

Uncovering hidden flows in physical networks

Chengwei Wang

Institute for Complex Systems and Mathematical Biology
University of Aberdeen, King's College, AB24 3UE Aberdeen, United Kingdom

Celso Grebogi

Institute for Complex Systems and Mathematical Biology
University of Aberdeen, King's College, AB24 3UE Aberdeen, United Kingdom

Murilo S. Baptista

Institute for Complex Systems and Mathematical Biology
University of Aberdeen, King's College, AB24 3UE Aberdeen, United Kingdom

September 11, 2017

Abstract

Understanding the interactions among nodes in a complex network is of great importance, since they disclose how these nodes are cooperatively supporting the functioning of the network. Scientists have developed numerous methods to uncover the underlying adjacent physical connectivity based on measurements of functional quantities of the nodes states. Often, the physical connectivity, the adjacency matrix, is available. Yet, little is known about how this adjacent connectivity impacts on the “hidden” flows being exchanged between any two arbitrary nodes, after travelling longer non-adjacent paths. In this Letter, we show that hidden physical flows in conservative flow networks, a quantity that is usually inaccessible to measurements, can be determined by the interchange of physical flows between any pair of adjacent nodes. Our approach applies to steady or dynamic state of either linear or non-linear complex networks that can be modelled by conservative flow networks, such as gas supply networks, water supply networks and power grids.

1 Introduction

Research on complex networks [9, 8, 43, 7, 37, 35, 11, 32, 24, 17, 29, 5, 40, 39, 13, 28, 42] and their applications to real world problems [36, 29] have been attracting the attention of many scientists. To understand large-scale behaviour of complex networks, it is imperative to calculate the amount of physical flow going from one node to another one, a quantity that we refer to in this work as “hidden” flow, since this quantity is usually inaccessible to measurements.

In this Letter, we avail from the flow tracing method, known in electrical engineering [16, 31, 18, 19, 4, 15, 44, 2, 27, 30], to calculate the hidden flow between any two nodes, by only requiring information about the adjacent flows between any two connected nodes. This work provides a rigorous way to calculate hidden flows, which in turn enables one to gauge the non-adjacent interactions among nodes in a network, for networks whose non-adjacent nodes are far apart. The applicability of the method is enormous since flow networks can be used as simple models of flow behaviour to many complex networks, such as transportation networks, water supply networks and power grids. We extend the method to provide an immediate picture of how nodes interact non-adjacently in non-linear networks by constructing linear equivalent models to these networks.

Flow networks describe a system that exchanges physical flows. Physical flows are usually recognised as the transference of a physical entity (such as the electric charge, a liquid, a solid, a gas volume, cars, airplanes, air, etc) from one node to another in a given unit of time. But they can also be, in a more general sense, probabilities or the information rate (in bits/s). In a flow network, there are source nodes that input physical flows (a generator in a power grid, for example) and sink nodes from which the physical flows leave the network (a consumer in a power grid, for example). Flow networks can have several configurations, and for each configuration there are several scientific challenges. This work deals with flow networks that are conservative (i.e., total inflow arriving at a node is equal to total outflow leaving it) and whose rule of flow exchange is linear, such as is the case of a direct current electric network. Moreover, the edges carrying the flows are uncapacitated, allowing any arbitrary flow intensity. A remarkable challenge in the area of flow networks is to trace the flow between two non-adjacent nodes (or edges). In lieu of studying flows provided by adjacent connections, tracing methods enable one to calculate the amount of flow exchanged from one node (or edge) to another node (or edge), after travelling through several different paths in the network, a quantity being referred in this work as the “hidden” flow. This computationally doable complex task in small flow networks becomes impractical in larger complex flow networks. The present work reduces this complicated tracing mathematical process into a trivial manipulation of the so called extended incidence matrix \mathbf{K} that can be easily calculated from information on the flows along the edges. We then demonstrate that the hidden flows between any arbitrary pair of nodes can be calculated by our result condensed in Eq. (14). This result, rigorously derived for directed flow networks (preferential direction of flows) and to networks without closed looping flows (where flows circle around a closed path loop) was also extended to the treatment of networks whose flows are undirected and networks that present closed loops. Finally, we also show how to extend this result to understand the non-adjacent interactions between any pair of nodes in more general dynamical networks, such as phase oscillator networks, whose behaviour can be well represented by a conservative flow network.

2 Flow Networks

A flow network is a digraph, $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where \mathcal{V} and \mathcal{E} are the sets of nodes and edges, respectively. A flow network normally contains three types of nodes: (i) the source node [e.g., node 1 or 2 in Fig. 1 (a)], which has a source injecting flow into the network; (ii) the sink node [e.g., node 3 or 4 in Fig. 1 (a)], which has a sink taking flow away from the network; (iii) the junction node [e.g., node 5 in Fig. 1 (a)], which distributes the flow. We define f_{ij} to be the *adjacent flow*, or simply the *flow* which is the measurable flow coming from nodes i to j through edge $\{i, j\} \in \mathcal{E}$. $f_{ij} = 0$ if nodes i and j are not physically connected. We begin our analysis with the conservative flow networks [1] satisfying: (i) $f_{ij} = -f_{ji}$; (ii) $\sum_{j \in \mathcal{V}} f_{ij} = 0$, where node i is a junction node; (iii) there is no loop flow representing a closed path in a flow network, where a loop flow is shown in Fig. 1 (b); (iv) every node must be connected to at least one other node in the network. A path in a digraph \mathcal{G} from node i to node j , $P(i, j) = i \{i, i'\} i' \{i', i''\} \dots \{j', j\} j$, is an alternating sequence of distinct nodes and edges in which the directions of all edges must coincide with their original directions in \mathcal{G} . The *hidden flow*, $f_{i \rightarrow j}$, is defined to be the summation of the flows going from node i to j through all possible paths from node i to j .

