L$  of matrix  $L$ , as

$$N_m = \max_i \{N - \text{rank}[\lambda_i^L I - L]\}, \quad (\text{S6})$$

which is exactly the same as that for discrete-time dynamical networks studied in the main text. Insofar as  $N_m$  is determined, the key to source localization is then to identify messengers to obtain the output matrix  $C$ . The method of identifying messengers is essentially the same as that for the discrete time case: by using the Popov-Belevitch-Hautus (PBH) test theory [S3], we obtain the output matrix  $C$  associated with  $N_m$  messenger nodes through  $\text{rank} \begin{pmatrix} \lambda^{\max} I - L \\ C \end{pmatrix} = N$ , where  $\lambda^{\max}$  is the eigenvalue corresponding to the maximum geometric multiplicity  $\mu(\lambda^{\max})$  of matrix  $L$ . To determine the output matrix  $C$ , we implement elementary row transformation on matrix  $\lambda^{\max} I - L$  to obtain the row canonical form of the matrix. The nodes whose numbers correspond to the linearly-dependent columns are the messenger nodes. Finally, combining with Eq. (S5), we can locate sources in continuous-time dynamic networks.

Therefore, our theoretical framework of source localization, including the minimum output theory for determining a minimum set of messenger nodes and identifying the messenger nodes is exactly the same for both discrete and continuous dynamical network systems.

## II. Proof of the minimum output theory

According to the exact controllability theory [S4] and the dual relation between network controllability and observability [S1], for system (2) in the main text,  $N_m$  is determined by the maximum geometric multiplicity of the eigenvalue  $\lambda_i$  of the matrix  $I + \beta L$ , i.e.,

$$N_m = \max_i \{N - \text{rank}[\lambda_i I - (I + \beta L)]\}. \quad (\text{S7})$$

We can prove that  $N_m$  is independent of  $\beta$  in the sense that  $\beta$  can be eliminated from Eq. (S7). For matrix  $I + \beta L$ , we have

$$\begin{aligned} (I + \beta L)\mathbf{v} &= \mathbf{v} + \beta L\mathbf{v}, \\ &= \mathbf{v} + \beta \lambda_i^L \mathbf{v}, \\ &= (1 + \beta \lambda_i^L)\mathbf{v}, \end{aligned} \quad (\text{S8})$$

where  $\lambda_i^L$  is the eigenvalue of matrix  $L$  and  $\mathbf{v}$  is the associated eigenvector. Equation (S8) gives the eigenvalue  $\lambda_i$  of matrix  $I + \beta L$ , i.e.,

$$\lambda_i = 1 + \beta \lambda_i^L. \quad (\text{S9})$$

Inserting the expression of  $\lambda_i$  into Eq. (S7), we have

$$\begin{aligned} N_m &= \max_i \{N - \text{rank}[(1 + \beta \lambda_i^L)I - (I + \beta L)]\}, \\ &= \max_i \{N - \text{rank}[\beta(\lambda_i^L I - L)]\}, \\ &= \max_i \{N - \text{rank}[\lambda_i^L I - L]\}, \end{aligned} \quad (\text{S10})$$

which indicates that the minimum output  $N_m$  is independent of the value of  $\beta$ . The exact minimum output theory for arbitrary network [Eq. (5) in the main text] is proved.

For system (2) in the main text with an arbitrary undirected network, according to the exact controllability theory [S4] and the dual relation between network controllability and observability,  $N_m$  is determined by the maximum eigenvalue degeneracy of matrix  $I + \beta L$ , i.e.,

$$N_m = \max_i \{\delta(\lambda_i)\}, \quad (\text{S11})$$

where  $\delta(\lambda_i)$  is the eigenvalue degeneracy of matrix  $I + \beta L$ . Equation (S9) demonstrates that there is a one-to-one correspondence between the eigenvalue  $\lambda_i$  of matrix  $I + \beta L$  and the eigenvalue  $\lambda_i^L$  of matrix  $L$ . Thus,  $I + \beta L$  and  $L$  have exactly the same eigenvalue degeneracy, which yields the exact minimum output theory [Eq. (6) in the main text] for undirected networks, i.e.,

$$N_m^{\text{undirect}} = \max_i \{\delta(\lambda_i^L)\}, \quad (\text{S12})$$

and the diffusion parameter  $\beta$  is eliminated.

Furthermore, Eq. (S10) indicates that the geometric multiplicity of the eigenvalues of  $L$  is equal to that of  $I + \beta L$ , and the output matrix  $C$  of  $L$  is identical to that of  $I + \beta L$  as well. Utilizing the Popov-Belevitch-Hautus (PBH) test theory [S3], we can get the output matrix  $C$  from

$$\text{rank} \begin{pmatrix} \lambda^{\max} I - L \\ C \end{pmatrix} = N, \quad (\text{S13})$$

instead of

$$\text{rank} \begin{pmatrix} \hat{\lambda}^{\max} I - (I + \beta L) \\ C \end{pmatrix} = N, \quad (\text{S14})$$

where  $\hat{\lambda}^{\max}$  is the maximum geometric multiplicity of the eigenvalues of  $I + \beta L$ . This implies that the diffusion parameter  $\beta$  has no influence on identifying messenger nodes based on PBH test as well, as described in the subsection *Identification of messenger node set*.

### III. Analytical treatment of locatability of ER and SF networks

In general, for an undirected network (symmetric matrix), eigenvalue degeneracy is exactly the same as its geometric multiplicity. Thus, according to the ET formulas, e.g., Eqs. (5) and (6) in the main text, the eigenvalue  $\lambda^{\max}$  with the maximum geometric multiplicity in Eq. (5) is nothing but the eigenvalue with the maximum degeneracy (the number of appearances in the eigenvalue spectrum). In this regard, if we are able to evaluate the eigenvalue with the maximum degeneracy a priori, the calculation of the all eigenvalues in Eq. (6) and that of matrix ranks for all possible eigenvalues in Eq. (5) can be saved, leading to a fast estimation of  $n_m$  by inserting the estimated eigenvalue into Eq. (5).

