Exploiting symmetry in network analysis

Rubén J. Sánchez-García

Virtually all network analyses involve structural measures between pairs of vertices, or of the vertices themselves, and the large amount of symmetry present in real-world complex networks is inherited by such measures. This has practical consequences that have not yet been explored in full generality, nor systematically exploited by network practitioners. Here we study the effect of network symmetry on arbitrary network measures, and show how this can be exploited in practice in a number of ways, from redundancy compression, to computational reduction. We also uncover the spectral signatures of symmetry for an arbitrary network measure such as the graph Laplacian. Computing network symmetries is very efficient in practice, and we test real-world examples up to several million nodes. Since network models are ubiquitous in the Applied Sciences, and typically contain a large degree of structural redundancy, our results are not only significant, but widely applicable.
Network models of real-world complex systems have been extremely successful at revealing structural and dynamical properties of these systems. The success of this approach is due to its simplicity, versatility, and surprising universality, with common properties and principles shared by many disparate systems.

One property of interest is the presence of structural redundancies, which manifest themselves as symmetries in a network model. Symmetries relate to system robustness, as they identify structurally equivalent nodes, and can arise from replicative growth processes such as duplication, evolution from basic principles, or functional optimisation, and can be arbitrarily generated in model graphs. It has been shown that real-world networks possess a large number of symmetries, and that this has important consequences for network structural, spectral, and dynamical properties for instance cluster synchronisation, and eigenvector centrality. To facilitate dissemination, we provide fast symmetry-based eigendecomposition algorithm. We achieve remarkable empirical results in our real-world test networks: For a more general study of arbitrary symmetry in networks: For a more general study of arbitrary symmetry in networks, we can partition the vertex set V into the symmetric core of fixed points V0 (an automorphism σ moves a vertex i ∈ V if σ(i) ≠ i, and fixes it otherwise), and the vertex sets Mi of the symmetric motifs, as shown in Fig. 1a for a toy example. Equation (2) is called the geometric decomposition of the network.

Results
Symmetry in complex networks. The notion of network symmetry is captured by the mathematical concept of graph automorphism. This is a permutation of the vertices (nodes) preserving adjacency, and can be expressed in matrix form using the adjacency matrix of the network. If a network (mathematically, a finite simple graph) G has n vertices, labelled 1 to n, its adjacency matrix A = (aij) is an n × n matrix with (i, j)-entry aij = 1 if there is an edge between nodes i and j, and zero otherwise. A graph automorphism σ is then a permutation, or relabelling, of the vertices v → σ(v) such that (σ(i), σ(j)) is an edge only if (i, j) is an edge, or, equivalently, aij = aσ(i)σ(j) for all i, j. In matrix terms, this can be written as

\[ AP = PA \]
However, each symmetry is the product (composition) of automorphisms permuting a very small number of vertices within a symmetric motif. For example, the toy graph in Fig. 1a has $2^7 \times 3! \times 4! = 18,432$ symmetries (size of the automorphism group) but they generated by (all combinations of) just ten permutations, each permuting a few vertices within a symmetric motif (one permutation per motif except two for $M_3$, $M_5$, and $M_6$).

Each symmetric motif can be further subdivided into orbits of structurally indistinguishable nodes (shown by colour in Fig. 1a), which play the same structural role in the network and, therefore, contribute to network redundancy and thus to the robustness of the underlying system. Our notion of structurally indistinguishable nodes (nodes in the same orbit of the automorphism group) extends the notion of structurally equivalent nodes found in the social sciences\textsuperscript{39}, that is, nodes with the same set of neighbours. It is not equivalent: nodes in the same orbit may not have the same neighbours (e.g., $M_1$, $M_4$, or $M_5$ in Fig. 1a).

Network symmetries (possibly very large) real-world networks can be effectively computed, stored, and manipulated (see "Methods"). For instance, we computed generators of the automorphism group, and the subsequent geometric decomposition, for real-world networks up to several million nodes and edges in a few seconds (see $t_1$ and $t_2$ in Table 1).

Most symmetric motifs in real-world networks (typically over 90\%, see the $bsm$ column in Table 1) are of a very specific type, called basic\textsuperscript{11}, they are made of one or more orbits of the same size, and every permutation of the vertices in each orbit is realisable, that is, can be extended to a network automorphism (see Fig. 1a). Basic symmetric motifs (BSMs) have a very constrained structure\textsuperscript{13}, which we will generalise to arbitrary network measures and exploit throughout this article. Non-basic symmetric motifs (typically branched trees, as $M_7$ in Fig. 1a) are called complex; they are rare and can either be studied on a case-by-case basis, or removed from the symmetry computation altogether (by ignoring the symmetries generated by them).

The definition of network automorphism Eq. (1) carries to an arbitrary $n \times n$ real matrix $A = (a_{ij})$. Any such matrix can be seen as the adjacency matrix of a network with $n$ vertices labelled $1$ to $n$, and an edge (link) from node $i$ to node $j$ with weight $a_{ij}$ if $a_{ij} \neq 0$, and no such edge if $a_{ij} = 0$. This means that an automorphism does not only preserve edges, but also their weights and directions. This may not be a realistic assumption for real-world weighted networks, where the weights often come from observational or experimental data, but it applies to the matrix representing a network structural measure, as we illustrate in Fig. 2 and explain next.

### Structural network measures

A (pairwise) structural network measure is a function $F(i, j)$ on pairs of vertices which satisfies

$$ F(\sigma(i), \sigma(j)) = F(i, j) \quad \text{for all} \ i, j \in V $$

for all automorphisms $\sigma \in \text{Aut}(G)$. Since automorphisms identify structurally indistinguishable vertices ($i$ and $\sigma(i)$) and, similarly, edges ($(i, j)$ and $(\sigma(i), \sigma(j))$), structural network measures are (edge) functions that depend on the network structure alone, and not, for example, on node or edge labels, or other meta-data. Most network measures are structural, including graph metrics (e.g., shortest path), and matrices algebraically derived from the adjacency matrix (e.g., Laplacian matrix). (We identify matrices $M$ with pairwise measures via $F(i, j) = [M]_{ij}$.) In particular, structural measures are independent of the ordering or labelling of the vertices. In contrast, functions depending, explicitly or implicitly, on some vertex ordering or labelling, are not structural, for example the shortest path length through a given node.

We can encode a structural measure $F$ as a network with adjacency matrix $[F(A)]_{ij} = F(i, j)$ (see Fig. 2a for the adjacency matrix and Fig. 2b, c for two examples of structural measures), and write (3) in matrix form as

$$ F(A) = P F(A), $$

where $P$ is the permutation matrix corresponding to the permutation $\sigma$. Theorem 1 can be formulated in the matrix form

$$ P F(A) = F(A) $$

with $Q_1$, $Q_2$, and $Q_3$ defined as

$$ Q_1 = \left[ \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{array} \right], \quad Q_2 = \left[ \begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ 1 \end{array} \right], \quad Q_3 = \left[ \begin{array}{c} 0 \\ 1 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{array} \right], $$

and $c_{bsm}$, $c_{full}$, and $e_{full}$ are used throughout the article.

### Algorithm

#### Pre-processing

1. **Compute automorphism group**: Compute the automorphism group of the network, which is a group of permutations that preserve the network structure.

2. **Generate generators**: From the automorphism group, generate a set of generators (functions that generate all automorphisms by composition) with a desired size.

#### Compression

1. **Reduce symmetry**: Apply a compression algorithm to reduce the number of symmetries in the network. This could be done by identifying and removing symmetries that do not significantly affect the network structure.

2. **Optimize symmetry**: Optimize the symmetry by finding a set of symmetries that best preserve the network structure.

#### Decomposition

1. **Decompose network**: Decompose the network into its constituent motifs. This can be done by identifying and separating the network into smaller, structurally equivalent components.

2. **Analyze motifs**: Analyze the motifs to extract structural information about the network.

#### Post-processing

1. **Measure network properties**: Measure various network properties and metrics using the decomposed network and the extracted structural information.

2. **Compare with original network**: Compare the measured properties with those of the original network to assess the effectiveness of the compression and decomposition methods.

### Results

1. **Symmetry in real-world networks**: Investigate the symmetry in real-world networks, such as social networks, biological networks, and communication networks.

2. **Effect of compression**: Evaluate the effect of compression on the network properties to determine if the compression methods preserve the essential structural information.

3. **Decomposition effectiveness**: Assess the effectiveness of the decomposition methods in identifying and separating structural motifs.

### Conclusion

Automorphism-based compression, decomposition, and analysis provide a powerful framework for understanding and studying complex networks. By identifying and exploiting symmetries, we can reduce the complexity of network analysis, leading to more efficient and insightful approaches for network properties and structure.
where $P$ is the permutation matrix corresponding to $\sigma$. Comparing this to Eq. (1), we see that the network representation of $F$, $F(\mathcal{G})$, with adjacency matrix $F(A)$, inherits all the symmetries of $\mathcal{G}$. In particular, the network $F(\mathcal{G})$ has the same decomposition into symmetric motifs Eq. (2), and orbits, as $\mathcal{G}$. The BSMs in $F(\mathcal{G})$ must occur on the same vertices $M_n$, although they are now all-to-all weighted subgraphs in general (Fig. 2b). Nevertheless, they have a very constrained structure: the intra and inter orbit connectivity depends on two parameters only. Namely, each orbit in a BSM is uniquely determined by $\beta = F(v_i, v_j)$ (the connectivity of a vertex with itself) and $\alpha = F(v_i, v_j)$, $i \neq j$ (the connectivity of a vertex with every other vertex in the orbit), for all $v_i, v_j$ in the orbit. Similarly, the connectivity between two orbits $\Delta_1$ and $\Delta_2$ in the same BSM also depends on two parameters: after a suitable reordering $\Delta_1 = [v_1, \ldots, v_n]$ and $\Delta_2 = [w_1, \ldots, w_m]$, we have $\delta = F(v_i, w_j)$ and $\gamma = F(v_i, w_j)$ for all $1 \leq i, j \leq n$. (For a proof, see Theorem 1 in “Methods.”) This can be observed in Fig. 2c and is represented schematically in Fig. 3a, b. In particular, each BSM takes a very constrained form in the quotient, as shown schematically in Fig. 3c, d.

