Supplementary Information

Structural controllability of unidirectional bipartite networks

by J.C. Nacher and T. Akutsu.

Contents

I. INTRODUCTION

Communication and technological networks have become the backbone of modern societies. Currently, economic transactions, transportation systems, web services such as Google, social interactions, including those from Facebook and Twitter, as well as the spread of biological and mobiles viruses are largely driven by networks. Similarly, our own bodies are composed of interdependent organs and cells, which depend on the fundamental molecular networks of highly interconnected neurons, proteins, metabolites and regulatory genes that cooperate to maintain an internal equilibrium and address significant changes in the environmental conditions and complex disorders. Although these networks perform specific functions in disparate systems, the idea of pursuing the controlling system's entire dynamics and being able to redirect its functionality at will is common among all of them, from reverting negative spreading opinions and rumours in social networks to suppressing a set of oncogenes that are expressed at high levels in a gene regulatory network. Although the processes that occur on real-world networks are mostly non-linear, canonical linear, time-invariant nodal dynamics [1] has been proposed for studying the controllability of networks

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) + B\mathbf{u}(t) \tag{1}$$

where $\mathbf{u}(t)$ is the vector of the input signals and B is the input matrix, which indicates how the signals are coupled to the network nodes. The state of a system $\mathbf{x}(t)$ of N nodes at time t, can indicate positive/negative opinions or high/low expression levels that change with time according to the transpose of the adjacency matrix A that represents the interaction between the system's nodes. The reason for this simplification in modelling non-linear systems is because the structural controllability [2] of a given system is equivalent to the controllability of a continuum of linearised systems; therefore, the analytical results could provide sufficient controllability conditions for most nonlinear systems [1, 3].

This system is controllable if the $N \times NM$ controllability matrix $C = (B, AB, A^2B, ..., A^{N-1}B)$ has a full rank of $rank(C) = N$. Several problems prevent the direct application of the controllability matrix to an arbitrary network. The weights of the edges (i.e., the elements of the adjacency matrix A) are not always known for real networks, and even if they were known, we would still compute the rank C for $2^N - 1$ combinations, which may not be computationally feasible for large networks.

To address network controllability, models based on nodal [1] and edge dynamics [4] were recently proposed. In nodal dynamics, the minimum number of input signals necessary to control the whole network is determined by finding the maximum matching in a bipartite graph obtained from the original network. The number of unmatched nodes is the number of driver nodes.

In this approach, the input signals (or driver nodes) tend to avoid the high-degree nodes [1]. As a result, random networks without hubs are easier to control. An alternative view tackles the problem by evaluating a dynamical process defined that is defined on the edges of a network rather than on the nodes [4]. In this edge dynamics approach, each node i acts as a simple switchboard-like device, mathematically represented as a mixing matrix M_i with the rows (columns) being equal to the out-degree (in-degree), which receives information through its inbound edges and transmits the outcome or decisions to the neighbouring nodes by means of the outbound edges. In sharp contrast to the nodal dynamics approach, this approach concludes that a scale-free degree distribution, in which hubs are present, is easier to control. However, despite the fact that bipartite networks represent a type of networks that is often used to represent the interactions of distinct units in real-world systems, both nodal and edge dynamics frameworks only address simple (unipartite) graphs.

Here we address the controllability of unidirectional bipartite networks, but instead of using nodal dynamics we attempt the problem from a different angle by, considering a modified version of the *minimum dominating set* (MDS) [5, 6]. A set $S \subseteq V$ of nodes in a graph $G = (V, E)$ is a dominating set if every node $v \in V$ is either an element of S or is adjacent to an element of S [7]. This *dominant set* (DS) of nodes plays the role of the set of driver nodes in the sense of [1, 4]. In our companion work, the MDS was suggested as a method to investigate the controllability of complex networks under the assumption that each node can control its outgoing edges separately [5]. The presented conceptual approach, based on edges rather than nodes, was similar to the edge dynamics that was independently proposed in [4]. Our findings showed that as the network degree distribution becomes increasingly heterogeneous, the entire system also becomes easier to control.

Here, we exploit the powerful framework of the MDS, which in bipartite graphs is known as the *Set Covering Problem*, to tackle the controllability of unidirectional bipartite networks. The developed analytical tools, combined with the evaluation of real-world networks from socio-technical and biological systems, offer a promising framework to control unidirectional

bipartite networks with the minimum number of driver nodes. Our analysis unveils the role of the maximum degree H in addressing the network controllability, and how this dependence significantly changes when the power-law degree exponent (γ_1) of the set of nodes that exert the control is above or below the value 2. The theoretical analysis shows that the maximum degree has a significant influence on the size of the DS. Additionally, the analysis also derives the order of nodes (*upper bound*) necessary to control the network. Among all of the topologies, unidirectional bipartite networks with scale-free degree distribution with γ_1=1.5 lead to a smaller upper bound of the number of nodes to be controlled. The dynamics model corresponding to the MDS approach for unidirectional bipartite networks is also presented.

This document is organised as follows. Section II describes the network datasets used in this study as well as the data sources. Section II also includes the main statistical features of the real-world bipartite networks analysed in the study. Section III describes in more detail some of the results derived in this study from the data analysis. Section IV contains the dynamics model and the details of the mathematical tools derived from the theoretical analysis for the controllability of bipartite networks. Finally, Section V describes the controllability problem in the case of bi-directional bipartite networks.

II. DATASETS

Here we introduce the network datasets used in this work. We have collected ten bipartite networks to investigate the minimum number of nodes required to control the entire network. These datasets correspond to the network representation of social and biological systems.

A. Fit of power-law distributions using real-world data

The collected networks were fitted to power-law distributions to determine the degree exponent γ and the lower bound of power-law behaviour x_{min}. By following [8, 9], let x represent a sequence of observations of some variable that can only take a discrete set of values whose distribution we wish to fit as a power-law. We then consider the case of integer values with a probability distribution that follows

$$p(x) = Pr(X = x) = Cx^{-\gamma} \tag{2}$$

Because this distribution is divergent at zero, a lower bound $x_{min} > 0$ should exist for the power-law behaviour. By calculating the normalising constant, we can write

$$p(x) = \frac{x^{-\gamma}}{\zeta(\gamma, x_{min})}, \tag{3}$$

where

$$\zeta(\gamma, x_{min}) = \sum_{n=0}^{\infty} (n + x_{min})^{-\gamma} \tag{4}$$

is the Hurwitz zeta function. Here, we note that the complementary cumulative distribution function defined as $P(x) = Pr(X \geq x)$ can take the form in the discrete case as

$$P(x) = \frac{\zeta(\gamma, x)}{\zeta(\gamma, x_{min})}. \tag{5}$$

The method for fitting power-law distributions to the observed data is the method of *maximum likelihood*. By assuming that the data set corresponds to a distribution that follows a power-law exactly for $x \geq x_{min}$, it is possible to derive the maximum likelihood estimators (MLEs) of the scaling parameter.

For the general case $x_{min} > 1$, we can write the log-likelihood function as

$$\mathcal{L} = \ln \prod_{i=1}^{n} \frac{x_i^{-\gamma}}{\zeta(\gamma, x_{min})} = -n \ln \zeta(\gamma, x_{min}) - \gamma \sum_{i=1}^{n} \ln x_i. \tag{6}$$

By taking $\frac{\partial \mathcal{L}}{\partial \gamma} = 0$, we obtain

$$\frac{-n}{\zeta(\alpha, x_{min})} \frac{\partial}{\partial \gamma} \zeta(\gamma, x_{min}) - \sum_{i=1}^{n} \ln x_i = 0. \tag{7}$$

Then, the MLE $\hat{\gamma}$ for the scaling parameter can be obtained as a solution of

$$\frac{\zeta'(\hat{\gamma}, x_{min})}{\zeta(\hat{\gamma}, x_{min})} = -\frac{1}{n} \sum_{i=1}^{n} \ln x_i. \tag{8}$$

This equation can be solved numerically. Here, the *hatted* symbol denotes the estimates derived from data and the *hatless* symbol denotes the true values, which may not be known in practice. Implementations for this algorithm can be found in [10]. Although there is no exact closed-form expression for $\hat{\gamma}$ in the discrete case, an approximate expression can be derived as

$$\hat{\gamma} \simeq 1 + n \left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{min} - \frac{1}{2}} \right]^{-1} . \quad (9)$$

When $x_{min} \geq 6$, it gives results that are accurate to approximately 1% or better. For details on the derivation, fitting procedure and implementations, we refer to [9, 10]. The maximum likelihood, as described above, estimates the scaling parameter γ for each possible value of x_{min}. Then, the Kolmogorov-Smirnov goodness-of-fit statistic KS is computed. This is achieved by computing the maximum distance between the cumulative distribution function (CDF) of the data and the fitted model. The estimate of x_{min} is determined as the value that gives the minimum value KS over all values of x_{min}. The results for the scaling exponent γ are shown in Figs. 1S-4S and Table 2, including the standard errors of γ which are derived from the width of the maximum likelihood as well as the x_{min} values [9].

B. Social networks

We analysed six social networks to evaluate our controllability approach based on dominating sets for bipartite networks. A Facebook-like social network was originated from an online community for students at the University of California, Irvine. Here we do not focus on the emails between users. Instead, we consider the users' activity in the forum [11]. This forum can be represented as a bipartite network, in which a user is linked to a topic based on whether the user posted messages to that topic. More messages will assign a higher weight to that edge. Because we are considering unweighted networks, we omit the weight of each edge and assume binary interactions among users and topics. This example fits well with the assumption of our DS approach. As stated before, a network is structurally controllable if a dominating set is selected as a set of control nodes under the assumption that each driver node can control its outgoing edges separately. In the Facebook network, $V_\top$ is the users set, and each user can decide the topic in which a new message is posted. The cover size for this network represents 10% of all users (see Table 2). This small set of nodes may have an influence on the opinions circulating in the forums. Although we have not considered the network that connects individuals, the analysed bipartite network can be considered to be complex. In fact, current analyses only focus on complex bipartite networks without considering the connections between the same type of nodes [12–18]. In

our Facebook network, two users (A and B) are not connected to each other (if they were, it would not be a bipartite graph); hence, it is true that the opinions do not circulate between users directly. However, user A can post a message in topic (1), and then this message is read by B, which triggers him/her to post a message in either the same (1) or different topic (2).

Note also that the bipartite graph representations and set cover algorithmic framework have already been used in many applications, some of which are related to clinical analysis. In a recent work, a minimal hitting set algorithm, which is equivalent to set cover, is applied on a bipartite strain-drug response network, which successfully targets the whole population of 60 tumour derived cell lines, uncovering 14 anticancer drug combinations [35].

Next, we focus on the firms-world cities network. These data represent the services (indicating the importance of a city in the office network of a firm) of 100 global service firms distributed across 315 cities worldwide. Each firm supplies a variety of producer services, such as accountancy, advertising, banking/finance, insurance, law, and management consultancy through offices in worldwide cities [19]. Although a weighted network was already available from the raw data where high weights were assigned to links between strong firms and highly populated cities (they tend to offer more services in larger cities), here we consider the unweighted network representation. The computation of the cover set in this bipartite network, where $V_\top$ represents the firms, shows that a very small set of firms (8%) offers services in all 315 cities. This shows that these firms have a prominent role in controlling the socio-economic development in the world. Additionally, each firm is able to establish its offered services separately and satisfies the structural controllability assumption for DS.

The cond-mat scientific collaboration consists of a network of scientists and research papers. For example, two scientists are considered connected to the same paper if they coauthored the paper together. The network corresponds to a subset of the Los Alamos Archive, covering the condensed matter physics (cond-mat) research field, including 16,726 scientists and 22,016 papers [12]. Here, the structural controllability assumption is reasonable because each scientist can choose to investigate each research subject independently. Therefore, the scientists are the $V_\top$ set in this network.