For a sparse undirected network (symmetric matrix), the diagonal elements of the network matrix often dominate eigenvalue spectrum [S5]. Hence, the diagonal elements with the maximum multiplicity (the largest number of appearances in the diagonal) could be a proxy of the eigenvalue with maximum degeneracy. However, for Laplacian matrix  $L$ , in addition to the diagonal elements, zero could dominate eigenvalue spectrum as well in the absence of any zero diagonal elements, because of the existence of isolated components. Note that each isolated component or node will contribute one null eigenvalue to the eigenvalue spectrum of a Laplacian matrix. Thus, diagonal elements and zero are possible candidates for formulating a fast estimation of the source locatability measure  $n_m$ :

$$n_m^{\text{sparse}} \approx 1 - \frac{\text{rank}(aI - L)}{N}, \quad (\text{S15})$$

where  $a$  is either zero or the diagonal element of  $L$  with the maximum multiplicity.

In an undirected ER network with small connection probability, there are a number of isolated nodes and isolated components, accounting for the dominance of zero in the eigenvalue spectrum. Thus, the source locatability  $n_m$  can be estimated by examining the isolated nodes without links and the nodal pairs. Using the degree distribution  $P(k)$  of ER networks,  $e^{-\langle k \rangle} \langle k \rangle^k / k!$ , we have

$$n_m^{\text{UER}} \approx \max\{1/N, P(k=0) + P(k=1)^2\}, \quad (\text{S16})$$

which gives

$$n_m^{\text{UER}} = \delta(0) \approx e^{-\langle k \rangle} + \langle k \rangle^2 e^{-2\langle k \rangle}. \quad (\text{S17})$$

For a directed network, let  $k_{\text{out}}$  and  $k_{\text{in}}$  be the out-degree and in-degree, respectively, and suppose that the links are unidirectional. The average degree of the network is  $\langle k \rangle = \langle k_{\text{out}} \rangle / 2 = \langle k_{\text{in}} \rangle / 2$ . A fast

estimation of the source locatability yields

$$n_m^{\text{sparse}} \approx 1 - \frac{\text{rank}(aI - L)}{N}, \quad (\text{S18})$$

where  $a$  is 0, -1 or -2, due to the fact that the diagonal element of matrix  $L$  is dominated by 0, -1 or -2 for small average degree  $\langle k \rangle$ . Numerical calculations suggest that the main contributions to  $n_m$  come from eigenvalues 0, -1 and -2.

For a directed ER network with a small connection probability  $2\langle k \rangle/N$ , analogous to the undirected case, we only need to consider isolated nodes and nodal pairs to obtain

$$n_m^{\text{DER}} \simeq \max\{1/N, P(k_{\text{out}} = 0, k_{\text{in}} = 0) + P(k_{\text{out}} = 1, k_{\text{in}} = 0)P(k_{\text{out}} = 0, k_{\text{in}} = 1)\}. \quad (\text{S19})$$

Since  $k_{\text{in}}$  is independent of  $k_{\text{out}}$ , we have

$$n_m^{\text{DER}} \approx e^{-\langle k \rangle} + \frac{\langle k \rangle^2 e^{-2\langle k \rangle}}{4}. \quad (\text{S20})$$

For a directed SF network, nodes of zero out-degree must be the messengers. In this case, we can estimate  $n_m$  as

$$n_m^{\text{DSF}} \simeq \max\{1/N, \sum_m^{N-1} P(k_{\text{out}} = 0|k)P(k)\}, \quad (\text{S21})$$

where  $P(k) = P(k_{\text{in}} + k_{\text{out}})$  follows a power law, and  $P(k_{\text{out}}|k)$  is the conditional probability that one node has out-degree  $k_{\text{out}}$  when its degree is  $k$ . According to binomial theorem, we have

$$P(k_{\text{out}}|k) = \binom{k}{k_{\text{out}}} \left(\frac{1}{2}\right)^{k_{\text{out}}} \left(\frac{1}{2}\right)^{k-k_{\text{out}}}, \quad (\text{S22})$$

which yields

$$P(k_{\text{out}} = 0|k) = 2^{-k} \quad (\text{S23})$$

and consequently,

$$n_m^{\text{DSF}} \approx \sum_{k=m}^{N-1} 2^{-k} P(k). \quad (\text{S24})$$

## IV. Cavity method for estimating locatability

In Ref. [S6], the cavity method was used to quantify the network controllability via the density of the driver nodes for directed networks. The method was subsequently extended to undirected networks [S4]. Because of the duality between controllability and observability, we can use the cavity method to estimate the locatability  $n_m$  for general complex networks. In particular, for a directed network with similar in- and out-degree distribution  $P(k)$ , the locatability  $n_m$  is given by

$$n_m = G(w_2) + G(1 - w_1) - 1 + \langle k \rangle w_1 (1 - w_2), \quad (\text{S25})$$

where  $G(x)$  is the generating function satisfying

$$G(x) = \sum_{k=0}^{\infty} P(k)x^k. \quad (\text{S26})$$

The quantities  $w_1$  and  $w_2$  in Eq. (S25) can be obtained through the following self-consistent equations:

$$w_1 = H[1 - H(1 - w_1)], \quad (\text{S27})$$

$$w_2 = 1 - H[1 - H(w_2)], \quad (\text{S28})$$

where  $H(x)$  is a generating function defined as

$$H(x) = \sum_{k=0}^{\infty} Q(k+1)x^k, \quad (\text{S29})$$

and  $Q(k) = kP(k)/\langle k \rangle$ .

In Ref. [S5], the cavity method was used to calculate the controllability measure for directed networks with multiple types of self loops, where the diagonal elements of matrix  $L$  were regarded as self loops. This is key to calculating the locatability measure  $n_m$  for matrix  $L - aI$ . Adopting the method in Ref. [S5], we have that, if a diagonal element of  $L - aI$  is nonzero, the in- and out-degree of the corresponding node is increased by 1, while the degrees of the other nodes remain the same in the original weighted network. For an undirected network, if a diagonal element of  $L - aI$  is nonzero, the degree of the node is increased by 2. Finally, based on the new node degrees, we can obtain  $n_m$  by combining Eqs. (S25-S29).