The results in this article apply to arbitrary structural measures, although the two most common cases in practice are the following. We call $F$ full if $F(i, j) \neq 0$ for all $i \neq j \in V$ (e.g., a graph metric), and sparse if $F(i, j) = 0$ if $a_{ij} = 0$, for all $i \neq j \in V$ (e.g., the graph Laplacian). The graph representation of $F(\mathcal{G})$ is an all-to-all weighted graph if $F$ is full, and has a sparsity similar to $\mathcal{G}$ if $F$ is sparse (cf. Fig. 2c).

From now on, we will assume that $\mathcal{G}$ is undirected and $F$ is symmetric, $F(i, j) = F(j, i)$, which may not be the case even if $\mathcal{G}$ is undirected (e.g., the transition probability of a random walker $F(i, j) = \frac{a_{ij}}{\deg(i)}$), and discuss directed networks and asymmetric measures in the “Methods” section.

**Quotient network.** The formal procedure to quantify and eliminate structural redundancies in a network is via its quotient network. This is the graph with one vertex per orbit or fixed point (see Fig. 1b) and edges representing average connectivity. Formally, if $A$ is the $n \times n$ adjacency matrix of a graph $\mathcal{G}$, the quotient network with respect to a partition of the vertex set $V = V_1 \cup \cdots \cup V_m$ is the graph $\mathcal{J}$ with $m \times m$ adjacency matrix the quotient matrix $Q(A) = (b_{ij})$ defined by

$$b_{ij} = \frac{1}{|V_i|} \sum_{v_j \in V_j} a_{ij},$$

the average connectivity from a vertex in $V_i$ to all vertices in $V_j$. There is an explicit matrix equation for the quotient. Consider the $n \times m$ characteristic matrix $S$ of the partition, that is, $[S]_{ik} = 1$ if $i \in V_k$, and zero otherwise, and the diagonal matrix $\Lambda = \text{diag}(n_1, \ldots, n_m)$, where $n_k = |V_k|$. Then

$$Q(A) = \Lambda^{-1} S^T A S.$$

The quotient network is a directed and weighted network in general. An alternative is to use the symmetric quotient, with adjacency matrix $Q_{sym}(A) = \Lambda^{-1/2} S^T A S \Lambda^{-1/2}$, which is weighted but undirected. Note that $Q(A)$ and $Q_{sym}(A)$ are spectrally equivalent matrices: they have the same eigenvalues, with eigenvectors related by the transformation $v \mapsto \Lambda^{1/2} w$.

In the context of symmetries, we will always refer to the quotient with respect to the partition of the vertex set into orbits. This quotient removes all the original symmetries from the network: if $\sigma(v_i) = v_j$, then $v_i$ and $v_j$ are in the same orbit and hence represented by the same vertex in the quotient network, which is then fixed by $\sigma$. We can, therefore, infer and quantify properties arising from redundancy alone by comparing a network to its quotient. The quotients of real-world networks are often significantly smaller (in vertex and edge size) than the original networks (see $\mathcal{J}_2$ and $\mathcal{J}_3$ in Table 1), and this reduction quantifies the structural redundancy present in an empirical network. Not every real-world network is equally symmetric, and, in our test networks, we give examples of network quotient reductions ranging from about 50% to just 2%. Computing the network quotient involves multiplication by very sparse matrices ($\Lambda$ is diagonal and $S$ has one non-zero element per row) and hence is computationally efficient (a few seconds in all our test networks).
found a remarkable amount of redundancy (up to 70%) due to sequences for network analysis, is whether we can easily eliminate the symmetry-induced redundancy. If a network has \( n \) vertices and \( n \) orbits, there are \( n^2 \) pairs of vertices but only \( n^2 \) pairs of orbits, achieving a reduction, or compression ratio, of

\[
\varepsilon_{\text{full}} = \left( \frac{n^2}{n^2} \right) = \frac{n^2}{n^2}.
\]

(7)

for a full network measure, typically much smaller than the ratio \( n^2 = n_{\text{full}}/n_{\text{full}} \). On the other hand, for a sparse network measure, we only need to consider edge values, hence the reduction is the ratio between the number of edges in the graph and in its quotient

\[
\varepsilon_{\text{sparse}} = \frac{n_{\text{sparse}}}{n_{\text{full}}}.
\]

(8)

For an arbitrary network measure, its compression ratio, which measures the redundancy present (zero values excluded), will range between \( \varepsilon_{\text{full}} \) and \( \varepsilon_{\text{sparse}} \). The compression ratios \( \varepsilon_{\text{full}} \) and \( \varepsilon_{\text{sparse}} \) are shown on Table 1 for our test networks. We found a remarkable amount of redundancy (up to 70%) due to symmetry alone (Fig. 4).

**Symmetry compression.** A natural question, with practical consequences for network analysis, is whether we can easily "eliminate" the symmetry-induced redundancies. This means storing only one value of a network function for each orbit of structurally indistinguishable nodes or edges, all sharing the same such value. Although this has been explored in particular cases, such as shortest path distances\(^\text{27}\), here we present a general treatment. A simple method is to use the quotient matrix

\[
B = S^T AS,
\]

(9)

which is easier to store than \( \Lambda^{-1} S^T AS \). This matrix achieves a compression ratio between \( \varepsilon_{\text{full}} \) and \( \varepsilon_{\text{sparse}} \) (by using a sparse representation of \( B \), as explained before. From this matrix, we can recover all but the internal connectivity inside a symmetric motif, which is replaced by the average connectivity. Namely, let us define

\[
\bar{a}_{ij} = \frac{1}{n_i} b_{ij},
\]

(10)

where \( n_i \) is the size of the orbit containing \( v_i \) (note that these orbit sizes can be obtained as the row sums of the characteristic matrix \( S \)). Then one can show ("Methods", Theorem 2) that

\[
\bar{a}_{ij} = \begin{cases} 
1 & \text{if } v_i \text{ and } v_j \text{ are external,} \\
\frac{1}{n_i} + \sum_{v_k, v_l} a_{kl} & \text{if } v_i \text{ and } v_j \text{ are internal,}
\end{cases}
\]

(11)

where we call a pair of vertices external if they belong to two different symmetric motifs, and internal otherwise, and \( v_i \in \Delta_i \) and \( v_j \in \Delta_j \) are orbits. Hence, if we are not interested in the exact internal connectivity (inside a symmetric motif), or it can be recovered easily by other means (e.g., one motif at a time), we can use this simple method to eliminate all the symmetry-induced redundancies on an arbitrary network measure encoded as a matrix \( A \). We have included simple average symmetry compression and decomposition algorithms (Algs. 1 and 2), where \( A_{\text{avg}} \) is the matrix with entries \( \bar{a}_{ij} \). The original \( n \times n \) matrix \( A \) is stored using the \( n \times n \) quotient matrix \( B \) plus a very sparse (\( n \) non-zero elements) characteristic matrix \( S \).

**Algorithm 1.**

**Average symmetry compression.**

\[\text{Input: } \text{adjacency matrix } A, \text{ characteristic matrix } S \]

\[\text{Output: } \text{quotient matrix } B \]

\[B \leftarrow S^T AS\]

**Algorithm 2.**

**Average symmetry decomposition.**

\[\text{Input: } \text{quotient matrix } B, \text{ characteristic matrix } S \]

\[\text{Output: } \text{adjacency matrix } A_{\text{avg}} \]

\[\Lambda \leftarrow \det(q(\lambda; S)) \]

\[R \leftarrow SA^{-1} \]

\[A_{\text{avg}} \leftarrow RB^T \]

The vast majority of edges in the network representation of a network measure are external (at least 99.999% for a full measure in our test networks, see \( \text{int}_f \) in Table 1), and hence the information loss by using \( A_{\text{avg}} \) instead of \( A \) is minimal. We can nevertheless enforce lossless compression, by storing the intra-motif connectivity separately. Indeed, we can exploit the fact that most symmetric motifs in empirical networks are basic, and hence each orbit, or pair of orbits, is uniquely determined by two parameters (Fig. 3). If we disregard the symmetries generated at non-basic symmetric motifs, the corresponding quotient, called basic quotient, written \( A_{\text{basic}} \), leaves non-basic motifs unchanged and retains most of the symmetry in a typical real-world network.
We can often recover the average values of a computational reduction ratio of between \( F \) partially quotient recoverable and \( F \) quotient recoverable.
where \((\lambda, \nu)\) is an eigenpair of \(A\) and \(P\) the permutation matrix of a network automorphism (Eq. (1)). This gives another eigenpair \((\lambda, \nu)\) whenever \(\nu \neq P\nu\) are linearly independent (obviously not always the case).

Let \(B = Q(A)\) be the \(m \times m\) quotient of \(A\) (Eq. (6)) with respect to the partition of the vertex set into orbits. This partition satisfies a regularity condition called equitability\(^{25,26}\), which can be written in matrix form as \(AS = SB\), where \(S\) is the characteristic matrix of the partition. In particular, if \((\nu, \varphi)\) is a quotient eigenpair, then \((\lambda, \varphi)\) is a parent eigenpair,

\[
A(S\varphi) = SB\nu = \lambda(S\varphi).
\]

In fact, one can show ("Methods", Theorem 3) that \(A\) has an eigenbasis of the form

\[
\{S\nu_1, \ldots, S\nu_m, w_1, \ldots, w_{n-m}\},
\]

where \(\nu_1, \ldots, \nu_m\) is any eigenbasis of \(B\), and \(S^i\nu_i = 0\) for all \(i\). We can think of a vector \(\nu \in \mathbb{R}^m\), respectively \(\mathbf{w} \in \mathbb{R}^n\), as a vector on (the vertices of) the quotient, respectively, the parent, network. Then, each vector \(S\nu_i\) equals the vector \(\nu_i\) lifted to the parent network by repeating the value on each orbit. Similarly, \(S^i\nu_i = 0\) means that the sum of the entries of \(\nu_i\) on each orbit is 0. All in all, we can always find an eigenbasis of \(A\) consisting of non-redundant eigenvectors \(\{S\nu_1, \ldots, S\nu_m\}\) arising from a quotient eigenbasis by repeating values on each orbit, and redundant eigenvectors \(\{w_1, \ldots, w_{n-m}\}\) arising from the network symmetries, which add up to zero on each orbit (hence "dissappering" in the quotient). Similarly, we call their respective eigenvalues redundant and non-redundant.

