

## Algebraic approach to small-world network models

Michelle Rudolph-Lilith\* and Lyle E. Muller

*Unité de Neurosciences, Information et Complexité, CNRS, 1 Avenue de la Terrasse, 91198 Gif-sur-Yvette, France*

(Received 14 October 2013; published 28 January 2014)

We introduce an analytic model for directed Watts-Strogatz small-world graphs and deduce an algebraic expression of its defining adjacency matrix. The latter is then used to calculate the small-world digraph's asymmetry index and clustering coefficient in an analytically exact fashion, valid nonasymptotically for all graph sizes. The proposed approach is general and can be applied to all algebraically well-defined graph-theoretical measures, thus allowing for an analytical investigation of finite-size small-world graphs.

DOI: [10.1103/PhysRevE.89.012812](https://doi.org/10.1103/PhysRevE.89.012812)

PACS number(s): 89.75.Da, 02.10.Ox, 02.30.Vv

In recent years, small-world graphs have gained considerable interest as models of real-world systems [1], which often display features residing between regularity and randomness. The most notable of these models is the Watts-Strogatz graph [2], though alternative models have been proposed [3]. The unifying characteristics of these models are that any two nodes are joined with a small number of links between them (i.e., short path length), while at the same time connected node pairs exhibit an abundance of triangular relations resulting in a high degree of local redundancy (i.e., high clustering).

Theoretical investigations of small-world graph models have generally applied asymptotic evaluations in the limit of large system size [4] or the continuum approximation [5,6] to the algorithmic definition of the graph, in the absence of an analytic representation. In this work, we introduce a generative model of small worlds that can be viewed as a directed version of a canonical Watts-Strogatz graph and provide an algebraic expression for the adjacency matrix defining the graph. We employ this formulation to calculate various properties of this graph in a nonasymptotic fashion. The resulting expressions are exact for all graph parameters and are valid at all system sizes.

This paper is organized as follows. We start by introducing the directed small-world model, which is then formulated algebraically. We apply this formulation in the calculation of various graph-theoretic measures, specifically, the asymmetry index and graph clustering coefficient. The paper ends with a discussion of the limitations and possible generalizations of the proposed approach.

### I. MODELS OF SMALL-WORLD DIGRAPHS

We start by describing the algorithmic construction of a non-self-looped directed small-world graph. Let  $\mathfrak{G}_{RG}$  denote a classical undirected ring graph with  $N_N$  nodes, degree  $k$ , and adjacency matrix  $a_{ij}^{RG}$ , in which each node is symmetrically connected to  $2k$  of its neighbors. A new graph  $\mathfrak{G}_{SW}$  is constructed by uniform random rewiring  $N_E^{\text{rew}}$  of the total  $N_E^{\text{max}RG} = 2kN_N$  edges of  $\mathfrak{G}_{RG}$  across  $N_E^{\text{max}} = N_N(N_N - 1)$  possible edges of the full graph. Let  $q = N_E^{\text{rew}}/N_E^{\text{max}RG}$

denote the rewiring probability. Then this rewiring will, on average, add  $N_E^{\text{rew}}N_N(N_N - 2k - 1)/N_E^{\text{max}}$  edges between nodes not connected within the initial ring graph and remove  $N_E^{\text{rew}}(1 - N_E^{\text{max}RG}/N_E^{\text{max}})$  edges from  $\mathfrak{G}_{RG}$ , thus ensuring that statistically the total number of edges in the constructed graph is conserved and independent of  $q$ . Moreover, whereas for  $q = 0$  the original symmetric ring graph remains unchanged, for  $q = 1$  only  $4k^2N_N/(N_N - 1)$  edges remain of  $\mathfrak{G}_{RG}$ , whereas  $2kN_N(N_N - 2k - 1)/(N_N - 1)$  edges are randomly distributed outside  $\mathfrak{G}_{RG}$ , yielding a classical Erdős-Rényi digraph  $\mathfrak{G}_{ER}$  with expected uniform connectedness  $\mathcal{C} = 2k/(N_N - 1)$ , where the connectedness of a graph is defined as

$$\mathcal{C} = \frac{N_E}{N_E^{\text{max}}},$$

with  $N_E$  denoting the total number of edges.

The total adjacency  $A$  and asymmetry index  $\mathcal{A}$  of a directed non-self-looped ( $a_{ii} = 0 \forall i \in [1, N_N]$ ) graph are defined as [7–9] (see also [10])

$$A = \sum_{i,j=1}^{N_N} a_{ij} \quad (1)$$

and

$$\mathcal{A} = \frac{N_A}{A - N_S}, \quad (2)$$

respectively. The asymmetry index is bound  $0 \leq \mathcal{A} \leq 1$  and quantifies the ratio between the number of nonsymmetrical edges  $N_A$ , defined as the number of node pairs ( $\{i, j\}, j \geq i$ ) for which  $a_{ij} \neq a_{ji}$ , and symmetrical edges  $N_S$ , the number of node pairs ( $\{i, j\}, j \geq i$ ) for which  $a_{ij} = a_{ji} = 1$ . It can easily be shown that for  $\mathfrak{G}_{SW}$  we have in the statistical limit

$$A = 2kN_N \forall q, \quad \mathcal{C} = \frac{2k}{N_N - 1} \forall q, \quad (3)$$

$$\mathcal{A} = \begin{cases} 0 & \text{for } q = 0 \\ \frac{N_N - 2k - 1}{N_N - k - 1} & \text{for } q = 1. \end{cases}$$

Moreover,  $\mathfrak{G}_{SW}$  gives rise to classical small-world characteristics, i.e., a narrow regime for intermediate values of the rewiring probability  $q$  with coexisting small average geodesic graph distance  $\langle d \rangle$  and large total global clustering coefficient  $\mathcal{C}^d$  (Fig. 1, black dots). Here

$$\langle d \rangle = \frac{1}{N_C} \sum_{i,j=1}^{N_N} d_{ij}, \quad (4)$$

\*rudolph@unic.cnrs-gif.fr

Published by the American Physical Society under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