## V. Characteristics of real networks studied

In the main text, results from a number of real networks are presented to test our ET and FE methods. The details of the real networks are listed in table S1, which include the names of the data sets, the data type, the number of nodes  $N$ , the number of links  $E$ , and brief descriptions of the networks.

## VI. Performance assessment of source localization

The area under a receiver operating characteristic (AUROC) for the source localization is defined in terms of true positive rate (TPR) and false positive rate (FPR). TPR and FPR are defined as follows:

$$\text{TPR}(s) = \frac{\text{TP}(s)}{P} \quad \text{and} \quad \text{FPR}(s) = \frac{\text{FP}(s)}{Q}, \quad (\text{S30})$$

where  $s$  is the cutoff (threshold) in the list of reconstructed state  $x_i(t)$  at time  $t$ ,  $\text{TP}(s)$  ( $\text{FP}(s)$ ) is the number of true (false) positives in the top  $s$  reconstructed values of  $x_i(t)$  and  $P$  ( $Q$ ) is the number of positives (negatives) in the gold standard. AUROC is the area under the TPR-FPR curve.



## VII. More examples of source localization

### A. An example of locating sources in undirected weighted ER network without noise

To be concrete, we set  $\text{Data} = 0.5$  and assume that the initial triggering time  $t_0$  is unknown. In Fig. S1A, we show that measurements from any single node are sufficient to locate the sources, as the network has a single connected component with random link weights. Fig. S1B shows that our method can accurately infer both the number of sources and their locations, as well as the initial triggering time  $t_0$ . Fig. S1C shows, for different initial observation time  $t_{\text{ini}}$ , that the number of sources and their locations can be determined in a large range after  $t_0$  is detected as in Fig. S1B.

The performance of our method can be assessed, as follows. For concreteness, we assume  $t_{\text{ini}} = t_0 + 10$ . Fig. S1D shows that the value of AUROC will reach unity at  $t_{\text{ini}} - 10$ , which is the triggering time  $t_0$ , and the AUROC value decreases more significantly when the inferred time is  $t < t_0$  as compared with the case of  $t > t_0$ . From Fig. S1E, we see that AUROC reaches unity when the observation time is  $t_{\text{ini}} \approx 3$  time steps before  $t_0$ , and AUROC is almost unchanged as  $t_{\text{ini}}$  is increased. These results are similar to those of SF networks (Fig. 3 in the main text).

### B. Examples of locating sources in undirected weighted SF and ER networks in presence of noise

We extend our source localization framework to cases where there is noise for SF and ER networks. Assuming  $\sigma = 0.5$  and setting  $\text{Data} = 0.5$ , we reconstruct  $\mathbf{x}(t_0)$  with four random sources. Fig. S2 and Fig. S3 show essentially the same results as Fig. S1, indicating the robustness of our localization framework.

## VIII. Robustness of our source localization framework

We systematically investigate source localization for undirected and weighted ER and SF networks in terms of data requirement and noise resistance. Supplementary Fig. S4A and Fig. S4B show, for  $\sigma = 0$ , relatively small data amount, and both ER and SF networks, that the value of AUROC increases when the number of sources is decreased. This means that the localization accuracy tends to increase with the sparsity of vector  $\mathbf{x}(t_0)$ . We also see that the value of AUROC exceeds 0.9 even though data amount is only 0.3, and unity value of AUROC can be achieved for  $\text{Data} \geq 0.5$ . For  $\sigma = 0.5$ , as shown in Supplementary Fig. S4C and Fig. S4D, similar results are obtained, except that the corresponding AUROC values are slightly smaller. As can be seen from Supplementary Fig. S4E and Fig. S4F, for  $\text{Data} = 0.5$ , the AUROC values are nearly indistinguishable for different values of  $N_s$ . For  $\sigma \leq 0.3$ , the value of AUROC reaches unity but will decrease with  $\sigma$  for  $\sigma > 0.3$ , regardless of the values of  $N_s$ . The AUROC reaches unity with error less than 5% despite that  $\sigma$  is as large as 0.5. All the results provide additional evidence for the robustness of our method against noise and insufficient data.

## IX. Effects of diffusion parameter

We investigate the effects of the diffusion parameter  $\beta$  on the accuracy of source localization for different data amounts and values of the noise variance. In principle, to ensure that system (1) in the main text can characterize a diffusion process, the parameter  $\beta$  should be less than certain critical value. Specifically, we rewrite system (1) in the main text as

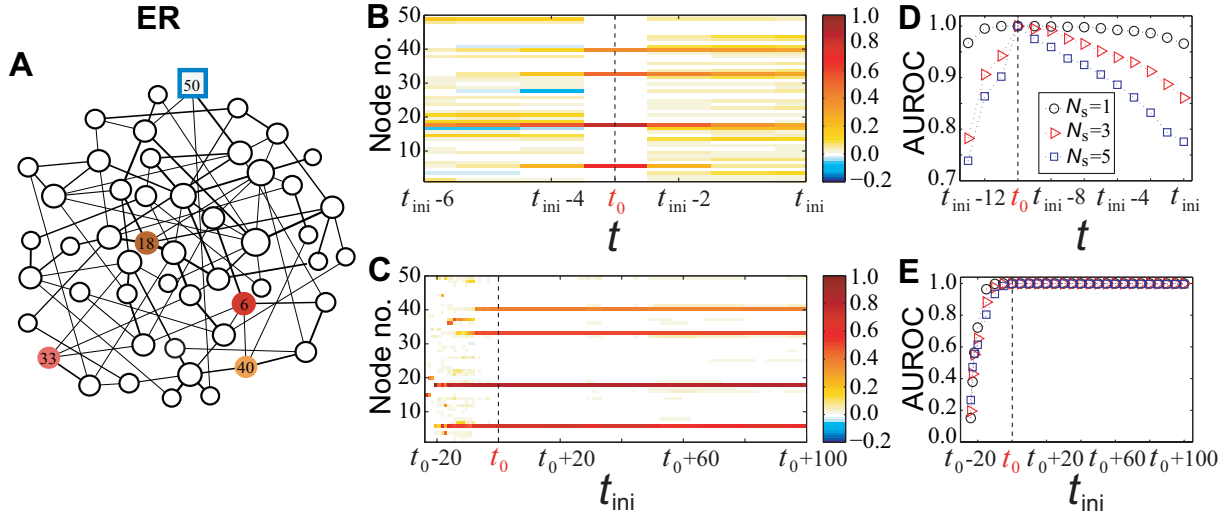
$$\begin{aligned} x_i(t+1) &= x_i(t) - \beta \sum_{j=1, j \neq i}^N w_{ji} x_i(t) + \beta \sum_{j=1, j \neq i}^N w_{ij} x_j(t), \\ &= \left( 1 - \beta \sum_{j=1, j \neq i}^N w_{ji} \right) x_i(t) + \beta \sum_{j=1, j \neq i}^N w_{ij} x_j(t). \end{aligned} \quad (\text{S31})$$