Analogous to the way that symmetry is generated at symmetric motifs, the redundant eigenvectors and eigenvalues arise directly from certain eigenvectors and eigenvalues of the symmetric motifs, considered as networks on their own (Fig. 6). In fact, each symmetric motif \(\mathcal{M}\) contributes the same (called redundant) eigenpairs to any network containing \(\mathcal{M}\) as a symmetric motif: One can show ("Methods", Theorem 4) that if \(\mathcal{M}\) is a symmetric motif of a network \(\mathcal{G}\) and \((\lambda, \nu)\) is a redundant eigenpair of \(\mathcal{M}\) (that is, the values of \(\nu\) add up to zero on each orbit of \(\mathcal{M}\)), then \((\lambda, \nu)\) is an eigenpair of \(\mathcal{G}\), where \(\nu\) is equal to \(\nu\) on (the vertices of) \(\mathcal{M}\), and zero elsewhere. We call such a vector localized on the motif \(\mathcal{M}\), as it is zero outside the motif. Moreover, if \(\mathcal{M}\) has \(n\) vertices and \(k\) orbits, then it has an eigenbasis consisting of \(n - k\) redundant eigenpairs, which are inherited by any network containing \(\mathcal{M}\) as a symmetric motif (Fig. 6, Theorem 4 in "Methods").

Furthermore, since most symmetric motifs in real-world networks are basic, thus have a very constrained structure (Fig. 3), we can in fact determine the redundant spectrum of BSMs with up to a few orbits, that is, we can predict where the most significant "peaks" in the spectral density of an arbitrary network function will occur. The formulae for the redundant spectra for BSMs of one or two orbits (which covers most BSMs, up to 99% of them in our test networks) is given on Table 2.

We now give more details of the computation of the redundant spectrum of BSMs up to two orbits (Table 2), with full details in the "Methods" section. A BSM with one orbit is an \((\alpha, \beta)\)-uniform graph \(K_{\alpha,\beta}^n\) with adjacency matrix \(A_{\alpha,\beta}^n = (a_{ij})\) given by \(a_{ij} = \alpha\) for \(i \neq j\) and \(a_{ii} = \beta\), for all \(i \neq j\), and some constants \(\alpha\) and \(\beta\). A BSM with two orbits consists of the \((\gamma, \delta)\)-uniform join of two uniform graphs \(K_{\alpha,\beta}^n\) and \(K_{\gamma,\delta}^m\), that is, the graph with \(2n + m\) vertices and block adjacency matrix (after a suitable labelling of the vertices)

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix},
\]

where \(A = A_{\alpha,\beta}^n\), \(B = A_{\alpha,\beta}^m\) and \(\mathbf{C} \equiv A_{\gamma,\delta}^m\), each defined as above. We write \(e_i\) for the vector with non-zero entries 1 at position 1, and \(-1\) at position \(i\) (\(2 \leq i \leq n\)), \(e_i\), and \(e_j\), for the two solutions of the quadratic equation \(\lambda^2 + (a_i + b)\lambda + c = 0\), where \(a = \alpha_1 - \beta_2\), \(b = \alpha_2 - \beta_2\) and \(c = \gamma - \delta\), and use \((\nu|\nu)\) to represent the concatenation of two vectors.

| BSM          | Eigenvalues | Multiplicity | Eigenvectors |
|--------------|-------------|--------------|--------------|
| \(K_{\alpha,\beta}^n\) | \(-\alpha + \beta\) | \(n-1\) | \(e_1\) |
| \(K_{\alpha,\beta}^n \ast K_{\gamma,\delta}^m\) | \(-\gamma + \delta\) | \(m-1\) | \(e_1\) |

### Table 2: Redundant spectra of basic symmetric motifs (BSMs) with one or two orbits.

A BSM with one orbit is a uniform graph \(K_{\alpha,\beta}^n\) with \(n\) vertices and adjacency matrix \(A_{\alpha,\beta}^n = (a_{ij})\), where \(a_{ij} = \alpha\) if \(i \neq j\) and \(a_{ii} = \beta\), for all \(i\), \(j\) and some constants \(\alpha\) and \(\beta\). A BSM with two orbits consists of the \((\gamma, \delta)\)-uniform join of two uniform graphs \(K_{\alpha,\beta}^n\) and \(K_{\gamma,\delta}^m\), that is, the graph with \(2n + m\) vertices and block adjacency matrix (after a suitable labelling of the vertices)

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix},
\]

where \(A = A_{\alpha,\beta}^n\), \(B = A_{\alpha,\beta}^m\) and \(\mathbf{C} \equiv A_{\gamma,\delta}^m\), each defined as above. We write \(e_i\) for the vector with non-zero entries 1 at position 1, and \(-1\) at position \(i\) (\(2 \leq i \leq n\)), \(e_i\), and \(e_j\), for the two solutions of the quadratic equation \(\lambda^2 + (a_i + b)\lambda + c = 0\), where \(a = \alpha_1 - \beta_2\), \(b = \alpha_2 - \beta_2\) and \(c = \gamma - \delta\), and use \((\nu|\nu)\) to represent the concatenation of two vectors.
become, in turn, eigenvectors of $A$ by repeating their values on each orbit, and can be obtained mathematically by left multiplying by the characteristic matrix $S$. Then, for each motif, we compute the redundant eigenpairs using a null space matrix (explained below), storing eigenvalues and localised (zero outside the motif) eigenvectors.

Only redundant eigenvectors of a symmetric motif (that is, those which add up to zero on each orbit) become eigenvectors of $A$ by extending them as zero outside the symmetric motif. Therefore, we need to construct redundant eigenvectors from the output of $e \Sigma g$ on each motif (the spectral decomposition of the corresponding submatrix). If $U_j = (v_1 \ldots v_k)$ are $\lambda$-eigenvectors of a symmetric motif with characteristic matrix of the orbit partition $S_{sm}$, we need to find linear combinations such that

$$S_{sm}^T(\alpha_1v_1 + \ldots + \alpha_kv_k) = 0 \iff S_{sm}^TU_\lambda \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = 0.$$  

(19)

Therefore, if the matrix $Z \neq 0$ represents the null space of $S_{sm}^TU_j$, that is, $S_{sm}^TU_jZ = 0$ and $Z^TZ = 0$, then the columns of $U_jZ$ are precisely the redundant eigenvectors. This is implemented in Algorithm 3 within the innermost for loop.

**Algorithm 3.**

Eigendecomposition algorithm.

**Input:** adjacency matrix $A$, characteristic motif $S$, list of motifs  
**Output:** spectral decomposition $A = UDUT^T$

initialize $U$, $D$ to zero matrices  
$A \leftarrow \Sigma g(\Sigma g(S))$  
$B_{sym} \leftarrow A^{-1/2}S^TAS^{-1/2}$  
$[U_q, D_q] \leftarrow e \Sigma g(B_{sym})$ so that $B_{sym} = U_qD_qU_q^{-1}$  
$U_q \leftarrow AU_q$  
$U \leftarrow (SU_q | 0)$  
$D \leftarrow (D_q | 0)$

for each motif $m$ do

$$A_{sm} \leftarrow A(\text{motif, motif})$$  
compute orbits from motif and $S$  
$S_{sm} \leftarrow S(\text{motif, motif})$  
$[U_{sm}, D_{sm}] \leftarrow e \Sigma g(A_{sm})$

for $\lambda$ in unique(eig(D_{sm})) do

$$U_j \leftarrow \lambda \text{-eigenvectors from } U_{sm}$$  
$Z \leftarrow \text{null}(S_{sm}^TU_j)$  
$d \leftarrow \text{null}(Z)$

if $d > 0$ then

store $U_jZ$ in $U$  
store $\lambda$ in $D$ with multiplicity $d$

end

end

**Vertex measures.** We have so far considered network measures of the form $F(i, j)$, where $i$ and $j$ are vertices. However, many important network measurements are vertex based, that is, of the form $G(i)$ for each vertex $i$. We say that a vertex measure $G$ is **structural** if it only depends on the network structure and, therefore, satisfies

$$G(i) = G(\sigma(i))$$  

(20)

for each automorphism $\sigma \in \text{Aut}(G)$, that is, it is constant on orbits (Fig. 1).

Although for vertex measures we do not have a network representation, we can still exploit the network symmetries. First, $G$ needs only to be computed/stored once per orbit, resulting on a reduction/compression ratio of $\bar{n} = n_j/n_g$ (Table 1).

Secondly, when quotient recovery holds (that is, we can recover $G$ from its values on the quotient and symmetry information alone), it amounts to a further computational reduction (Fig. 5), depending on the computational complexity of $G$. Finally, many vertex measures arise nevertheless from a pairwise function, such as $G(i) = F(i, i)$ (subgraph centrality from communicability), or $G(i) = 1/\beta \sum F(i, j)$ (closeness centrality from shortest path distance), allowing the symmetry-induced results on $F$ to carry over to $G$.

**Applications.** We illustrate our methods on several popular pairwise and vertex-based network measures. Although novel and of independent interest, these are example applications: Our methods are general and the reader should be able to adapt our results to the network measure of their interest.

**Adjacency matrix:** the methods in this paper can be applied to the network itself, that is, to its adjacency matrix. We recover the structural and spectral results in refs. [11,13] and the quotient compression ratio reported in ref. [12], here $c_{\text{spare}} = m_j$ in Table 1. The network (adjacency) eigendecomposition can be significantly sped up by exploiting symmetries (Fig. 5).

**Communicability:** communicability is a very general choice of structural measure, consisting on any analytical function $f(x) = \Sigma_{n=0}^{\infty} a_n x^n$ applied to the adjacency matrix, $f(A) = \Sigma_{n=0}^{\infty} a_n A^n$, and it is a natural measure of network connectivity, since the matrix power $A^k$ counts walks of length $k^{[37]}$. The most common choice of coefficients is $a_n = \frac{1}{n!}$, which gives the exponential matrix $e^A = \Sigma_{n=0}^{\infty} \frac{A^n}{n!}$. Communicability is a structural network measure and its network representation, the graph $f(G)$ with adjacency matrix $f(A)$, inherits all the symmetries of $G$ and thus it has the same symmetric motifs and orbits. The BSMs are uniform joints of orbits, and each orbit is a uniform graph (Figs. 3 and 2b) characterised by the communicability of a vertex to itself (a natural measure of centrality$^{[36]}$), and the communicability between distinct vertices. As a full network measure, the compression ratio $c_{\text{null}}$ applies (Table 1), indicating the fraction of storage needed by using the quotient to eliminate redundancies (Fig. 4). Moreover, average quotient recovery holds for communicability since $f(\Sigma G(A)) = f(\Sigma A)$ (Methods, Theorem 6). Alternatively, we can use the spectral decomposition algorithm on the adjacency matrix ($A = UDUT^T$) implies $f(A) = \Sigma f(D)U^T$ reducing the computation, typically cubic on the number of vertices, by $sp = \bar{n}_g$ (Table 1 and Fig. 5). For the spectral results, note that $f(A) = \Sigma f(D)U^T$ has eigenvalues $f(\lambda)$, and same eigenvectors, as $A$. Thus,

$$f(-2), f(-\varphi), f(-1), f(0), f(\varphi - 1), \text{ and } f(1)$$  

(21)

account for most of the discrete part of the spectrum $f(A)$, for the adjacency matrix $A$ of a typical (undirected, unweighted) real-world network (Eq. (18)).

