

## Network quotients: Structural skeletons of complex systems

Yanghua Xiao,<sup>1</sup> Ben D. MacArthur,<sup>2</sup> Hui Wang,<sup>3</sup> Momiao Xiong,<sup>4,5</sup> and Wei Wang<sup>1</sup>

<sup>1</sup>*Department of Computing and Information Technology, Fudan University, Shanghai 200433, People's Republic of China*

<sup>2</sup>*School of Mathematics, University of Southampton, Southampton, SO17 1BJ, United Kingdom*

<sup>3</sup>*Business School, University of Shanghai for Science and Technology, Shanghai 200093, People's Republic of China*

<sup>4</sup>*Theoretical Systems Biology Lab, School of Life Science, Fudan University, Shanghai 200433, People's Republic of China*

<sup>5</sup>*Human Genetics Center, University of Texas Health Science Center at Houston, Houston Texas 77225, USA*

(Received 17 March 2008; published 3 October 2008)

A defining feature of many large empirical networks is their intrinsic complexity. However, many networks also contain a large degree of structural repetition. An immediate question then arises: can we characterize essential network complexity while excluding structural redundancy? In this article we utilize inherent network symmetry to collapse all redundant information from a network, resulting in a coarse graining which we show to carry the essential structural information of the “parent” network. In the context of algebraic combinatorics, this coarse-graining is known as the “quotient.” We systematically explore the theoretical properties of network quotients and summarize key statistics of a variety of “real-world” quotients with respect to those of their parent networks. In particular, we find that quotients can be substantially smaller than their parent networks yet typically preserve various key functional properties such as complexity (heterogeneity and hub vertices) and communication (diameter and mean geodesic distance), suggesting that quotients constitute the essential structural skeletons of their parent networks. We summarize with a discussion of potential uses of quotients in analysis of biological regulatory networks and ways in which using quotients can reduce the computational complexity of network algorithms.

DOI: [10.1103/PhysRevE.78.046102](https://doi.org/10.1103/PhysRevE.78.046102)

PACS number(s): 89.75.Fb, 05.40.-a, 02.20.-a

### I. INTRODUCTION

Many physical systems—from the world-wide web to scientific collaborations and biochemical reactions inside cells—can be modeled as networks. The ubiquity of empirical networks has generated increasing interest in their study over the last decade during which much progress has been made toward elucidating general network organizational principles beyond the specific details of individual systems [1–6]. Structural properties which are commonly found in many disparate networks include: the “small-world” property [5], the scale-free distribution of vertex degrees [6], hierarchical modularity [7], network construction from motifs [8], assortative mixing [9], and self-similarity [10] amongst others. Together, investigation of generic structural properties such as these may be thought of as an attempt to understand network complexity [11].

In order to find simplicity in this complexity some authors have attempted to extract network “skeletons:” related networks which capture essential structural features of the system from which they are derived, but are simpler in some quantitative way. Existing network skeletons include for instance, the fractal skeleton [12], which is responsible for fractal scaling, the communication skeleton [13], which is responsible for the majority of communication flow through the network, and the modularity skeleton [14], which reduces the size of the network while preserving the value of modularity. Such skeletons are generally formed with respect to a given property, for example, fractality, communication, or modularity, and thus do not represent a structural skeleton in the strongest sense. In this article we propose an alternative skeleton which formally captures all essential structural information, and which can be significantly smaller than the

original network from which it is derived. The method we use is based upon utilizing inherent network symmetry.

Although almost all large random networks are asymmetric [15], many empirical real-world networks are surprisingly richly symmetric [16–19]. This symmetry commonly results from the presence of locally treelike or bicliquelike structures [17,19] which are present in many empirical networks, and derive naturally from elementary growth processes such as growth with preferential attachment [17] and growth with similar linkage pattern [19]. However, despite a rich abstract theory of graph symmetry [20–23], the symmetry structure of complex real-world networks has not yet been explored extensively.

Intuitively, a network is symmetric if two or more of its vertices can be permuted without altering vertex adjacency. Symmetric networks therefore necessarily contain a certain degree structural redundancy in that they possess multiple vertices which play the same structural role. Thus network symmetry is strongly related to network redundancy. In this paper we use this relationship to show how symmetry also provides a natural means to formally exclude redundancy while still preserving essential network structure, by factoring out structurally identical elements.

The structure of the remainder of this paper is as follows. First we introduce essential background material concerning network automorphism groups, and show how a network's automorphism group may be used to produce a coarse-grained skeleton of the network called the “quotient.” We also introduce a variation on the classical quotient which we call the  $s$ -quotient. Then we show that quotients can be substantially smaller than the network from which they are derived and explore ways in which key structural properties are inherited by the quotient and the  $s$ -quotient from the “parent” network. In particular, we shall examine how network het-

erogeneity, degree distribution and communication properties are carried from the parent to its quotients.