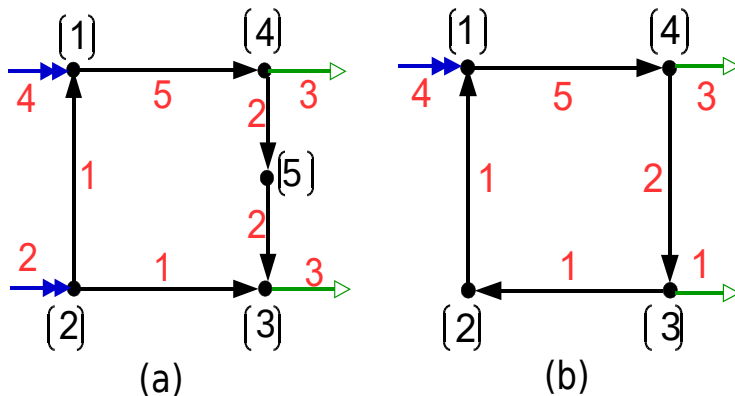


Figure 1: (colour online) (a) A flow network without loop flow. (b) A flow network with loop flow. The black numbers in square brackets are labels of nodes, the red numbers are adjacent flows, the blue lines with double filled arrows are flow sources, the green lines with unfilled arrows are flow sinks, and the black lines with single filled arrows are directed adjacent flows between nodes.

Normally, we can measure or calculate the adjacent flows in a flow network, but it is not easy to obtain the hidden flows, a quantity typically not accessible through measurements. We find the calculation of hidden flows based on the information of adjacent flows, in a conservative flow network, by the “flow tracing” method.

Define the *node-net exchanging flow* at node i by

$$f_i = \sum_{j=1}^N f_{ij}. \quad (1)$$

If node i is a source node, we have $f_i > 0$; we denote f_i by f_i^s as the amount of the *source flow* being injected into the network from a source at node i . We set $f_i^s = 0$ if node i is a sink node or a junction node. If node i is a sink node we have $f_i < 0$; we denote $f_i^t = -f_i > 0$ to indicate the amount of the *sink flow* leaving the network from the sink at node i . We set $f_i^t = 0$ if node i is a source node or a junction node.

Assume there is a positive flow from node i to node j , denoted by $f_{ij} > 0$. We use f_{ij}^{out} to indicate f_{ij} as an *outflow* from node i arriving at node j , and f_{ij}^{in} to represent f_{ij} as an *inflow* at node j coming from node i . Thus, $f_{ij}^{out} = f_{ij}^{in} > 0$. f_{ij} can be positive, negative or zero in a flow network. However, we restrict any outflow or inflow at a node to be a non-negative number. This means that, if $f_{ij} < 0$, we force f_{ij}^{out} and f_{ij}^{in} to be zeros. Analogously, $f_{ij} < 0$ means $f_{ji} > 0$, we have $f_{ji}^{out} > 0$ to denote the outflow from node j to node i and $f_{ji}^{in} > 0$ to be the inflow at node i from node j .

Define the *total inflow* at node i by

$$f_i^{in} = f_i^s + \sum_{f_{ji} > 0} f_{ji} = f_i^s + \sum_{j=1}^N f_{ji}^{in}, \quad (2)$$

and the *total outflow* at node i by

$$f_i^{out} = f_i^t + \sum_{f_{ij} > 0} f_{ij} = f_i^t + \sum_{j=1}^N f_{ij}^{out}. \quad (3)$$

In a conservative flow network, the total inflow of a node is equal to its total outflow, i.e., $f_i^{out} = f_i^{in}$. We assume $f_i^{out} = f_i^{in} > 0, \forall i$, meaning that each node in a flow network must exchange flow with other nodes, i.e., no node is isolated.

3 Flow tracing by proportional sharing principle

The *proportional sharing principle (PSP)* [25, 15] states that for an arbitrary node, a , with m inflows and n outflows (Fig. 2) in a conservative flow network, (i) the outflow on each outflow edge is proportionally fed by all inflows, and (ii) by assuming that node i injects a flow f_{ia}^{in} to node a , and node j takes a flow f_{aj}^{out} out of node a , we have that the *node-to-node hidden flow* from node i to node j via node a is calculated by

$$f_{i \rightarrow j} = f_{ia}^{in} \frac{f_{aj}^{out}}{f_a^{out}}, \quad (4)$$

or by

$$f_{i \rightarrow j} = f_{aj}^{out} \frac{f_{ia}^{in}}{f_a^{in}}. \quad (5)$$

Equations (4) and (5) result in the same value of $f_{i \rightarrow j}$, since $f_a^{out} = f_a^{in}$. Equation (4) represents the *downstream flow tracing* method, where we start tracing the hidden flow from a source node i to a sink node j , by using the percentage, f_{aj}^{out}/f_a^{out} , to indicate the percentage of f_{ia}^{in} that goes to j . Equation (5) denotes the *upstream flow tracing* method, where we trace the flow from a sink node j to a source node i , by knowing the proportion of f_{aj}^{out} is provided by f_{ia}^{in} .

The percentage f_{aj}^{out}/f_a^{out} in Eq. (4) and f_{ia}^{in}/f_a^{in} in Eq. (5) are related to the flows on edges. They are similar to the probability of jumping from a node to one of its neighbours in a biased random walk process [33, 12, 23], where a similar percentage is related to the weight of edges.