**Shortest path distance:** this is the simplest metric on a (connected) network, namely the length of a shortest path between vertices. A path of length $n$ is a sequence $v_1, v_2, \ldots, v_{n+1}$ of distinct vertices, except possibly $v_1 = v_{n+1}$, such that $v_i$ is connected to $v_{i+1}$ for all $1 \leq i \leq n - 1$. The shortest path distance $d^*(u, v)$ is the length of the shortest (minimal length) path from $u$ to $v$. If $p = (v_1, v_2, \ldots, v_n)$ is a path and $\sigma \in \text{Aut}(G)$, we define $\sigma(p) = (\sigma(v_1), \sigma(v_2), \ldots, \sigma(v_n))$, also a path since $\sigma$ is a bijection.
One can show that (i) automorphisms preserve shortest paths and their lengths; (ii) shortest paths between vertices in different symmetric motifs do not contain intra-orbit edges; and (iii) shortest path distance is a partially quotient recoverable structural measure (“Methods”, Theorem 7). In particular, automorphisms σ preserve the shortest path metric, \( d(i, j) = d(σ(i), σ(j)) \), and we can compute shortest distances from the quotient,

\[
d^Q(α, β) = d^Q(i, j), \quad α ∈ V_i, β ∈ V_j,
\]

whenever \( V_i \) and \( V_j \) are orbits in different symmetric motifs. This accounts for all but the (small) intra-motif distances and reduces the computation as shown in Fig. 5.

Distances between points within the same motif cannot in general be directly recovered from the quotient, not even for BSMs. (Consider for instance the double star, motif \( M_1 \), in Fig. 1: The distance from the top red to the bottom blue vertex is three, while in the quotient is one.) In general, therefore, the shortest path distance is partially, but not average, quotient recoverable. Intra-motif distances, if needed, could still be recovered one motif at a time.

Note that these results can be exploited for other graph-theoretic notions defined in terms of distance, for example eccentricity (and thus radius or diameter), which only depends on maximal distances and thus it can be computed directly in the quotient.

In terms of symmetry compression, the compression ratio \( e_{\text{full}} \) applies, accounting for the amount of structural redundancy due solely to symmetries. The spectral results, although perhaps less relevant, still apply for \( d(δ) \), the graph encoding pairwise shortest path distances. The adjacency matrix \( d(A) = (d^Q(i, j)) \) is non-zero outside the diagonal, hence \( d(δ) \) is a all-to-all weighted network without self-loops and integer weights, and so is each symmetric motif. Using the formula in Table 2, we can easily compute values of the most significant part of the discrete spectrum (redundant eigenvalues) of \( d(A) \), namely \(-3, -2, -1, 0, -2 \pm \sqrt{2}, -3 \pm \sqrt{2}, -\frac{1}{2} \sqrt{5}, -\frac{3}{2} \sqrt{5} \), and \(-\frac{5}{2} \sqrt{13} \).

**Laplacian matrix:** The Laplacian matrix of a network \( L = D − A \), where \( D \) is the diagonal matrix of vertex degrees, is a (sparse) network measure and therefore inherits all the symmetries of the network. The matrix \( L \) can be seen as the adjacency matrix of a network \( \mathcal{L} \) with identical symmetric motifs, except that all edges are weighted by \(-1\) and all vertices have self-loops weighted by their degrees in \( \mathcal{G} \) (Fig. 2c). In particular, the motif structure (namely, the self-loop weights) depends on the how the motif is embedded in the network \( \mathcal{G} \).

Quotient compression and computational reduction are less useful in this case, however the spectral results are more interesting. The spectral decomposition applies, and we can compute redundant Laplacian eigenvalues directly from Table 2, for instance positive integers for BSMs with one orbit (“Methods”, Corollary 2). This explains and predicts most of the “peaks” (high-multiplicity eigenvalues) in the Laplacian spectral density, confirmed on our test networks (Fig. 7). Using the formula in Table 2, one can similarly compute the redundant spectrum for 2-orbit BSMs, and for other versions of the Laplacian (e.g., normalised, vertex weighted). Finally, observe that the spectral decomposition applies, thus Algorithm 3 provides an efficient way of computing the Laplacian eigendecomposition with an expected \( sp = 3n^2 \) (see Table 1) computational time reduction.

**Commute distance and matrix inversion:** The commute distance is the expected time for a random walker to travel between two vertices and back44. In contrast to the shortest path distance, it is a global metric, which takes into account all possible paths between two vertices. The commute distance is equal up to a constant (the volume of the network) to the resistance metric \( r^P \), which can be expressed in terms of \( L^P = (l^P)v \), the pseudo-inverse (or Moore-

\[
\begin{align*}
\text{HumanDisease} & \quad \text{HumanPPI} & \quad \text{OpenFlights} \\
0.2 & \quad 0.2 & \quad 0.2 \\
0.1 & \quad 0.1 & \quad 0.1 \\
0.0 & \quad 0.0 & \quad 0.0 \\
5 & \quad 5 & \quad 5 \\
10 & \quad 10 & \quad 10 \\
20 & \quad 20 & \quad 20 \\
\end{align*}
\]

**Fig. 7 Spectral signatures of network symmetry.** Laplacian spectrum of six test networks (blue) and of their quotient (red), given as relative probability of eigenvalue count, with multiplicity, in bins of size 0.1. Only the most significant part of the spectrum is shown. Most of the “peaks” observed in the spectral density occur at positive integers, as predicted. (Insets) Percentage of \( m_i \), for the quotient eigenvalues, and for the Laplacian eigenvalues, where \( m_i \) is the multiplicity of an eigenvalue \( λ \) rounded to 8 decimal places.

**Penrose inverse** of the Laplacian, as \( r(i, j) = l^P_{ii} + l^P_{jj} − 2l^P_{ij} \). The commute (or resistance) distance is a (full) structural measure, and all our structural and spectral results apply. Crucially, we can use eigendecomposition algorithm to obtain \( L = UDU^T \) (and hence \( I^T = U^D^T U^T \) ), and \( r \) from the quotient and symmetric motifs, resulting in significant computational gains (Fig. 5). More generally, if \( M_p \) is the matrix representation of a network measure, its pseudo-inverse \( M_p^+ \) is also a network measure, and the comments above apply. Note that \( M_p^+ \) is generally a full measure even if \( M_p \) is sparse (the inverse of a sparse matrix is not generally sparse).

**Vertex symmetry compression:** as a vertex measure \( G \) is constant on orbits, we only need to store one value per orbit. Let \( S \) be the characteristic matrix of the partition of the vertex set into orbits, and \( \Lambda \) the diagonal matrix of orbit sizes (column sums of \( S \)). If \( G \) is represented by a vector \( v = (G(i)) \) of length \( n_g \), we can store one value per orbit by taking \( w = \Lambda^{-1} S^Tv \), a vector of length \( n_j \), and recover \( v = S^T w \) (“Methods”, Theorem 9).

**Degree centrality:** the degree of a node \( (in- \) or out-degree if the network is directed) is a natural measure of vertex centrality. As expected, the degree is preserved by any automorphism \( σ \), which can also be checked directly,

\[
d_j = \sum_{\mathcal{J} \in V} a_{\mathcal{J}(j)} = \sum_{\mathcal{J} \in V} a_{\mathcal{J}(i)} = d(\sigma(i)), \quad (23)
\]
as automorphisms permute orbits (so \( j ∈ V \) and \( σ(j) ∈ V \) are the same elements but in a different order). In particular, the degree is constants on orbits. We recover the degree centrality from the quotient as the out-degree (“Methods”, Proposition 2).

**Closeness centrality:** the closeness centrality of a node \( i \) in a graph \( \mathcal{G} \), \( cc(\sigma(i)) \), is the average shortest path length to every node in the graph. As symmetries preserve distance, they also preserve closeness centrality, explicitly,

\[
cc(i) = \frac{1}{n_g} \sum_{\mathcal{J} \in V} d(i, j) = \frac{1}{n_g} \sum_{\mathcal{J} \in V} d(\sigma(i)),
\]

| HumanDisease | Yeast | OpenFlights |
|--------------|-------|-------------|
| 90.6%        | 94.2% | 89.2%       |
| 0.1          | 0.1   | 0.1         |
| 0.0          | 0.0   | 0.0         |
| 5            | 5     | 5           |
| 10           | 10    | 10          |
| 20           | 20    | 20          |

and centrality is constant on each orbit, as expected. Moreover, closeness centrality can be recovered from the quotient (shortest
paths does not contain intra-orbit edges, except between vertices in the same symmetric motif, see above), as
\[
cc^\text{F}(i) = \sum_{l,k} n_l n_g d_k k(l, i) + n_l n_g d_k k(l, i)
\]
if \(i\) belongs to the orbit \(V_k\) and \(d_k\) is the average intra-motif distance, that is, the average distances of a vertex in \(V_k\) to any vertex in \(H\), the motif containing \(V_k\). By annotating each orbit by \(d_k\), we can recover betweenness centrality exactly. Alternatively, as \(d_k \ll n\) (note that \(d_k \leq m\) if \(H\) has \(m\) orbits), we can approximate \(cc^\text{F}(i)\) by the first summand, or simply by the quotient centrality \(cc^\text{Q}(i)\), in most practical situations.

Betweenness centrality: this is the sum of proportions of shortest paths between pairs of vertices containing a given vertex. It can be computed from shortest path distances and number of shortest paths between pairs of vertices containing a given vertex. It can be made an inexpensive pre-processing step. We showed that the effects of network symmetry on arbitrary network measures, and the algorithmic aspects throughout, and provide pseudocode and full implementations. Since real-world network models and data are very common, and typically contain a large degree of structural redundancy, our methods should be relevant to any network practitioner.

