Gradient networks

2008 J. Phys. A: Math. Theor. 41 155103

(http://iopscience.iop.org/1751-8121/41/15/155103)

J. Phys. A: Math. Theor. **41** (2008) 155103 (13pp) doi:10.1088/1751-8113/41/15/155103

Gradient networks

Zoltán Toroczkai[1], Balázs Kozma[2], Kevin E Bassler[3], N W Hengartner[4] and G Korniss[5]

[1] Department of Physics and Center for Complex Network Research, University of Notre Dame, Notre Dame, IN 46556, USA
[2] Laboratoire de Physique Théorique, Bâtiment 210, Université Paris-Sud, Orsay 91405, France
[3] Department of Physics, 617 Science and Research bld. I, University of Houston, Houston, TX 77204, USA
[4] Computational, Computer, and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[5] Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, 110, 8th Street, Troy, NY 12180, USA

Received 18 December 2007, in final form 6 March 2008
Published 2 April 2008
Online at stacks.iop.org/JPhysA/41/155103

Abstract
Gradient networks are defined (Toroczkai and Bassler 2004 *Nature* **428** 716) as directed graphs formed by local gradients of a scalar field distributed on the nodes of a substrate network G. We present the derivation for some of the general properties of gradient graphs and give an exact expression for the in-degree distribution $R(l)$ of the gradient network when the substrate is a binomial (Erdős–Rényi) random graph, $G_{N,p}$, and the scalars are independent identically distributed (i.i.d.) random variables. We show that in the limit $N \to \infty, p \to 0, z = pN = \text{const} \gg 1, R(l) \propto l^{-1}$ for $l < l_c = z$, i.e., gradient networks become scale-free graphs up to a cut-off degree. This paper presents the detailed derivation of the results announced in Toroczkai and Bassler (2004 *Nature* **428** 716).

PACS numbers: 89.75.Fb, 89.75.Hc, 89.20.Hh, 89.75.Da

(Some figures in this article are in colour only in the electronic version)

1. Introduction

It is well known [2–5] that a large number of systems are organized into structures best described by complex networks, or massive graphs. Many of these networks are strongly heterogeneous, possessing a power-law degree distribution $P(k) \sim k^{-\gamma}$, also called *scale-free networks* [5]. These networks are very different from pure random graphs [6] which are homogeneous structures, and have a 'bell curve' Poisson degree distribution. Since many of the real-world networks are scale-free, the question naturally arises: what mechanisms can lead to these structures? There are a number of models in the literature, which introduce the required heterogeneity for the scale-free degree distribution, such as the preferential attachment

model [7] and its variants [8–10], the copying model [11, 12], fitness-based models [13–15] and graph optimization models [16, 17]. All these are *network evolution models* where the structure of the graph changes in time by changing its nodes and edges according to some rules, eventually resulting in scale-free structures.

Here we present a different mechanism that can generate power-law degree distributed networks. The motivation behind our study lies with the main function of networks, namely, that of transport. Most real-world networks transport entities such as information, material goods, power, water, oil, gas, forces, etc. Such transport processes are often driven by local gradients of a scalar. Physical examples include the electric current which is driven by an electric potential gradient, and heat flow which is driven by a temperature gradient. The existence of gradients has also been shown to play an important role in social systems, economics and certainly in biology (e.g., cell migration [18]).

Naturally, the same mechanism will generate flows on complex networks as well. Besides the obvious examples of traffic flows, power distribution on the grid and waterways, we recall two, less-known examples, where gradient-induced transport on complex networks plays a key role: (1) diffusive load balancing schemes used in distributed computation [19] (and also employed in packet routing on the internet), and (2) reinforcement learning on social networks with competitive dynamics [20]. In the first example, a computer (or a router) asks its neighbors on the network for their current job load (or packet load), and then the router balances its load with the neighbor that has the *minimum* number of jobs to run (or packets to route). In this case the scalar at each node is the negative of the number of jobs at that node, and the *flow* occurs in the direction of the gradient of this scalar in the node's network neighborhood. In the second example, a number of agents/players who are all parts of a social network, compete in an iterated game based with limited information [20]. At every step of the game each agent has to decide whose advice to follow before taking an action (such as placing a bet), in its circle of acquaintances (network neighborhood). Typically, an agent will try to follow that neighbor who in the past proved to be the most reliable. That neighbor is recognized using a reinforcement learning mechanism: a score is kept for every agent measuring its past success at predicting the correct outcome of the game, and then each agent follows the advice of the agent in its network neighborhood which has the *highest score* [20] accumulated up to that point in time. In this case the scalar is the past success score kept for each agent, and the gradient from an agent points toward the chosen neighbor (with the highest score in its neighborhood).