Note that the coefficient of the first term on the right hand side of Eq.(S31) should be positive, for otherwise, the state  $x_i(t+1)$  may become negative if  $x_i(t)$  is large and positive and the value of the second term of Eq.(S31) is small. An example is shown in Fig. S5, where  $\beta$  is 0.15 and the coefficient  $1 - \beta \sum_{j=1, j \neq i}^N w_{ji}$  associated with some nodes is negative. As a result, the state of a node, say  $x_7(t)$ , exhibits a negative value, causing the system to diverge. However, this scenario is not physically meaningful for describing a diffusion process. Thus, the following constraint on the coefficient  $1 - \beta \sum_{j=1, j \neq i}^N w_{ji} > 0$  should be imposed:

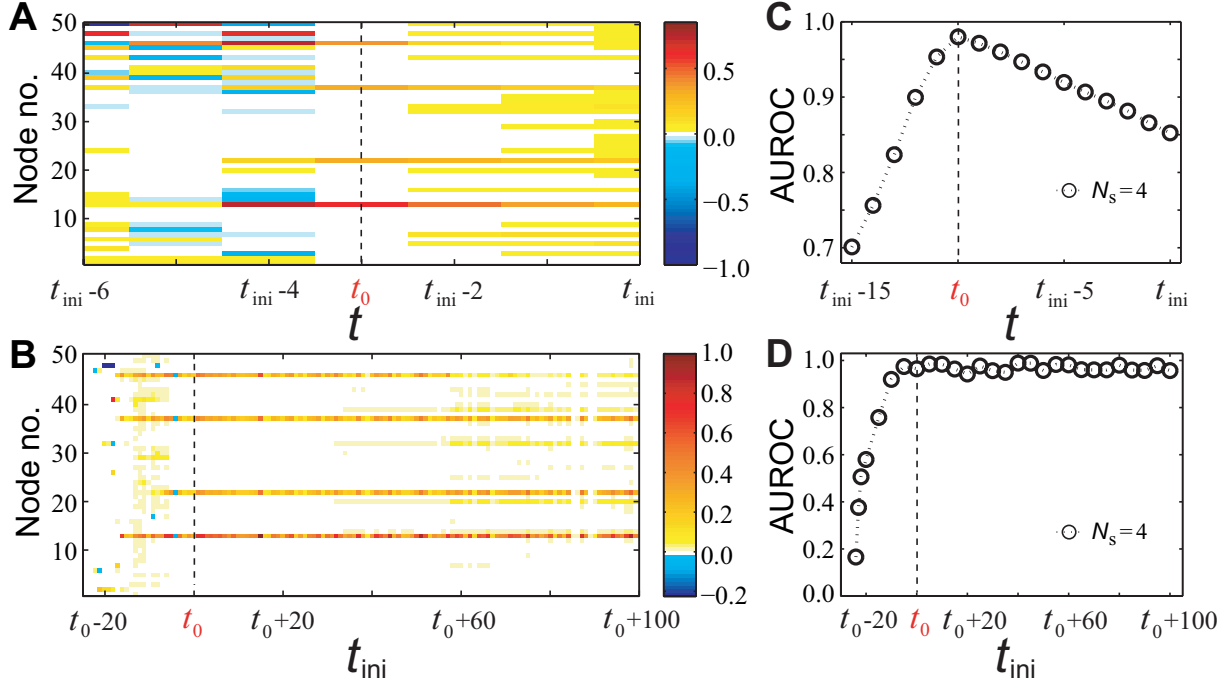
$$\beta < \frac{1}{\sum_{j=1, j \neq i}^N w_{ji}}. \quad (\text{S32})$$

We find that, under the constraint, the choice of different values of  $\beta$  has little influence on the accuracy of source localization with respect to different data amounts and values of noise variance, as shown in Figs. S6 and S7. We see that, as  $\beta$  is decreased, e.g.,  $\beta = 0.01$ , the localization accuracy is slightly reduced (Fig. S6), due to the computational errors associated with the iterative process in the implementation of the compressive sensing algorithm.