## II. BACKGROUND AND DEFINITIONS

### A. Preliminaries

Formally, a network is a graph  $G=(V,E)$  with vertex set  $V$  and edge set  $E$  in which two vertices are adjacent if there is an edge between them. An automorphism is a permutation of the vertices of the network which preserves adjacency, and the set of automorphisms under composition forms a group  $\text{Aut}(G)$ . The automorphism group of a network compactly describes its symmetry structure. Automorphism groups can be efficiently calculated with the use of an appropriate graph isomorphism algorithm such as the NAUTY algorithm [24] which we use in this study. We say that a network which possesses a nontrivial automorphism group is symmetric. Previous studies have highlighted the fact that many real-world networks possess nontrivial (and often quite large) automorphism groups [16–19].

The vertices of a symmetric network can be partitioned into disjoint equivalence classes called orbits: for every vertex  $v \in V(G)$ ,  $v$  belongs to the orbit

$$\Delta(v) = \{g \cdot v \in V; g \in \text{Aut}(G)\}.$$

We refer to the partitioning of the network vertices into disjoint orbits as the automorphism partition [18]. Note that since they can be permuted without altering network structure, two vertices in the same orbit are equivalent in the strongest possible structural sense: they play precisely the same structural role in the network and cannot be distinguished from one another by any meaningful structural measure (more formally, a vertex property which is preserved under isomorphism is known as a vertex invariant; vertices in the same orbit are indistinguishable by vertex invariants [25]). Vertices in the same orbit therefore possess many of the same structural properties, including the same degree, eigenvector centrality, and clustering coefficient [16] (for more examples, see Ref. [25]). We therefore say that vertices in the same group orbit are structurally equivalent. Since many real-world empirical networks possess a nontrivial automorphism partition they therefore carry a significant amount of redundant information in which more than one vertex plays the same structural role. In addition to elucidating the precise nature of structural repetitions in a network, the automorphism partition also provides a convenient way to factor out these structural repetitions by “gluing together” structurally equivalent vertices to create a coarse-graining of the network, known as the quotient.

### B. Quotients

More formally, let  $\Delta = \{\Delta_1, \Delta_2, \dots, \Delta_s\}$  be the automorphism partition of a network  $G$ . A significant property of this partition is that it is equitable [23]: the number of neighbors in  $\Delta_j$  of a vertex  $v \in \Delta_i$  is a constant  $q_{ij}$  ( $i, j = 1, 2, \dots, s$ ), which depends upon  $i$  and  $j$  but is independent of the choice of  $v \in \Delta_i$ . The quotient  $Q$  of  $G$  under the action of  $\text{Aut}(G)$  is the multidigraph with vertex set  $\Delta$  and adjacency matrix  $q_{ij}$ .

We refer to the network  $G$  as the parent of  $Q$ , and note that network quotients may be easily calculated using the NAUTY algorithm [24].

The quotient contains all the structural information of its parent network but, by associating structurally equivalent vertices, formally excludes all structural repetitions. Crucially, this means that many characteristic properties of the parent network are preserved in the quotient (any differences are due to the fact that the quotient only carries the unique structural features of its parent without repetitions). Consequently, while they are often very similar, it is the properties of the quotient, and not those of the parent network *per se*, that describe core system complexity. For this reason, the quotient may be thought of as the structural skeleton of its parent.

In the context of algebraic graph theory [20–23], certain properties of quotients are well-known including, for example, the fact that the eigenvalues of the quotient are a subset of those of its parent [23]. However, previous studies of graph quotients have been largely mathematical in nature, and have tended to focus on properties of quotients of completely regular graphs. An investigation of the properties of quotients of real-world networks—which typically contain both regular and random elements—has not as yet been undertaken.

### C. $s$ -quotients

As we noted above, quotients are generally multidigraphs (that is, their edges are weighted and directed). This is the case even when the parent network is simple (that is, the edges are not weighted or directed).

When a given network is a multidigraph it is often convenient to consider properties of the simple underlying network, in which edge weights and directions are removed. Such underlying networks carry the adjacency information of the full network, and so retain many key network properties. Therefore as well as examining properties of the quotient we shall also focus on properties of the simple underlying quotient (or  $s$ -quotient for short), denoted  $Q_s$ , in which edge directions, edge weights, and loops are removed from the quotient. Figure 1 shows a network, its quotient and its  $s$ -quotient. The  $s$ -quotient has the advantage that it retains the adjacency information of the quotient, yet has a binary symmetric adjacency matrix and thus is more computationally efficient to work with.

## III. PROPERTIES OF QUOTIENTS

### A. Relative size

Since quotients and  $s$ -quotients are formed by factoring out network redundancy they can be significantly smaller than their parent networks. Table I shows that many empirical  $s$ -quotients are less than 50% the size of their parent network, illustrating that much real-world network structure is due to repetition of structurally identical elements.