We only deal with the downstream flow tracing in the Letter and explain the upstream flow tracing in the Supplementary Material.

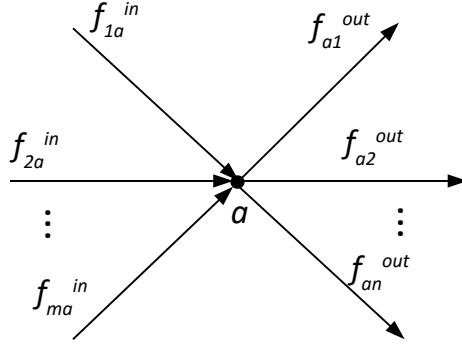


Figure 2: A node a with m inflows and n outflows.

Define the *downstream coefficient* at node a for the outflow f_{aj}^{out} by

$$\kappa_{aj}^d = \frac{f_{aj}^{out}}{f_a^{out}}, \quad (6)$$

to indicate the proportion of the outflow at edge $\{a, j\}$ to the total outflow at node a . Define the *upstream coefficient* at node a for the inflow f_{ia}^{in} by

$$\kappa_{ai}^u = \frac{f_{ia}^{in}}{f_a^{in}}, \quad (7)$$

denoting the proportion of the inflow at edge $\{i, a\}$ to the total inflow at node a . Then the calculation of $f_{i \rightarrow j}$ can be simply expressed by $f_{i \rightarrow j} = f_{ia}^{in} \kappa_{aj}^d$ or $f_{i \rightarrow j} = f_{aj}^{out} \kappa_{ai}^u$.

Define the *sink proportion* and *source proportion* at node a by

$$\iota_a^t = \frac{f_a^t}{f_a^{out}} \text{ and } \iota_a^s = \frac{f_a^s}{f_a^{in}}, \quad (8)$$

respectively, where the sink proportion, ι_a^t , indicates the proportion of the sink flow to the total outflow at node a , and the source proportion, ι_a^s , indicates the proportion of the source flow to the total inflow at node a . By defining the sink proportion and source proportion, we are now able to calculate the *source-to-sink hidden flow* from a source at node i to a sink at node j denoted by $f_{si \rightarrow tj}$. From Eq. (2), we know that f_i^s is a part of f_i^{in} , where f_i^s is the source flow at node i . From Eq. (8), we know the proportion of f_i^s to f_i^{in} . According to the PSP, we can then calculate the source-to-sink hidden flow by $f_{si \rightarrow tj} = \frac{f_i^s}{f_i^{in}} f_{i \rightarrow j} \frac{f_j^t}{f_j^{out}} = \iota_i^s f_{i \rightarrow j} \iota_j^t$.

It is possible to trace (calculate) the hidden flows from any arbitrary pair of nodes in a flow network using either the downstream or the upstream approach. However, all the paths connecting a pair of nodes must be considered. In particular, the hidden flow from two adjacent nodes will include the flow exchanged along the adjacent connection and all the flows travelling along other longer paths connecting these two adjacent nodes. Suppose one wants to calculate the hidden flow $f_{i \rightarrow j}$ from two non-adjacent nodes i and j , and there are two possible paths, $P_1(i, j) = i\{i, k\}\{k, j\}j$ and $P_2(i, j) = i\{i, l\}\{l, g\}\{g, j\}j$, P_1 with length 2 and P_2 with length 3. Each path produces a hidden flow, $f_{i \rightarrow j}^{(1)}$ and $f_{i \rightarrow j}^{(2)}$, respectively. The total hidden flow from i to j is thus calculated using that $f_{i \rightarrow j} = f_{i \rightarrow j}^{(1)} + f_{i \rightarrow j}^{(2)}$, where $f_{i \rightarrow j}^{(1)} = f_i^{in} \kappa_{ik}^d \kappa_{kj}^d$ and $f_{i \rightarrow j}^{(2)} = f_i^{in} \kappa_{il}^d \kappa_{lg}^d \kappa_{gj}^d$. This process is feasible when dealing with small flow networks, as illustrated in the Supplementary Material, where we show how to trace hidden electric current flows in a direct current (DC) electric network. But it becomes impractical when dealing with large networks, for which the number of paths carrying flows can grow exponentially fast with the size of the network. To circumvent this challenging calculation, the use of the extended incidence matrix, \mathbf{K} , proposed in Refs. [44, 2, 27], is taken forward.

4 Flow tracing by extended incidence matrix

The *downstream extended incidence matrix*, \mathbf{K} , in a flow network with N nodes is an $N \times N$ dimensional matrix, defined by

$$K_{ij} = \begin{cases} -f_{ji}^{in}/f_j^{out} & \text{if } i \neq j, \text{ and } f_{ji} > 0, \\ 1 & \text{if } i = j, \\ 0 & \text{else.} \end{cases} \quad (9)$$

Transform Eq (2) to $f_i^{in} - \sum_{j=1}^N f_{ji}^{in}/f_j^{out} \cdot f_j^{out} = f_i^s$. Considering $f_i^{in} = f_i^{out}$,

we have

$$f_i^{out} - \sum_{j=1}^N f_j^{in} / f_j^{out} \cdot f_j^{out} = f_i^s. \quad (10)$$

From Eqs. (9) and (10), we have

$$\mathbf{K}\mathbf{F}^{out} = \mathbf{F}^s, \quad (11)$$

where $\mathbf{F}^{out} = [f_1^{out}, f_2^{out}, \dots, f_N^{out}]^T$, and $\mathbf{F}^s = [f_1^s, f_2^s, \dots, f_N^s]^T$. \mathbf{K} is an $N \times N$ square invertible matrix [44, 27, 30], thus, $\mathbf{F}^{out} = \mathbf{K}^{-1}\mathbf{F}^s$, implying that,

$$f_i^{out} = \sum_{j=1}^N [\mathbf{K}^{-1}]_{ij} f_j^s, \quad (12)$$

$[\mathbf{K}^{-1}]_{ij}$ being an entry (i^{th} row, j^{th} column) of the matrix \mathbf{K}^{-1} . Equation (12) indicates that the outflow of node i , f_i^{out} , is fed by every source f_j^s . More specifically, $[\mathbf{K}^{-1}]_{ij}$ represents the proportion of the source inflow in the source node j that goes to node i .