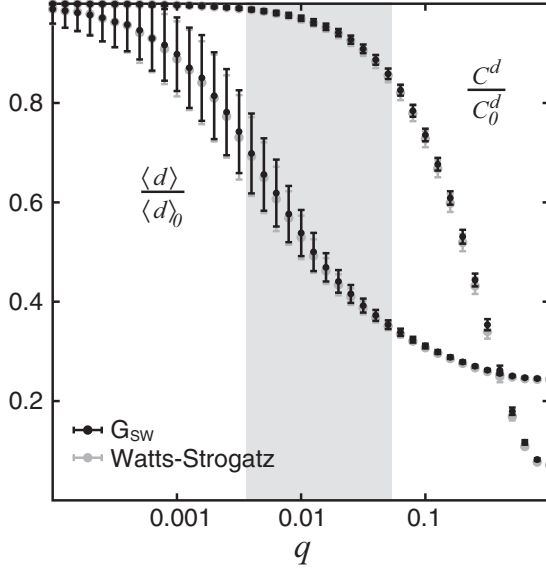


FIG. 1. Small-world digraph models. The dependence (mean  $\pm$  SD) of the average geodesic graph distance  $\langle d \rangle$  [Eq. (4)] and total global clustering coefficient  $C^d$  [Eq. (5)] on the rewiring probability  $q$  are shown for the generic model of small worlds  $\mathfrak{G}_{SW}$  (black dots) and the canonical Watts-Strogatz small-world digraph model (gray dots). Here  $\langle d \rangle_0$  and  $C_0^d$  denote the distance and clustering for the original ring graph, respectively. The gray region marks the small-world regime in which a small average geodesic distance and large clustering coefficient coexist. The analysis was performed on representative graph models with  $N_N = 200$  and  $k = 5$ , with 1000 random realizations for each parameter set.

where  $d_{ij}$  denotes the number of edges in the shortest path connecting nodes  $i$  and  $j$  and  $N_C$  is the total number of connected node pairs, and

$$C^d = \frac{1}{\mathcal{N}} \sum_{i,j,h=1}^{N_N} (a_{ij} + a_{ji})(a_{ih} + a_{hi})(a_{jh} + a_{hj}), \quad (5)$$

where

$$\mathcal{N} = 2 \sum_{i=1}^{N_N} [a_i^{\text{tot}}(a_i^{\text{tot}} - 1) - 2a_i^b],$$

$$a_i^{\text{tot}} = \sum_{j=1}^{N_N} (a_{ji} + a_{ij}),$$

$$a_i^b = \sum_{j=1}^{N_N} a_{ij} a_{ji}.$$

The algorithm introduced above yields a graph  $\mathfrak{G}_{SW}$  that can be shown to correspond to a canonical model for Watts-Strogatz small-world [2] digraphs. The latter are constructed from an undirected non-self-looped ring graph  $\mathfrak{G}_{RG}$  with  $N_N$  nodes, degree  $k$ , and subsequent rewiring of all edges with probability  $q$ . To that end, for each edge  $a_{ij}$ , a new target node  $j'$  with  $a_{ij'} = 0$  is chosen and with probability  $q$  the original edge is removed ( $a_{ij} = 1 \rightarrow a_{ij} = 0$ ) and a new edge added ( $a_{ij'} = 0 \rightarrow a_{ij'} = 1$ ), thus leading to a strict conservation of the graph's total adjacency. This algorithmic

approach is equivalent to constructing a “reduced” ring graph with  $2kN_N - q2kN_N(N_N - 2k - 1)/(N_N - 1)$  edges uniformly distributed randomly across the  $2kN_N$  edges of  $\mathfrak{G}_{RG}$  and distributing the remaining  $q2kN_N(N_N - 2k - 1)/(N_N - 1)$  edges outside  $\mathfrak{G}_{RG}$ . We note, however, that  $\mathfrak{G}_{SW}$  preserves the total adjacency only in the statistical average, i.e., over many realizations.

Figure 1 (gray dots) shows the numerical results for  $\langle d \rangle$  and  $C^d$  as a function of the rewiring probability  $q$  for a representative Watts-Strogatz small-world digraph. The results match sufficiently well that for the corresponding  $\mathfrak{G}_{SW}$  (Fig. 1; compare black and gray dots), thus justifying the introduction of  $\mathfrak{G}_{SW}$  as a generic model for small worlds. Whereas a nonasymptotic analytical treatment of the canonical Watts-Strogatz small-world digraph model remains elusive, we will present an analytically exact consideration of  $\mathfrak{G}_{SW}$  in the remainder of this study.

## II. ALGEBRAIC APPROACH

To analytically explore the small-world graph model  $\mathfrak{G}_{SW}$  introduced above, we now construct explicitly its adjacency matrix  $a_{ij}^{SW}$ . To that end, we define two disjunct sets of indices

$$\begin{aligned} \mathcal{A} &= (\{i, j\}; i, j \in [1, N_N] : [(i - k \leq j \leq i + k) \wedge (i \neq j)] \\ &\quad \vee (i + N_N - k \leq j) \vee (j + N_N - k \leq i)), \\ \mathcal{B} &= (\{i, j\}; i, j \in [1, N_N] : [(i + k < j < i + N_N - k) \\ &\quad \vee (j + k < i < j + N_N - k)]). \end{aligned} \quad (6)$$

With these, we further define the following matrices, which will later serve as building blocks for  $a_{ij}^{SW}$ :

$$\begin{aligned} \delta_{ij}^{\mathcal{A}} &= \begin{cases} 1 & \text{if } \{i, j\} \in \mathcal{A} \\ 0 & \text{otherwise} \end{cases} \\ &= \sum_{l=1}^k (\delta_{i+l, j} + \delta_{j+l, i} + \delta_{i+N_N-l, j} + \delta_{j+N_N-l, i}) \\ &= \text{circ} \left[ \left( \sum_{l=1}^k (\delta_{1+l, j} + \delta_{1+N_N-l, j}) \right)_j \right] \\ &= \text{circ} \left( \left\{ 0, \overbrace{1, \dots, 1}^k, 0, \dots, 0, \overbrace{1, \dots, 1}^k \right\} \right), \end{aligned} \quad (7)$$

$$\begin{aligned} \delta_{ij}^{\mathcal{B}} &= \begin{cases} 1 & \text{if } \{i, j\} \in \mathcal{B} \\ 0 & \text{otherwise} \end{cases} \\ &= \sum_{l=1}^{N_N-2k-1} (\delta_{i+k+l, j} + \delta_{j+k+l, i}) \\ &= \text{circ} \left[ \left( \sum_{l=1}^{N_N-2k-1} (\delta_{1+k+l, j}) \right)_j \right] \\ &= \text{circ} \left( \left\{ \overbrace{0, \dots, 0}^{k+1}, 1, \dots, 1, \overbrace{0, \dots, 0}^k \right\} \right), \end{aligned} \quad (8)$$

where  $\text{circ}(c_j)$  denotes the circulant matrix constructed from vector  $c_j$  and  $\delta_{ij}$  the Kronecker delta. Note that both  $\delta_{ij}^{\mathcal{A}}$  and