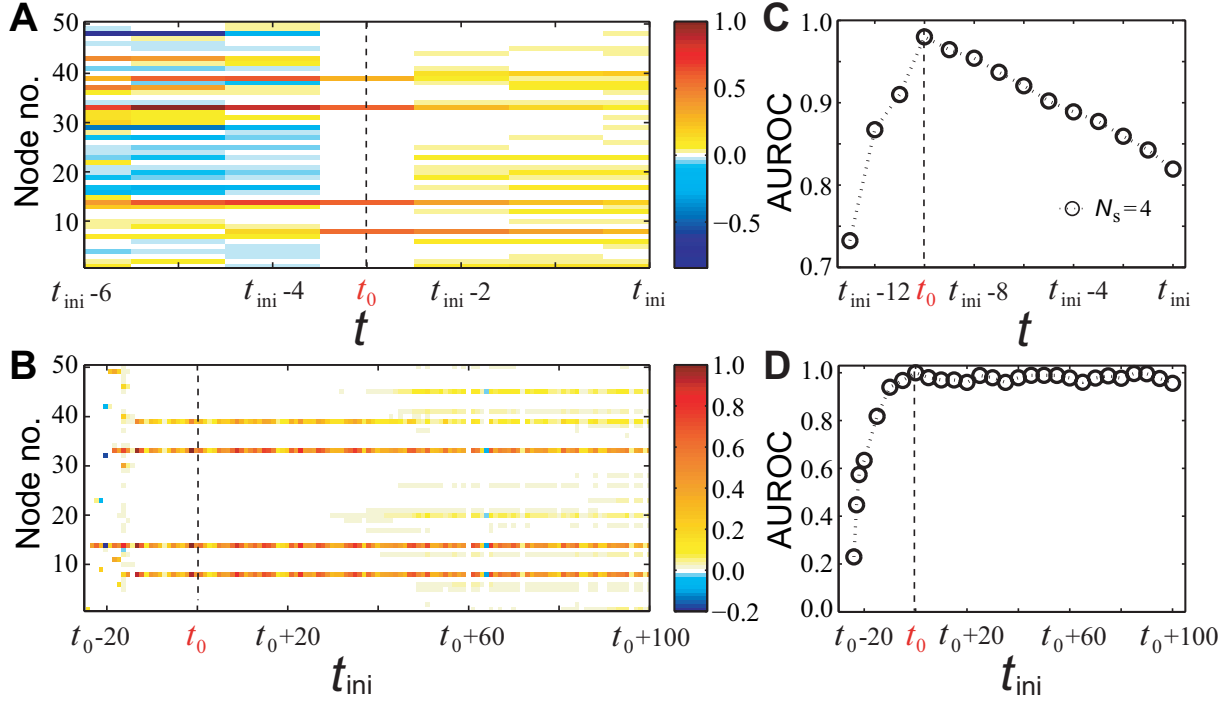
## Supplementary Figures



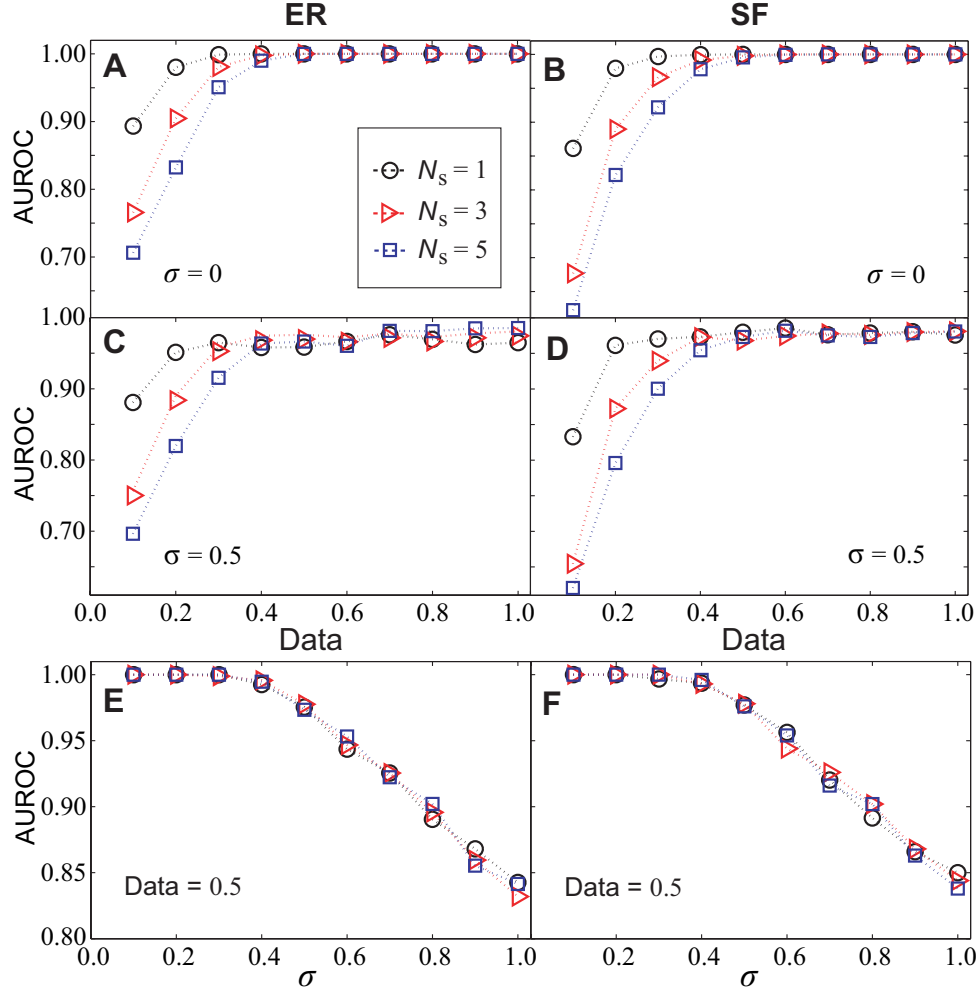
**Fig. S1. An example of locating sources in undirected weighted ER networks.** (A) Illustration of a ER network with four sources with colors representing the initial state values. One messenger node is specified as a blue square. The thickness of links represents their weight and the size of the nodes corresponds to their degrees. (B) Reconstructed state  $x_i(t)$  of each node for  $t \leq t_{ini}$ , where the initial observation time is  $t_{ini}$  ( $t_{ini} \geq t_0$ ). Colors represent the values of  $x_i(t)$  with  $t \leq t_{ini}$ . (C) Reconstructed initial state  $x_i(t_0)$  of each node from different initial observation time  $t_{ini}$  when  $t_0$ , the true triggering time, is being successfully inferred. Colors represent the reconstructed values of  $x_i(t_0)$ . The colors have the same meanings as those in (A). The four sources are randomly selected and their  $x_i(t_0)$  values are larger than zero. (D) AUROC as a function of  $t$  ( $t \leq t_{ini}$ ) for a fixed initial observation time  $t_{ini}$ . (E) AUROC versus  $t$  for different initial observation time  $t_{ini}$  and different number of sources ( $N_s$ ). For both (D) and (E), there is no noise, cases for different number of sources are illustrated,  $N_s$ , and  $t_0$  is the true triggering time. In all cases, the network size is  $N = 50$ , the average degree is  $\langle k \rangle = 4$ , the link weights are uniformly distributed in  $(0, 2)$ , the diffusion parameter  $\beta = 0.1$ , and  $Data = 0.5$ . The results in (D) and (E) are obtained by averaging over 30 independent simulations. The other parameters are the same as in Fig. 5 in the main text.