**Methods**

**Geometric decomposition and symmetric motifs.** We write \(\text{Aut}(\mathcal{G})\) for the automorphism group of an (unweighted, undirected, possibly very large) network \(\mathcal{G} = (V, E)\) (see below for a discussion of directed and weighted networks). Each automorphism (symmetry) \(\sigma \in \text{Aut}(\mathcal{G})\) is a permutation of the vertices and its support is the set of vertices moved by \(\sigma\),
\[
\text{supp}(\sigma) = \{i \in V \text{ such that } \sigma(i) \neq i\}.
\]
Two automorphisms \(\sigma\) and \(\tau\) are support-disjoint if the intersection of their supports is empty, \(\text{supp}(\sigma) \cap \text{supp}(\tau) = \emptyset\). The orbit of a vertex \(i\) is the set of vertices to which \(i\) can be moved by an automorphism, that is,
\[
\{\sigma(i) \text{ such that } \sigma \in \text{Aut}(\mathcal{G})\}.
\]
One can show that there is a partition of generators of \(\text{Aut}(\mathcal{G})\) into its finest support-disjoint classes \(X = X_1 \cup \ldots \cup X_n\), which is unique up to permutation of the sets \(X_i\). The vertex sets \(M_i = \cup_{\sigma \in X_i} \text{supp}(\sigma)\) give the geometric decomposition Eq. (2), and the subgraphs induced by them are, by definition, the symmetric motifs of \(\mathcal{G}\). (The next section explains how to compute the geometric decomposition in practice.) Since support-disjoint automorphisms must commute (the order in which they are composed is irrelevant), the subgraphs of \(\text{Aut}(\mathcal{G})\) generated by \(X_i\) to \(X_n\), call them \(H_1\) to \(H_n\), give a direct product decomposition \(\text{Aut}(\mathcal{G}) = H_1 \times \ldots \times H_n\). The geometric decomposition is defined from the finest support-disjoint partition of a special set of generators (called essential), as explained in ref. 13. However, the results in this article are valid for any support-disjoint decomposition of any set of generators (essential or not) of \(\text{Aut}(\mathcal{G})\).

If all the orbits of a symmetric motif have the same size \(k\) and every permutation of the vertices in each orbit can be extended to a network automorphism supported on the motif, we call the symmetric motif basic (or BSM) of type \(k\). (In particular, the corresponding subgroup \(H_i\) must be Sym(\(k\)), the symmetric group of all permutations of \(k\) elements.) If a symmetric motif is not basic, we call it complex or of type 0.

**Network symmetry computation.** First, we compute a list of generators of the automorphism group from an edge list (we use saucy37, which is extremely fast for the large but sparse networks typically found in applications). Then, we partition the set of generators \(X\) into support-disjoint classes \(X = X_1 \cup \ldots \cup X_n\), that is, \(\sigma\) and \(\tau\) are support-disjoint whenever \(\sigma \in X_i\) and \(\tau \in X_j\) for \(i \neq j\). To find the finest such partition, we use a bipartite graph representation of vertices \(V\) and generators \(X\). Namely, let \(\mathcal{A}\) be the graph with vertex set \(V \times X\) and edges between \(i\) and \(\sigma\) whenever \(i \in \text{supp}(\sigma)\). Then \(X_1, \ldots, X_n\) are the connected components of \(\mathcal{A}\) (as vertex sets intersected with \(X\)). Each \(X_i\) corresponds to the vertex set \(M_i\) of a symmetric motif \(\mathcal{G}_i\), as \(M_i = \cup_{\sigma \in X_i} \text{supp}(\sigma)\). Finally, we use GAP48 to compute the orbits and type of each symmetric motif (Algorithm 5). Full implementations of all the procedures outlined above are available at a public repository.26

**Algorithm 5.**

Orbits and type of a symmetric motif.

Discussion

We have presented a general theory to describe and quantify the effects of network symmetry on arbitrary network measures, and explained how this can be exploited in practice in a number of ways.

Network symmetry of the large but sparse graphs typically found in applications can be effectively computed and manipulated, making it an inexpensive pre-processing step. We showed that the amount of symmetry is amplified in a pairwise network measure but can be easily discounted using the quotient network. We can for instance eliminate the symmetry-induced redundancies, or use them to simplify the calculation by avoiding unnecessary computations. Symmetry has also a profound effect on the spectrum, explaining the characteristic “peaks” observed in the spectral densities of empirical networks, and occurring at values we are able to predict.

Our framework is very general and apply to any pairwise or vertex-based network measure beyond the ones we discuss as examples. We emphasised practical and algorithmic aspects

**Input:** adjacency matrix \(A\), characteristic matrix \(\Sigma\)
**Output:** (right) Perron–Frobenius eigenvector \((\lambda, v)\) of \(A\)
\[
\begin{align*}
\Lambda & \leftarrow \text{diag}(\text{sum}(\Sigma)) \\
\Sigma & \leftarrow SA^{-1/2} \\
B_{\text{sym}} & \leftarrow \Sigma^{1/2}AR \\
(\lambda, w) & \leftarrow \text{eig}(B_{\text{sym}}, 1) \quad \text{eigenpair of the largest eigenvalue} \\
v & \leftarrow Rw
\end{align*}
\]

**Algorithm 4.**

Eigenvector centrality from the quotient network.

**Input:** \(X\) a set of permutations of a symmetric motif
**Output:** \(O_1, \ldots, O_k\) orbits, and type \(m\), of the symmetric motif
\[
\begin{align*}
H & \leftarrow \text{Group}(X) \\
(O_1, \ldots, O_k) & \leftarrow \text{Orbits}(H) \\
m & \left\{ \min(\text{size}(O_1), \ldots, \text{size}(O_k)) \right. \\
\text{if } m & \left\{ \max(\text{size}(O_1), \ldots, \text{size}(O_k)) \right. \\
\text{for } i & \left\{ 1 \text{ to } k \right. \\
\text{if } \not\exists \text{NaturalSymmetricGroup}(\text{Action}(H, O_i)) & \left. \text{then} \\n\quad m & \left. \left\{ 0 \right. \\
\quad \text{break} \\
\quad \text{end} \\
\quad \text{end} \\
\quad \text{end} \\
\text{else} & \\
\quad m & \left. 0 \right. \\
\quad \text{end} \\
\text{end} \\
\end{align*}
\]

**Algorithm 5.**

Orbits and type of a symmetric motif.
Structural network measures. We prove below the structural result for BSMs for arbitrary graphs and network measures. The proof is a generalisation of the argument on ref. 49 (p. 48) to weighted directed graphs with symmetries.

**Theorem 1** Let M be the vertex set of a BSM of a network $\mathcal{G}$, and F a structural network measure. Then the graph induced by M in $F(\mathcal{G})$ is a BSM of $F(\mathcal{G})$, and satisfies:

(i) for each orbit $\Delta = \{v_1, \ldots, v_n\}$, there are constants $\alpha$ and $\beta$ such that the orbit internal connectivity is given by $= F(v_1, v_2)$ for all $i$ and $j$.

(ii) if every orbit $\Delta_i$ and $\Delta_j$ is a labelling $\Delta_i = \{v_1, \ldots, v_n\}$, $\Delta_j = \{w_1, \ldots, w_m\}$ and constants $\alpha_{ij}$, $\beta_{ij}$ such that $F(v_1, v_i) = F(v_i, v_j)$, $F(v_i, v_j) = F(w_1, v_i)$, $\alpha_{ij} = F(w_i, v_j)$, $\beta_{ij} = F(v_i, w_j)$, for all $i,j$.

(iii) every vertex $v$ not in the BSM is joined uniformly to all the vertices in each orbit $\{v_1, \ldots, v_n\}$ in the BSM, that is, $F(v, v_1) = F(v, v_2) = F(v, v_3) = F(v, v_4) = F(v, v_5)$ for all $v$.

Moreover, property (iii) holds in general for any symmetric motif.

If $\mathcal{G}$ is undirected and F is symmetric, $\alpha_{ij}=\beta_{ij}$ and each orbit is a $(\alpha, \beta)$-uniform graph $K^\alpha_\beta$ and each pair of orbits form a $(\gamma, \delta)$-uniform join, explaining Fig. 3a, b.

**Proof** As $F(\mathcal{G})$ inherits all the symmetries of $\mathcal{G}$, M has the same orbit decomposition and the symmetric group $S_n$ acts in the same way, hence M induces a BSM in $F(\mathcal{G})$ too. For the internal connectivity, note that every permutation of the vertices $v_1$ is realisable. Thus, given arbitrary $1 \leq i, j, k, l \leq n$, we can define $\sigma \in Aut(\mathcal{G})$ such that $\sigma(v_i) = v_j$, and, if $j \neq k$ additionally satisfies $\sigma(v_l) = v_k$. This gives $F(v_i, v_j) = F(\sigma(v_i), \sigma(v_j)) = F(v_k, v_l)$, as $F$ is a structural network measure. The other case, $i = j$ and $k = l$, gives $F(v_i, v_l) = F(\sigma(v_i), \sigma(v_l))$. For the orbit connectivity result (iii), we generalise the argument in ref. 49 (p. 48) to weighted directed graphs with symmetries, particularly ($F(\mathcal{G})$). We assume some basic knowledge and terminology about graph actions\(^{50}\) and symmetric groups $S_n$. Given two orbits $\Delta = \{v_1, \ldots, v_n\}$ and $\Delta_2 = \{w_1, \ldots, w_m\}$ and $1 \leq i < n$, define