In order to investigate relative sizes of  $s$ -quotients we examined the correlation between various measures of network symmetry and the ratio of the size of the  $s$ -quotient to that of

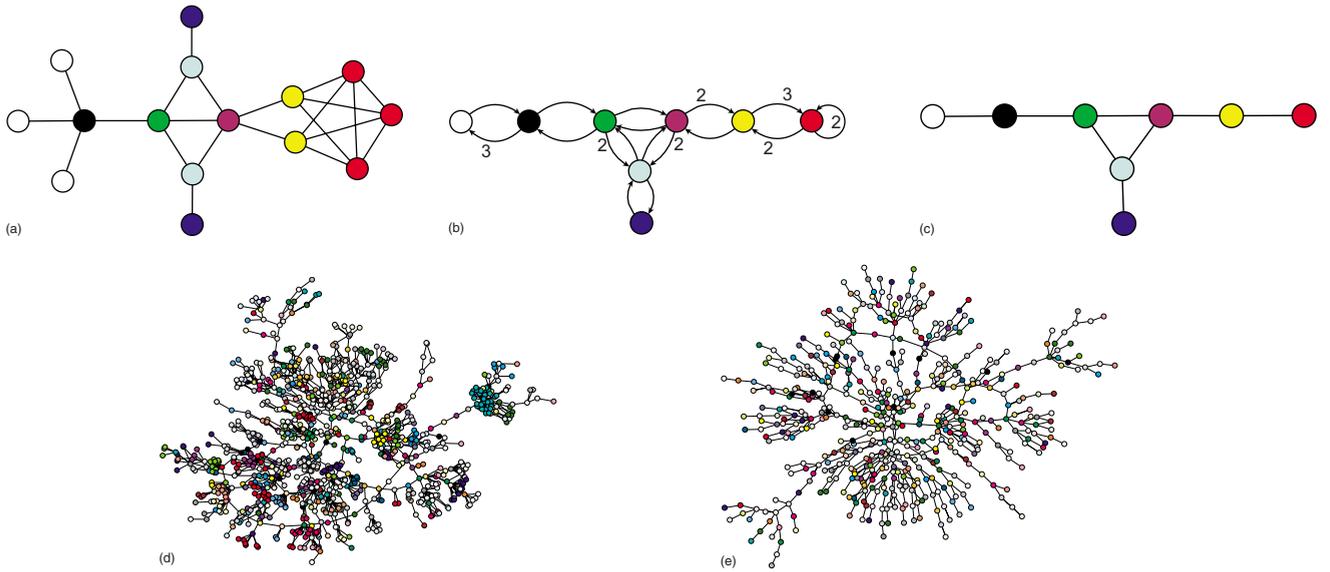


FIG. 1. (Color online) Networks and their quotients. (a) A hypothetical network, (b) its quotient, and (c) its  $s$ -quotient. (d) A real-world example: the network of ties between Ph.D. students and their advisers in theoretical computer science [29]. Each edge links an adviser to a student. Vertices in the same orbit are the same color. (e) The  $s$ -quotient of the theoretical computer science network: vertices are colored as the orbits in the parent network. There are 1025 vertices in the parent and only 511 in the  $s$ -quotient (a reduction of 50.15%); similarly there are 1043 edges in the parent and only 525 in the  $s$ -quotient (a reduction of 49.67%). Empirical networks were visualized using Pajek [30].

its parent (the reduction ratio). The most direct measure of network symmetry is the size of automorphism group, namely,  $|\text{Aut}(G)|$ . However, to compare symmetry in networks with different sizes  $\beta_G$ , the  $n$ th root of the ratio of automorphism group size to that of the (maximally symmetric) complete graph of the same size [16,17]:

$$\beta_G = \left( \frac{|\text{Aut}(G)|}{N!} \right)^{1/N}$$

is often used. An intuitive observation concerning network symmetry is that the more structurally equivalent vertices a network has, the more symmetric it is. Based on this obser-

TABLE I. Statistics for representative networks and their  $s$ -quotients. Summarized statistics for each network are the number of nodes  $N$ , the number of edges  $M$ , the mean vertex degree  $Z$ , assortative mixing coefficient  $r$  [9], the mean geodesic distance  $m$ , the diameter  $D$ , and the clustering coefficient  $C$  [5]. The subscript  $s$  indicates quantities for  $s$ -quotients. We also calculate the ratio of  $N_s$  to  $N$ , the ratio of  $M_s$  to  $M$ , and the ratio of  $M - M_s$  to  $N - N_s$ , denoted  $z$ . In all cases, we consider properties of the underlying graph of largest connected component of the parent networks. Real networks tested here include biological networks (PPI, Yeast and Homo), information networks (California and Epa), social networks (DeutchElite, Erdos02, Eva, Geom, and P-fei1738) and technological networks (InternetAS). Except for PPI, InternetAS and Homo, all network data can be downloaded from <http://vlado.fmf.uni-lj.si/pub/networks/data/>