Let $\mathbf{C} = \mathbf{K}^{-1}$ be the *downstream contribution matrix*. Considering $\iota_j^s = f_j^s / f_j^{in}$, we have

$$f_i^{out} = \sum_{j=1}^N C_{ij} f_j^{in} \iota_j^s. \quad (13)$$

Knowing that the source-to-node hidden flow from source node j to node i is given by $f_{sj \rightarrow i} = \iota_j^s f_{j \rightarrow i}$, Eq. (13) thus implies that for a source node j with $\iota_j^s \neq 0$, $C_{ij} f_j^{in}$ represents the node-to-node hidden flow from node j to node i , i.e., $f_{j \rightarrow i} = C_{ij} f_j^{in}$. The tracing of flows from source to nodes, previously known in the literature, only applied to source nodes. To extend it to any other general situation, including the tracing of flows from and to edges, sinks and junction nodes, we introduce an *equivalence principle*. We treat any sink or junction node as a hypothetical source node, without altering the original network topology and flows. If node j is a sink or junction node with a total inflow $f_j^{in} > 0$ and $\iota_j^s = 0$, we treat node j as a hypothetical source node with $f_j^s = f_j^{in} > 0$, where the hypothetical source takes the place of all the edges injecting flows into j . By this treatment, we can hypothetically treat node j as a source node with $\iota_j^s = f_j^s / f_j^{in} = 1$, in Eq. (13), such that the node-to-node hidden flow from node j to node i can also be calculated by

$$f_{j \rightarrow i} = C_{ij} f_j^{in}. \quad (14)$$

Thus, from our analysis, $C_{ij} = [\mathbf{K}^{-1}]_{ij}$ is a *downstream contribution factor* indicating how much hidden flow goes from node j to i , i.e., $f_{j \rightarrow i} = C_{ij} f_j^{in}$ for any pair of nodes.

Now, we show how non-adjacent hidden flows can be traced in conservative flow networks. Notice for networks whose non-adjacency nodes are far apart

from each other, the hidden flows can gauge how non-adjacent interactions emerge in the studied system. Let i, j, m, n, p, q be different nodes in a conservative flow network, where node i has a source, node j has a sink, nodes m, n are connected by edge $\{m, n\}$ with $f_{mn} > 0$, and nodes p, q are connected by edge $\{p, q\}$ with $f_{pq} > 0$. The non-adjacent interaction includes: (i) the node-to-node hidden flow from node i to j is $f_{i \rightarrow j} = C_{ji} f_i^{in}$; (ii) the source-to-node hidden flow from source node i to node j is $f_{si \rightarrow j} = \iota_i^s f_{i \rightarrow j}$; (iii) the node-to-sink hidden flow from source node i to sink node j is $f_{i \rightarrow tj} = f_{i \rightarrow j} \iota_j^t$; (iv) the source-to-sink hidden flow from node i to j is $f_{si \rightarrow tj} = \iota_i^s f_{i \rightarrow j} \iota_j^t$; (v) the node-to-edge hidden flow from node i to edge $\{m, n\}$ is $f_{i \rightarrow \{m, n\}} = f_{i \rightarrow m} \cdot \kappa_{mn}^d$; (vi) the edge-to-node hidden flow from edge $\{m, n\}$ to node j is $f_{\{m, n\} \rightarrow j} = \kappa_{nm}^u \cdot f_{n \rightarrow j}$; and (vii) the edge-to-edge hidden flow from edge $\{p, q\}$ to $\{m, n\}$ is $f_{\{p, q\} \rightarrow \{m, n\}} = \kappa_{qp}^u \cdot f_{q \rightarrow m} \cdot \kappa_{mn}^d$.

To illustrate the calculation of these hidden flows, as well as the calculation of the matrices involved in it, in the Supplementary Material we trace the flows in an electric network using our downstream extended incidence matrix approach.

5 Extension to flow networks with closed loops and with undirected flows

Loops: If the closed loop (or loops) is inside a larger network, one needs first to identify the existence of a loop. A closed loop at the node i with a length P exists in a network if $[A^P]_{ii} > 0$, where $[A^P]_{ii}$ represents the term ii in the power to P of the adjacency matrix of the network. The source node of the loop is any node receiving input flow, and the sink node is the one containing an edge with an outflow, and whose path length connecting it to the source node is the longest.

We consider a network with 4 nodes, with a loop flow as in Fig. 1(b). Let us call it network N . Denote the input flow as $f_1^s(N)$, the output flows as $f_4^t(N)$ and $f_3^t(N)$, and the adjacent flows as $f_{14}(N)$, $f_{43}(N)$, $f_{32}(N)$, and $f_{21}(N)$. A loop in a flow network is broken down into subnetworks in which the flows are directed. Merging the flows of all subnetworks must preserve edge, source and sink flows of the original network N . In Fig. 1(b), the loop is formed by $1\{1, 4\}\{4, 3\}\{3, 2\}\{2, 1\}1$. To break up the loop, one firstly choose a source and a sink node, where flows enter and leave the closed loop, respectively. Node 1 is the only source node. The sink node to be chosen must be the one whose length of a direct path connecting it to the source node is the longest one. We choose node 3 as the sink node. Then, one needs to determine all the directed paths connecting the source node (node 1) and to the sink node (node 3), and all the directed paths connecting the sink to the source nodes. Among all paths, one takes only the paths that have the same flow directions as the original network N . These directed paths form the subnetworks whose net flow represents the original network flow and from which the hidden flows are calculated.