Since gradients signify directions where maximum flow is expected, collecting all gradients on a complex network results in a directed graph that potentially carries the *maximum flow* (see below for precise definitions) in the system, and which we call the *gradient network*. Naturally, the gradient network can be thought of as the backbone for transport on the underlying 'substrate' network. In this paper we uncover some fundamental properties of gradient networks on general substrate graphs, and show that they can be scale-free networks even in the case when the substrate graph is a homogeneous, scaled structure, such as a binomial (Erdős–Rényi) random graph. These results were first presented in [1] without mathematical derivations. This paper presents the proofs and derivations in detail, and in particular for the in-degree distribution of the gradient network it shows two solutions, a combinatorial and one based on integral transforms. Note that here we only consider the case of independent identically distributed (i.i.d.) random scalars associated with the nodes of the substrate graph. As we will see, even in this case, the mathematical derivations are fairly involved. The more general case of correlated scalar fields has recently been applied to the study of folding pathways on the energy landscape of polypeptide chains (proteins) [21]. In this case, a node represents a 3D spatial conformation of the polypeptide and the

associated scalar is the potential energy of the conformation. Naturally, the potential energy of the polypeptide conformation is not independent of the conformation and its neighbors on the conformation network, so the i.i.d. models presented here do not apply. As shown via numerical simulations in [21], the gradient network there is also a scale-free graph (indicating that this is a general property), with a connectivity exponent that is sensitive to correlations. The formalism presented here, however, lays the foundation for more complex and correlated situations, to be presented in future publications. Network constructions similar to gradient networks have been studied in [22].

The paper is organized as follows. In section 2, we formulate a mathematical framework for analyzing the collective properties of gradients on networks, which, as will be demonstrated, typically organize themselves into a directed network structure without loops (collection of directed trees). In section 3, we obtain the exact expression for the in-degree distribution of the gradient network on binomial (Erdős–Rényi) random graphs and i.i.d. random scalar fields, and show that in the natural scaling limit the gradient network becomes a scale-free graph with connectivity exponent of $\gamma = -1$ (up to a cut-off degree). Section 4 is devoted to conclusions.

2. Definition of a gradient network

In this section we give the definition of a gradient network as was first introduced in [1], and highlight some of its general properties. Let us consider that transport can take place on a fixed network $G = G(V, E)$ which we will call in the remainder, the *substrate graph*. It has N nodes, $V = \{0, 1, \ldots, N-1\}$ and the set of edges E is specified by the adjacency matrix $\mathbf{A} = \{a_{ij}\}$ ($a_{ij} = 1$ if i and j are connected, $a_{ij} = 0$ otherwise, and $a_{ii} = 0$). Given a node i, we will denote its set of neighbors in G by $S_i^{(1)} = \{j \in V | a_{ij} = 1\}$. Let us also consider a scalar field (which could just as well be called potential landscape) $\mathbf{h} = \{h_0, \ldots, h_{N-1}\}$ defined on the set of nodes V, so that every node i has a scalar value h_i associated with it. We define the *gradient* ∇h_i of the field $\mathbf{h}$ in the node i to be the *directed edge* $\nabla h_i = (i, \mu(i))$ which points from i to that neighbor, $\mu(i) \in S_i^{(1)} \cup \{i\}$ on G at which the scalar field has the maximum value in $S_i^{(1)} \cup \{i\}$, i.e.:

$$\mu(i) = \underset{j \in S_i^{(1)} \cup \{i\}}{\operatorname{argmax}} \ (h_j), \tag{1}$$

see figure 1. According to its classical definition, a gradient vector points in the direction of the steepest ascent at a point on a continuous (d-dimensional) landscape. The above definition is a natural generalization to the case when the continuous landscape is replaced by a graph.

Note that $\mu(i) = i$, if i has the largest scalar value in its neighborhood (i.e., in the set $S_i^{(1)} \cup \{i\}$), and in this case the gradient edge is a *self-loop* at that node. Since $\mathbf{h}$ always has a global maximum, there is always at least one self-loop. It is possible that equation (1) has more than one solution (several equal maxima) in the case of which we say that the scalar field is degenerate. In this paper we deal only with non-degenerate fields, which is typical when for example $\mathbf{h}$ is a continuous stochastic variable.

This allows us to define the set $\mathbf{F}$ of gradient edges on G, together with the vertex set V form the *gradient network*, $\nabla G = \nabla G(V, \mathbf{F})$.