$$\Gamma_i = \{w_j \in \Delta_2 : F(v_i, w_j) \neq 0\},$$

the vertices in $\Delta_2$ joined to $v_i$ in $F(\mathcal{G})$. If a finite group G operates on a set X, the stabiliser of a point $G_x = \{g \in G : gx = x\}$ is a subgroup of G of index $[G : H] = \frac{|G|}{|H|}$ equals to the size of the orbit of x. Hence, the stabilizers $G_{v_i}$ or $G_{w_i}$ are subgroups of $S_n$ of size $n$, for all $i,j$. The group $S_n$ has a unique, up to conjugation, subgroup of index $n$ if $n \neq 6$. In this case, $G_{v_i}$ is conjugate to $G_{w_i}$, so $G_{v_i} = G_{w_i} = e^{-1}G_{v_i}$. For some $\sigma \in S_n$. Relabelling $w_i$ as $v_i$ we have $G_{v_i} = G_{w_i}$. Similarly, we can relabel the remaining vertices in $\Delta_2$ so that $G_{v_i}$ is the same for all $i$: write $v_i = v_i$, $v_j = v_j$, $v_k = v_k$, and relabel $w_j = w_j$, $w_k = w_k$, $w_l = w_l$, noticing there cannot be repetitions as $\sigma(v_i) = \sigma(v_k)$, for $k \neq l$ implies $\sigma(v_j) \neq v_j$, a contradiction. Fix $1 \leq i < n$. The stabilizer $G_{v_i}$ fixes $v_i$, but it may permute vertices in $\Delta_2$. In fact, the set $\Gamma_i$ above must be a union of orbits of $G_{w_j}$ on $\Delta_2$, if $w_j \in \Gamma_i$ and $\sigma \in G$, then, $0 \neq F(\sigma(v_i), w_j) = F(\sigma(v_i), \sigma(w_j))$ so also belongs to $\Gamma_i$. The orbits of $G_{v_i}$ on $\Delta_2$ are $\{w_j\}$ and $\Delta_2 \setminus \{w_j\}$, as $G_{v_i}$ fixes $w_j$ and freely permutes all other vertices in $\Delta_2$. The case $n = 6$ is similar, except that $S_6$ has two conjugacy classes of subgroups of index 6, one as above, and the other a subgroup acting transitively on the 6 vertices, which gives a unique orbit $\Delta_2$. In all cases, the set $\Delta_2 \setminus \{w_j\}$ is part of an $\mathcal{G}$-orbit, which gives the connectivity result, as follows. Fix $1 \leq i < n$. For $1 \leq j, k, l \leq n$ different from $i$, the vertices $w_j$ and $w_k$ are in the same $G_{v_i}$-orbit so there is $w_l \in G_{v_i}$ with $w_l = w_l$ and, therefore, $F(v_i, w_k) = F(\sigma(v_i), \sigma(w_k)) = F(v_i, w_k)$. The argument is general, so we have shown $a_{ij} = A_{i,j} = A_{i,j}$ for all $i \neq j$. It is enough to show $a_{ii} = \alpha_{ii}$. Choose $j \neq i$, then

$$a_{ij} = F(v_i, v_j) = F(\sigma(v_i), \sigma(v_j)) = F(v_i, \sigma^{-1}(\sigma(v_j)) = t_{ij},$$

as long as $t_{ij} = t_{ij}$, which cannot happen as otherwise $t_{ij} = t_{ij}$ implies $t_{ij} = t_{ij} = t_{ij}$, a contradiction. Hence, we have shown $F(v_i, v_j)$ is a constant, call it $\gamma_{ij}$, for all $i, j$. In addition, $F(v_i, w_j) = F(\sigma(v_i), \sigma(w_j)) = F(v_i, w_j)$ is also a constant, call it $\delta_{ij}$, for all $i, j$. The cases $y_j \geq F(v_i, v_j)$ and $\delta_{ij} \leq F(v_i, v_j)$ are identical, reversing the roles of $\Delta_1$ and $\Delta_2$. Property (iii) holds for any symmetric motif, not necessarily basic, as follows. By the definition of orbit, for each pair (i, j), we can find an automorphism $\sigma$ in the geometric factor such that $\sigma(v_i) = v_j$. Since $v$ is not in the support of that geometric factor, it is fixed by $\sigma$, that is, $\sigma(v_i) = v_j$. Therefore, $F(v_i, v_j) = F(\sigma(v_j), \sigma(v_i)) = F(v_j, v_i)$, and similarly $F(v_i, v_j) = F(\sigma(v_j), \sigma(v_i))$. $\square$

**Average compression.** Theorem 2 Let $A = (a_{ij})$ be the $n \times n$ adjacency matrix of a network with vertex set $V$. Let $S$ be the $n \times n$ characteristic matrix of the restriction of $\mathcal{G}$ to $\Delta_1 = \{v_1, \ldots, v_n\}$ and $\Delta_2 = \{w_1, \ldots, w_m\}$ and $S = \frac{1}{2}F(v_i, v_j)$. Then

(i) if $i, j \in V$ belong to different symmetric motifs, $\bar{a}_{ij} = a_{ij}$;

(ii) if $i, j \in V$ belong to orbits in $\Delta_1$ and $j \in \Delta_2$ in the same symmetric motif, $\\frac{1}{\Delta_1} \sum_{i,j \in \Delta_1} a_{ij}$

$$= \frac{1}{\Delta_2} \sum_{i,j \in \Delta_2} a_{ij}.$$

**Lossless compression.** We can achieve lossless compression by exploiting the structure of BSMs, which account for most of the symmetry in real-world networks. If the motif is basic, we can preserve the exact parent network connectivity in an annotated quotient, as follows. Each orbit in a BSM is a uniform graph $K^\alpha_\beta$, which appears in the quotient as a single vertex with a self-loop weighted by $(\alpha - 1) + \beta$ (Fig. 3c). Hence, if we annotate this vertex in the quotient by not only $n$ but also $\alpha$, or $\beta$, we can recover the internal connectivity. Similarly, the connectivity between orbits in the same symmetric motif is given by two parameters $\gamma$, and $\delta$, appears in the quotient as an edge weighted by $(n - \gamma) + \delta$ (Fig. 3d) and thus can also be recovered from a quotient with edges annotated by $\gamma$, or $\delta$.

Since there is no general formula for an arbitrary non-basic symmetric motif, we can work with the basic quotient $\mathcal{G}$ and instead, that is, the quotient with respect to the partition of the vertex set into orbits in BSMs only (vertices in non-basic symmetric motifs become fixed points hence part of the asymmetric core). The annotated (as above) basic quotient achieves most of the symmetry reduction in a typical empirical network (Fig. 2). To maintain the same number of vertices in the parent network, we record, for each pair of orbits in the same symmetric motif, the corresponding permutation of the second orbit (else we recover the adjacency matrix only up to permutations of the orbits).

Algorithms for lossless compression and recovery based on the basic quotient are shown below (Algorithms 6 and 7), and MATLAB implementations for BSMs up to two orbits are available at a public repository\(^{59}\). The results reported in Fig. 4 are with respect to these implementations, and the actual compression ratios reported include the size of the annotation data for lossless compression with vertex identity (a very small fraction of the size of the quotient in practice, adding at most 0.02% to the basic full compression ratio in all our test cases).
Algorithm 6.
Lossless symmetry compression.

Input: adjacency matrix $A$, characteristic matrix for the basic quotient $S$, list of BSMs motifs
Output: quotient matrix $B$, annotation structure $a$

$B \leftarrow S^T A S$
extract orbits from $S$
foreach orbit in orbits do
  $\beta \leftarrow A(\text{rep}, \text{rep})$
store $\beta$ in annotation structure $a$
end
$k_{\text{max}} \leftarrow \max(\text{size(molfs)})$ maximal number of orbits in a motif
for $k \leftarrow 2$ to $k_{\text{max}}$ do
  extract $k$-BSM (list of BSMs with $k$ orbits) from motifs
  foreach pairs of distinct orbits $V_1, V_2 \in \text{bsm}$ do
    compute $\delta$ and permutation of $V_1$ perm such that $A(k, \text{perm}(k)) = \delta$ for all $k \in V_1$
    store orbit numbers (with respect to $S$), $\delta$ and perm in annotation structure $a$
  end
end

Algorithm 7.
Lossless symmetry decompression.

Input: quotient matrix $B$, characteristic matrix $S$, annotation structure $a$
Output: adjacency matrix $A$

$A \leftarrow \text{diag}(\text{sum}(S))$
$R \leftarrow S A^{-1}$
$A \leftarrow R B R^T$
extract orbits from $S$
foreach orbit in orbits do
  $n \leftarrow \text{size(orbit)}$
  extract $\beta$ from $\alpha$
  compute $\alpha \leftarrow A, \beta$ and $n$ (using $||B||_{\text{orbit}} = n/((n-1)\alpha + \beta)$)
  construct adjacency matrix of the orbit $A_{\alpha, \beta}^{a, \beta}$
  $A(\text{orbit, orbit}) \leftarrow A_{\alpha, \beta}^{a, \beta}$
end
extract pairs of orbits in the same BSM from $a$
foreach ($V_1, V_2$) in pairs do
  $n \leftarrow \text{size}(V_1)$
  extract $\delta$, perm from $\alpha$
  compute $\gamma$ from $\beta$, $\delta$ and $n$ (using $||B||_{V_1, V_2} = n/((n-1)\gamma + \delta)$)
  construct matrix $A_{\alpha, \beta}^{\gamma, \delta}$
  $A(V_1, \text{perm}) \leftarrow A_{\alpha, \beta}^{\gamma, \delta}$
  $A(\text{perm, V1}) \leftarrow A_{\alpha, \beta}^{\gamma, \delta}$
end

Spectral signatures of symmetry. The partition into orbits satisfy the following regularity condition\cite{14, 35}. A partition of the vertex set $V = V_1 \cup \ldots \cup V_m$ is equitable if

$$\sum_{i,j \in V_i} a_{ij} = \sum_{j,k \in V_k} a_{jk} \quad \text{for all } i, j \in V_i, \text{ for all } 1 \leq k, l \leq m,$$

Proposition 1 Let $V = V_1 \cup \ldots \cup V_m$ be a partition of the vertex set of a graph with adjacency matrix $A = (a_{ij})$, and let $S$ be the characteristic matrix of the partition. Write $Q(A)$ for the quotient with respect to the partition.

(i) The partition is equitable if and only if $A = SQ(A)$;
(ii) The partition into orbits of the automorphism group is equitable.