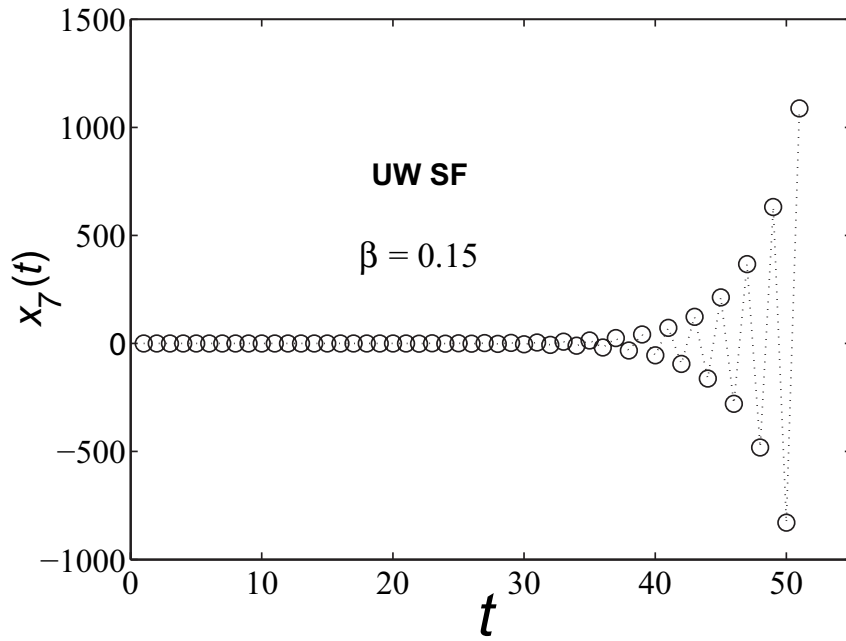
**Fig. S2. An example of locating sources in undirected weighted SF network with noise.** (A) Reconstructed state  $x_i(t)$  of each node at time step  $t$ , for  $t \leq t_{ini}$  and initial observation time  $t_{ini}$  ( $t_{ini} \geq t_0$ ). (B) Reconstructed state  $x_i(t)$  of each node for different initial observation time  $t_{ini}$  when  $t_0$  is known. (C-D) AUROC as a function of (C) time  $t$  ( $t \leq t_{ini}$ ) when the initial observation time is  $t_{ini}$  and as a function of (D) initial observation time  $t_{ini}$ . Network size is  $N = 50$ , the average degree is  $\langle k \rangle = 4$ , and the link weights are uniformly distributed in  $(0, 2)$ . Four sources are randomly selected and their initial states  $x_i(t_0)$  assume positive values. We set  $\beta = 0.05$ ,  $\sigma = 0.5$ , and  $Data = 0.5$ . The results in (C) and (D) are obtained by averaging over 30 independent simulations. The other parameters are the same as in Fig. 5 in the main text.



**Fig. S3. An example of locating sources in undirected weighted ER network with noise.** (A) Reconstructed state  $x_i(t)$  of each node at time step  $t$  for  $t \leq t_{ini}$  and for initial observation time  $t_{ini}$  ( $t_{ini} \geq t_0$ ). (B) Reconstructed state  $x_i(t)$  of each node for different initial observation time  $t_{ini}$  and for known  $t_0$ . (C-D) AUROC as a function of (C) time  $t$  ( $t \leq t_{ini}$ ) when the initial observation time is  $t_{ini}$  and as a function of (D) initial observation time  $t_{ini}$ . Network size is  $N = 50$ , the average degree is  $\langle k \rangle = 4$ , and the link weights are uniformly distributed in  $(0, 2)$ . Four sources are randomly selected and their initial state values  $x_i(t_0)$  are positive. We set  $\sigma = 0.5$ ,  $\beta = 0.1$ , and  $Data = 0.5$ . The results in (C) and (D) are obtained by averaging over 30 independent simulations. The other parameters are the same as in Fig. 5 in the main text.