The cover size is not large and shows that 25% of scientists may induce research opinions and new scientific routes by leading and participating in all of the research conducted in

the field. It also should noted that in this case, the $V_\perp$ set of nodes does not follow a clear enough power-law but instead shows faster exponential decay (see Fig. S1 for the degree distributions). As shown in Table 1, the cover size tends to grow with the degree exponent γ_2. Because the cond-mat network undergoes rapid exponential decay, its equivalent degree exponent would be relatively large leading to a higher percentage for the cover set.

The Davis' Southern Women Club dataset was collected by Davis and colleague in the 1930s [20]. It contains the observed attendance at 14 social events by 18 Southern women. In this case, although the size of the network is small, both degree distribution seems more compatible with a random network with exponential decay rather than with a scale-free network (see Fig. S4). In this network, only two women dominate the entire network.

The network corresponding to the bibliography of the book *Graph Products* by Imrich and Klavzar also was investigated [21]. This network represents an author-by-paper bipartite network, in which edges e_{ij} indicate that an author i is the co-author of the paper j. While the $V_\top$ set of nodes (authors) follows a power-law, the $V_\perp$ follows a faster exponential decay. Here, a bit more than half of the authors (174) are necessary to cover all of the papers cited in the bibliography.

Finally, we analysed the network data on the administrative elite in The Netherlands in April 2006 [22]. Here, individuals are connected to administrative bodies through memberships. This topology is the opposite case of the bibliography network discussed above. Here, the $V_\top$ (persons) network decays exponentially, and the $V_\perp$ follows a clear power-law distribution showing that few key administrative bodies have memberships for many persons. In this case, only 14% of the persons are necessary to cover all of the administrations.

The rest of the social networks show asymmetric behavior characterised by a power-law for $V_\top$ and faster exponential decay for $V_\perp$. Therefore, we performed extensive computer simulations, which allowed us to investigate with higher accuracy the cover size in the more suited large-scale computer-generated scale-free networks, i.e., the networks with up to 100,000 nodes (see Table 1). The lower bound is shown in Table 1 and it is always smaller than the observed cover size.

C. Biological networks

On the biological side, we analysed four bipartite networks, including a transcriptional regulatory network, ncRNA-protein interactions and the human drug-target protein network.

The transcriptional network can be represented as a bipartite network with sets of nodes as transcriptional factors (TFs) and regulated (or controlled) genes. A direct interaction between these molecules indicates that a TFs regulates a gene. We analysed the *S. cerevisiae* transcriptional network [23] and computed the cover set of this network, which represents approximately 57% of the TFs. The structure of this network resembles that of the non-coding RNA (ncRNA)-protein interaction network, where only the degree of $V_\top$ nodes (TFs or ncRNAs) follows a power-law and the degree of regulated molecules in $V_\perp$ decays approximately exponentially (see Fig. S3). A non-coding molecule is a transcribed molecule that is not translated into a protein. Recent studies have shown that non-coding molecules may play relevant roles in the genetic regulatory level [24]. The cover set analysis of the human ncRNA-protein network reveals that 71% of the ncRNAs are needed to dominate the whole network. This ratio does not change significantly (68%) when we combine six organisms (*E. coli, S. cerevisiae, C. elegans, D. melanogaster, M. musculus and H. sapiens*) within the network to enlarge the dataset. All of these datasets were downloaded from NPInter Database [25].

Because of its importance and potential implications, a more extensive statistical analysis is performed on the drug-target network. Network medicine is emerging as a new framework to not only identify novel disease gene candidates and distinct disease-specific functional modules but also elucidate the deeper roots of a human disease or complex disorder [26]. The recent works have offered rich and promising findings by not only including the analysis the genes and molecules corresponding to a single disease but also considering the whole set of human disease genes (*disease genome*) and their associated disorders (*disease phenome)* [27]. The analysis of the *diseasome* network showed that distant and, in principle, different diseases share an unexpectedly high number of disease genes, suggesting that apparently distinct pathophenotypes are more linked at the molecular level than expected by common medical knowledge. This finding not only opens new avenues to understand the deeper origins of complex diseases but could also aid drug discovery because newly developed drugs could be aimed for targeting the disease gene-products that molecularly link distinct

disorders.

In this bipartite network of drug-protein interactions, a drug and a protein are connected to each other if the protein is a known target of the drug. Here, the $V_{\top}$ set of nodes is represented by the drugs that can alter the activity of the targeted protein. The background data corresponds to the DrugBank database and the same dataset shown in [26, 28] was used to construct the drug-target network with 888 approved drugs that target 394 human proteins in total.

A small fraction of validated disease genes encodes the drug-target proteins. The drug targets were assigned a human disorder class if the protein was a disease-gene product. Each gene was assigned to a disorder class as shown in SI, Table S3 in [27]. This information is available in the OMIM database, which reports on human disorders and disease-related genes. The target proteins encoded by disease genes are coloured based on the disorder class to which they belong. A complete map of the giant component of the bipartite network with a mapping of the identified dominating set of drugs is shown in Figure 4. Figure S5 shows the isolated components of the network. To satisfy the controllability assumption for DS, we have assumed that each drug is designed to interact with specific targets and that these interactions are independent to some extent. The computation of the cover size shows that only 21% of the approved drugs could control the entire known druggable proteome.

The network features of the relationships between all drugs and drug targets are important for organising the current knowledge of the relationships between the drug targets and disease-gene products and even then human therapies [26, 28, 37, 38]. This kind of network representation could aid drug discovery because newly developed drugs could target the disease-gene products that molecularly link the roots of distinct disorders [27]. Here, we raise questions on the controllability of the human drug-target protein (DT) network, the system's minimum number of *driver drugs* as well as their topological role in the network.

III. RESULTS OF THE DATA ANALYSIS

A. Cover set of bipartite drug-target protein network targets high-degree drugs

The computation of the cover size showed that only 21% of the drugs could control the entire known druggable proteome. Based on the linkage of the genes to disparate disease

pathophenotypes [27] and the reduced size of *driver drugs* identified, we suggest that a relatively small number of the drugs could address the common genetic origin of these diseases. Morever, if we consider only the giant connected component, the fraction of drugs required to control the target proteins is significantly reduced to 8%. Although the average degree of the drugs is 1.81, the drugs cover set shows $< k >$=2.2, showing that the DS targets, on average, the more highly connected drugs. Furthermore, this value increases to 3.59 when only the giant component is considered. Unfortunately, most of the current drugs are not etiology-specific, and therefore, they do not act on the actual disease-associated protein. In contrast, most of the drugs that we use are palliative and tend to act on the proteins in the network neighbourhood of the disease genes [28].

The analysis of the drug-target interactions indicates that high-degree drugs tend to belong to the dominating set (see Fig. S6). However, this analysis also shows that the DS does not exclusively include hubs. Low-degree drugs also play an important role in the control of the whole network (see Fig. 1 and Fig. S5).

B. Disease-gene products are inhomogeneously targeted by the cover set of drugs

A classification of disorders based on the fraction of the disease-gene products targeted by the cover set is shown in Figure S7. A high fraction of the proteins that belong to dermatological, neurological and psychiatric disorders tend to be targeted by the 12 most highly connected drugs in the DS. In contrast, cancer, immunology and renal disorders tend to be targeted by the low-degree drugs in the DS.

C. Drug projection of the bipartite drug-target network

To shed light on the topological role of the cover set in the drug-target network, we projected the giant component of the network onto the drug space. In the resulting drug network, two drugs are connected if they target at least one common protein. Here, we first mapped the identified cover set of drugs (DS) onto this unipartite network (see Fig. S8).

D. The DS of drugs hits bridged locations in the network

Next, we computed various topological metrics, such as the closeness centrality, the betweenness centrality, the degree correlation and the node degree. The results are again mapped in the drug network (See Figs. S9-S12). The computation of the average values of these metrics for the full network and for the set of nodes that belong to the DS indicates that there is no specific correlation between the cover set of drugs and their degree correlation and closeness centrality in the projected network (see Fig. S13). However, this computation clearly reveals that the cover set tends to select the nodes with a high betweenness centrality. A similar correlation is shown in Figure S14, where the degree correlations and the closeness and betweenness centralities are plotted against the degree. Although the set of drugs in the DS and the rest of drugs tend to show a similar pattern for the degree correlation and the closeness centrality, the distribution shows significantly differences when the betweenness function is considered, showing that drugs with a high betweenness only belong to the DS (see Fig. S14). These drugs contribute to a higher average betweenness centrality (see Fig. S13). It also can be seen that nodes belonging to the DS (see Fig. S10) play the role of bridges, connecting the groups of drugs in the unipartite network.

E. Drugs in the DS are also among the most influential spreaders

The previous finding connecting the driver nodes with a high betweenness centrality also suggests that the topological features shown by the cover set and the so-called influential spreaders could be similar [29]. It was shown that many shortest path crossings through one node (high betweenness) indicate a high spreading capability if the node does not exists at the end of a branch at the periphery of the network or lower shells. The node location can then be classified using the k-shell decomposition method. The analysis assigns an integer index k_s to each node. A high value indicates a high shell; i.e., the node is in the core of the network [30]. A low value indicates a low shell; i.e., the node occupies peripheral locations.

Our findings then suggest that the dominating sets in networks can not only control the entire network but could also be the most influential spreaders of information in a network. Because the DS represents drugs in this case, it could be thought that the DS represents the most influential nodes for eradicating the disorders or perturbations across the network. To

verify this unexpected finding in more detail, we also performed a k-shell decomposition of the network. The result shown in Fig. S15 indicates that many of the most highly connected drugs in the DS occupy the core (or higher shells) of the network. In contrast, fewer drugs are located in the periphery (or lower shells) of the network. As shown in [29], the nodes in the higher shells of several social networks tend to have higher betweenness values.

F. Implications of dominating set of drugs for network medicine

The application of structural controllability to social and biological systems could be of significant importance. The control of neuronal signals and transcription factor concentrations could revolutionise neurology, molecular biology and medicine. Despite the breathtaking advances in our molecular understanding of cellular processes and new drug targets, the number of drugs that receive approval by the US Food and Drug Administration (FDA) is significantly decreasing each year. The screening for single-target drugs is in part responsible for this slow development [31, 32]. While the genes that are linked to two distant diseases were thought to be important to understand the deep roots of complex disorders [27], here we have found that the minimum set of drugs controlling disease-gene products share some unique properties, which could be used to develop future drugs. For example, the drugs that belong to the DS occupy core locations in the network such that they bridge multiple disease-gene products, with many shortest paths crossing between these drugs, showing that these drugs have specific chemical features for treating distant disorders.

This property should be examined in detail because it may aid in the design of influential drugs that disable multiple network pathways in distant parts of the network that are responsible for different disorders with a few drugs.

Use of the network framework in medicine has become a necessity because diseases are not caused by single genes but by larger disease-specific functional modules and pathways. Multi-target and optimal combinations of drugs in cancer and HIV have been suggested as new courses of action to address with such complex disease modularity and to aid in drug design [33–35]. The drawbacks could be in the possible antagonistic drug combinations, where the strength of two drugs in the same treatment is weaker than that of either drug alone in addition to the unwanted side effects [36].

In this work, although we have assumed that no interactions exist between drugs in the

drug-target network, it would be possible to adapt the MDS algorithm to include constraints such that certain drugs cannot appear in the MDS at the same time. With respect to clinical applications, a recent work has already explored the optimal drug combination using a minimal hitting set algorithm, which is equivalent to cover set, that successfully targets the whole population of 60 tumour-derived cell lines, uncovering 14 anticancer drug combinations [35]. Our approach also could be applied to future drug-patient networks in the context of personalised medicine [37, 38]. Interestingly, given the significant amount of data related to genomics and medicine, it is expected that systems medicine data integration may lead to novel and larger bipartite networks, the controllability of which could be predicted by the presented methodology.