$\delta_{ij}^B$  are symmetric matrices, i.e.,  $\delta_{ij}^A = \delta_{ji}^A$  and  $\delta_{ij}^B = \delta_{ji}^B$ . The elementwise (Hadamard) product obeys

$$\begin{aligned} (\delta^A \circ \delta^A)_{ij} &\equiv \delta_{ij}^A \delta_{ij}^A = \delta_{ij}^A, \\ (\delta^B \circ \delta^B)_{ij} &\equiv \delta_{ij}^B \delta_{ij}^B = \delta_{ij}^B, \\ (\delta^A \circ \delta^B)_{ij} &\equiv \delta_{ij}^A \delta_{ij}^B = 0 \end{aligned} \quad (9)$$

$\forall i, j \in [1, N_N]$  and we define

$$\begin{aligned} N_E^{\max A} &= \sum_{i,j=1}^{N_N} \delta_{ij}^A = 2kN_N, \\ N_E^{\max B} &= \sum_{i,j=1}^{N_N} \delta_{ij}^B = N_N(N_N - 2k - 1) \end{aligned} \quad (10)$$

as the number of nonzero elements in  $\delta^A$  and  $\delta^B$ , respectively. Furthermore, we have

$$\begin{aligned} \sum_{j=1}^{N_N} \delta_{ij}^A &= 2k, \\ \sum_{j=1}^{N_N} \delta_{ij}^B &= N_N - 2k - 1 \end{aligned} \quad (11)$$

$\forall i \in [1, N_N]$ . Finally, due to construction,  $\delta_{ij}^A$  corresponds to the adjacency matrix  $a_{ij}^{RG}$  of a ring graph of degree  $k$ .

As the construction of small-world graphs requires the random rewiring and assignment of edges, we next introduce formally a binomial random operator  $\hat{r}^p(x)$  that returns its argument  $x$  with probability  $p$  and zero otherwise, i.e.,

$$\hat{r}^p(x) = \begin{cases} x & \text{with probability } p \\ 0 & \text{with probability } 1 - p. \end{cases} \quad (12)$$

Due to its definition,  $\hat{r}^p(x)$  can be viewed as a random annihilation operator and must be considered statistically. Specifically, the sum over  $n$  applications of this operator on  $x$  returns  $np x$ , i.e.,

$$\sum_{i=1}^n \hat{r}^p(x) = np x. \quad (13)$$

It can easily be demonstrated that the set of these operators forms a linear algebra that is commutative and associative under multiplication, as well as distributive. For instance, both  $\hat{r}^p[\hat{r}^q(x)] = \hat{r}^p \circ \hat{r}^q(x)$  and  $\hat{r}^q \circ \hat{r}^p(x)$  return  $x$  with probability  $qp$ , thus

$$\hat{r}^p \circ \hat{r}^q(x) = \hat{r}^q \circ \hat{r}^p(x) \quad (14)$$

(commutative). In a similar fashion, one can show

$$\begin{aligned} \hat{r}^p(x) &= x \hat{r}^p(1), \\ \hat{r}^p \circ [\hat{r}^q \circ \hat{r}^s(x)] &= (\hat{r}^p \circ \hat{r}^q) \circ \hat{r}^s(x), \\ \hat{r}^p \circ \hat{r}^q(x) &= \hat{r}^{pq}(x), \\ (\hat{r}^p)^n(x) &= \overbrace{\hat{r}^p \circ \dots \circ \hat{r}^p}^{n \text{ times}}(x) = \hat{r}^{p^n}(x), \\ \hat{r}^p(x + y) &= \hat{r}^p(x) + \hat{r}^p(y), \\ (\hat{r}^p + \hat{r}^q)(x) &= \hat{r}^p(x) + \hat{r}^q(x). \end{aligned} \quad (15)$$

We note that all properties (14) and (15) have to be understood statistically. For instance, the last equation of (15) yields

$$\sum_{i=1}^n (\hat{r}^p + \hat{r}^q)(x) = \sum_{i=1}^n \hat{r}^p(x) + \sum_{i=1}^n \hat{r}^q(x) = n(p + q)x.$$

With Eq. (12), we can now proceed with formalizing the generation of  $\mathfrak{G}_{SW}$ . First we consider the elements of the adjacency matrix defined by  $\mathcal{A}$ , that is, the initial ring graph  $a_{ij}^{RG} = \delta_{ij}^A$ . According to the definition (12), the Hadamard product  $\hat{r}_{ij}^{pA} \delta_{ij}^A$ , where  $\hat{r}_{ij}^p$  denotes a matrix of random annihilation operators  $\hat{r}^p$ , will remove with probability  $1 - p_A$  entries of  $\delta_{ij}^A$ , leaving a matrix with, on average,  $p_A N_E^{\max A}$  nonzero elements confined to region  $\mathcal{A}$ . Similarly,  $\hat{r}_{ij}^{pB} \delta_{ij}^B$  will remove  $1 - p_B$  elements from  $\delta_{ij}^B$ , yielding a matrix with, on average,  $p_B N_E^{\max B}$  nonzero elements confined to region  $\mathcal{B}$ . As  $\delta_{ij}^A$  and  $\delta_{ij}^B$  are disjunct and span a whole square matrix of dimension  $N_N$  (excluding the diagonal), the matrix

$$r_{ij} = \hat{r}_{ij}^{pA} \delta_{ij}^A + \hat{r}_{ij}^{pB} \delta_{ij}^B \quad (16)$$

acts as a randomization when elementwise multiplied with a given adjacency matrix.