Network	$N$	$N_s$	$N_s/N$	$M$	$M_s$	$M_s/M$	$z$	$Z$	$Z_s$	$r$	$r_s$	$m$	$m_s$	$D$	$D_s$	$C$	$C_s$
PPI [31]	1458	1019	69.89%	1948	1469	75.41%	1.09	2.67	2.88	-0.21	-0.05	6.80	6.68	19	19	0.07	0.07
Yeast [32]	2224	1843	82.87%	6609	6133	92.80%	1.25	5.94	6.66	-0.11	-0.04	4.37	4.21	11	11	0.14	0.14
Homo [33]	7020	6066	86.41%	19811	18575	93.76%	1.30	5.64	6.12	-0.06	-0.01	4.86	4.77	14	14	0.10	0.11
California [34]	5925	4009	67.66%	15770	12882	81.69%	1.51	5.32	6.43	-0.23	-0.18	5.02	4.66	13	13	0.08	0.09
Epa [35]	4253	2212	52.01%	8897	6545	73.56%	1.15	4.18	5.92	-0.30	-0.16	4.50	4.11	10	10	0.07	0.10
DutchElite [36]	3621	1907	52.67%	4310	2576	59.77%	1.01	2.38	2.70	-0.24	-0.04	8.56	7.71	22	22	0.00	0.00
Erdos02 [37]	6927	2365	34.14%	11850	7034	59.36%	1.06	3.42	5.95	-0.12	-0.08	3.78	3.41	4	4	0.12	0.29
Eva [38]	4475	898	20.07%	4652	1056	22.70%	1.01	2.08	2.35	-0.19	0.00	7.53	7.43	18	18	0.01	0.05
Geom [39]	3621	2803	77.41%	9461	7346	77.65%	2.59	5.23	5.24	0.17	0.19	5.31	5.15	14	14	0.54	0.43
P-fei1738 [40]	1738	1176	67.66%	1876	1312	69.94%	1.00	2.16	2.23	-0.27	0.07	10.22	10.39	29	29	0.00	0.00
InternetAs [41]	22442	11392	50.76%	45550	29564	64.90%	1.45	4.06	5.19	-0.20	-0.19	3.86	3.86	10	10	0.22	0.20

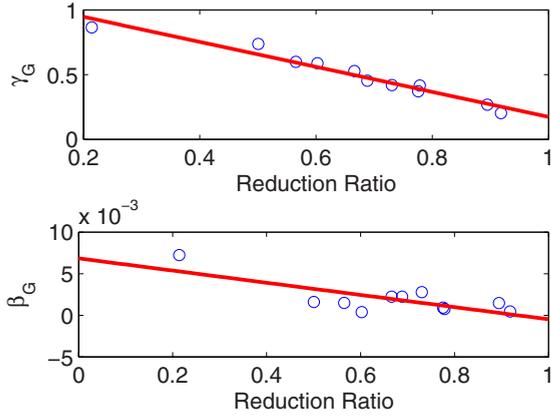


FIG. 2. (Color online) Network symmetry and  $s$ -quotient size are inversely correlated.

variation  $\gamma_G$  the ratio of the number of vertices in nontrivial orbits to  $N$ , the number of vertices in the network

$$\gamma_G = \frac{\sum_{|\Delta_i| > 1} |\Delta_i|}{N}$$

has also been used to measure network symmetry [18,19]. Here, we use both  $\beta_G$  and  $\gamma_G$  to measure the relative degree of symmetry in networks of various sizes. We define the size of a network as  $|G|=N+M$ , where  $M$  is the number of edges in  $G$ . The quotient reduction ratio is defined as  $r_G = |Q_s|/|G|$ .

Figure 2 shows the correlation between these two measures of symmetry and the quotient reduction ratio for eleven representative real-world networks (further details of these networks are given in Table I). The correlation coefficient between  $r_G$  and  $\beta_G$  is  $-0.7567$ , the correlation coefficient between  $r_G$  and  $\gamma_G$  is  $-0.9767$ , illustrating that the degree of symmetry, and the relative size of the  $s$ -quotient are strongly negatively correlated over a variety of networks.

### B. Heterogeneity

Network heterogeneity—that is, the degree to which different vertices play different roles or possess different properties in a network—is important in determining many dynamic network properties such as robustness [26] and synchronization [27]. A completely heterogeneous network is one in which all vertices play a unique structural role (that is, the network has a trivial automorphism group), while a completely homogeneous network is one in which all vertices play the same structural role (that is, the network has a transitive automorphism group) [18]. Since structurally equivalent elements are removed in the quotient while structurally nonequivalent elements are preserved, network quotients are generically completely heterogeneous: all vertices in the quotient play a different structural role (see the quotient in Fig. 1, for example) [45]. However, since edge weights, directions and loops are removed in the  $s$ -quotient, some vertices may still play the same structural role (for example, in the  $s$ -quotient in Fig. 1 the red and white vertices are structurally equivalent, as are the yellow and black vertices; as are