IV. METHODS

In this section, we explain in detail the derived mathematical tools used to study controllability in bipartite networks. For the sake of clarity, we introduce once more the main notation and the proposition presented in the main text. We first introduce the dynamic model for our Minimum Dominating Set approach (MDS) on bipartite networks.

Next, we explain in detail the analytical results for structural controllability in bipartite networks and how the results are linked to the controllability of networks in the sense of [1]. Then, we then make predictions on the minimum number of nodes necessary to control the entire bipartite network for the cases $\gamma_1 < 2$ and $\gamma_1 > 2$. Finally, we perform a theoretical analysis for the case of random bipartite networks.

A. Dynamics model

Let $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_{n_2}(t))^T$ be the state of nodes in $V_\perp$ at time t, where M^T denotes the transposed matrix of M.

Suppose that $U = \{u_1, u_2, \ldots, u_h\}$ be the set of driver nodes (in our sense) selected from

$V_\top$ and each u_i has d_i edges. We define a state vector $\mathbf{u}$ for all edges from U by

$$\begin{aligned}\mathbf{u}(t) \;=\; & (u_{1,1}(t), u_{1,2}(t), \ldots, u_{1,d_1}(t),\\ & u_{2,1}(t), u_{2,2}(t), \ldots, u_{2,d_2}(t),\\ & \ldots,\\ & u_{h,1}(t), u_{h,2}(t), \ldots, u_{h,d_h}(t))^T.\end{aligned}$$

We rename $u_{i,j}$s by $u'_1, u'_2, \ldots, u'_l$ and and rewrite $\mathbf{u}(t)$ by

$$\mathbf{u}(t) = (u'_1(t), u'_2(t), \ldots, u'_l(t))^T,$$

where each node u'_i has only one outgoing edge that corresponds to the original edge.

Here, we assume that the dynamics is given by

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) + B\mathbf{u}(t)$$

where A is the diagonal matrix (i.e., $A_{i,j} = 0$ for all $i \neq j$) and B satisfies that $B_{i,j} \neq 0$ only if u'_j is connected to w_i. It is to be noted that A can be the null matrix (i.e., any node in $V_\perp$ does not have a self-loop).

Then, it is clear that this system is structurally controllable if each node $w \in V_\perp$ has at least one incoming edge (i.e., U is a dominating set).

Furthermore, there can be connections between nodes in $V_\perp$ because addtion of non-zero elements to A does not impair the structural controllability. However, in this case, the number of required driver nodes may be fewer than that derived from our model.

In the above, we assumed that $u_i \notin V_\top - U$ does not have any effect on $v_j \in V_\perp$. However, this assumption can be removed if we can know the signals from nodes in $V_\top - U$ to $V_\perp$. Suppose that v_j has incoming edges from $u_{i_0}, u_{i_1}, \ldots, u_{i_k}$ where $u_{i_0} \in U$ and $u_{i_1}, \ldots, u_{i_k} \notin U$ (this case can be trivially extended for the case where there exist edges from multiple nodes in U). Let the state vector from these nodes to v_j be $(u_0^j(t), u_1^j(t), \ldots, u_k^j(t))^T$ and the vector of corresponding weights be $(b_0, b_1, \ldots, b_k)^T$. Then, in order to remove the effects from $u_{i_1}, \ldots, u_{i_k}$ to v_j, it is enough to add the following term to $u_0^j(t)$:

$$-\frac{1}{b_0}(b_1 u_1^j(t) + \cdots b_k u_k^j(t)).$$

It is worth mentioning that apart from dynamics, the MDS approach has control applications in several areas, ranging from engineering to systems biology, including clinical studies focused on uncovering anticancer drug combinations to target specific tumour cells [35].

Additionally, the control aspects of the problem follow directly from the fact that a controlled edge applying a unique signal to a single integrator node is controllable. Consequently, structural controllability is not even needed.

B. Dominating set and structural controllability in bipartite networks

We consider a bipartite graph $G(V_\top, V_\bot; E)$, where $V_\top$ is a set of top nodes, $V_\bot$ is a set of bottom nodes, and E is a set of edges ($E \subseteq V_\top \times V_\bot$). It should be noted that the directions of all edges are from $V_\top$ to $V_\bot$ in this definition. This assumption is reasonable for networks such as drug-target networks because the activities of nodes in $V_\top$ are usually not affected by those in $V_\bot$. We discuss later in a different section the controllability of bi-directional bipartite networks, which require a specific analysis.

In this work, we use a modified version of the dominating set, in which a set must be selected from $V_\top$ and it is sufficient to dominate all nodes in $V_\bot$ (i.e., for all node $w \in V_\bot$, there exists a node $v \in V_\top$ such that $(v, w) \in E$). This corresponds to a *set cover* problem by associating a set $S_v = \{w|(v, w) \in E\}$ for each $v \in V_\top$. We use MDS to denote the minimum dominating set (i.e., the dominating set with the minimum number of nodes) in the sense described above.

As proved in [5, 6], a network is structurally controllable if a dominating set is selected as a set of control nodes under the assumption that each control node can control its outgoing edges separately.

We also can consider structural controllability under the assumption in [1] that each driver node can control only its own value. Figure 2 (*i*) shows a bipartite graph $G(V_\top, V_\bot; E)$ that can be transformed into $G'(V^L, V^R; E')$. To contruct G', for each node $v \in V_\top$, we create two nodes v^L and v^R. Similarly, for each node $v \in V_\bot$, we create two nodes v^R and v^L. For each edge $(u, v) \in E$, we create an edge between u^L and v^R. In such a case, the number of driver nodes is determined by the number of nodes in V^R not appearing in a maximum matching of the adjunct bipartite graph $G'(V^L, V^R; E')$ [1]. However, in this case, all nodes in V^R corresponding to $V_\top$ remain unmatched because there is no edge that connects any of these nodes (see Fig. 2 (*ii*)). Therefore, we have

Proposition 1 *The number of driver nodes in the sense of [1] is at least* $|V_\top|$ *for a bipartite network* $G(V_\top, V_\bot; E)$ *such that* $E \subseteq V_\top \times V_\bot$.

Because $|V_\top|$ is usually a very large number, in this study, we focus on the structural controllability in terms of MDS.

Note that $\min(|V_\top|, |V_\perp|)$ driver nodes are still required even if we only aspire to control the nodes in $V_\perp$ because there is no outgoing edge from $V_\perp$.

C. Theoretical analysis of the MDS size in bipartite networks

First, we describe a relationship between DS and the structural controllability of a bipartite network. As a direct consequence of Theorem 1 in [5], it can be seen that the following proposition holds.

Proposition 2 *Suppose that we need to control the states of nodes only in $V_\perp$ and that every node in the DS ($\subseteq V_\top$) can control all of its outgoing links separately. Then, the network is structurally controllable by selecting the nodes in the DS as the driver nodes.*

Here we assume that the degree distributions of $V_\top$ and $V_\perp$ follow $P_\top(k) \propto k^{-\gamma_1}$ and $P_\perp(k) \propto k^{-\gamma_2}$, respectively. We let $n_1 = |V_\top|$ and $n_2 = |V_\perp|$.

We assume that all of the nodes in a dominating set DS must be selected from $V_\top$ and that it is necessary to dominate all nodes in $V_\perp$ (we need not dominate the nodes in $V_\top$), which means that DS is a *set cover* for $V_\perp$. In [5, 6], structural controllability was studied in terms of MDS for unipartite graphs. In what follows, we present the analytically derived predictions for the minimum number of drivers using the MDS controllability approach for bipartite networks by considering two cases, $\gamma_1 > 2$ and $\gamma_1 < 2$.

Here, we note that in our theoretical analysis we use a standard mean-field approach which assumes a continuum approximation, in which the degree k becomes a continuous real variable. This approach is often used in the field of complex networks to evaluate mathematical models and obtain relatively easily analytical expressions [39–44]. However, it should also be noted that this approximation may not give the exact value for the normalising constant α. The exact normalising constant α considering a discrete degree distribution can be obtained from $(\zeta(\gamma, 1) - \zeta(\gamma, k_{max}))^{-1}$ [9, 44]. In this study, we discuss the order of number of driver nodes, not the exact number of driver nodes. Because the value of normalising constant α does not affect the order of n, we did not use the exact value of α. Moreover,

as shown in [44], in the tail of the distribution the sum over k is well approximated by an integral so that the normalising constant can be written as $\alpha \simeq \frac{1}{\int_{k_{min}}^{\infty} k^{-\gamma}} = (\gamma - 1){k_{min}}^{\gamma-1}$.

1. Case of $\gamma_1 > 2$

We assume $P_\top(k)$ follows $\alpha_1 k^{-\gamma_1}$ with cut off at $k = n_1$, where $\gamma_1 > 2$. From $\alpha_1 n_1 \int_1^{n_1} k^{-\gamma} dk = n_1$, we have $\alpha_1 \approx \gamma_1 - 1$.

For $S \subseteq V_\top$, $\Gamma(S)$ denotes the set of edges between S and $V_\perp$ (i.e., $\Gamma(S) = \{(u, v) \mid u \in S \text{ and } v \in V_\perp\}$). The following property is trivial

$$\text{if } |\Gamma(S)| < n_2, \quad S \text{ can not dominate } V_\perp.$$

Let S be the set of nodes whose degree is greater than or equal to K. It is to be noted that S is chosen so that the total degree (i.e., the number of edges incident to S) is maximized among the sets with the same cardinality.

We estimate the size of $\Gamma(S)$ as follows.

$$\begin{aligned} |\Gamma(S)| &< \alpha n_1 \int_K^{n_1} k \cdot k^{-\gamma_1} dk \approx n_1(\gamma_1 - 1) \int_K^{n_1} k^{-\gamma_1+1} dk \\ &= n_1 \cdot \left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \left(\frac{1}{K^{\gamma_1-2}} - \frac{1}{n_1^{\gamma_1-2}}\right) < n_1 \cdot \left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \frac{1}{K^{\gamma_1-2}}. \end{aligned} \tag{10}$$

If S is a dominating set, the last term should be no less than n_2. Therefore, the following inequality should be satisfied:

$$n_1 \cdot \left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \frac{1}{K^{\gamma_1-2}} \geq n_2. \tag{11}$$

By solving this inequality, we have

$$K \leq \left[\left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right) \cdot \left(\frac{n_1}{n_2}\right)\right]^{1/(\gamma_1-2)}. \tag{12}$$

Then, the size of S is estimated as

$$\begin{aligned} |S| &\approx \alpha n_1 \int_K^{n_1} k^{-\gamma_1} dk \approx n_1 \left(\frac{1}{K^{\gamma_1-1}} - \frac{1}{{n_1}^{\gamma_1-1}}\right) \approx n_1 \cdot \frac{1}{K^{\gamma_1-1}} \\ &> \left[\left(\frac{\gamma_1 - 1}{\gamma_1 - 2}\right)\right]^{-\frac{\gamma_1-1}{\gamma_1-2}} \cdot \left(\frac{n_2}{n_1}\right)^{\frac{\gamma_1-1}{\gamma_1-2}} \cdot n_1. \end{aligned} \tag{13}$$

From this inequality and the fact that $V_\top$ is a trivial dominating set, we can see that the size of the minimum dominating set is $\Theta(n_1)$ (for fixed γ_1 if n_2 is the same order as n_1). As a

difference from the case of $\gamma_1 < 2$, the maximum degree does not largely influence the above lower bound. If the maximum degree is H, it is enough to simply replace $\frac{1}{n_1^{\gamma_1-2}}$ with $\frac{1}{H^{\gamma_1-2}}$ and $\frac{1}{n_1^{\gamma_1-1}}$ with $\frac{1}{H^{\gamma_1-1}}$ in Eqs. (10) and (13), respectively which would result in negligible changes.