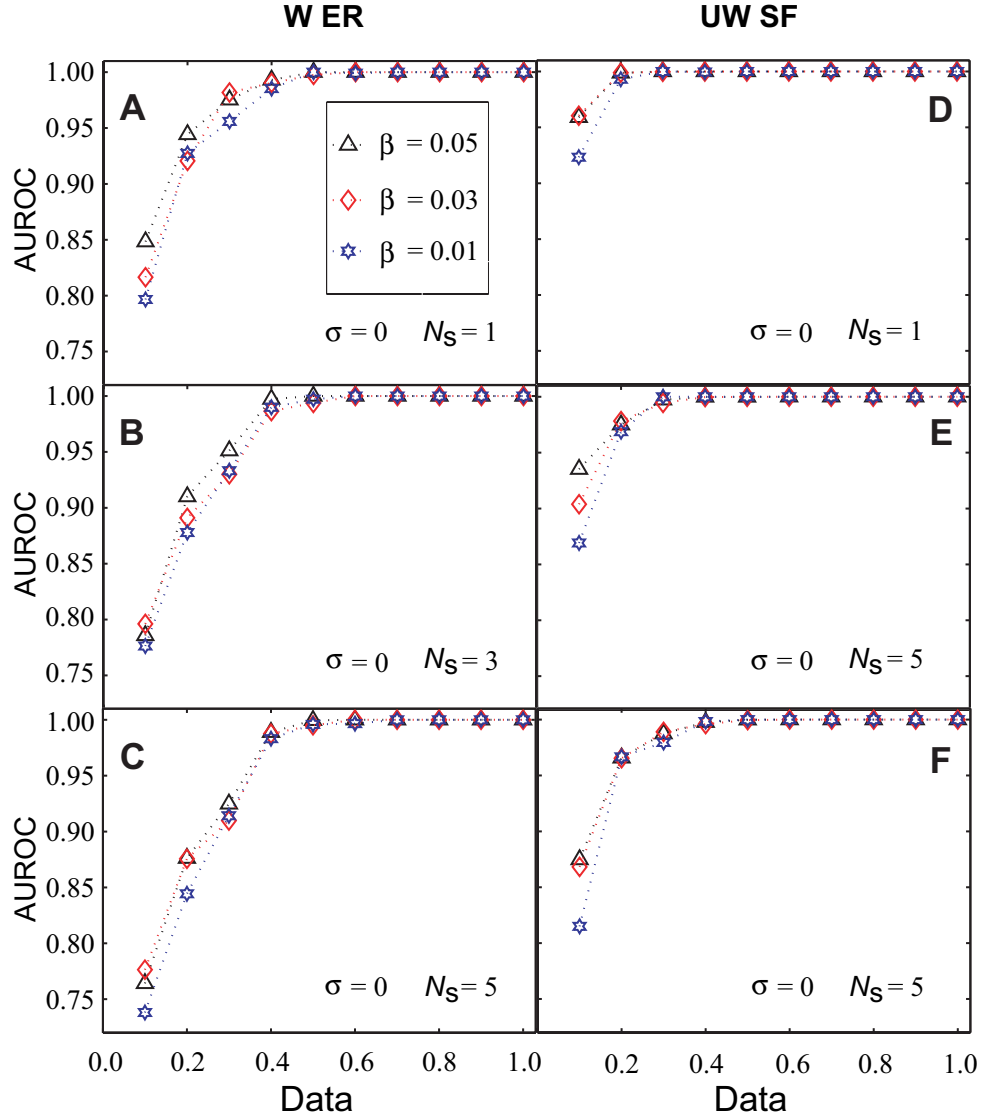


**Fig. S4. Locating sources in weighted ER and SF networks.** (A-B) In absence of noise, AUROC as a function of Data for different source number  $N_s$  for ER and SF networks, respectively. (C-D) The corresponding plots but with noise of amplitude  $\sigma = 0.5$ . (E-F) For Data= 0.5, AUROC as a function of  $\sigma$  for different values of  $N_s$  for ER and SF networks, respectively. The observational noise is modeled as  $\mathbf{y}(t)[1 + \mathcal{N}(0, \sigma^2)]$ , where  $\mathcal{N}(0, \sigma^2)$  is the Gaussian distribution. The baseline of AUROC is 0.5 (corresponding to random identification). The average degree is  $\langle k \rangle = 4$  for both ER and SF networks and the link weights are uniformly distributed in  $(0, 2)$ . The network size is  $N = 50$ . We set  $\beta = 0.1$  for ER networks and  $\beta = 0.05$  for SF networks. The results are obtained by averaging over 1000 independent simulations. The other parameters are the same as in Fig. 5 in the main text.

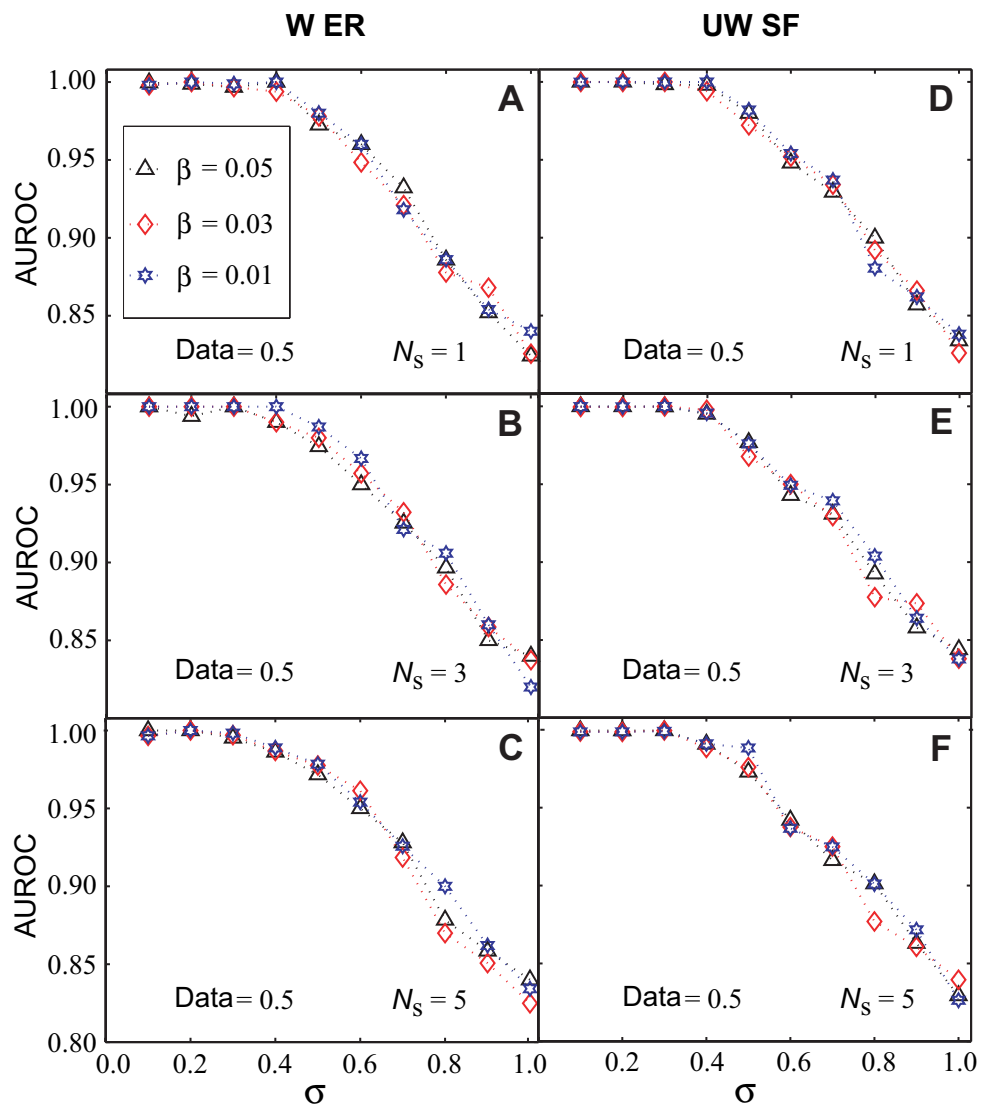


**Fig. S5. Illustration of nodal state if the constraint on  $\beta$  is violated.** The state  $x_7(t)$  of node No. 7 as a function of time step  $t$  for  $\beta = 0.15$  that violates the constraint on  $\beta$  (Eq. (S32)).  $x_7(t)$  presents negative value and tends to diverge as  $t$  increases. The degree  $k_i$  of node No. 7 is 10, the average degree  $\langle k \rangle$  of the unweighted SF network is 4 and the network size  $N$  is 50. The number of sources  $N_s$  is 3.