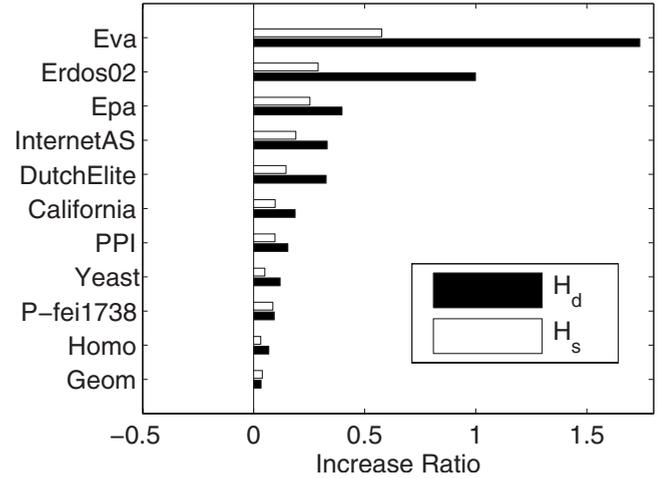


FIG. 3. Heterogeneity of empirical networks and their  $s$ -quotients. The horizontal axis shows the ratio of heterogeneity in the  $s$ -quotient to that of its parent, defined as  $\bar{H}_{d,s}(Q_s)/\bar{H}_{d,s}(G)-1$ .

the green and purple vertices). Thus, although  $s$ -quotients may not be completely heterogeneous we expect that they will be more heterogeneous than their parent networks.

In order to assess network heterogeneity we used two distinct measures: degree-based entropy [28]  $H_d(G)$  and symmetry-based entropy [18]  $H_s(G)$ . These two entropies have a common algebraic form

$$H_{d,s}(G) = - \sum_i p_i \ln p_i,$$

where  $p_i$  is the probability that a vertex has degree  $i$  when calculating  $H_d(G)$  and  $p_i$  is the probability that  $v \in \Delta_i$  when calculating  $H_s(G)$ . In order to compare networks of different sizes we normalized these measures as follows:

$$\bar{H}_{d,s}(G) = \frac{H_{d,s}(G) - \min(H_{d,s}, N)}{\max(H_{d,s}, N) - \min(H_{d,s}, N)},$$

where  $\max(H_{d,s}, N)$  and  $\min(H_{d,s}, N)$  are the maximal and minimal entropy values for a network with  $N$  vertices, respectively. Figure 3 summarizes these two entropy measures for the 11 empirical networks in Table I. As expected in all cases the  $s$ -quotient is more heterogeneous than its parent indicating that structural features which contribute to network homogeneity are factored out in the  $s$ -quotient while structural features which contribute to network heterogeneity are preserved in the  $s$ -quotient.

### C. Vertex degree distributions

A quotient’s vertex degree distribution is strongly related to that of its parent. Recall that all vertices in the same orbit have the same degree [20]. Thus,  $n_{\text{out}}(k, Q) = O_k$ , where  $O_k$  is the number of orbits of degree  $k$  in  $G$  and  $n_{\text{out}}(k, Q)$  is the number of vertices in  $Q$  with out-degree  $k$ . We may think of the vertex out-degree distribution in the quotient as being formed by measuring the degree of one representative vertex from each orbit in  $G$ . Thus, the quotient vertex out-degree distribution represents the “essential” vertex degree distribu-

tion of its parent and is dependent upon both the vertex degree distribution and the symmetry structure of its parent.

Hub vertices (those with high degree) often dominate real-world network topology and consequently crucially affect network properties such as robustness [26] and traffic along geodesics [42]. Hence, in order to accurately preserve network properties, quotients should preserve hub vertices. Since they generally connect many disparate regions of a network, hub vertices are more likely to be fixed by the automorphism group than are vertices of low degree, and consequently we expect that generically this is indeed the case.

Much of the symmetry present in many real world networks is due to the presence of bicliques [17,19] and in particular, the presence of stars [16,17] (a  $k$  star is a subgraph consisting of a central vertex of degree  $>k$  adjacent to  $k$  vertices of degree 1). In  $k$  stars, the  $k$  vertices of degree 1 are structurally equivalent to each other and collapse to a single vertex in the quotient. Thus, each  $k$  star reduces the vertex order of the quotient by  $k-1$ . Figure 1 shows how a 3-star (in white on the left) collapses to a single vertex in the quotient and  $s$ -quotient. In networks which contain a significant number of bicliques or stars, the  $s$ -quotient is formed largely by “pruning” appropriate vertices of small degree from the parent network while fixing hubs.