2. Case of $\gamma_1 < 2$

In this section, we focus on degree distribution for $V_\top$ and thus we let $\gamma = \gamma_1$, $n = n_1$, and $m = n_2$.

We assume that the maximum degree is H. Then, we have

$$n \;=\; \alpha n \int_1^H k^{-\gamma} dk \;=\; \frac{\alpha n}{1-\gamma}\left[k^{-\gamma+1}\right]_1^H \;=\; \frac{\alpha n}{\gamma-1}(1-H^{1-\gamma}) \;\approx\; \frac{\alpha n}{\gamma-1} \tag{14}$$

from which $\alpha = \gamma - 1$ follows.

Let DS be the set of nodes with degree between B and H. Then, the number of nodes N_{DS} in DS is given by

$$N_{DS} \;=\; \alpha n \int_B^H k^{-\gamma} dk \;=\; \frac{\alpha}{-\gamma+1} \cdot n \cdot \left[k^{-\gamma+1}\right]_B^H \;=\; n\left(B^{1-\gamma} - H^{1-\gamma}\right)$$
$$=\; O(nB^{1-\gamma}). \tag{15}$$

On the other hand, the total number of edges E_G is

$$E_G \;=\; \alpha n \int_1^H k \cdot k^{-\gamma} dk \;=\; \frac{\gamma-1}{2-\gamma} \cdot n \cdot \left[k^{-\gamma+2}\right]_1^H \;=\; \frac{\gamma-1}{2-\gamma} \cdot n \cdot (H^{2-\gamma}-1) \;\approx\; \frac{\gamma-1}{2-\gamma} \cdot nH^{2-\gamma}$$
$$=\; < k > n, \tag{16}$$

from which $< k > \approx \frac{\gamma-1}{2-\gamma} \cdot H^{2-\gamma}$ follows.

The number of edges E_{NDS} not covered by DS is

$$E_{NDS} \;=\; \alpha n \int_1^B k \cdot k^{-\gamma} dk \;=\; \frac{\gamma-1}{2-\gamma} \cdot n \cdot (B^{2-\gamma}-1) \;\approx\; \frac{\gamma-1}{2-\gamma} \cdot n \cdot B^{2-\gamma} \tag{17}$$

Therefore, the probability that an arbitrary edge is not covered by DS is

$$\frac{E_{NDS}}{E_G} \;\approx\; \left(\frac{B}{H}\right)^{2-\gamma} \tag{18}$$

Let $V_\perp \ominus DS$ denote the set of nodes in $V_\perp$ that are not dominated by DS. Since a node is dominated by DS if at least one edge connecting to the node is covered by DS, the expected number of nodes (denoted by $N_{V_\perp \ominus DS}$) of $V_\perp \ominus DS$ is bounded as

$$N_{V_\perp \ominus DS} \le O\left(m \cdot \left(\frac{B}{H}\right)^{2-\gamma}\right), \tag{19}$$

where m is the number of nodes in $V_\perp$.

In order to dominate $V_\perp \ominus DS$, it is enough to select at most $N_{V_\perp \ominus DS}$ nodes from $V_\top$. Therefore, the size of a minimum dominating set is bounded by

$$|DS| + N_{V_\perp \ominus DS} \ \leq \ O\left(nB^{1-\gamma} + m\left(\frac{B}{H}\right)^{2-\gamma}\right). \tag{20}$$

Then, in order to find B minimizing this order, we let

$$nB^{1-\gamma} = m\left(\frac{B}{H}\right)^{2-\gamma}, \tag{21}$$

which results in $B = \frac{n}{m} \cdot H^{2-\gamma}$.

Therefore, an upper bound of the size of the dominating set is estimated as

$$O\left(\frac{n^{2-\gamma} \cdot m^{\gamma-1}}{H^{(2-\gamma)(\gamma-1)}}\right). \tag{22}$$

It is to be noted that $(2-\gamma)(\gamma-1) \leq 0.25$.

By using $< k > \approx \frac{\gamma-1}{2-\gamma} \cdot H^{2-\gamma}$, this upper bound can also be written as

$$O\left(\frac{n^{2-\gamma} \cdot m^{\gamma-1}}{< k >^{(\gamma-1)} \left(\frac{2-\gamma}{\gamma-1}\right)^{\gamma-1}}\right). \tag{23}$$

On the other hand, if $H = n$, the upper bound becomes

$$O\left(n^{(2-\gamma)^2} \cdot m^{\gamma-1}\right). \tag{24}$$

If $m = cn$ where c is a constant, this order is $O(n^{\gamma^2-3\gamma+3})$, which takes the minimum order $(O(n^{0.75}))$ when $\gamma = 1.5$. Note that this upper bound might not be accurate when γ is close to 1 or 2 because $\gamma - 1$ and $2 - \gamma$ appear as denominators in some equations.

Readers might wonder why γ_2 is ignored in the above analysis. We briefly discuss this point. The number of nodes with degree 1 or 2 in $V_\perp$ is approximated by

$$\alpha_2 n_2 \int_1^2 k^{-\gamma} dk \ = \ \frac{\alpha_2 n_2}{\gamma_2 - 1}\left(1 - \frac{1}{2^{\gamma_2 - 1}}\right). \tag{25}$$

Since we have $\alpha_2 = \gamma_2 - 1$ as in the case of $V_\top$, this number is equal to

$$n_2\left(1 - \frac{1}{2^{\gamma_2 - 1}}\right). \tag{26}$$

Therefore, a constant fraction of nodes in $V_\perp$ have degree 1 or 2 if γ_2 is a constant. Here, we note that in the analysis of the MDS size, we only used $V_\perp$, a property that every node

in $V_\perp$ has a degree of at least 1. If most nodes in $V_\top$ were of a high degree, the size of MDS would be much significantly less. However, as mentioned above, a constant fraction of nodes in $V_\perp$ have degree 1 or 2; thus, we can only expect the reduction of a constant factor (i.e., not an exponent) of the MDS size, even if we make extensive use of γ_2.

The finite size of the networks limits the maximum degree value for each node. Some real networks show faster exponential decay in the tail that can be mathematically represented by an exponential cutoff as $k^{-\gamma}e^{-k/k_c}$ [40, 42]. However, this is not the case of all the real-world networks and some networks conserve a well-defined power-law decay until the very end of the tail and the cuffoff is simply the maximum degree (see also [43]). Here, we have considered the effects of the maximum degree H in the MDS, as showed in Figure. 3. We leave as a future work the computation of the effects of the exponential cutoff in the MDS.

D. Theoretical analysis of the MDS size in random bipartite networks

In this section, we consider random bipartite networks, that is, we assume that edges are randomly drawn between $V_\top$ and $V_\perp$ where $n_1 = V_\top$ and $n_2 = V_\perp$. Let $\langle k \rangle$ be the average degree of nodes in $V_\top$, which means that there exist $\langle k \rangle n_1$ edges. Recall that $S \subseteq V_\top$ is a dominating set for $V_\perp$ if $(\forall w \in V_\perp)(\exists v \in S)((v, w) \in E)$ holds.

We use a well-known result on the *coupon collector* problem [45], which states that we can collect all n kinds of coupons with probability greater than $1 - (1/n)$ if we randomly draw $2n \ln n$ coupons. In our case, since we are assuming that edges are drawn randomly, nodes in $V_\perp$ correspond to the kinds of coupons and each edge corresponds to a coupon. Therefore, if $S \in V_\top$ has $2n_2 \ln n_2$ edges, it is highly expected that S is a dominating set for $V_\perp$.

First we estimate an upper bound of the size of DS. Since $\langle k \rangle$ is the average degree for any integer $C > 0$, there must exist $S \subseteq V_\top$ such that $|\Gamma(S)| \geq \langle k \rangle |S|$ and $|S| = C$ (otherwise, the average degree would be smaller than $\langle k \rangle$). As discussed above, if $|\Gamma(S)| > 2n_2 \ln n_2$, S is a dominating set with high probability. Therefore, an upper bound of the size of DS is estimated as

$$\frac{2n_2 \ln n_2}{\langle k \rangle}. \tag{27}$$

Next we estimate a lower bound of the size of DS. We show that the degree of any node in $V_\top$ does not become much larger than the average. Since edges are randomly drawn, we

can regard the assignment process of edges to each node $v \in V_\top$ as a Poisson process with the success probability (per trial) $1/n_1$. Then, from *Chernoff bound* (see Thm. 4.4 in [45]), we have

$$Prob(deg(v) \ \geq \ 6\langle k \rangle) \ \leq \ 2^{-6\langle k \rangle}. \tag{28}$$

Here we consider two cases: $\langle k \rangle \geq \log_2 n_1$ and $\langle k \rangle < \log_2 n_1$. First we consider the former case. Since there exist n_1 nodes, the probability that at least one node has degree more than $6\langle k \rangle$ is bounded by

$$n_1 \cdot 2^{-6\langle k \rangle}, \tag{29}$$

which is a negligibly small probability for sufficiently large n_1 because $\langle k \rangle \geq \log_2 n_1$. Therefore, $|\Gamma(S)| \leq 6\langle k \rangle |S|$ holds for any $S \subseteq V_\top$ with high probability. Since $|\Gamma(S)|$ should be no less than $2n_2 \ln n_2$, a lower bound of the size of DS is estimated as

$$\frac{2n_2 \ln n_2}{6\langle k \rangle} = \frac{n_2 \ln n_2}{3\langle k \rangle}. \tag{30}$$

Therefore, the size of DS is estimated as $\Theta(\frac{n_2 \ln n_2}{\langle k \rangle})$ for $\langle k \rangle \geq \log_2 n_1$. Next we consider the case of $\langle k \rangle \ < \ \log_2 n_1$. Using *Chernoff bound* again, we have

$$Prob(deg(v) \ \geq \ 2\langle k \rangle \log_2 n_1) \ \leq \ 2^{-2\langle k \rangle \log_2 n_1} \ = \ \frac{1}{(n_1)^{2\langle k \rangle}} \tag{31}$$

and thus the probability that at least one node has degree more than $2\langle k \rangle \log_2 n_1$ is bounded by

$$n_1 \cdot 2^{-2\langle k \rangle \log_2 n_1} \ = \ \frac{n_1}{(n_1)^{2\langle k \rangle}} \ = \ \frac{1}{(n_1)^{2\langle k \rangle - 1}}, \tag{32}$$

which is a negligibly small probability for sufficiently large n_1 (assuming that $\langle k \rangle \geq 1$). Then, as in the former case, a lower bound of the size of DS is estimated as

$$\frac{2n_2 \ln n_2}{2\langle k \rangle \log_2 n_1} = \frac{n_2 \ln n_2}{\langle k \rangle \log_2 n_1}, \tag{33}$$

which matches the upper bound except a factor of $\log_2 n_1$.

E. Computational results

We consider the case of $\gamma_1 < 2$, and we evaluate the relationship between H and the cover size under the condition that $\gamma_1 = \gamma_2$ and $|V_\top| = |V_\perp| \approx 20,000$, The simulation results shown in Figure 3 show that the maximum degree of nodes in the network has a significant impact on the cover size.

By observing Figure 3, we also note that the exponent takes its minimum value at $\gamma = 1.5$, which is in good agreement with our theoretical result. The reason for this is as follows: As shown above, an upper bound on the size of the dominating set is estimated as

$$O\left(\frac{n^{2-\gamma} \cdot m^{\gamma-1}}{H^{(2-\gamma)(\gamma-1)}}\right). \tag{34}$$

It should be noted that $(2-\gamma)(\gamma-1) \leq 0.25$.