In order to construct the desired small-world graph  $\mathfrak{G}_{SW}$ , let us define a function  $f(r_{ij}, a_{ij}^{RG})$ . According to the algorithmic model presented in the preceding section, edges are removed from the initial ring graph. This is achieved by requiring  $f(r_{ij}, a_{ij}^{RG}) = 0$  if and only if  $r_{ij} = 1 \wedge a_{ij}^{RG} = 1$  and  $f(r_{ij}, a_{ij}^{RG}) = 1$  if and only if  $r_{ij} = 0 \wedge a_{ij}^{RG} = 1$ , which in effect will delete with probability  $p_A$  edges from  $a_{ij}^{RG}$ . Similarly, requiring  $f(r_{ij}, a_{ij}^{RG}) = 1$  if and only if  $r_{ij} = 1 \wedge a_{ij}^{RG} = 0$  and  $f(r_{ij}, a_{ij}^{RG}) = 0$  if and only if  $r_{ij} = 0 \wedge a_{ij}^{RG} = 0$  will add with probability  $p_B$  edges between nodes not connected within the original ring graph. These requirements are summarized in Table I and correspond to the logic XOR operation. It can easily be shown that the function

$$f(r_{ij}, a_{ij}^{RG}) = a_{ij}^{RG} + r_{ij} - 2a_{ij}^{RG} r_{ij} \quad (17)$$

yields the desired algebraic formulation.

With this, we can now construct the adjacency matrix  $a_{ij}^{SW}$  of a small-world graph  $\mathfrak{G}_{SW}$ ,

$$\begin{aligned} a_{ij}^{SW} &= a_{ij}^{RG} + r_{ij} - 2a_{ij}^{RG} r_{ij} \\ &= (1 - \hat{r}_{ij}^{pA}) \delta_{ij}^A + \hat{r}_{ij}^{pB} \delta_{ij}^B. \end{aligned} \quad (18)$$

TABLE I. Requirements imposed on  $f(r_{ij}, a_{ij}^{RG})$  for constructing a small-world graph  $\mathfrak{G}_{SW}$  from a given ring graph  $a_{ij}^{RG}$ . These requirements correspond to the logic XOR and are analytically described by  $a_{ij}^{RG} + r_{ij} - 2a_{ij}^{RG} r_{ij}$ .

$r_{ij}$	$a_{ij}^{RG}$	$f(r_{ij}, a_{ij}^{RG})$	$r_{ij} \oplus a_{ij}^{RG}$	$a_{ij}^{RG} + r_{ij} - 2a_{ij}^{RG} r_{ij}$
1	1	0	0	0
0	1	1	1	1
1	0	1	1	1
0	0	0	0	0

Here we used the fact that  $a_{ij}^{RG} = \delta_{ij}^A$  and applied Eqs. (9). What remains is to calculate the probabilities  $p_A$  and  $p_B$ . As  $q2kN_N(N_N - 2k - 1)/(N_N - 1)$  edges will be deleted from  $\mathcal{A}$ , which has a total of  $N_E^{\max \mathcal{A}}$  edges,

$$p_A = q \frac{N_N - 2k - 1}{N_N - 1}. \quad (19)$$

Equivalently,  $q2kN_N(N_N - 2k - 1)/(N_N - 1)$  edges will be added to  $\mathcal{B}$ , which is distributed among  $N_E^{\max \mathcal{B}}$  possible edges, thus

$$p_B = q \frac{2k}{N_N - 1}. \quad (20)$$

In the remainder of this section, we will list a number of identities that will be employed in the application of Eq. (18) to various graph-theoretic measures. We have

$$\begin{aligned} \sum_{i,j=1}^{N_N} \hat{r}_{ij}^{r_A} \delta_{ij}^A &= \sum_{i,j=1}^{N_N} \hat{r}_{ji}^{p_A} \delta_{ij}^A = p_A N_E^{\max \mathcal{A}}, \\ \sum_{i,j=1}^{N_N} \hat{r}_{ij}^{r_B} \delta_{ij}^B &= \sum_{i,j=1}^{N_N} \hat{r}_{ji}^{p_B} \delta_{ij}^B = p_B N_E^{\max \mathcal{B}}, \\ \sum_{i,j=1}^{N_N} \hat{r}_{ij}^{r_B} \delta_{ij}^A &= \sum_{i,j=1}^{N_N} \hat{r}_{ji}^{p_B} \delta_{ij}^A = p_B N_E^{\max \mathcal{A}}, \\ \sum_{i,j=1}^{N_N} \hat{r}_{ij}^{r_A} \delta_{ij}^B &= \sum_{i,j=1}^{N_N} \hat{r}_{ji}^{p_A} \delta_{ij}^B = p_A N_E^{\max \mathcal{B}}. \end{aligned} \quad (21)$$

Moreover,

$$\begin{aligned} \sum_{j=1}^{N_N} \hat{r}_{ij}^{r_A} \delta_{ij}^A &= \sum_{j=1}^{N_N} \hat{r}_{ji}^{p_A} \delta_{ij}^A = 2kp_A, \\ \sum_{j=1}^{N_N} \hat{r}_{ij}^{r_B} \delta_{ij}^B &= \sum_{j=1}^{N_N} \hat{r}_{ji}^{p_B} \delta_{ij}^B = (N_N - 2k - 1)p_B \end{aligned} \quad (22)$$

$\forall i \in [1, N_N]$ .

### III. APPLICATION

With Eq. (18), we have an algebraic formulation of small-world graphs  $\mathfrak{G}_{SW}$  that, together with Eqs. (9)–(11), (21), and (22), can be used to calculate analytically exact various graph-theoretic measures. In this section, we will demonstrate the application of this formalism by considering the total adjacency  $A$ , asymmetry index  $\mathcal{A}$ , and total global graph clustering coefficient  $C^d$ , defined in (1), (2), and (5), respectively.

#### A. Total adjacency

We start with the simplest graph-theoretic measure, the total adjacency  $A$ . Inserting  $a_{ij}^{SW}$  [Eq. (18)] into definition (1)

yields

$$A^{SW} = (1 - p_A)N_E^{\max \mathcal{A}} + p_B N_E^{\max \mathcal{B}},$$

where identities (10) and (21) were applied. Using further the expressions for  $N_E^{\max \mathcal{A}}$  and  $N_E^{\max \mathcal{B}}$  in Eq. (10) and for  $p_A$  and  $p_B$  in Eqs. (19) and (20), respectively, we finally arrive at

$$A^{SW} = 2kN_N \quad (23)$$

for the total adjacency of a small-world graph with  $N_N$  nodes and degree  $k$ . As expected,  $A^{SW}$  is independent of the rewiring probability  $q$  and identical to the result expected from the algorithmic construction of  $\mathfrak{G}_{SW}$  [Eq. (3)].