In order to assess the degree to which hub vertices are preserved in quotients, and the degree to which quotients are formed by pruning vertices of low degree, we investigated the degree distributions of those vertices that have been factored out in  $s$ -quotients for a variety of real-world networks. In particular, we considered two distinct quantities: (1)  $P_k$  the number of vertices of degree  $k$  factored out in the  $s$ -quotient as a percentage of the total number of vertices in the parent network

$$P_k = \frac{N_k - O_k}{N_k} \times 100\% ,$$

where  $N_k$  is the number of vertices with degree  $k$  and  $O_k$  is the number of orbits in which each vertex has degree  $k$  and (2)  $R_k$  the number of vertices of degree  $k$  factored out in the  $s$ -quotient as a percentage of the total number of vertices factored out:

$$R_k = \frac{N_k - O_k}{N - |\Delta|} \times 100\% .$$

Figure 4 shows that generally in empirical networks only vertices with small degree are factored out in the  $s$ -quotient (in all tested cases the maximum degree of any factored vertex was 29), vertices with higher degree tended to be fixed by the automorphism group and thus retained in the  $s$ -quotient.

In order to identify the proportion of vertices which are factored out by degree we considered two further quantities. Let  $d(v)$  be the degree of a vertex  $v$ . Consider now the total network degree set:  $\text{Deg} = \{d(v) | v \in G\}$  (i.e., the set of all vertex degrees), and the nontrivial orbit degree set:  $\text{Deg}'$

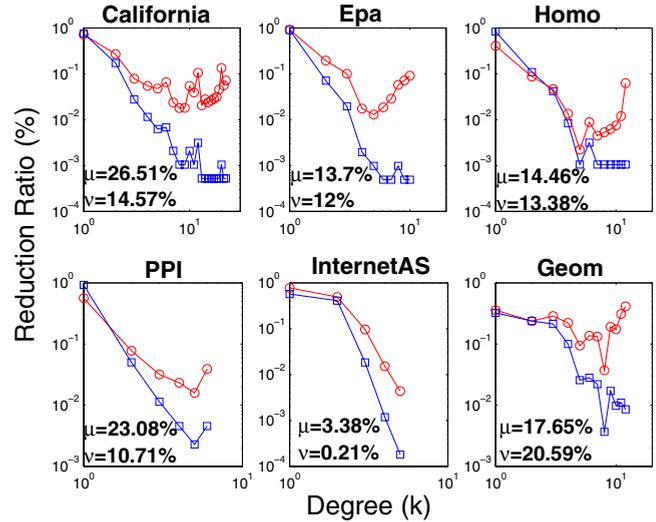


FIG. 4. (Color online) Factored degree distributions in  $s$ -quotients. The symbol  $\circ$  gives the reduction ratio  $P_k$ ; the symbol  $\square$  gives the reduction ratio  $R_k$ .

$= \{d(v) | v \in \Delta_i, |\Delta_i| > 1\}$  (i.e., the set of degrees of those vertices in nontrivial orbits). Note that  $\text{Deg}' \subseteq \text{Deg}$ . We define two quantities based upon these sets:

$$\mu = \frac{|\text{Deg}'|}{|\text{Deg}|} \times 100\% ,$$

the percentage of the degree-set factored out in the  $s$ -quotient, and

$$\nu = \frac{\max(\text{Deg}')}{\max(\text{Deg})} \times 100\% ,$$

the maximum vertex degree factored out in the  $s$ -quotient as a percentage of the maximum vertex degree in the parent network. Figure 4 shows these measures for six real-world networks. It is clear that vertices which are factored out in the  $s$ -quotient constitute only a minority of the whole network-degree set (the maximum value for  $\mu$  we found was 26.51%) and that only vertices with relatively low degree are factored out (the maximum value for  $\nu$  we found was 20.59%). Furthermore, Table I also shows that it is common for  $s$ -quotients to have an average degree larger than that of their parent, demonstrating that vertices of small degree are more likely to lie in a nontrivial orbits (and thus be factored out in the quotient) than are hub vertices (which are generically retained).

#### D. Communication properties

Many empirical large complex networks are “small-world” meaning that there exists a relatively short path between any two vertices in the network [2]. The shortest path between a pair of vertices is known as a geodesic and the length of the longest geodesic is known as the diameter of the network, which we denote  $D(G)$ . Distribution of geodesic distances and network diameter both significantly effect dynamic network properties such as information transfer [13] and tolerance to attack [26].

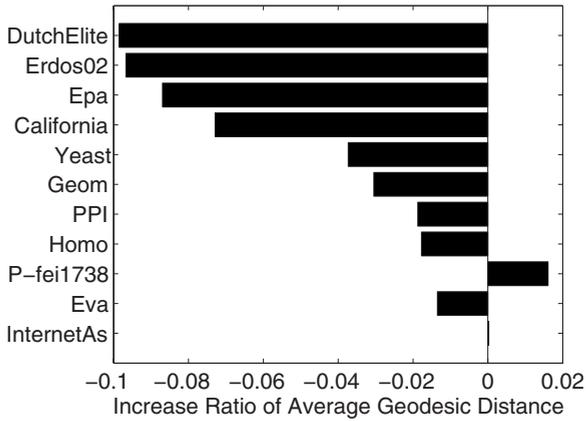


FIG. 5. Mean geodesic distances in  $s$ -quotients. The ratio  $(m_s/m)-1$  is plotted, where  $m_s$  and  $m$  are the mean geodesic distances in the  $s$ -quotient and its parent, respectively.