From here, and by taking $\gamma = \gamma_1$, it can be seen that the dependence of the minimum number of drivers on H (maximum degree) is $1/(H^{(2-\gamma_1)(\gamma_1-1)})$. When $\gamma_1 = 1.5$, the exponent of H becomes -0.25, which is a minimum value. From Figure 3, we see that for γ_1=1.3,1.5,1.6 the exponents are -0.20, -0.21, -0.18, respectively and among them γ_1=1.5 corresponds to the exponent with the minimum value. A comparison of the analytically predicted H exponent and that calculated with computer simulations is shown in Figure S16. Because Figure 3 only considers the case $\gamma_1 < 2$, we also performed computations for the case of $\gamma_1 > 2$, as shown in Figure S17. In this case, there is almost no dependence on the maximum degree H. This different behaviour contrasts with the scaling-law observed in Figure 3 for $\gamma_1 < 2$. As discussed in SI IV-C, our analytical results indicate that the dependence on H is $\frac{1}{H^{\gamma_1-2}}$, which vanishes for large H and $\gamma_1 > 2$, in agreement with the computer simulations.

Here, we should note that we have analytically computed an upper bound; therefore the exact value of γ_1 attaining the minimum size MDS may differ from 1.5, although it will be in the vicinity of this value. This shows that our finding is far from trivial and that the computation of the exact value remains as future work.

In Figure S18, we investigate the effect of the size of $|V_\perp|$ on the controllability of bipartite networks. The relationship between the cover size and n_2 under the condition that $H = 100$ also shows a scaling-law for a variety of degree exponents $\gamma_1 = 1.1$ to $\gamma_1 = 1.9$ in fair agreement with the theoretical predictions, except for γ_1=1.1, where the observed exponent 0.45 for $\gamma_1 = 1.1$ is significantly larger than the theoretically estimated exponent (0.1).

However, it can be observed that for $m > 5000$, the exponent is smaller than 0.45 and is closer to 0.10. If γ_1 is close to 1 and m is small, $B = \frac{n}{m} \cdot H^{2-\gamma_1}$ might be larger than H. Because we assumed in our theoretical analysis that B is no larger than H, such a large B might lead to a non-accurate estimate.

Finally, we also compared the cover set computed for the scale-free networks with that from the random networks obeying the Poisson degree distribution. We generated sale-free

networks with a variety of degree exponents as shown in Table S1. The results are discussed in the main text.

V. CONTROLLABILITY OF BIDIRECTIONAL BIPARTITE NETWORKS

In this section, we briefly discuss on the case of controllability on bidirectional bipartite networks. We focused on bipartite networks in which all of the edges are directed from one side of the nodes ($V_\top$) to the other side of nodes ($V_\perp$); thus, the driver nodes can be selected only from $V_\top$. However, other types of bipartite graphs exist. One typical example is the metabolic network. In this network, there exist two kinds of nodes [46, 47], including nodes that correspond to chemical reactions and nodes that correspond to chemical compounds. In this case, it also is reasonable to assume that only the values (activities) of chemical reactions can be controlled because the activities of the chemical reactions might be modified by controlling the concentrations of the corresponding enzymes via knockout or overexpression of the genes, whereas it seems difficult to directly control the concentrations of chemical compounds within a cell. However, edges exist in both directions; hence, we require a smaller number of driver nodes. In the example shown in Figure S19, the transformed bipartite network has a complete matching, and thus, we require only one driver node. However, if we assume that all of the edges are directed from left to right, then there exist three unmatched nodes in the transformed network, and hence, we need three driver nodes. Therefore, the structural controllability in bipartite networks strongly depends on the edge directions.

The efficiency (the small number of driver nodes in non-heterogeneous networks) of nodal dynamics relies on the existence of many long paths. However, in unidirectional bipartite networks, all paths are of length 1 (or length 2 if self-loops are allowed in $V_\perp$). Therefore, nodal dynamics cannot be effectively applied to this kind of networks. However, nodal dynamic may be effectively applied to the control of bidirectional networks, such as metabolic networks.

The theoretical and simulation analyses of the bidirectional bipartite networks are expected to be more complicated than those of unidirectional bipartite networks because we must consider four degree distributions (the indegree and outdegree of $V_\top$; the indegree and

outdegree of $V_\perp$). Hence, we leave this analysis for future work.

---

[1] Liu, Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. *Nature* **473**, 167-173 (2011).

[2] Lin, C. Structural controllability. *IEEE Trans. Automat. Contr.* **19**, 201-208 (1974).

[3] Slotine, J.-J. & Li, W. *Applied Nonlinear Control* (Prentice-Hall, 1991).

[4] Nepusz, T. & Vicsek, T. Controlling edge dynamics in complex networks. *Nature Physics* **8**, 568-573 (2012).

[5] Nacher, J.C. & Akutsu, T. Dominating scale-free networks with variable scaling exponent: heterogeneous networks are not difficult to control. *New Journal of Physics* **14**, 073005 (2012).

[6] Nacher, J.C. & Akutsu, T. Analysis on controlling complex networks based on dominating sets. *Journal of Physics: Conf. Ser.* **410**, 012104 (2013).

[7] Haynes, T.W., Hedetniemi, S.T. & Slater P.J. *Fundamentals of Domination in graphs* (Chapman and Hall/CRC Pure Applied Mathematics, New York, 1998).

[8] Newman, M.E.J. Power laws, pareto distributions and Zipf's law. *Contemporary Physics* **46**, 323-351 (2005).

[9] Clauset, A., Shalizi, C.R. & Newman, M.E.J. Power-law distributions in empirical data. *SIAM Review* **51**(4), 661-703 (2009).

[10] *http://tuvalu.santafe.edu/ aaronc/powerlaws*

[11] Opsahl, T. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Social Networks* **34**, doi:10.1016/j.socnet.2011.07.001. (2012)

[12] Newman, M.E.J. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA* **98**, 404-409 (2001).

[13] Newman, M.E.J., Strogatz, S.H. & Watts, D.J. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118 (2001).

[14] Guillaume, J.-L. & Latapy M. Bipartite graphs as models of complex networks. *Physica A* **471**, 795-813 (2004)

[15] Battiston, S. & Catanzaro, M. Statistical properties of corporate board and director networks. *Eur. Phys. J. B* **38**, 345-352 (2004).

[16] Caldarelli, G. & Catanzaro, M. The corporate boards networks. *Physica A* **338**, 98-106 (2004).

[17] Saavedra, S., Reed-Tshochas F. & Uzzi, B. A simple model for bipartite cooperation for ecological and organization networks. *Nature* **457**, 463-466 (2009).

[18] Nacher, J.C. & Akutsu T. On the degree distribution of projected networks mapped from bipartite networks. *Physica A* **390**, 4636-4651 (2011).

[19] Taylor, P.J. *World city network: a global urban analysis* (London: Routledge, 2000).

[20] Wasserman, S. & Faust, K. *Social Network Analysis* (Cambridge University Press, 1994).

[21] Imrich, W. & Klavzar, S. *Products Graphs: Structure and Recognition* (Wiley, New York, 1999).

[22] *http://vlado.fmf/uni-lj.si/pub/networks/data*

[23] Costanzo, M.C. et al., YPDTM, PombePDTM and WormPDTM: model organisms volumes of BioKnowledgeTM Library, an integrated resource for protein information. *Nucleic Acids Research* **29**, 75-79 (2001).

[24] Amaral, P.P., Dinger, M.E., Mercer, T.R. & Mattick, J.S. The eukaryotic geneome as an RNA machine. *Science* **319**, 1787-1789 (2008).

[25] Wu, T. et al., NPInter: the noncoding RNAs and protein related biomacromolecules interaction database. *Nucleic Acids Research* **34**, D150-D152 (2006).

[26] Barabási, A.-L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics* **12**, 56-68 (2011).

[27] Goh, K-I, et al., The human disease network. *Proc. Natl. Acad. Sci. USA* **104**, 865-8690 (2007).

[28] Yildirim, M.A. et al., Drug-target network. *Nature Biotech.* **25**, 1119-1126 (2007).

[29] Kitsak, M. et al., Identification of influential spreaders in complex networks. *Nature Physics* **6**, 888-893 (2010).

[30] Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y. & Shir, E. A model of Internet topology using k-shell decomposition. *Proc. Natl. Acad. Sci. USA* **104**, 11150-11154 (2007).

[31] Nolan, G.P. What's wrong with drug screening today. *Nat. Chem. Biol.* **3**, 187-191 (2007).

[32] Hopkins, A.L. Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* **4**, 682-690 (2008).

[33] Yang, K. et al., Finding multiple target optimal intervention in disease related molecular networks. *Mol. Syst. Biol.* **4**, 228 (2008).

[34] Csermely, P., Agoston, V. & Pongor, S. The efficiency of multi-target drugs: the network

approach might help drug design. *Trends Pharmacol. Sci.* **26**, 178-182 (2005).

[35] Vázquez, A. Optimal drug combinations and minimal hitting sets. *BMC Systems Biology* **3**, 81 (2009).

[36] Motter, A.E. Improved network performance *via* antagonism: From synthetic rescues to multi-drug combinations. *Bioessays* **32**, 236-245 (2010).

[37] Spiro, Z., Kovacs I.A., & Csermely P. Drug therapy networks and the prediction of novel drug targets. *Journal of Biology* **7**, 20 (2008).

[38] Nacher, J.C. & Schwartz J.-M. A global view of drug-therapy interactions. *BMC Pharmacology* **8**, 5 (2008).

[39] Barabási, A.-L., Albert, R. & Jeong, H. Mean-field theory for scale-free random networks. *Physica A*, **272**, 173-187 (1999).

[40] Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Reviews of Modern Physics* **74**, 47-97 (2002).

[41] Dorogovtsev, S.N. & Mendes J.F.F. Evolution of networks. *Adv. Phys.* **51**, 1079-1187 (2002).

[42] Boguñá, M., Pastor-Satorras, R. & Vespignani, A. Cut-offs and finite size effects in scale-free networks. *The European Physical Journal B* **38**, 205-209 (2004).

[43] Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. Resilience of the Internet to random networks. *Physical Review Letters* **85**, 4626 (2000).

[44] Newman, M.E.J. *Networks: An introduction* (Oxford Univ. Press, 2011).

[45] Mitzenmacher, M. & Upfal, E. *Probability and Computing. Randomized Algorithms and Probabilistic Analysis* (Cambridge Univ. Press, 2005).

[46] Smart, A.G., Amaral, L.A.N & Ottino, J.M. Cascading failure and robustness in metabolic networks, *Proc. Natl. Acad. Sci. USA* **105**, 13223-13228 (2008).

[47] Takemoto, K., Tamura, T., Cong, Y., Ching, W.-K., Vert, J.-P. & Akutsu, T. Analysis of the impact degree distribution in metabolic networks using branching process approximation. *Physica A* **391**, 379-387 (2012).