#### B. Asymmetry index

Next we calculate the asymmetry index  $\mathcal{A}$  [Eq. (2)], which involves the second Hadamard power of the adjacency matrix  $a_{ij}^{SW}$ . The number of symmetrical edges  $N_S^{SW}$  is given by

$$N_S^{SW} = \frac{1}{2} \sum_{i,j=1}^{N_N} a_{ij}^{SW} a_{ji}^{SW}. \quad (24)$$

Inserting (18) yields

$$\begin{aligned} N_S^{SW} &= \frac{1}{2} \sum_{i,j=1}^{N_N} [(1 - \hat{r}_{ij}^{p_A})\delta_{ij}^A + \hat{r}_{ij}^{p_B}\delta_{ij}^B][(1 - \hat{r}_{ji}^{p_A})\delta_{ji}^A + \hat{r}_{ji}^{p_B}\delta_{ji}^B] \\ &= \frac{1}{2} \sum_{i,j=1}^{N_N} [\delta_{ij}^A - 2\hat{r}_{ij}^{p_A}\delta_{ij}^A + (\hat{r}_{ij}^{p_A})^2\delta_{ij}^A + (\hat{r}_{ij}^{p_B})^2\delta_{ij}^B] \\ &= \frac{1}{2}(1 - p_A^2)N_E^{\max \mathcal{A}} + \frac{1}{2}p_B^2N_E^{\max \mathcal{B}}, \end{aligned} \quad (25)$$

where in the last two steps identities (9), (10), and (21) were used. Similarly, the number of nonsymmetrical edges  $N_A^{SW}$  is given by

$$N_A^{SW} = \frac{1}{2} \sum_{i,j=1}^{N_N} (a_{ij}^{SW} + a_{ji}^{SW}) - \sum_{i,j=1}^{N_N} a_{ij}^{SW} a_{ji}^{SW} \quad (26)$$

and yields

$$\begin{aligned} N_A^{SW} &= \sum_{i,j=1}^{N_N} (\delta_{ij}^A - \hat{r}_{ij}^{p_A}\delta_{ij}^A + \hat{r}_{ij}^{p_B}\delta_{ij}^B) - 2N_S^{SW} \\ &= p_A(1 - p_A)N_E^{\max \mathcal{A}} + p_B(1 - p_B)N_E^{\max \mathcal{B}}. \end{aligned} \quad (27)$$

Inserting Eqs. (25) and (27) into (2) and using (23) for  $A^{SW}$ , as well as (10) for  $N_E^{\max \mathcal{A}}$  and  $N_E^{\max \mathcal{B}}$  and (19) and (20) for  $p_A$  and  $p_B$ , respectively, we finally obtain for the asymmetry index of a small-world graph  $\mathfrak{G}_{SW}$

$$\mathcal{A}^{SW} = \frac{2q(2 - q)(N_N - 2k - 1)}{N_N - 1 + q(2 - q)(N_N - 2k - 1)}. \quad (28)$$

We emphasize that (28) is an exact result and holds in the nonasymptotic limit, i.e., for arbitrary  $N_N$  and  $q$ . Figure 2 compares the analytic solution with the corresponding numerical result for a representative graph of small size.

For  $q = 0$ , the original symmetric ring graph with  $\mathcal{A}^{SW} = 0$  is retained. For  $q = 1$ ,  $\mathfrak{G}_{SW}$  is equivalent to a classical

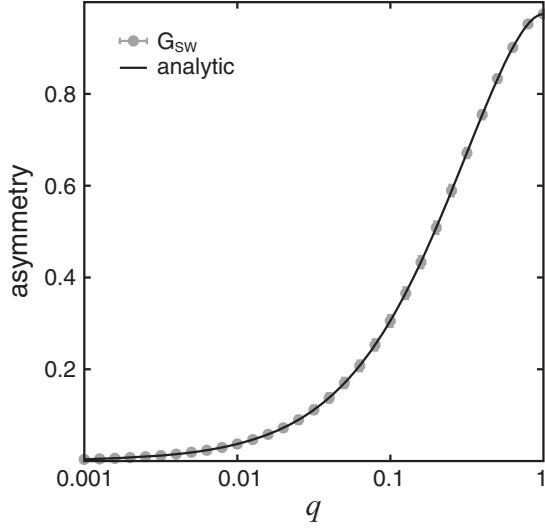


FIG. 2. Graph asymmetry. Compared are the numerical result (gray) and analytic solution (28) (black) for the asymmetry index  $\mathcal{A}^{SW}$  of a representative small-world graph  $\mathcal{G}_{SW}$  with  $N_N = 200$  and  $k = 5$  as function of the rewiring probability  $q$ . In the limit  $q = 1$  (classical Erdős-Rényi digraph),  $\mathcal{A}^{SW}$  takes the value 0.9742 in the shown example [see Eq. (29)], whereas for  $q = 0$  (ring graph),  $\mathcal{A}^{SW} = 0$ . For the numerical model, 1000 random realizations were used for each parameter set.

Erdős-Rényi digraph. In this case, Eq. (28) yields

$$\mathcal{A}^{SW}|_{q=1} = \frac{N_N - 2k - 1}{N_N - k - 1}. \quad (29)$$

This limit corresponds to the expectation for small-world graphs [Eq. (3)]. Finally, in the asymptotic limit for  $N_N \rightarrow \infty$ ,  $k$  finite, we have

$$\mathcal{A}^{SW}|_{N_N \rightarrow \infty} = \frac{2q(q-2)}{q^2 - 2q - 1}, \quad (30)$$

which yields a completely asymmetric Erdős-Rényi digraph for  $q = 1$ .