Table I shows the network diameter and  $s$ -quotient diameter for a variety of empirical networks. In all cases, network diameter is preserved exactly in the  $s$ -quotient. For example, in the Eva network, a telecommunications and media ownership network [38], the vertex and edge numbers of the  $s$ -quotient are 20% and 22.7% that of the original network, respectively, yet network diameter is maintained in the  $s$ -quotient. In this case the  $s$ -quotient is substantially smaller than its parent, yet it preserves the communication properties of its parent. In fact, this empirical observation is true for all “locally symmetric” networks.

Intuitively, a network is globally symmetric if the longest geodesic is between vertices in the same orbit (that is, there are automorphisms which permute distant vertices); otherwise the network is locally symmetric (that is, all automorphisms act on local vertex subsets). Since many real-world networks are commonly subject to continuous stochastic fluctuations in topology, we do not expect—neither did we find—that any large real-world networks are globally symmetric.

The  $s$ -quotient describes the orbit adjacency structure of its parent network. Thus, network diameter is exactly preserved in the  $s$ -quotient as long as the parent network is not globally symmetric. For an illustration of this see the network shown in Fig. 1(a). In this network the longest geodesic is between any of the red vertices on the right and any of the white vertices on the left, and the network has diameter=5. The  $s$ -quotient of this network [shown in Fig. 1(c)] also has diameter=5, and diameter is preserved in the  $s$ -quotient since orbit adjacency is preserved.

While the diameter of a network is related to the maximum information transfer cost in the network, mean geodesic distance is related to the average transit cost. Empirical measurements show that the disparity between mean geodesic distance in the  $s$ -quotient and its parent is usually quite small. As shown in Fig. 5 for all tested networks the mean geodesic distance of the  $s$ -quotient is within 10% of that of its parent network irrespective of the relative size of the  $s$ -quotient to its parent. Since both network diameter and mean geodesic distance are robustly inherited by the  $s$ -quotient

from its parent, we conclude that the  $s$ -quotient forms the communication skeleton of its parent.

#### IV. CONCLUSIONS

The quotient of a network is formed by associating structurally equivalent vertices into disjoint equivalence classes and considering adjacency relationships between these equivalence classes. Thus quotients capture all essential network complexity, yet formally exclude all structural redundancy. Quotients may therefore be thought of as the structural skeletons of the systems from which they are derived. Consequently properties of the quotient, and not those of the parent network *per se*, describe core system complexity. Observation of the statistics of real-world networks verifies that elements which contribute to network homogeneity (or simplicity) are removed in network quotients; while the elements which contribute to the heterogeneity (or complexity) are completely retained in quotients.

Many biological networks are thought to form by growth with vertex duplication or partial duplication [43]. Vertex duplication is useful in a biological context since it naturally endows biological regulatory systems with functional redundancy, thus reinforcing against damage. Quotients of biological regulatory networks therefore encode core relationships between biochemical control motifs, minus any repetitions due to redundancy. While the large-scale properties of analogous biological regulatory networks are often remarkably similar across a broad range of species, their detailed properties can differ significantly [33,44]. Thus it may be of particular interest to explore the similarities between structural properties of quotients of regulatory networks for various different species, since this could provide a new means to analyze functional conservation of regulatory motifs across species.

Finally, since quotients carry the structure of their parents, yet are often substantially smaller, performing analysis directly on quotients, rather than on the corresponding parent networks, can reduce the complexity of network algorithms. For example, average shortest path length computation time can be reduced from  $\Theta(NM)$  to  $\Omega(r_N r_M NM)$  if calculated on the  $s$ -quotient (where  $r_N = N_s/N$  and  $r_M = M_s/M$ ). We anticipate that further investigation of properties of network quotients will be both of theoretical and pragmatic interest.

#### ACKNOWLEDGMENTS

The work was supported by the National Natural Science Foundation of China under Grants No. 60673133, No. 60703093; the National Grand Fundamental Research 973 Program of China under Grant No. 2005CB321905; NIH-NIAMS Grant No. P01 AR052915-01A1; Shanghai Commission of Science and Technology Grant 04dz14003 in China and by Shanghai Leading Academic Discipline Project Under Project No. B114. This work was partly funded by the EPSRC, UK. This work was conducted while BDM was at The Bone and Joint Research Group, Centre for Human Development, Stem Cells and Regeneration, Institute of Developmental Sciences, University of Southampton, Southampton, SO16 6YD, UK.