We show, in Fig. 3, the subnetworks of the network in Fig. 1(b). Panel (a1) represents a directed path and its flows from node 1 to node 3. Panels

(a2) and (a3), with the same directed path subnetwork, show the directed paths connecting nodes 3 to 1. Notice that a negative source and sink, in nodes 1 and 3, respectively, in panel (a2), is equivalent to a positive sink and source nodes, respectively, as represented in panel (a3). In panels (b1)-(b3), we show another practical way to determine the break up of the network with a closed loop. Once a loop, and a source and a sink nodes, are identified, we remove it from the network. Panel (b1) is the subnetwork after the loop removal. The closed loop is formed by merging the flows represented in panels (b2) and (b3), and it has a constant flow of 1 unit. One restores the original network by adding the subnetworks in panel (a1) and (a3), or by adding the subnetworks in panels (b1), (b2), and (b3). Calculating hidden flows of the original network needs to take into consideration of hidden flows in all subnetworks. One subnetwork [panel (a1)], let us call it $N1$, is formed by the nodes 1, 3, and 4. Node 2 is absent and, therefore, to preserve edge flows one is required to make $f_1^s(N1) = f_1^s + f_{21}(N)$ and $f_3^t(N1) = f_3^t + f_{32}(N)$. From this network, $f_{1 \rightarrow 4} = 5$, $f_{s1 \rightarrow t4} = 3$, $f_{s1 \rightarrow t3} = 2$. The other network [panel (a2)], let us call it $N2$, is formed by the nodes 1, 2, and 3, so node 4 is now absent and therefore, to preserve edge flows we are required to make $f_1^s(N2) = f_1^s + f_{41}(N) = f_1^s - f_{14}(N)$ and $f_3^t(N2) = f_3^t + f_{34}(N) = f_3^t - f_{43}(N)$. These equations lead to $f_1^s(N2) < 0$ and $f_3^t(N2) < 0$, whose flows are indicated in panel (a2). The hidden flow from node 2 and 4 is zero, since no subnetworks contribute to a hidden flow from node 2 to 4.

Undirected flow networks: Similarly, our method can also be applied to an undirected flow network if the network can be split into two independent unidirectional networks. For example, under the assumption that all traffic roads are bidirectional, we can separate the transportation network of a city into two networks. One network includes all the left-hand roads and the other one contains all the right-hand roads. Thus, both separated networks become unidirectional networks.

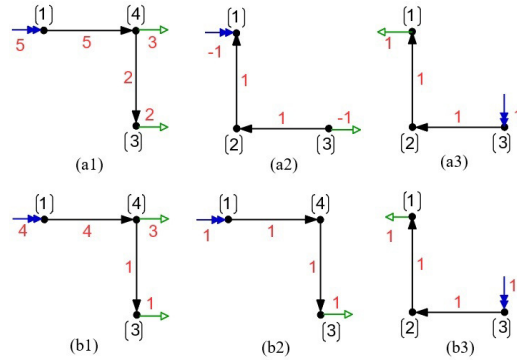


Figure 3: (colour online) Illustrations of two approaches to break up a flow network with a closed loop flow into smaller subnetworks with only directed flows.

6 Non-adjacent interaction in non-linear networks

Next, we extend our tracing hidden flow approach to study non-linear systems by constructing linear model analogous to the non-linear networks. Let the equation

$$\dot{x}_i = S(x_i) - \sum_{j=1}^N L_{ij} \cdot H(x_i, x_j) \quad (15)$$

indicate a dynamic scheme describing the behaviour of N coupled nodes, where x_i is the dynamical variable of each node, $S(x_i)$ is the isolated dynamic function, L_{ij} is the element of the Laplacian matrix, and $H(x_i, x_j)$ is an arbitrary coupled dynamic function. We treat the system as a flow network by interpreting $f_i(t) = S(x_i) - \dot{x}_i$ as the node-net exchanging flow at node i . The value and sign of $f_i(t)$ may change over time. If $f_i(t) > 0$ (or $f_i(t) < 0$), we treat node i as a source (or sink) node at time t and the source (or sink) flow is $f_i^s(t) = f_i(t)$ (or $f_i^t(t) = -f_i(t)$). If $f_i(t) = 0$, we treat node i as a junction node at time t . Let $f_{ij}(t) = L_{ij}H(x_i, x_j)$ be the adjacent flow from node i to node j . If $f_{ij}(t) > 0$, we have $f_{ij}^{out}(t) > 0$ as the outflow from node i and $f_{ij}^{in}(t) > 0$ as the inflow at node j at time t . If $f_{ij}(t) < 0$, we have $f_{ji}^{out}(t) > 0$ as the outflow from node j and $f_{ji}^{in}(t) > 0$ as the inflow at node i at time t . By doing this interpretation, we are constructing an equivalent linear conservative flow network that behaves in the same way as the non-linear network described by Eq. (15). This enables us to calculate the non-adjacent interactions in the equivalent linear flow network which informs us about the non-adjacent interactions in the original non-linear network.