Supplementary Figure captions and Tables

Fig. S1 The cumulative degree distributions of bipartite social networks. The distribution for the top (bottom) set of nodes is shown in the left (right) figures. (a-b) Firms-world cities network with the degre exponents $\gamma_1 = 3.27 \pm 0.36$ and $\gamma_2 = 2.40 \pm 0.31$. (c-d) Facebook-like forum users-topics with the degree exponents $\gamma_1 = 3.25 \pm 0.22$ and $\gamma_2 = 2.98 \pm 0.29$. (e-f) Cond-mat Scientific-collaboration network with the degree exponent $\gamma_1 = 3.49 \pm 0.02$. The distribution for the bottom set of nodes shows exponential decay. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound on power-law behavior x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

Fig. S2 The cumulative degree distributions of the human drug-target protein network. (a-b) The Drug-Target network with degre exponents $\gamma_1 = 2.96 \pm 0.23$ and $\gamma_2 = 1.84 \pm 0.08$. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound of the power-law behavior x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

Fig. S3 The cumulative degree distributions of bipartite biological networks. The distribution for the top (bottom) set of nodes is shown in the left (right) figures. (a) The degree of the top set of nodes (TFs) for the transcriptional network for *S. cerevisiae* follows a power-law with the degre exponent $\gamma_1 = 2.45 \pm 0.38$. The degree of the top set of nodes (ncRNAs) for the ncRNA-protein network for (c) human and (e) six organisms follows a power-law with degree exponents $\gamma_1 = 2.04 \pm 0.46$ and $\gamma_1 = 1.97 \pm 0.31$, respectively. (b,d,f) The bottom set of nodes shows a faster exponential decay. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound of the power-law behavior x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

Fig. S4 The cumulative degree distributions of bipartite social networks. The distribution for the top (bottom) set of nodes is shown in the left (right) figures. (a) The degree of

the top set of nodes (authors) follows a power-law with the degree exponent $\gamma_1 = 2.31 \pm 0.10$. (b) The degree of the bottom set of nodes (papers) shows exponential decay. (c) The degree of the Dutch Elite (persons) decays exponentially. (d) The degree of the bottom set of nodes (administrative bodies) scales as a power-law with degree exponent $\gamma_2 = 3.33 \pm 0.16$. (e-f) The degree distributions of (e) the top nodes (Southern women) and (f) the bottom nodes (events) tend to decrease faster than a power-law, although the network size is small. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound of the power-law behaviour x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

Fig. S5 The small isolated components in the drug-target protein network. Drugs (hexagons) belonging to the DS are indicated in red. The interactions between DS drugs and target proteins (circles) are denoted by wavy red arrows. Colored nodes have the same meaning as shown in Fig. 4.

Fig. S6 The fraction of drugs in the DS with degree k computed in the bipartite drug-target protein network. The drugs with high degree tend to belong to the DS.

Fig. S7 The statistics of the twelve most highly connected drugs in the DS. The histogram shows an inhomogenous distribution because the proteins belonging to specific disorder classes, such as neurological or psychiatric disorders, generally can be controlled by highly connected drugs in the DS. In contrast, other gene-products belonging to complex disorders, such as cancer, immunology or renal disorders, can be controlled by weakly connected drugs in the DS, which exhibits specialised mechanisms of drug action. Disorder classes are the same as shown in the label of Fig. 4.

Fig. S8 The projection of the drug network constructed from the giant component of the human bipartite drug-target protein network. Each node corresponds to a drug and two drugs are connected to each other if they share at least one target protein. The names of the drugs in the DS with highest degree in the bipartite network are highlighted (red (highest) and orange (high-average) nodes). Green nodes represent drugs in the DS with low connectivity in the bipartite network. Blue nodes denote drugs that not belong to the

DS.

Fig. S9 The closeness centrality strength for each drug is indicated by colors in the projected drug network.

Fig. S10 The betweeness centrality strength for each drug is indicated by colors in the projected drug network. Note that nodes belonging to the DS (see Fig. S6) exhibit high betweenness centrality and occupy bridged locations in the network.

Fig. S11 The degree correlation (average neighboor degree) for each drug is indicated by colors in the projected drug network.

Fig. S12 The node degree for each drug is indicated by colors in the projected drug network.

Fig. S13 The histogram with the average values of the analysed centrality measures for drugs belonging (green) and not belonging (red) to the DS computed in the projected drug network. From left to right, average degree, average neighborhood (i.e., average number of links of the adjacent nodes), average betweeness and average closeness values.

Fig. S14 The functions of the neighborhood degree, closeness and betweenness centralities versus node degree in the projected network. Triangles indicate drugs belonging to the DS. Circles show the values for drug nodes that not belong to the DS. Drugs in DS shows higher deviations for the betweenness centrality if compared with drugs that not belong to the DS.

Fig. S15 the k-shell decomposition of the drug network. Names of the ten drugs in the DS with highest degree computed in the bipartite network are highlighted. Each color denotes a shell and shell number k_s is denoted in figure legend. The analysis includes drugs in the isolated components, which are shown in white color.

Fig. S16 The analytically predicted H exponent and the one calculated using computer

simulations. The error bars, shown in the figure, are in some cases smaller than the symbols.

Fig. S17 The dependence between the cover size and H under the condition that $\gamma_1 = \gamma_2 > 2$ and $|V_\top| = |V_\bot| \approx 20,000$ nodes. The error bars are smaller than symbols.

Fig. S18 The dependence between the cover size and m under the condition that $H = 100$, where n_2 is controlled by varying γ_2. Note that $m = n_2$ and it corresponds to the number of nodes of $V_\bot$, and $V_\top$=20,000 nodes. The fitted functions are $m^{(0.453\pm0.017)}$ ($r = 0.9863$) for $\gamma_1 = 1.1$, $m^{(0.560\pm0.003)}$ ($r = 0.9937$) for γ_1=1.3, $m^{(0.706\pm0.022)}$ ($r = 0.9930$) for $\gamma_1 = 1.5$, $m^{(0.798\pm0.012)}$ ($r = 0.9987$) for γ_1=1.7 and $m^{(0.878\pm0.016)}$ ($r = 0.9990$) for $\gamma_1 = 1.9$. Statistical errors of the exponents are shown together with correlation coefficients r in parentheses. The error bars are smaller than symbols.

Fig. S19 The relationship between a bidirectional bipartite network (left) and its transformed network (right). In this example, $G'(V^L, V^R; E')$ has a complete matching and thus we need only one driver node.

TABLE I: Computational results of cover set size for computer generated poisson bipartite networks (RAND) and scale-free networks (SF) with the same average degree. The network size is n_1=n_2=100 and $H = 90$. The results were averaged over 10 realizations. The standard error of the mean (s.e.m.) is also shown.

Type	γ_1	γ_2	Edges	Cover size	s.e.m.(+/-)
SF	1.2	1.2	261	31	4
RAND			261	26	1
SF	1.5	1.5	121	42	2
RAND			121	75	3
SF	1.8	1.8	191	52	2
RAND			191	54	2
SF	2.2	2.2	155	63	2
RAND			155	63	5
SF	3.0	3.0	191	82	1
RAND			191	54	2

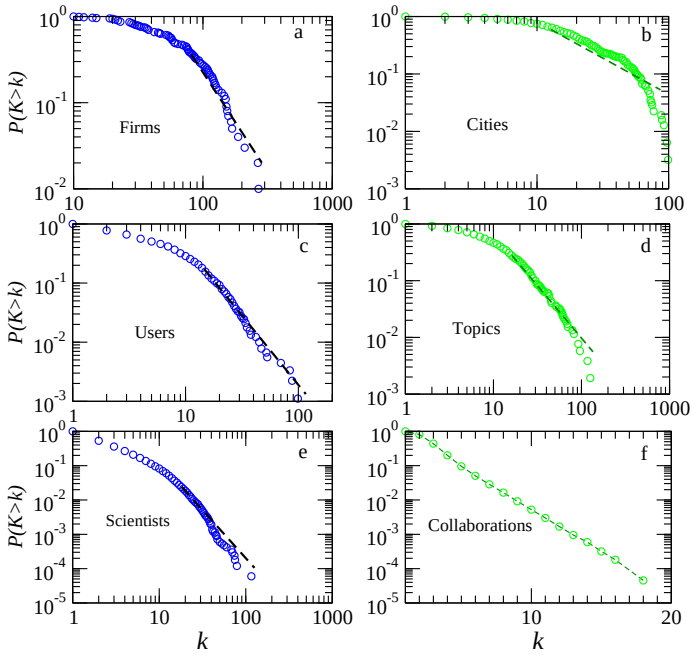

FIG. 1: The cumulative degree distributions of bipartite social networks. The distribution for the top (bottom) set of nodes is shown in the left (right) figures. (a-b) Firms-world cities network with the degre exponents $\gamma_1 = 3.27 \pm 0.36$ and $\gamma_2 = 2.40 \pm 0.31$. (c-d) Facebook-like forum users-topics with the degree exponents $\gamma_1 = 3.25 \pm 0.22$ and $\gamma_2 = 2.98 \pm 0.29$. (e-f) Cond-mat Scientific-collaboration network with the degree exponent $\gamma_1 = 3.49 \pm 0.02$. The distribution for the bottom set of nodes shows exponential decay. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound of the power-law behavior x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

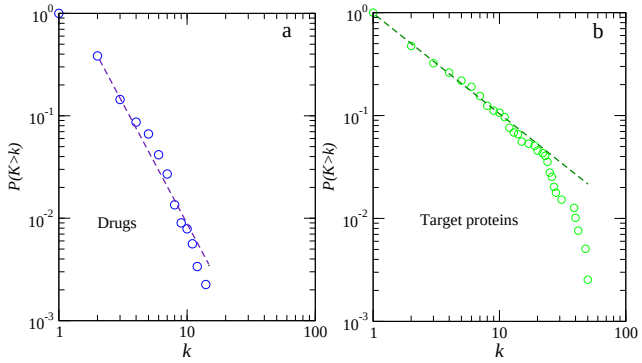

FIG. 2: The cumulative degree distributions of the human drug-target protein network. (a-b) The Drug-Target network with the degre exponents $\gamma_1 = 2.96 \pm 0.23$ and $\gamma_2 = 1.84 \pm 0.08$. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound of the power-law behavior x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

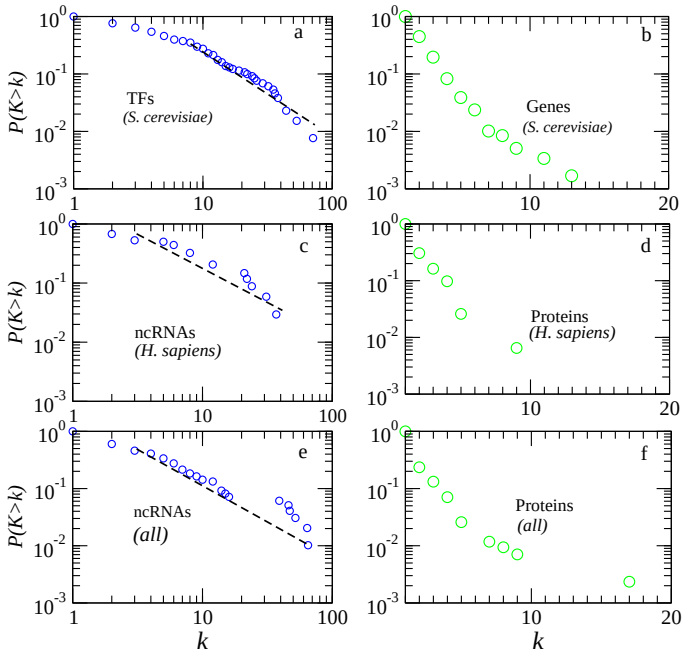

FIG. 3: The cumulative degree distributions of bipartite biological networks. The distribution for the top (bottom) set of nodes is shown in the left (right) figures. (a) The degree of the top set of nodes (TFs) for the transcriptional network for *S. cerevisiae* follows a power-law with the degre exponent $\gamma_1 = 2.45 \pm 0.38$. The degree of the top set of nodes (ncRNAs) for the ncRNA-protein network for (c) human and (e) six organisms follows a power-law with degree exponents $\gamma_1 = 2.04 \pm 0.46$ and $\gamma_1 = 1.97 \pm 0.31$, respectively. (b,d,f) The bottom set of nodes shows a faster exponential decay. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound of the power-law behavior x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