### C. Graph clustering coefficient

As a last example, we deduce the analytic form of the total graph clustering coefficient  $C^d$ , defined in Eq. (5), which involves the third matrix power of the adjacency matrix  $a_{ij}^{SW}$ . To that end, we first calculate the normalization term

$$\mathcal{N}^{SW} = 2 \sum_{i=1}^{N_N} [a_i^{\text{totSW}} (a_i^{\text{totSW}} - 1) - 2a_i^{bSW}]. \quad (31)$$

With (24), we have

$$\sum_{i=1}^{N_N} a_i^{bSW} = \sum_{i,j=1}^{N_N} a_{ij}^{SW} a_{ji}^{SW} = 2N_S^{SW}. \quad (32)$$

Similarly, with (26) one obtains

$$\sum_{i=1}^{N_N} a_i^{\text{totSW}} = \sum_{i,j=1}^{N_N} (a_{ji}^{SW} + a_{ij}^{SW}) = 2N_A^{SW} + 4N_S^{SW}. \quad (33)$$

Finally,

$$\begin{aligned} \sum_{i=1}^{N_N} (a_i^{\text{totSW}})^2 &= \sum_{i=1}^{N_N} \left( \sum_{j=1}^{N_N} (a_{ji}^{SW} + a_{ij}^{SW}) \right)^2 \\ &= 4 \sum_{i=1}^{N_N} \left( \sum_{j=1}^{N_N} a_{ij}^{SW} \right)^2 \\ &= 4N_N [(1 - p_A)2k + p_B(N_N - 2k - 1)]^2, \end{aligned} \quad (34)$$

where we used in the last step Eq. (22) to obtain  $\forall i \in [1, N_N]$

$$\sum_{j=1}^{N_N} a_{ij}^{SW} = (1 - p_A)2k + p_B(N_N - 2k - 1).$$

Inserting Eqs. (32)–(34) into (31) yields for the normalization term

$$\begin{aligned} \mathcal{N}^{SW} &= 4 \{ 2N_N [(N_N - 1)p_B - 2k(p_A + p_B - 1)]^2 \\ &\quad - 2N_E^{\max A} (1 - p_A)^2 - N_E^{\max A} (1 - p_A) p_A \\ &\quad - 2N_E^{\max B} p_B^2 - N_E^{\max B} (1 - p_B) p_B \} \\ &= \frac{8kN_N}{N_N - 1} [2k(q^2 - 2q + 2N_N - 2) \\ &\quad - (N_N - 1)(q^2 - 2q + 2)]. \end{aligned} \quad (35)$$

In order to arrive at an explicit expression for the numerator in Eq. (5), we require the third matrix power of the adjacency matrix (18). To that end, we first calculate matrix products of  $\delta_{ij}^A$  and  $\delta_{ij}^B$  defined in Eqs. (7) and (8), respectively. As both  $\delta^A$  and  $\delta^B$  are circulant matrices, we can utilize the circulant diagonalization theorem. The latter states that all circulants  $c_{ij} = \text{circ}(c_j)$  constructed from an arbitrary  $N$ -dimensional vector  $c_j$  are diagonalized by the same unitary matrix  $\mathbf{U}$  with components

$$u_{rs} = \frac{1}{\sqrt{N}} \exp \left[ -\frac{2\pi i}{N} (r-1)(s-1) \right], \quad (36)$$

$r, s \in [1, N]$ . Moreover, the  $N$  eigenvalues are explicitly given by

$$E_r(\mathbf{C}) = \sum_{j=1}^N c_j \exp \left[ -\frac{2\pi i}{N} (r-1)(j-1) \right] \quad (37)$$

such that

$$c_{ij} = \sum_{r,s=1}^N u_{ir} e_{rs} u_{sj}^* \quad (38)$$

with  $e_{rs} = \text{diag}[E_r(\mathbf{C})] \equiv \delta_{rs} E_r(\mathbf{C})$  and  $u_{rs}^*$  denoting the complex conjugate of  $u_{rs}$ .

Using (7), the eigenvalues of  $\delta^A$  are

$$\begin{aligned} E_r(\delta^A) &= \sum_{j=1}^{N_N} \sum_{l=1}^k (\delta_{1+l,j} + \delta_{1+N_N-l,j}) \exp \left[ -\frac{2\pi i}{N_N} (r-1)(j-1) \right] \\ &= \sum_{l=2}^{1+k} \exp \left[ -\frac{2\pi i}{N_N} (r-1)(l-1) \right] + \sum_{l=N_N-k+1}^{N_N} \exp \left[ -\frac{2\pi i}{N_N} (r-1)(l-1) \right] \\ &= 2 \frac{\sin \left[ \frac{k(r-1)}{N_N} \pi \right] \cos \left[ \frac{(1+k)(r-1)}{N_N} \pi \right]}{\sin \left[ \frac{(r-1)}{N_N} \pi \right]}. \end{aligned} \quad (39)$$

As the  $n$ th power of a diagonal matrix is the diagonal matrix spanned by the  $n$ th power of its eigenvalues, we arrive, using (36), at

$$(\delta^A)_{rs}^n = \frac{(2k)^n}{N_N} + \frac{2^n}{N_N} \sum_{m=1}^{N_N-1} \left\{ \left( \frac{\sin \left[ \frac{km}{N_N} \pi \right] \cos \left[ \frac{(1+k)m}{N_N} \pi \right]}{\sin \left[ \frac{m}{N_N} \pi \right]} \right)^n \exp \left[ -\frac{2\pi i}{N_N} m(r-s) \right] \right\}. \quad (40)$$

In a similar fashion, the  $n$ th power of  $\delta^B$  can be calculated, yielding

$$(\delta^B)_{rs}^n = \frac{(N_N - 2k - 1)^n}{N_N} + \frac{(-1)^n}{N_N} \sum_{m=1}^{N_N-1} \left\{ \left( \frac{\sin \left[ \frac{(1+2k)m}{N_N} \pi \right]}{\sin \left[ \frac{m}{N_N} \pi \right]} \right)^n \exp \left[ -\frac{2\pi i}{N_N} m(r-s) \right] \right\}. \quad (41)$$

Finally,  $(\delta^A)^n (\delta^B)^{n'}$  is given by

$$\begin{aligned} [(\delta^A)^n (\delta^B)^{n'}]_{rs} &= \frac{(2k)^n (N_N - 2k - 1)^{n'}}{N_N} \\ &+ \frac{(-1)^{n'} 2^n}{N_N} \sum_{m=1}^{N_N-1} \left\{ \left( \frac{\sin \left[ \frac{km}{N_N} \pi \right] \cos \left[ \frac{(1+k)m}{N_N} \pi \right]}{\sin \left[ \frac{m}{N_N} \pi \right]} \right)^n \left( \frac{\sin \left[ \frac{(1+2k)m}{N_N} \pi \right]}{\sin \left[ \frac{m}{N_N} \pi \right]} \right)^{n'} \exp \left[ -\frac{2\pi i}{N_N} m(r-s) \right] \right\}. \end{aligned} \quad (42)$$