We consider a revised Kuramoto model [20, 21, 22] as an example, which is given by

$$\dot{\theta}_i = \omega_i - K \sum_{j=1}^N L_{ij} \sin(\theta_i - \theta_j), \quad (16)$$

where K is the coupling strength, L_{ij} is the entry of the Laplacian matrix, θ_i and ω_i indicate the phase angle and natural frequency in a rotating frame, respectively. In this rotating frame, $\dot{\theta}_i = \dot{\theta}_j = 0$, $\forall i \neq j$, when the oscillators emerge into *frequency synchronisation (FS)* for a large enough K [34]. In the FS state, all the node-net exchanging flows $f_i = \omega_i - \dot{\theta}_i = \omega_i$ and all the adjacent flows $f_{ij} = KL_{ij} \sin(\theta_i - \theta_j)$ are constants, since $\sin(\theta_i - \theta_j)$ are constants.

Let $\alpha_{ij} = |f_{ij}| / \max\{|f_{ij}| : \forall i, j\}$ be a normalised variable in $[0, 1]$ indicating the adjacent interaction strength between oscillator i and j , where $\max\{|f_{ij}| : \forall i, j\}$ is the maximum of all absolute values of adjacent flows. Since $f_{ij} = -f_{ji}$, we have $\alpha_{ji} = \alpha_{ij}$. Every hidden flow is traced by considering that flows are directed. This implies that all the calculated hidden flows are non-negative and at least one of $f_{i \rightarrow j}$ and $f_{j \rightarrow i}$ is 0. We let $\beta_{ij} = \beta_{ji} = \max\{f_{i \rightarrow j}, f_{j \rightarrow i}\} / \max\{f_{i \rightarrow j} : \forall i, j\}$ be the non-adjacent interaction strength between oscillator i and j , where $\max\{f_{i \rightarrow j}, f_{j \rightarrow i}\}$ is the non-zero one between $f_{i \rightarrow j}$ and $f_{j \rightarrow i}$, and $\max\{f_{i \rightarrow j} : \forall i, j\}$ is the maximum of all hidden flows. This

definition of the non-adjacent interaction strength allows us to compare α_{ij} and β_{ij} for the same pair of nodes in a network.

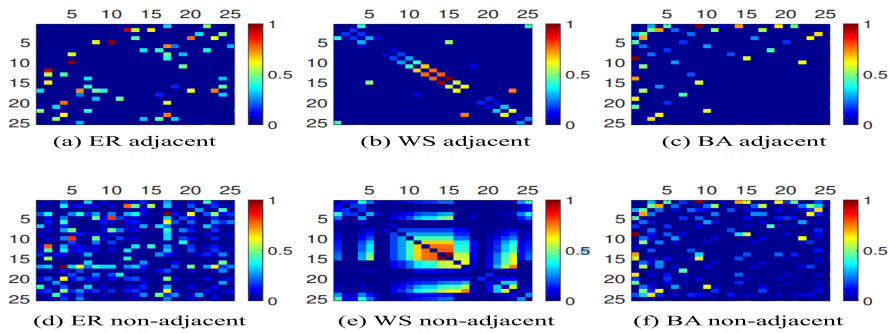


Figure 4: (colour online) Comparison of adjacent interactions and non-adjacent interactions in different networks described by the Kuramoto model after the occurrence of frequency synchronisation. (a), (b) and (c) demonstrate the adjacent interactions in ER network, WS network and BA network, respectively, compared with the non-adjacent interactions shown in (d), (e) and (f) for these networks. The numbers on axes are labels of nodes. The colours on maps indicate the interacting strength between nodes.

We construct three types of networks with 25 nodes, namely the Erdős-Rényi (ER) [9, 10], Watts-Strogatz (WS) [41] and Barabási-Albert (BA) models [3]. The dynamic behaviour of the nodes in these networks follows Eq. (16). Figure 4 shows the comparison of the adjacent interactions and the non-adjacent interactions when the oscillators emerge into FS with a large enough K . Figures 4 (a), (b) and (c) show the adjacent interaction strengths, α_{ij} , for ER, WS and BA networks, respectively. Figures 4 (d), (e) and (f) demonstrate the non-adjacent interaction strengths, β_{ij} , for ER, WS and BA networks, respectively. Figure 4 (d) exposes some hidden interactions that Fig. 4 (a) does not show to exist in an ER network. By comparing Figs. 4 (b) and (e), we see that a randomly rewired edge in a WS network not only produces interaction between the two adjacent nodes connected by this edge, but also creates functional clusters among nodes close to the two adjacent nodes. So, complex systems can in fact be better connected than previously thought. We constructed the BA network by assigning smaller labels to nodes with larger degrees. Both Figs. 4 (c) and (f) illustrate the strong interactions among the nodes with large degrees (small labels). Figure 4 (c) shows that the interactions between unconnected nodes with small degrees (large labels) are weak or inexistent, though, such interactions are revealed in Fig. 4 (f). Through this comparison, we understand that two nodes in a network may strongly interact with each other even if they are not connected by an edge.

Figure 5 shows the simulations results of the adjacent interaction strength and non-adjacent interaction strength for these networks when FS is not present.

Final results are taken by averaging the results of 100 time-points that are uniformly chosen in the time scale $[10,20]$, i.e., $\alpha_{ij} = \sum_k^{100} \alpha_{ij}(t_k)/100$ and $\beta_{ij} = \sum_k^{100} \beta_{ij}(t_k)/100$, where $\alpha_{ij}(t_k)$ and $\beta_{ij}(t_k)$ are the values of α_{ij} and β_{ij} at the k^{th} time-point. The dynamic behaviour of the oscillators in these networks is described by the Kuramoto model by assigning a small coupling strength, such that the oscillators are in an incoherent state.