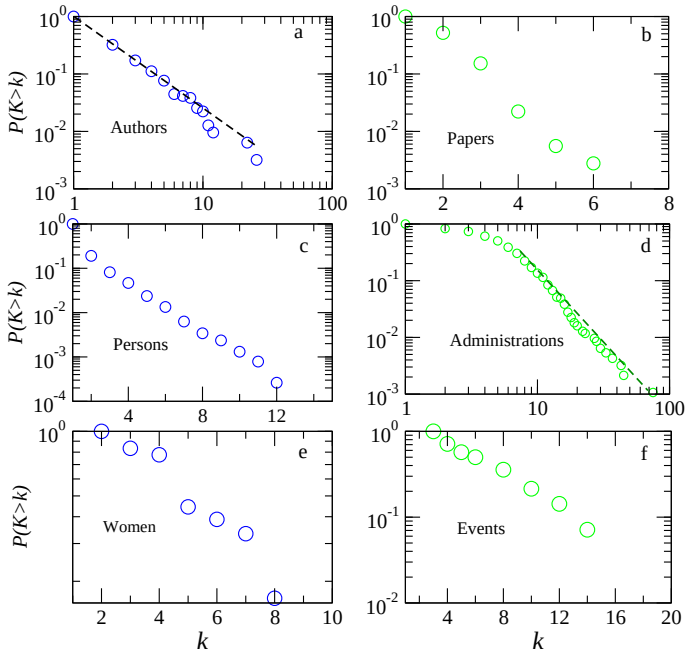

FIG. 4: The cumulative degree distributions of bipartite social networks. The distribution for the top (bottom) set of nodes is shown in the left (right) figures. (a) The degree of the top set of nodes (authors) follows a power-law with the degree exponent $\gamma_1 = 2.31 \pm 0.10$. (b) The degree of the bottom set of nodes (papers) shows exponential decay. (c) The degree of the Dutch Elite (persons) decays exponentially. (d) The degree of the bottom set of nodes (administrative bodies) scales as a power-law with degree exponent $\gamma_2 = 3.33 \pm 0.16$. (e-f) The degree distributions of (e) the top nodes (Southern women) and (f) the bottom nodes (events) tend to decrease faster than a power-law, although the network size is small. The standard errors of γ are derived from the width of the maximum likelihood method as shown in SI Section IIA. The lower bound of the power-law behaviour x_{min} is the starting point of each fit (dashed-line). The exact values and the errors for each network are shown in Table 2.

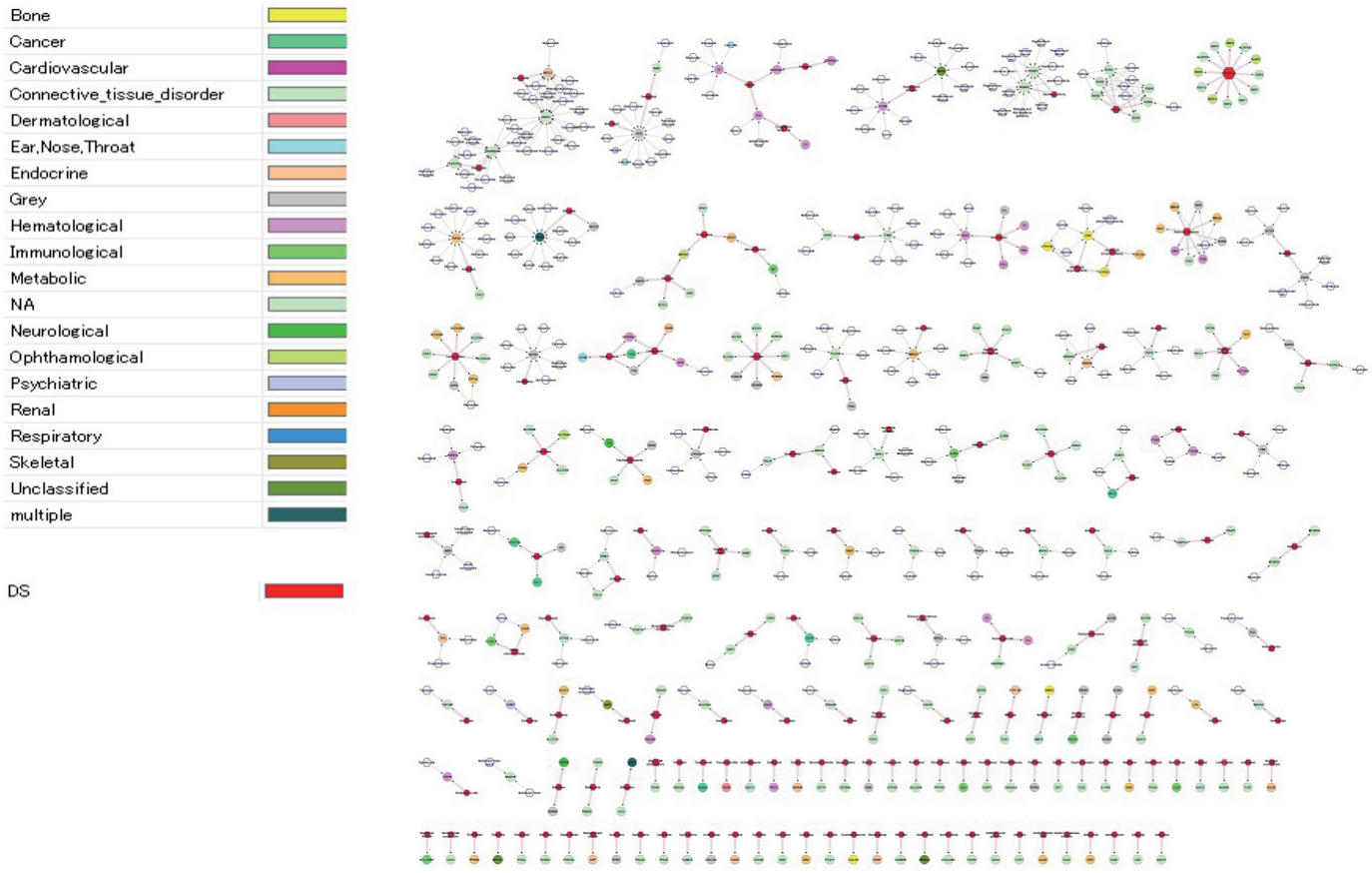

FIG. 5: The small isolated components in the drug-target protein network. Drugs (hexagons) belonging to the DS are indicated in red. The interactions between DS drugs and targets (circles) are denoted by wavy red arrows. Colored nodes have the same meaning as shown in Fig. 4.

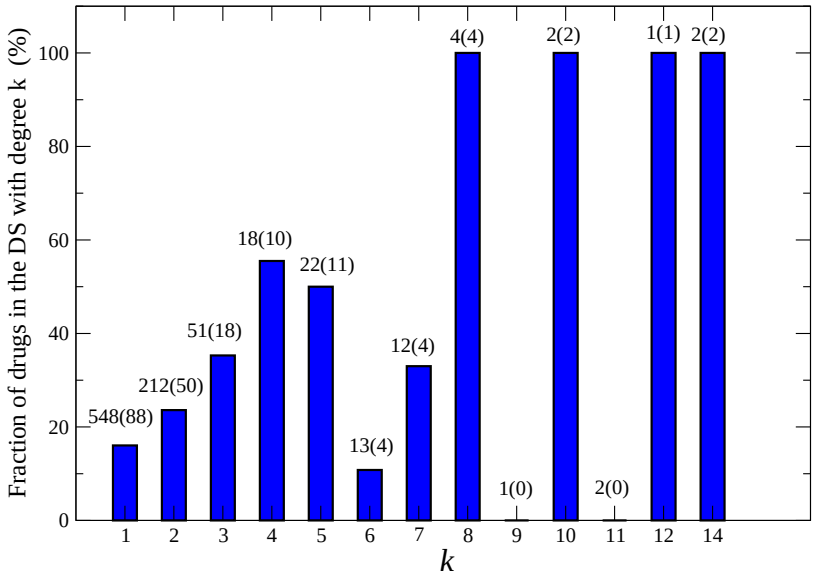

FIG. 6: The fraction of drugs in the DS with degree k computed in the bipartite drug-target protein network. The drugs with high degrees tend to belong to the DS.

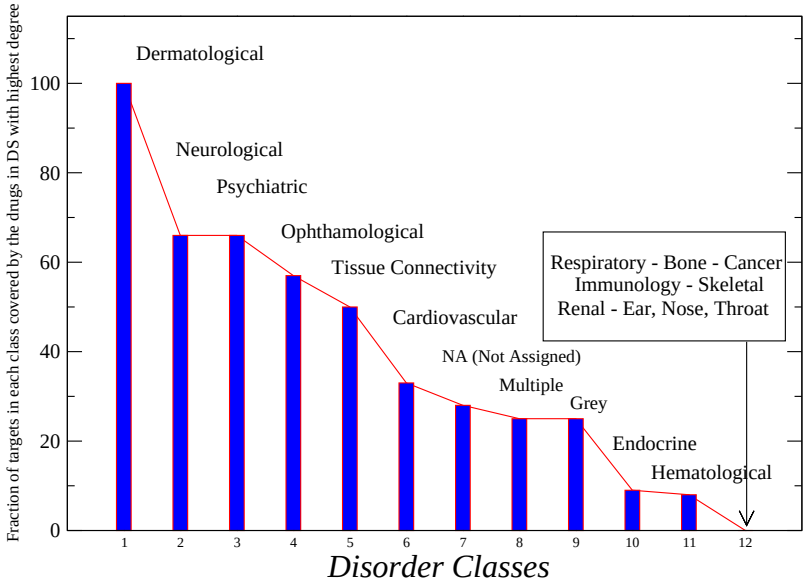

FIG. 7: The statistics of the twelve most highly connected drugs in the DS. The histogram shows an inhomogenous distribution because the proteins belonging to specific disorder classes, such as neurological or psychiatric disorders, generally can be controlled by highly connected drugs in the DS. In contrast, other gene-products belonging to complex disorders, such as cancer, immunology or renal disorders, can be controlled by weakly connected drugs in the DS, which exhibits specialised mechanisms of drug action. Disorder classes are the same as shown in the label of Fig. 4.

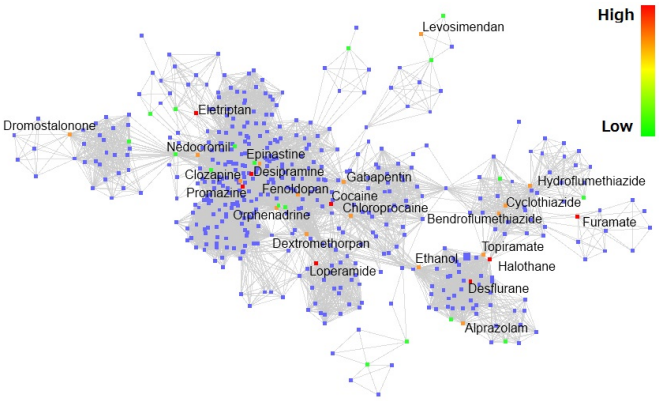

FIG. 8: The projection of the drug network constructed from the giant component of the human bipartite drug-target protein network. Each node corresponds to a drug and two drugs are connected to each other if they share at least one target protein. The names of the drugs in the DS with highest degree in the bipartite network are highlighted (red (highest) and orange (high-average) nodes). Green nodes represent drugs in the DS with low connectivity in the bipartite network. Blue nodes denote drugs that not belong to the DS.