**Fig. S6. Effect of  $\beta$  on source localization in networks from different amounts of data.** (A-F) AUROC as a function of Data for different number  $N_S$  of sources for (A-C) weighted ER networks and (D-F) for unweighted SF networks, respectively. For ER networks,  $\langle k \rangle = 2$  and for SF networks  $\langle k \rangle = 4$ . For weighted networks, the link weights are randomly selected from an uniform distribution in the range  $(0, 2)$ . The network size  $N$  is 50 and noise variance  $\sigma = 0$ . The results are obtained by averaging over 100 independent simulations. The other parameters are the same as in Fig. 5 in the main text.



**Fig. S7. Effect of  $\beta$  on source localization in networks from noisy data.** (A-F) AUROC as a function of noise variance  $\sigma$  for (A-C) weighted ER and (D-F) unweighted SF networks, respectively. The white Gaussian noise is in the form  $\mathbf{y}(t)[1 + \mathcal{N}(0, \sigma^2)]$ , where  $\mathcal{N}(0, \sigma^2)$  is the Gaussian distribution. Data= 0.5 and the results are obtained by averaging over 300 independent simulations. The other parameters are the same as in fig. S6.

## Supplementary Table

**Table S1. Summary of the real networks used in Fig. 3 in the main text.** The quantities  $N$  and  $E$  denote the network size and the number of links, respectively. UD (D) in the Type column indicates undirected (directed) networks. The structural data of all the networks are available online. The data of Erdős971 can be downloaded via <http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm>, and the data of USAir can be downloaded via <http://vlado.fmf.uni-lj.si/pub/networks/data/map/USAir97.net>.

Data sets Name	Type	$N$	$E$	Description
ca-GrQc [S7]	UD	5242	14496	collaboration network of Arxiv General Relativity category
ca-HepTh [S7]	UD	9877	25598	collaboration network of Arxiv High Energy Physics Theory category
Erdős971	UD	433	1314	all of Paul Erdős's coauthors and their respective coauthors
dolphin [S8]	UD	62	159	associations between dolphins in a community living off Doubtful Sound
football [S9]	UD	115	613	American football games between Division IA colleges during regular season Fall 2000
Jazz [S10]	UD	198	2742	links of the network of Jazz musicians
Zachary's karate club [S11]	UD	34	78	social network of friendships of a karate club at a US university in the 1970s
Political blogs [S12]	D	1224	19025	hyperlinks between weblogs on US politics in 2005 by Adamic and Glance
Wiki-Vote [S13, 14]	D	7115	103689	all the Wikipedia voting data from the inception of Wikipedia till January 2008
E-mail [S15]	UD	1133	5451	interchanges between members of the Univeristy Rovira i Virgili (Tarragona)
p2p-Gnutella [S16]	D	6301	20777	Gnutella peer-to-peer network on August 8 2002
PGP [S17]	UD	10680	24236	links of the giant component of the network of users of the Pretty-Good-Privacy algorithm for secure information interchange
USAir	UD	332	2126	US Air flights, 1997

## References

- [S1] R. Kalman, *On the general theory of control systems*, IRE Trans. Automat. Contr. **4**, 110–110 (1959).
- [S2] R. E. Kalman, *Mathematical description of linear dynamical systems*, J. Soc. Indus. Appl. Math. Ser. A **1**, 152–192 (1963).
- [S3] M. Hautus, *Controllability and observability conditions of linear autonomous systems*, Ned. Akad. Wetenschappen, Proc. Ser. A **72**, 443 (1969).
- [S4] Z. Yuan, C. Zhao, Z. Di, W.-X. Wang, and Y.-C. Lai, *Exact controllability of complex networks*, Nat. Commun. **4**, 1 (2013).
- [S5] C. Zhao, W.-X. Wang, Y.-Y. Liu, and J.-J. Slotine, *Intrinsic dynamics induce global symmetry in network controllability*, Sci. Rep. **5**, 1 (2015).
- [S6] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, *Controllability of complex networks*, Nature **473**, 167–173 (2011).
- [S7] J. Leskovec, J. Kleinberg, and C. Faloutsos, *Graph evolution: Densification and shrinking diameters*, ACM TKDD **1,2** (2007).
- [S8] D. Lusseau, *et al.*, *The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations*, Behav. Ecol. Sociobiol. **54**, 396–405 (2003).
- [S9] M. Girvan, and M. E. Newman, *Community structure in social and biological networks*, Proc. Natl Acad. Sci. U.S.A. **99**, 7821–7826 (2002).
- [S10] P. M. Gleiser, and L. Danon, *Community structure in jazz*, Adv. Complex Syst. **6**, 565–573 (2003).
- [S11] W. W. Zachary, *An information flow model for conflict and fission in small groups*, J. Anthropol. Res. **33**, 452–473 (1977).
- [S12] L. A. Adamic, and N. Glance, *The political blogosphere and the 2004 us election: divided they blog*, In Proceedings of the 3rd international workshop on Link discovery 36–43. ACM, 2005.
- [S13] J. Leskovec, D. Huttenlocher, and J. Kleinberg, *Signed networks in social media*, In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 1361–1370. ACM, 2010.

- [S14] J. Leskovec, D. Huttenlocher, and J. Kleinberg, *Predicting positive and negative links in online social networks*, In Proceedings of the 19th international conference on World wide web 641–650. ACM, 2010.
- [S15] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas, *Self-similar community structure in a network of human interactions*, Phys. Rev. E **68**, 065103 (2003).
- [S16] M. Ripeanu, I. Foster, and A. Iamnitchi, *Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system design*, arXiv preprint cs/0209028, 2002.
- [S17] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, *Models of social networks based on social distance attachment*, Phys. Rev. E **70**, 056122 (2004).