With (40)–(42), one can now proceed to calculate

$$\begin{aligned} &\sum_{i,j,h=1}^{N_N} (a_{ij}^{SW} + a_{ji}^{SW})(a_{ih}^{SW} + a_{hi}^{SW})(a_{jh}^{SW} + a_{hj}^{SW}) \\ &= 8 \sum_{i,j,h=1}^{N_N} a_{ij}^{SW} a_{jh}^{SW} a_{hi}^{SW} \\ &= 8(1 - p_A)^3 \text{Tr}(\delta^A)^3_{ij} + 24(1 - p_A)^2 p_B \text{Tr}[(\delta^A)^2 (\delta^B)]_{ij} + 24(1 - p_A) p_B^2 \text{Tr}[(\delta^A) (\delta^B)^2]_{ij} + 8 p_B^3 \text{Tr}(\delta^B)^3_{ij} \\ &= 64(1 - p_A)^3 \left( k^3 + \sum_{m=1}^{N_N-1} a_m^3(N_N, k) \right) + 96(1 - p_A)^2 p_B \left( k^2(N_N - 2k - 1) - \sum_{m=1}^{N_N-1} a_m^2(N_N, k) b_m(N_N, k) \right) \\ &+ 48(1 - p_A) p_B^2 \left( k(N_N - 2k - 1)^2 + \sum_{m=1}^{N_N-1} a_m(N_N, k) b_m^2(N_N, k) \right) + 8 p_B^3 \left( (N_N - 2k - 1)^3 - \sum_{m=1}^{N_N-1} b_m^3(N_N, k) \right), \end{aligned}$$

where

$$a_m(N, k) = \frac{\sin \left[ \frac{km}{N} \pi \right] \cos \left[ \frac{(1+k)m}{N} \pi \right]}{\sin \left[ \frac{m}{N} \pi \right]}, \quad b_m(N, k) = \frac{\sin \left[ \frac{(1+2k)m}{N} \pi \right]}{\sin \left[ \frac{m}{N} \pi \right]}. \quad (43)$$

This finally yields for the total global clustering coefficient  $C^{dSW}$  of a small-world graph  $\mathfrak{G}_{SW}$

$$\begin{aligned} C^{dSW} &= \frac{64}{N^{SW} (N_N - 1)^3} \left\{ [N_N - 1 - q(N_N - 2k - 1)]^3 \left( k^3 + \sum_{m=1}^{N_N-1} a_m^3(N_N, k) \right) \right. \\ &\quad \left. + 3kq [N_N - 1 - q(N_N - 2k - 1)]^2 \left( k^2(N_N - 2k - 1) - \sum_{m=1}^{N_N-1} a_m^2(N_N, k) b_m(N_N, k) \right) \right\} \end{aligned}$$

$$\begin{aligned}
& + 3k^2q^2[N_N - 1 - q(N_N - 2k - 1)] \left( k(N_N - 2k - 1)^2 + \sum_{m=1}^{N_N-1} a_m(N_N, k) b_m^2(N_N, k) \right) \\
& + q^3k^3 \left( (N_N - 2k - 1)^3 - \sum_{m=1}^{N_N-1} b_m^3(N_N, k) \right) \Bigg\}, \tag{44}
\end{aligned}$$

with  $\mathcal{N}^{SW}$  given in Eq. (35).

The analytical result for the clustering coefficient [Eq. (44)] is again an exact result and holds in the nonasymptotic limit, i.e., for arbitrary  $N_N$  and  $q$ . Figure 3 compares the analytic solution with the corresponding numerical result for a representative graph of small size.

For  $q = 0$ , i.e., in the case of the initial symmetric ring graph, we obtain

$$C^{dSW}|_{q=0} = \frac{4}{k(2k-1)N_N} \left( k^3 + \sum_{m=1}^{N_N-1} a_m^3(N_N, k) \right). \tag{45}$$

By careful inspection, it can be shown that

$$\sum_{m=1}^{N_N-1} a_m^3(N_N, k)|_{N_N \rightarrow \infty} = \gamma_1 k(k-1)N_N \tag{46}$$

with  $\gamma_1 = 3/8$ , which yields for (45) in the asymptotic limit  $N_N \rightarrow \infty$  and finite  $k$

$$C^{dSW}|_{\substack{q=0 \\ N_N \rightarrow \infty}} = \frac{3(k-1)}{2(2k-1)}. \tag{47}$$

This asymptotic limit corresponds to the result obtained previously [4].

Similarly,  $q = 1$  (classical Erdős-Rényi digraph) yields

$$\begin{aligned}
C^{dSW}|_{q=1} = & \frac{8k^2}{N_N(N_N-1)^2[1-N_N+k(4N_N-6)]} \left( (N_N-1)^3 + 8 \sum_{m=1}^{N_N-1} a_m^3(N_N, k) \right. \\
& \left. - 12 \sum_{m=1}^{N_N-1} a_m^2(N_N, k) b_m(N_N, k) + 6 \sum_{m=1}^{N_N-1} a_m(N_N, k) b_m^2(N_N, k) - \sum_{m=1}^{N_N-1} b_m^3(N_N, k) \right). \tag{48}
\end{aligned}$$

In order to obtain the asymptotic limit  $N_N \rightarrow \infty$  in this case, we observe that all sums over  $a_m$ ,  $b_m$ , and their products are linearly divergent in  $N_N$ . As the denominator in (48) is of fourth order in  $N_N$  (the numerator, however, is only a polynomial of third order), we have

$$C^{dSW}|_{\substack{q=1 \\ N_N \rightarrow \infty}} = 0, \tag{49}$$

which matches the asymptotic limit obtained in [4].

Finally, we consider the limit  $N_N \rightarrow \infty$  and  $k$  finite for arbitrary  $q$ . As the sums over  $a_m$ ,  $b_m$ , and their products are linearly divergent in  $N_N$ , all terms in the curly brackets in Eq. (44), with the exception of the first, will contribute polynomials of at most third order in  $N_N$  and thus vanish in the asymptotic limit. This leaves the first term proportional to  $\sum_{m=1}^{N_N-1} a_m^3(N_N, k)$ , which yields

$$C^{dSW}|_{N_N \rightarrow \infty} = \frac{3(k-1)}{2(2k-1) + q(2-q)} (q-1)^3 \tag{50}$$

for the clustering coefficient of a small-world graph  $\mathfrak{G}_{SW}$  in the asymptotic limit  $N_N \rightarrow \infty$  for finite  $k$ . Interestingly, this solution differs by the occurrence of an additional term  $q(2-q)$  in the denominator from the asymptotic solution proposed in [4], obtained by interpolating between the clustering values

of a ring graph and infinite-size random graph. However, it can easily be shown that for  $k > 1$  this term becomes negligible as  $q$  is strictly bound.