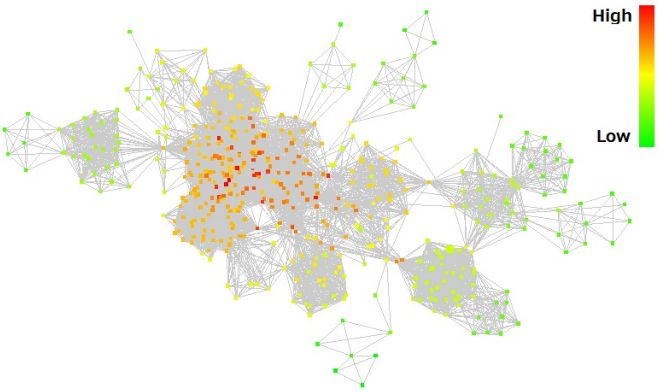

FIG. 9: The closeness centrality strength for each drug is indicated by colors in the projected drug network.

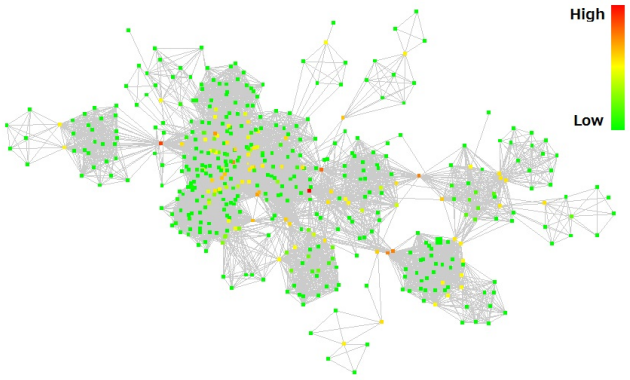

FIG. 10: The betweeness centrality strength for each drug is indicated by colors in the projected drug network. Note that nodes belonging to the DS (see Fig. S6) exhibit high betweenness centrality and occupy bridged locations in the network.

FIG. 11: The degree correlation (average neighboor degree) for each drug is indicated by colors in the projected drug network.

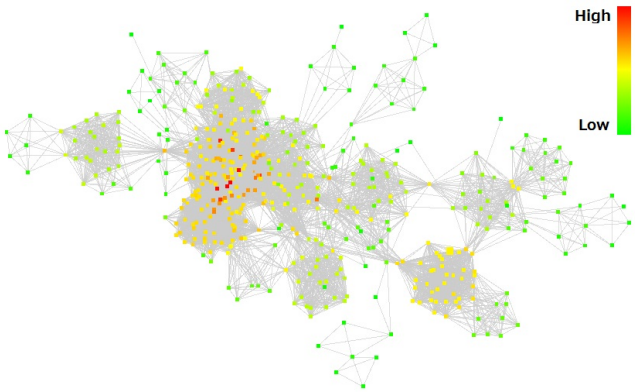

FIG. 12: The node degree for each drug is indicated by colors in the projected drug network.

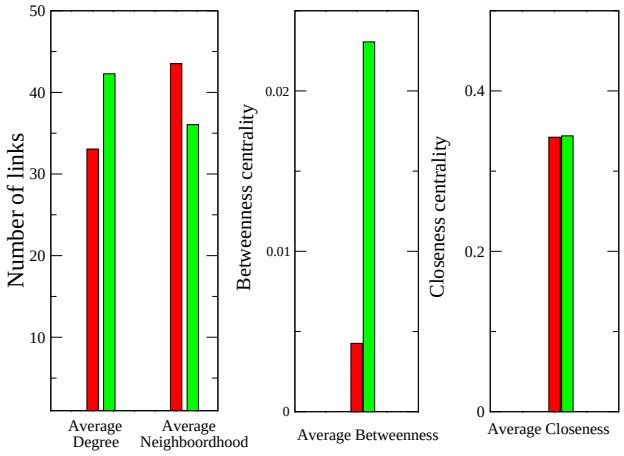

FIG. 13: The histogram with the average values of the analysed centrality measures for drugs belonging (green) and not belonging (red) to the DS computed in the projected drug network. From left to right, average degree, average neighborhood (i.e., average number of links of the adjacent nodes), average betweeness and average closeness values.

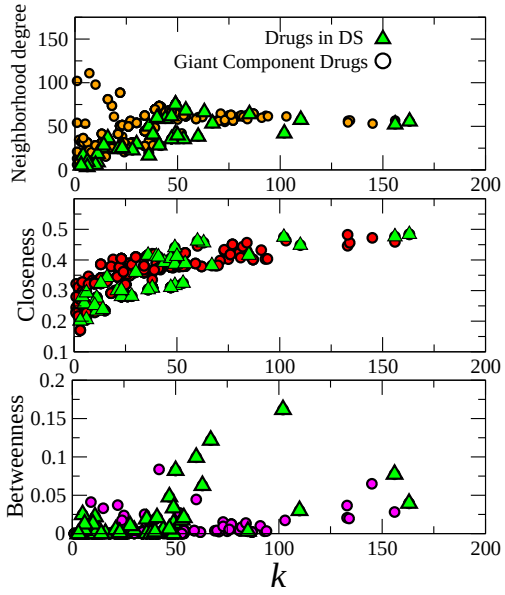

FIG. 14: The functions of the neighborhood degree, closeness and betweenness centralities versus node degree in the projected network. Triangles indicate drugs belonging to the DS. Circles show the values for drug nodes that not belong to the DS. Drugs in DS shows higher deviations for the betweenness centrality if compared with drugs that not belong to the DS.

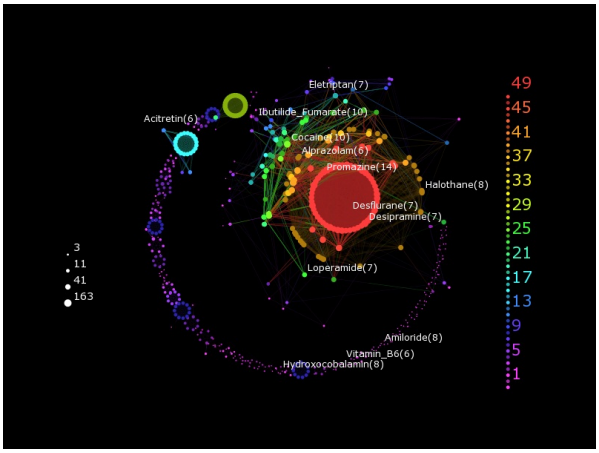

FIG. 15: The k-shell decomposition of the drug network. Names of the ten drugs in the DS with highest degree computed in the bipartite network are highlighted. Each color denotes a shell and shell number k_s is denoted in figure legend. The analysis includes the drugs in the isolated components, which are shown in white color.

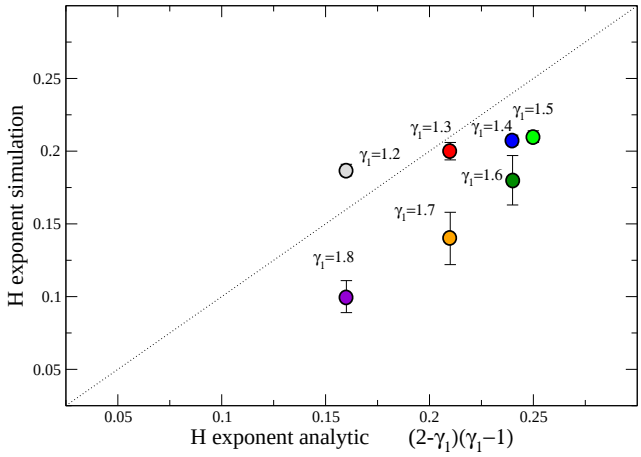

FIG. 16: The analytically predicted H exponent and the one calculated using computer simulations. The error bars, shown in the figure, are in some cases smaller than the symbols.

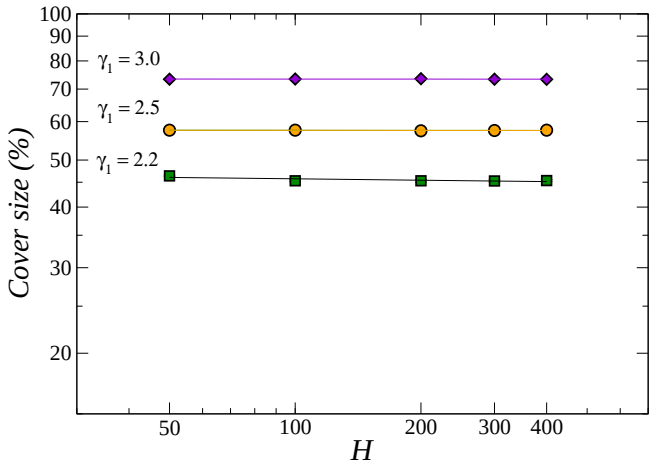

FIG. 17: The dependence between the cover size and H under the condition that $\gamma_1 = \gamma_2 > 2$ and $|V_\top| = |V_\perp| \approx 20,000$ nodes. The error bars are smaller than symbols.

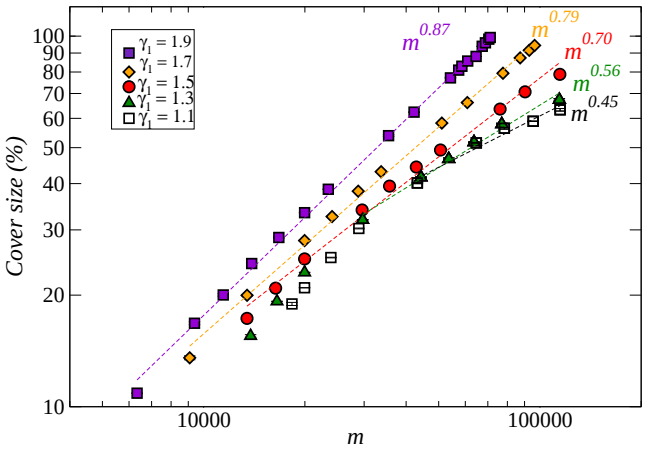

FIG. 18: The dependence between the cover size and m under the condition that $H = 100$, where n_2 is controlled by varying γ_2. Note that $m = n_2$ and it corresponds to the number of nodes of $V_\perp$, and $V_\top$=20,000 nodes. The fitted functions are $m^{(0.453\pm0.017)}$ ($r = 0.9863$) for $\gamma_1 = 1.1$, $m^{(0.560\pm0.003)}$ ($r = 0.9937$) for γ_1=1.3, $m^{(0.706\pm0.022)}$ ($r = 0.9930$) for $\gamma_1 = 1.5$, $m^{(0.798\pm0.012)}$ ($r = 0.9987$) for γ_1=1.7 and $m^{(0.878\pm0.016)}$ ($r = 0.9990$) for $\gamma_1 = 1.9$. Statistical errors of the exponents are shown together with correlation coefficients r in parentheses. The error bars are smaller than symbols.

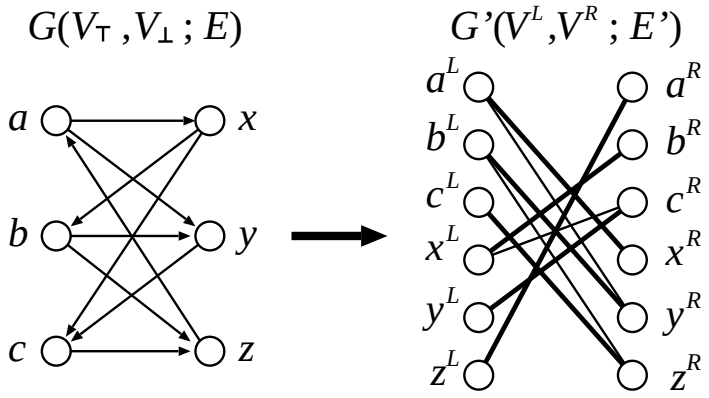

FIG. 19: The relationship between a bidirectional bipartite network (left) and its transformed network (right). In this example, $G'(V^L, V^R; E')$ has a complete matching and thus we need only one driver node.