Proof (i) Fix $1 \leq i \leq n$ and $1 \leq k, l \leq m$, and suppose $i \in V_l$. Then

$$[AS]_{ii} = \sum_{j \in V_l} a_{ij},$$

and, using the equitable condition,

$$[SQ(A)]_{ii} = [Q(A)]_{ii} = \frac{1}{|V_l|} \sum_{j \in V_l} a_{ij} = \frac{1}{|V_l|} |V_l| \sum_{j \in V_l} a_{ij} = \sum_{j \in V_l} a_{ij}.$$

For the converse, note that $[AS]_{ii}$ does not depend on $i$ but on the orbit of $i$. Namely, given $i_1, i_2 \in V_i$, we have

$$\sum_{j \in V_l} a_{ij} = [AS]_{ij} = [Q(A)]_{ij} = \sum_{j \in V_l} a_{ij}.\text{ for all } i, j,$$

where the latter equality follows from the fact that an element in a group permutes orbits, in this case, $|j : j \in \Delta_i| = (|0 : j \in \Delta_i|)$. Hence, the partition into orbits is equitable. □

It follows immediately that the quotient eigenvalues are a subset of the eigenvalues of the parent network,

$$Q(A)v = \lambda v = A(Sv) = SQ(A)v = \lambda Sv.$$

(31)

(Note that $Sv = 0$ if $v = 0$.) That is, the spectrum of the quotient is a subset of the spectrum of the graph, with eigenvectors lifted from the quotient by repeating entries on orbits. Moreover, we can complete an eigenvector with eigenvectors orthogonal to the partition (adding up to zero on each orbit).

Theorem 3 Suppose that $A$ is an $m \times m$ real symmetric matrix and $B$ the $m \times m$ quotient matrix with respect to an equitable partition $V_1 \cup \ldots \cup V_m$ of the set $\{1, 2, \ldots, n\}$. Let $S$ be the characteristic matrix of the partition. Then $A$ has an eigenbasis of the form

$$\{Sv_1, \ldots, Sv_m, w_1, \ldots, w_{m-1}\},$$

where $\{v_1, \ldots, v_m\}$ is any eigenbasis of $B$, and $S^T w_i = 0$ for all $i$. □

Proof Recall that $Sv_i = 0$ if $v_i = 0$ (5 lifts the vertex $v$ from the quotient by repeating entries on each orbit) so the linear map

$$\mathbb{R}^m \rightarrow \mathbb{R}^n, v \rightarrow Sv$$

has trivial kernel and hence it is an isomorphism onto its image. In particular, $S = \{Sv_1, \ldots, Sv_m\}$ is also a linearly independent set, and they are all eigenvectors of $A$, since $AS = B$ as the partition is equitable. To finish the proof we need to complete $S$ to a basis $\{Sv_1, \ldots, Sv_m, w_1, \ldots, w_{m-1}\}$ such that each $w_i$ is an $A$-eigenvector orthogonal to all $Sv_i$. As $S$ is a basis of $\text{Im}(S)$, this would imply $w_i \in \text{Im}(S^T)$, giving $S^T w_i = 0$ for all $i$, as desired. Since $A$ is diagonalizable, $E^R$ decomposes as an orthogonal direct sum of eigenspaces, $E^R = \bigoplus E_\lambda$. In each $E_\lambda$, we can find vectors $w_i$ such that they complete $V_i = \{Sv_i \in \mathbb{R}^m \mid \lambda - \text{eigenvector}\}$ to a basis of $E_\lambda$ and that are orthogonal to all vectors in $V_j$ (consider the orthogonal complement of the subspace generated by $V_i$ in $E_\lambda$). Repeating this procedure on each $E_\lambda$, we find vectors $\{w_1, \ldots, w_{m-1}\}$ as needed. □

The statement and proof above holds for arbitrary matrices $A$ by replacing “eigenvectors” by “maximally linearly independent set” and removing the condition $S^T w_i = 0$. It would be interesting to know whether the condition $S^T w_i = 0$ holds for motif eigenvectors in the directed case as well (the proof above is no longer valid).

Theorem 4 Let $\mathcal{M}$ be a symmetric motif of a (possibly weighted) undirected graph $G$. If $(\lambda, w)$ is a redundant eigenpair of $\mathcal{M}$ then $(\lambda, \tilde{w})$ is an eigenpair of $G$, where $w$ is equal to $\tilde{w}$ on the vertices of $\mathcal{M}$, and zero elsewhere.

Proof Since $(\lambda, v)$ is an $\mathcal{M}$-eigenpair,

$$\sum_{i \in V(\mathcal{M})} [A_{\mathcal{M}}]_{i} v_i = \lambda v,$$

where $A_{\mathcal{M}}$ is the adjacency matrix of $\mathcal{M}$. We can decompose $\mathcal{M}$ into orbits, $V(\mathcal{M}) = V_1 \cup \ldots \cup V_m$, and, by the spectral decomposition theorem above applied to $\mathcal{M}$, $w$ is orthogonal to each orbit, that is,

$$\sum_{i \in V(\mathcal{M})} w_i = 0, \quad \forall 1 \leq i \leq m.$$

We need to show that $(\lambda, \tilde{w})$ is a $\mathcal{G}$-eigenpair. Let us write $A$ for the adjacency matrix of $\mathcal{G}$ (recall $\mathcal{M}$ is a subgraph so $A$ restricts to $A_{\mathcal{M}}$ on $\mathcal{M}$). We need to show
The jth entry of the vector $A\nu$ is
\[\begin{align*}
&k\beta_j - \kappa\alpha_j + \delta - (\kappa\alpha_j + c) 
&= j = 1 \\
&\kappa\alpha_j - k\beta_j + \delta = \kappa\alpha_j + c 
&= i \\
&k\delta - k\nu_j + \beta_j - \alpha_j = - (\kappa\beta_j + b) 
&= j = n + 1 \\
&\kappa\nu_j - k\delta + \alpha_j - \beta_j = \kappa\beta_j + b 
&= j = n + i \\
&0 
&\text{otherwise.}
\end{align*}\]
Comparing these with the entries of the vector $\lambda\nu_j$, we obtain
\[A\nu_j = \lambda\nu_j \iff (k\alpha_j + c) - \lambda\kappa\nu_j + \kappa\beta_j + b = \lambda\kappa\nu_j + \kappa\beta_j + b - (k\alpha_j + c), \tag{34}\]

The two equations on the right-hand side are satisfied if and only if $\lambda = -\kappa\beta_j + b$ and $k$ is a solution of the quadratic equation
\[ck^2 + (b - a)k - c = 0, \tag{35}\]
which has two distinct real solutions
\[
\begin{align*}
(a - b) + \sqrt{(a - b)^2 + 4c^2} \\
2c
\end{align*}
\]
when $c < 0$, as explained above. Altogether with the lemma, we have shown the following.

**Theorem 5** The redundant spectrum of a symmetric motif with two orbits $K_{\alpha, \beta}^{+1} \leftrightarrow K_{\alpha, \beta}^{-1}$ is given by the eigenvalues
\[\lambda_i = -b - c\kappa_i = -(a + b) + \sqrt{(a - b)^2 + 4c^2} = \frac{2c}{2}, \tag{36}\]
with each multiplicity $n - 1$, and eigenvectors $(k_1, e(c_1), e(c_2), \ldots, e(c_{n-1}))$, respectively, where $k_1$ and $k_{n-1}$ are the solutions of the quadratic equation $c^2 + (b - a)k - c = 0$, $a = a - \beta_j, b = \beta_j - \beta_j$ and $c = \gamma - \delta = 0$. For unweighted graphs without loops, we recover the redundant eigenvalues for BSMs with two orbits predicted in ref. \[12\], as follows. We have $\beta_1 = \beta_j = 0, a = a_j, \beta_j, \delta \in [0, 1]$, and thus $a_j > 0$, $b_j > 0$ and $\gamma - \delta = 0$. In either case, the redundant eigenvalues $\lambda = (b - c) - c\kappa = -\lambda - \psi - 1$. Altogether, the redundant eigenvalues for 2-orbit BSMs are $\{-b, -\psi - 1, 0, -\gamma - 1\}$, which equals the redundant eigenvalues $\lambda_{\text{Spec}}$ in the notation of ref. \[13\]. We omit the calculation of the redundant spectrum of BSMs with three (or more) orbits, as it becomes much more elaborate, and its relevance in real-world networks is less justified (for example, $\leq$% of BSMs in each of our test networks, Table 1, has three or more orbits).

**Applications**

**Theorem 6** (Communicability) Let $Q(A)$ be the quotient of the adjacency matrix $A$ of a network with respect to the partition into orbits of the automorphism group. Let $f(x) = x^{\omega}e^{\chi x}$ be an analytic function. Then $Q(f(A)) = f(\lambda_{\text{Spec}})$.

**Proof** Call $B = QA$ and recall that $AS = SB$ by Proposition 1(i), where $S$ is the characteristic matrix of the partition. Therefore, $A^{\omega}S = SB^\omega$ for all $n \geq 0$ and
\[
Q(f(A)) = A^{-1}S^{-1}B^\omega = \sum_{n=0}^{\infty} a_n A^\omega S B^\omega = \sum_{n=0}^{\infty} a_n S B^\omega = f(B), \tag{38}\]

since $A^{\omega}S$ is the identity matrix.

**Theorem 7** (Shortest path distance) Let $A = \langle a_j \rangle$ be as above. Then
\[\begin{align*}
&\text{if } \langle v_1, v_2, \ldots, v_s \rangle \text{ is a shortest path from } v_1 \text{ to } v_s \text{ and } \sigma \in \text{Aut } (\mathcal{G}), \text{ then} \\
&\langle \sigma(v_1), \sigma(v_2), \ldots, \sigma(v_s) \rangle \text{ is a shortest path from } \sigma(v_1) \text{ to } \sigma(v_s) \text{ for all } \\
&\text{symmetric motifs, then } v_1 \text{ and } v_s \text{ belong to different orbits, for all} \\
&\text{such paths} \langle w_1, w_2, \ldots, w_n \rangle \text{ with } \sigma \in \text{Aut } (\mathcal{G}). \text{ Then} \\
&\text{if } u \text{ and } v \text{ belong to orbits } U \text{ and } V, \text{ respectively, in different symmetric motifs, then} \\
&\text{the distance from } u \text{ to } v \text{ is equal to the distance from } U \text{ to } V \text{ in the unweighted (or quotient) graph} \mathcal{G}. \text{ Then} \\
&\text{if } u \text{ and } v \text{ belong to orbits } U \text{ and } V, \text{ respectively, in different symmetric motifs, then} \\
&\text{the distance from } u \text{ to } v \text{ is equal to the distance from } U \text{ to } V \text{ in the unweighted (or quotient) graph} \mathcal{G}. \text{ Then} \\
&\text{if } u \text{ and } v \text{ belong to orbits } U \text{ and } V, \text{ respectively, in different symmetric motifs, then} \\
&\text{the distance from } u \text{ to } v \text{ is equal to the distance from } U \text{ to } V \text{ in the unweighted (or quotient) graph} \mathcal{G}. \text{ Then} \\
&\text{if } u \text{ and } v \text{ belong to orbits } U \text{ and } V, \text{ respectively, in different symmetric motifs, then} \\
&\text{the distance from } u \text{ to } v \text{ is equal to the distance from } U \text{ to } V \text{ in the unweighted (or quotient) graph} \mathcal{G}. \text{ Then}
\end{align*}\]