#### IV. DISCUSSION

In this study, we have introduced a generative model of directed small-world graphs  $\mathfrak{G}_{SW}$ , a canonical model of Watts-Strogatz digraphs, and have proposed an approach that yields the graph's defining adjacency matrix in algebraic terms, with the goal to provide mathematically rigorous access to the study of finite-size small-world graphs. The proposed approach makes use of random annihilation operators whose algebraic properties can be utilized to assess algebraically-well-defined graph-theoretic measures in an analytically exact framework, valid nonasymptotically for all graph sizes. We demonstrated the application of our approach by calculating the asymmetry index  $\mathcal{A}$  and total clustering coefficient  $C^d$  of small worlds.

Since its introduction [2,11], the idea of small worlds has dominated the literature as many real-world systems appear to display properties captured by the small-world model. Given their importance, in recent years, a number of theoretical studies have focused on providing mathematically more rigorous access in order to allow for an analytical

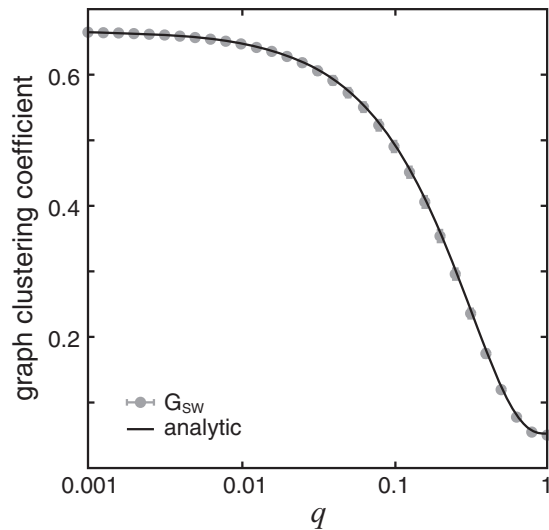


FIG. 3. Graph clustering coefficient. Compared are the numerical result (gray) and analytic solution (44) (black) for the total global graph clustering coefficient  $C^{dSW}$  of a representative small-world graph  $\mathcal{G}_{SW}$  with  $N_N = 200$  and  $k = 5$  as a function of the rewiring probability  $q$ . In the example shown,  $C^{dSW} = 0.6667$  in the limit  $q = 0$  (ring graph) and  $C^{dSW} = 0.0527$  for  $q = 1$  [classical Erdős-Rényi digraph; see Eqs. (45) and (48), respectively]. For the numerical model, 1000 random realizations were used for each parameter set.

study of the small-world phenomenon. Among these, heuristic expressions have been proposed for the clustering of a small-world graph, based on an interpolation between the known clustering coefficient of a ring graph and that expected for an infinite-size random graph [4]. Although this model provides an excellent fit to the true clustering coefficient even for medium-size graphs, it remains a heuristic approximation whose error scales only inversely proportionally to the graph size. A second approach followed in the recent literature is based on considering a small-world graph in the continuum limit, allowing for a formulation in the mean-field framework [6]. Although mathematically more sound, this approach is only valid in the large- $N$  limit and does not capture the clustering coefficient expected for an infinite-size Erdős-Rényi digraph in the case in which all edges of the small-world graph are randomly rewired.

Our proposed approach differs from the aforementioned mathematical models by explicit construction of the adjacency matrix of a small-world graph within an operational algebraic framework. By considering the properties of the underlying

algebra, the so-constructed adjacency matrix can be utilized to directly calculate all graph-theoretic measures algebraically defined in terms of the adjacency matrix. As the constructed adjacency matrix is valid at all system sizes, the obtained results provide an analytically exact description of small-world properties valid nonasymptotically for all system sizes.

Although the introduced approach allows for an analytically exact treatment of various graph-theoretic measures, it requires that the latter are well defined algebraically. Unfortunately, this excludes geodesic distance measures, which are crucial in the characterization of the small-world phenomenon. Specifically, the  $m$ th power of the adjacency matrix will generally yield the number of path (walks) of length  $m$  in a given graph and can be calculated exactly within our framework. However, the number of walks of length  $m$  that are also geodesics connecting two nodes constitutes an optimization problem that cannot currently be captured within the approach introduced here, as it involves a minimization of the entries of the  $m$ th power of the adjacency matrix. Possible solutions are to replace the concept of the geodesic distance as a defining graph-theoretic measure with an algebraically well-defined quantity not involving optimization or to extend our approach by including analytical tools that allow for algorithmically defined measures such as the geodesic graph distance  $\langle d \rangle$ .

A further extension to this framework could allow, in general, for the calculation of higher-order moments, such as the variance, of algebraically-well-defined graph-theoretic measures. Preliminary work along these lines shows, however, that calculating the variance of a higher-order graph-theoretic measure requires either the exchange of a sum and product in order to apply the rules (9) and (15) or an extension of the algebra considered here.

Finally, we emphasize that the operator framework utilized here can be viewed as a step towards an alternative conception of graphs as dynamical mathematical objects. Here we demonstrated the potential power of this operator graph-theoretic framework by calculating an analytically exact clustering coefficient of an important class of small-world graphs.

#### ACKNOWLEDGMENTS

The authors wish to thank O. D. Little for comments. This work was supported by CNRS, the European Community (BrainScales Project No. FP7-269921), and École des Neurosciences de Paris Ile-de-France.

- 
- [1] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
  - [2] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
  - [3] R. Monasson, *Eur. Phys. J. B* **12**, 555 (1999).
  - [4] A. Barrat and M. Weigt, *Eur. Phys. J. B* **13**, 547 (2000).
  - [5] M. E. J. Newman, C. Moore, and D. J. Watts, *Phys. Rev. Lett.* **84**, 3201 (2000).
  - [6] M. E. J. Newman, *Comput. Phys. Commun.* **147**, 40 (2002).
  - [7] S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, 1994).
  - [8] M. E. J. Newman, S. Forrest, and J. Balthrop, *Phys. Rev. E* **66**, 035101(R) (2002).
  - [9] M. A. Serrano and M. Boguñá, *Phys. Rev. E* **68**, 015101 (2003).
  - [10] D. Garlaschelli and M. I. Loffredo, *Phys. Rev. Lett.* **93**, 268701 (2004).
  - [11] S. Milgram, *Psychol. Today* **1**(1), 61 (1967).