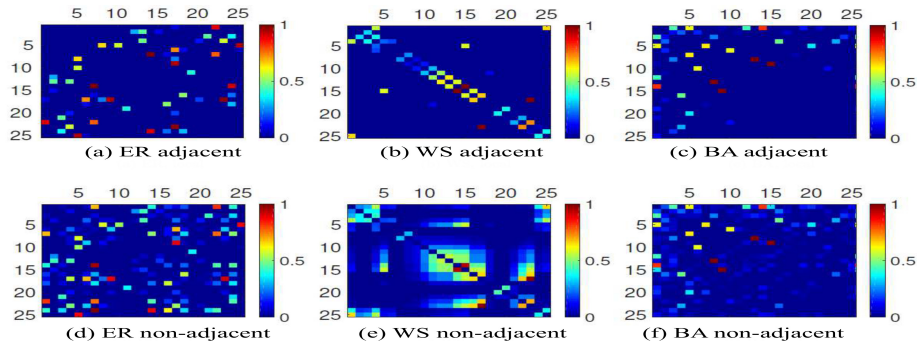


Figure 5: (colour online) Comparison of adjacency interactions and non-adjacency interactions in different types of networks described by the Kuramoto model when frequency synchronisation is inexistent. (a), (b) and (c) demonstrate the adjacent interaction strength in ER network, WS network and BA network, respectively. (d), (e) and (f) show the non-adjacent interaction strength for these networks. The numbers on axes are labels of nodes. The colour on map indicates the interacting strength between nodes.

Comparing the results in Fig. 5 with that when FS is present, we find that those pairs of nodes which are not interacting through hidden flows when FS is not present, also present no evident non-adjacency interactions when FS is present. This suggests that the existence of non-adjacent interaction between a pair of nodes strongly depends on the network topological features of the network rather than the coupling strength.

7 Conclusion

In this Letter, we introduced the proportional sharing principle and the extended incidence matrix to calculate the hidden flows in flow networks, and further extended this approach to trace the non-adjacent hidden flows in non-linear complex systems which can analogously be represented by linear flow networks. This allows us to understand the non-adjacency interactions among nodes either under a steady state (e.g., when FS is present in the Kuramoto model) or a dynamic state (e.g., when FS is not present in the Kuramoto model) in such a complex system. Our study illustrated that the nodes in a network not only interacts with their neighbours, but can also strongly influence those who are not

directly connected to them. By comparing the results of the non-adjacent study for the Kuramoto model when FS is present and that when FS is not present for different topological networks, we concluded that the emergence of non-adjacent interaction between a pair of nodes strongly depends on the topological features of the networks rather than the coupling strength between nodes.

We have extended our analysis to flow networks that present closed loops and for those that present undirected flows. The solution for these challenging problems is to break the network into subnetworks that only contain directed flows. The method can also be applied to weighted networks, as long as the weighted network can be modelled as a conservative flow network.

This work opens up a new area of research into non-adjacent interactions in complex networks, facilitating and enabling research that aims at unravelling complex behaviour as a function of the network topology. There is also great potential to link this work to other works in the area of complex networks, such as the link prediction problem [26], and to the study of information and energy transmission in complex networks [6, 14, 38]. These potential extensions will further widen the applicability of the method in the real world. It is worth mentioning that our work assumed at the outset that the adjacency matrix of the system as well as the adjacency physical flows is known *a priori*. Therefore, works such as those in Ref. [26] predicting the existence of a physical link should be used prior to our method.

8 Acknowledgments

Chengwei Wang is supported by a studentship funded by the College of Physical Sciences, University of Aberdeen.

References

- [1] Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. Network flows: theory, algorithms, and applications. 1993.
- [2] MA Anuar and T Hiyama. Observability Analysis Using Extended Incidence Matrix. In *2010 Asia-Pacific Power and Energy Engineering Conference*, pages 1–4. IEEE, 2010.
- [3] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [4] Janusz Bialek. Tracing the flow of electricity. In *Generation, Transmission and Distribution*, volume 143, pages 313–320. IET, 1996.
- [5] R. Carareto, M. S. Baptista, and C. Grebogi. Natural synchronization in power-grids with anti-correlated units. *Commun. Nonlinear Sci. Numer. Simul.*, 18:1035–1046, 2013.

- [6] Yu-Zhong Chen, Le-Zhi Wang, Wen-Xu Wang, and Ying-Cheng Lai. Energy scaling and reduction in controlling complex networks. *R. Soc. Open Sci.*, 3(4):160064, 2016.
- [7] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. *Phys. Rev. Lett.*, 85:4633–4636, 2000.
- [8] Paul Erdős and A Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.
- [9] P Erdos and A Rényi. On random graphs i. *Publ. Math. Debrecen*, 6:290–297, 1959.
- [10] Edgar N Gilbert. Random graphs. *Ann. Math. Stat.*, 30:1141–1144, 1959.
- [11] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA*, 99:7821, 2002.
- [12] Jesús Gómez-Gardeñes and Vito Latora. Entropy rate of diffusion processes on complex networks. *Phys. Rev. E*, 78(6):065102, 2008.
- [13] Dirk Helbing, Stefan Lämmer, Thomas Seidel, Pétr Šeba, and Tadeusz Płatkowski. Physics, stability, and dynamics of supply networks. *Physical Review E*, 70(6):066116, 2004.
- [14] Liang Huang, Kwangho Park, and Ying-Cheng Lai. Information propagation on modular networks. *Phys. Rev. E*, 73(3):035103, 2006.
- [15] Zhaoxia Jing and Fushuan Wen. Discussion on the proving of proportional sharing principle in electricity tracing method. In *Transmission and Distribution Conference and Exhibition: Asia and Pacific*, pages 1–5. IEEE, 2005.
- [16] Baseem Khan, Ganga Agnihotri, Gaurav Gupta, and Pawan Rathore. A Power Flow Tracing based Method for Transmission Usage, Loss & Reliability Margin Allocation. *AASRI Procedia*, 7:94–100, 2014.
- [17] A. R. Kiremire, M. R. Brust, and V. V. Phoha. Using network motifs to investigate the influence of network topology on PPM-based IP traceback schemes. *Comput. Netw.*, 72:14, 2014.
- [18] Daniel Kirschen, Ron Allan, and Goran Strbac. Contributions of individual generators to loads and flows. *IEEE Trans. Power Syst.*, 12(1):52–60, 1997.
- [19] Daniel Kirschen and Goran Strbac. Tracing active and reactive power between generators and loads using real and imaginary currents. *IEEE Trans. Power Syst.*, 14(4):1312–1319, 1999.