**Proof** (i) Since automorphisms are bijections and preserve adjacency, $(\sigma(v_1), \sigma(v_2), \ldots, \sigma(v_n))$ is a path from $\sigma(u)$ to $\sigma(v)$ of the same length. If there were a shorter path $\sigma(u) = w_1, w_2, \ldots, w_\ell = v$, then the same argument applied to $\sigma^{-1}$ gives a shorter path $u = \sigma^{-1}(w_1), \sigma^{-1}(w_2), \ldots, v = \sigma^{-1}(w_\ell)$ from $u$ to $v$, contradicting the minimality of the path. (ii) Any path of minimal length is also of minimal length between its endpoints. Arguing by contradiction, there exists a subpath $p = (w_1, w_2, \ldots, w_m)$ (or $p = (w_m, w_{m-1}, \ldots, w_1)$), such that $w_1$ and $w_m$ belong to the same orbit, and $w_i$ belongs to a different symmetric motif. Hence, we can find $\sigma_1 \in \text{Aut } (\mathcal{G})$
Let us choose essential \(^11\) sets of generators \(S\), respectively \(S'\), of \(\text{Aut}(\mathcal{G})\), respectively \(\text{Aut}(\mathcal{G}')\), with support-disjoint partitions 

\[ X = X_1 \cup \ldots \cup X_m, \quad \text{respectively} \quad X' = X_1' \cup \ldots \cup X_m' \]

It is enough to prove the statement for these sets: given, there is unique \(j\) such that \(X_j \subseteq X\). Let \(x_j' \in X_j' \subseteq \text{Aut}(\mathcal{G}')\) thus we can write \(x_j' = x_1' \cdots x_m'\), with \(h_i \in H_i\). Since \(\mathcal{X}'\) is an essential set of generators, there is an index \(j\) such that \(h_j = 1\) (the identity, or trivial permutation) for all \(k \neq j\), so that \(x_j' = h_j\).

Given any other \(x_j' \in \mathcal{X}'\), the same argument gives \(y' = x_j'\) for some \(1 \leq j \leq m\).

We claim \(j = j\), as follows. The partition of \(X\), respectively \(X'\), above are the equivalence classes of the equivalence relation generated by \(\sigma \sim \tau\) if \(\sigma\) and \(\tau\) are not support-disjoint permutations. Since \(x_j'\), \(y'\) are in the same equivalence class, so are \(h_j\) and \(y\) and thus \(j = j\).

This result applies to networks with other additional structure, not necessarily expressed in terms of the adjacency matrix, such as arbitrary vertex or edge labels, by restricting to automorphisms preserving the additional structure.

We obtain fewer symmetries, and a refinement of the geometric decomposition, symmetric motifs, and orbits as \(\mathcal{G}'\), and the structural results in this paper apply verbatim.

### Data availability

The data sets analysed during the current study are available at the locations stated in the caption to Table 1. The data sets generated during the current study can be found at [https://doi.org/10.6084/m9. figshare.11619792](https://doi.org/10.6084/m9.figshare.11619792).

### Code availability

The code used to process the data sets can be found at [https://bitbucket.org/ rubenjesan/garcia/networksymmetry](https://bitbucket.org/rubenjesan/garcia/networksymmetry).

Received: 12 July 2018; Accepted: 30 March 2020; Published online: 15 May 2020

### References

1. Newman, M. Networks: An Introduction (Oxford University Press, Oxford, 2010).
2. Watts, D. & Strogatz, S. Collective dynamics of small-world networks. Nature 393, 440–442 (1998).
3. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. Science 286, 509–512 (1999).
4. Milo, R. et al. Network motifs: simple building blocks of complex networks. Science 298, 824–827 (2002).
5. Tononi, G., Sporns, O. & Edelman, G. M. Measures of degeneracy and redundancy in biological networks. Proc. Natl Acad. Sci. USA 96, 3527–3532 (1999).
6. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47 (2002).
7. Chung, F., Lu, L., Dewey, T. G. & Galas, D. J. Duplication models for biological networks. J. Comput. Biol. 10, 677–687 (2003).
8. Xiao, Y., Xiong, M., Wang, W. & Wang, H. Emergence of symmetry in complex networks. Phys. Rev. E 77, 066108 (2008).
9. Nikishkov, T. & Motter, A. E. Network-complement transitions, symmetries, and cluster synchronization. Chaos 26, 094818 (2016).
10. Klickstein, I. & Sorrentino, F. Generating symmetric graphs Chaos: an interdisciplinary. J. Nonlinear Sci. 28, 121102 (2018).
11. MacArthur, B. D., Sánchez-Garcia, R. J. & Anderson, J. W. Symmetry in complex networks. Discret. Appl. Math. 156, 3525–3531 (2008).
12. Xiao, Y., MacArthur, B. D., Wang, H., Xiong, M. & Wang, W. Network quotients: structural skeletons of complex systems. Phys. Rev. E 78, 046102 (2008).
13. MacArthur, B. D. & Sánchez-Garcia, R. J. Spectral characteristics of network redundancy. Phys. Rev. E 80, 026117 (2009).
14. Pecora, L. M., Sorrentino, F., Hagerstrom, A. M., Murphy, T. E. & Roy, R. Cluster synchronization and isolated desynchronization in complex networks with symmetries. Nat. Commun. 5, 4079 (2014).
15. Golubitsky, M. & Stewart, I. The Symmetry Perspective: from Equilibrium to Chaos in Phase Space and Physical Space. Vol. 208 (Springer Science & Business Media, Basel, 2003).
16. Stewart, I., Golubitsky, M. & Ivato, M. Symmetry groupoids and patterns of synchrony in coupled networks. SIAM J. Appl. Dynamical Syst. 2, 69–74 (2003).
17. Golubitsky, M., Stewart, I. & Török, A. Patterns of synchrony in coupled cellular networks with multiple arrows. SIAM J. Appl. Dynamical Syst. 4, 78–100 (2005).
18. Aguiar, M. A. & Dias, A. P. S. Synchronization and equitable partitions in weighted networks Chaos: An Interdisciplinary. J. Nonlinear Sci. 28, 073105 (2018).
19. Sorrentino, F., Sduitique, A. B. & Pecora, L. M. Symmetries in the time-averaged dynamics of networks: reducing unnecessary complexity through minimal network models Chaos: An Interdisciplinary. J. Nonlinear Sci. 29, 011101 (2019).
20. Nicosia, V., Valencia, M., Chavez, M., Diaz-Guilera, A. & Latora, V. Remote synchronization reveals network symmetries and functional modules. Phys. Rev. Lett. 110, 174102 (2013).
21. Sorrentino, F., Pecora, L. M., Hagerstrom, A. M., Murphy, T. E. & Roy, R. Complete characterization of the stability of cluster synchronization in complex dynamical networks. Sci. Adv. 2, e1501737 (2016).
22. Sorrentino, F. & Pecora, L. Approximate cluster synchronization in networks with symmetries and parameter mismatches Chaos: An Interdisciplinary. J. Nonlinear Sci. 26, 094823 (2016).
23. Schaub, M. T. et al. Graph partitions and cluster synchronization in networks of oscillators Chaos: An Interdisciplinary. J. Nonlinear Sci. 26, 094821 (2016).
24. Pecora, L. M., Sorrentino, F., Hagerstrom, A. M., Murphy, T. E. & Roy, R. in Advances in Dynamics, Patterns, Cognition. 145–160 (Springer, 2017).
25. Siddique, A. B., Pecora, L., Hart, J. D. & Sorrentino, F. Symmetry- and input- cluster synchronization in networks. Phys. Rev. E 97, 042217 (2018).
26. Bitbucket repository. https://bitbucket.org/rubenbsanchezgarcia/ networksymmetry/ (Bitbucket repository, 2019).
27. Xiao, Y., Wu, W., Pei, J., Wang, W. & He, Z. Efficiently indexing shortest paths by exploiting symmetry in graphs. In Proc. 12th International Conference on Extending Database Technology: Advances in Database Technology, 493–504 (ACM, 2009).
28. Wang, J., Huang, Y., Wu, F.-X. & Pan, Y. Symmetry compression method for discovering network motifs. IEEE/ACM Trans. Comput. Biol. Bioinf. 9, 1776–1789 (2012).
29. Karalus, S. & Krug, J. Symmetry-based coarse-graining of evolved dynamical networks. Europhys. Lett. 111, 38003 (2015).
30. Nyberg, A., Gross, T. & Bassler, K. E. Mesoscopic structures and the laplacian spectra of random geometric graphs. J. Complex Netw. 3, 543–551 (2015).
31. Do, A.-L., Höfener, J. & Gross, T. Engineering mesoscale structures with distinct dynamical implications. N. J. Phys. 14, 115022 (2012).
32. Biggs, N. Algebraic Graph Theory (Cambridge University Press, Cambridge, 1993).
33. Godsil, C. & Royle, G. F. Algebraic Graph Theory (Springer, New York, 2013).
34. Cvetkovic, D. M., Rowlinson, P. & Simic, S. An Introduction to the Theory of Graph Spectra (Cambridge University Press, Cambridge, 2010).
35. Brouwer, A. E. & Haemers, W. H. Spectra of Graphs (Springer, New York, 2011).
36. Estrada, E. & Rodriguez-Velazquez, J. A. Subgraph centrality in complex networks. Phys. Rev. E 71, 056103 (2005).
37. Estrada, E. & Hatano, N. Communicability in complex networks. Phys. Rev. E 77, 036111 (2008).
38. Erdős, P. & Rényi, A. Asymmetric graphs. Acta Math. Hung. 14, 295–315 (1963).
39. Borgatti, S. & Gonsier, T. Structural Equivalence: meaning and measures. International Encyclopedia of the Social & Behavioral Sciences (2015).
40. Internet topology network dataset—KONECT. http://konect.uni-koblenz.de/networks/openflights/opshahl_powergridCA-AstroPh[topology[wordnetwords]com-amazon[actor-collaboration[as-skitter,roadNet-CAlivejournal-lin](2016).