- [1] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
- [2] M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).
- [3] J. Kleinberg, *Nature (London)* **406**, 845 (2000).
- [4] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11149 (2000).
- [5] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
- [6] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [7] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, *Science* **297**, 1551 (2002).
- [8] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, *Science* **298**, 824 (2002).
- [9] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
- [10] Ch. Song, S. Havlin, and H. A. Makse, *Nature (London)* **433**, 392 (2005).
- [11] A.-L. Barabási, *Nat. Phys.* **1**, 68 (2005).
- [12] K.-I. Goh, G. Salvi, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **96**, 018701 (2006).
- [13] D.-H. Kim, J. D. Noh, and H. Jeong, *Phys. Rev. E* **70**, 046126 (2004).
- [14] A. Arenas, J. Duch, A. Fernandez, and S. Gomez, *New J. Phys.* **9**, 176 (2007).
- [15] B. Bollobás, *Random graphs*, 2nd ed., Vol. 73 of *Cambridge Studies in Advanced Mathematics* (Cambridge University Press, Cambridge, 2001).
- [16] B. D. MacArthur and J. W. Anderson, e-print arXiv:cond-mat/0609274.
- [17] B. D. MacArthur, R. J. Sánchez-García, and J. W. Anderson, *Discr. Appl. Math.*, doi:10.1016/j.dam.2008.04.008 (2008).
- [18] Yang-Hua Xiao, Wen-Tao Wu, Hui Wang, Momiao Xiong, and Wei Wang, *Physica A* **387**, 2611 (2008).
- [19] Y. Xiao, M. Xiong, W. Wang, and H. Wang, *Phys. Rev. E* **77**, 066108 (2008).
- [20] J. Lauri and R. Scapellato, *Topics in Graph Automorphisms and Reconstruction* (Cambridge University Press, Cambridge, 2003).
- [21] N. Biggs, *Algebraic Graph Theory* (Cambridge University Press, London, 1974).
- [22] L. W. Beineke, R. J. Wilson, and P. J. Cameron, *Topics in Algebraic Graph Theory, Encyclopedia of Mathematics and its Applications* (Cambridge University Press, Cambridge, 2004), Vol. 102.
- [23] C. Godsil and G. Royle, *Algebraic Graph Theory*, Vol. 207 of *Graduate Texts in Mathematics* (Springer, Berlin, 2001).
- [24] B. D. McKay, *Congr. Numer.* **30**, 45 (1981).
- [25] G. Tinhofer and M. Klin (unpublished).
- [26] R. Albert, H. Jeong, and A.-L. Barabási, *Nature (London)* **406**, 378 (2000).
- [27] T. Nishikawa, A. E. Motter, Y. C. Lai, and F. C. Hoppensteadt, *Phys. Rev. Lett.* **91**, 014101 (2003).
- [28] B. Wang, H. W. Tang, C. H. Guo, and Z. L. Xiu, *Physica A* **363**, 591 (2005).
- [29] D. S. Johnson, *SIGACT News* **16**, 36 (1984).
- [30] V. Batagelj and A. Mrvar, *Connections* **21**, 47 (1998).
- [31] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, *Nature (London)* **411**, 41 (2001).
- [32] S. Sun, L. Ling, N. Zhang, G. Li, and R. Chen, *Nucleic Acids Res.* **31**, 2443 (2003).
- [33] C. Stark, B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, and M. Tyers, *Nucleic Acids Res.* **34**, D535 (2006).
- [34] J. Kleinberg, *J. ACM* **46**, 604 (1999).
- [35] Pages linking to [www.epa.gov](http://www.epa.gov). Obtained from Jon Kleinberg's web page, <http://www.cs.cornell.edu/courses/cs685/2002fa/>
- [36] Wouter de Nooy, <http://vlado.fmf.uni-lj.si/pub/networks/data/2mode/DutchElite.pdf>.
- [37] V. Batagelj and A. Mrvar, *Soc. Networks* **22**, 173 (2000).
- [38] K. Norlen, G. Lucas, M. Gebbie, and J. Chuang (unpublished).
- [39] B. Jones, *Computational Geometry Database*, February 2002, <http://compgeom.cs.uiuc.edu/jeffe/compgeom/biblios.html>.
- [40] Genealogical Data from NSF Project on Empirical Kinship Networks, <http://eclectic.anthrosciences.org/drwhite/linkages/datasets/student.html>.
- [41] The CAIDA Group, *The caida as relationships dataset*, 2003–2007, <http://www.caida.org/data/active/as-relationships/>.
- [42] K. I. Goh, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **87**, 278701 (2001).
- [43] F. Chung, L. Lu, T. G. Dewey, and D. J. Galas, *J. Comput. Biol.* **10**, 677 (2003).
- [44] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, *Nature (London)* **407**, 651 (2000).
- [45] Note: counterexample graphs for which the quotient is not completely heterogeneous can be constructed, but these examples are degenerate: this is not generically the case.