- [20] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422. Springer, 1975.
- [21] Yoshiki Kuramoto. *Chemical oscillations, turbulence and waves*. Springer, Berlin, 1984.
- [22] Yoshiki Kuramoto and Ikuko Nishikawa. Statistical macrodynamics of large dynamical systems. case of a phase transition in oscillator communities. *Journal of Statistical Physics*, 49(3-4):569–605, 1987.
- [23] Renaud Lambiotte, Roberta Sinatra, J-C Delvenne, Tim S Evans, Mauricio Barahona, and Vito Latora. Flow graphs: Interweaving dynamics and structure. *Phys. Rev. E*, 84(1):017102, 2011.
- [24] K. Langendoen and N. Reijers. Distributed localization in wireless sensor networks: a quantitative comparison. *Comput. Netw.*, 43:499, 2003.
- [25] FB Liu, Y Li, and GQ Tang. Proof and extension of ”proportional sharing principle” on energy market. In *2000 International Conference on Advances in Power System Control, Operation and Management.*, volume 1, pages 91–95. IET, 2000.
- [26] Linyuan Lü, Liming Pan, Tao Zhou, Yi-Cheng Zhang, and H Eugene Stanley. Toward link predictability of complex networks. *PNAS*, 112(8):2325–2330, 2015.
- [27] A. A Malik and N. D Ghawghawe. Determination of Generator Participation in Loads and Lines using Extended Incidence Matrix method. *IJSER*, 6(2):1124–1128, 2015.
- [28] Nora Molkenhain, Kira Rehfeld, Norbert Marwan, and Jürgen Kurths. Networks from flows-from dynamics to topology. *Scientific reports*, 4:4119, 2014.
- [29] P. H.J. Nardelli, N. Rubido, C. Wang, M. S. Baptista, C. Pomalaza-Raez, P. Cardieri, and M. Latva-aho. Models for the modern power grid. *Eur. Phys. J. Spec. Top.*, 223:2423, 2014.
- [30] N Prakash Reddy and P Sangameswara Raju. Tracing and Transmission Cost Allocation of Active Power Flow Based On Extended Incidence Matrix. *IJMER*, 3(6):3433–3440, 2013.
- [31] R Reta and A Vargas. Electricity tracing and loss allocation methods based on electric concepts. In *Generation, Transmission and Distribution*, volume 148, pages 518–522. IET, 2001.
- [32] G. Robins, P. Pattison, Y. Kalish, and D. Lusher. An introduction to exponential random graph (p^*) models for social networks. *Soc. Netw.*, 29:173, 2007.

- [33] Roberta Sinatra, Jesús Gómez-Gardenes, Renaud Lambiotte, Vincenzo Nicosia, and Vito Latora. Maximal-entropy random walks in complex networks with limited information. *Phys. Rev. E*, 83(3):030103, 2011.
- [34] S. H. Strogatz. From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Phys. D*, 143:1–20, 2000.
- [35] Chengwei Wang, Celso Grebogi, and Murilo S Baptista. One node driving synchronisation. *Sci. Rep.*, 5:18091, 2015.
- [36] Chengwei Wang, Celso Grebogi, and Murilo S Baptista. Control and prediction for blackouts caused by frequency collapse in smart grids. *Chaos*, 26(9):093119, 2016.
- [37] Chengwei Wang, Nicolás Rubido, Celso Grebogi, and Murilo S. Baptista. Approximate solution for frequency synchronization in a finite-size Kuramoto model. *Phys. Rev. E*, 92:062808, 2015.
- [38] Le-Zhi Wang, Yu-Zhong Chen, Wen-Xu Wang, and Ying-Cheng Lai. Physical controllability of complex networks. *Sci. Rep.*, 7:40198, 2017.
- [39] Wen-Xu Wang and Ying-Cheng Lai. Abnormal cascading on complex networks. *Physical Review E*, 80(3):036109, 2009.
- [40] Wen-Xu Wang, Ying-Cheng Lai, and Celso Grebogi. Data based identification and prediction of nonlinear and complex dynamical systems. *Phys. Rep.*, 644:1–76, 2016.
- [41] D. J. Watts and S. H Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.
- [42] Duncan J Watts. *Small worlds: the dynamics of networks between order and randomness*. Princeton university press, 1999.
- [43] Duncan J Watts. *Six degrees: The science of a connected age*. WW Norton & Company, 2004.
- [44] Kaigui Xie, Jiaqi Zhou, and Wenyan Li. Analytical model and algorithm for tracing active power flow based on extended incidence matrix. *Electr. Pow. Syst. Res.*, 79(2):399–405, 2009.