Note that the definitions above do not refer to actual *flows* through the network. The definitions so far present a 'pressure'-like quantity rather than a dynamical quantity such as current. However, if there are gradients of a scalar along edges of a graph, it will generate flows through the network. Assuming that all edges have the same 'conductance', or transport properties, the gradient network will be the substructure of the original network which at a

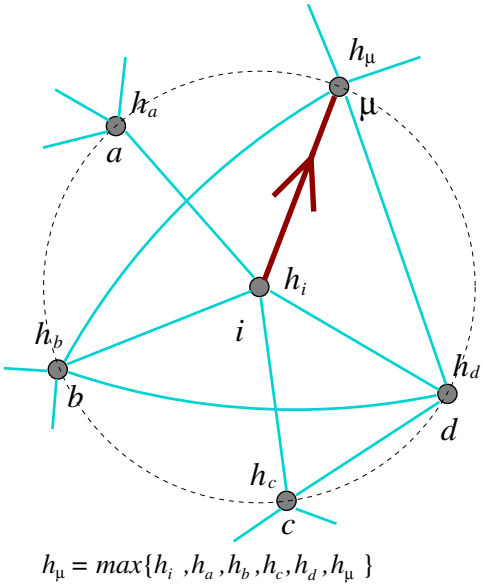

Figure 1. Definition of a gradient on a network, schematic drawing. The gradient at node i is a directed edge pointing toward the maximum value of the scalar (node μ) in the node's neighborhood.

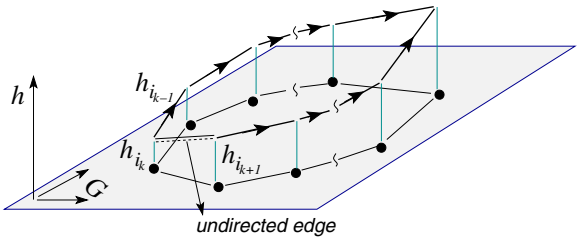

Figure 2. There cannot be loops in a non-degenerate ∇G (schematic drawing).

given instant will *channel the bulk of the flow*, and thus alternatively can be called as the *instantaneous maximum flow subgraph*. In general, the scalar field will be evolving in time, due to the gradients generated currents, and also to possible external sources and sinks on the network (for example packets are generated and used up at nodes, but they can also be lost). As a result, the gradient network ∇G will be *time-dependent, highly dynamic*.

2.1. Some general properties of gradient networks

Here we present a few fundamental *structural* properties valid for all non-degenerate gradient networks. The first observation we make about gradient networks is non-degenerate gradient networks form forests (i.e., there are no loops in ∇G, and it is a union of trees, more exactly of in-directed, planted pines).

To prove this statement, assume that in contrast, there is a closed path $\gamma = \{\nabla h_{i_1}, \nabla h_{i_2}, \ldots, \nabla h_{i_m}\}, m \geqslant 3$ made up only of directed edges from **F**, see figure 2. Let i_k be the node on this path for which $h_{i_k} = \min\{h_{i_1}, h_{i_2}, \ldots, h_{i_m}\}$. Node i_k has exactly two neighbors on γ, nodes $i_{k\pm1}$, but only one gradient direction, ∇h_{i_k} pointing away from i_k.

Since $h_{i_{k\pm1}} > h_{i_k}$, none of the neighbors $i_{k\pm1}$ will have their gradient edges pointing into i_k. Since there are two edges, (i_k, i_{k-1}) and (i_k, i_{k+1}), and only one gradient edge from i_k, one of the edges must not be a gradient edge, and thus the loop is not closed, in contradiction with the assumption that γ is a loop with only gradient edges. Using a similar reasoning we can show that for non-degenerate scalar fields, there is no continuous path in ∇G connecting two local maxima of the scalar field **h**. This means that on a given tree of ∇G there is only one local maximum of the scalar, and it is the only node with a self-loop on that tree. As a consequence, the number of trees in the forest equals the number of local maxima of the scalar field **h** on G. The fact that ∇G is made of trees (no loops), is advantageous for analytical techniques, especially if we take into consideration that ∇G is the most important substructure driving the flow in the network. Note that unless there is exactly one local maximum (and thus global as well) of **h** on G, ∇G is disconnected into a number of trees and thus ∇G is *not* a spanning tree. Since every node has exactly one gradient direction from it, the out-degree of every node on the gradient network is unity. It also means that ∇G has exactly N nodes and N edges (with at least one edge being a self-loop). However, the *in-degree* of a node i, which is the number of gradient edges *pointing into* i, can be anything in the range $k_i^{(in)} \in \{0, 1, \ldots, k_i\}$, where k_i is the degree of node i on G. Note that not all edges will necessarily have to have a gradient direction.

3. The in-degree distribution of a gradient network on random graphs and random fields

3.1. Summary of the results

In this section, we show that when the substrate graph G is a binomial random graph $G = G_{N,p}$ [6], and **h** is an i.i.d. random field over V, given by a distribution $\eta(h)$, the in-degree distribution

$$R(l) = \text{Prob.}\left\{k_i^{(in)} = l\right\}$$

of ∇G obeys the exact expression [1]:

$$R(l) = \frac{1}{N}\sum_{n=0}^{N-1}\binom{N-1-n}{l}[1-p(1-p)^n]^{N-1-n-l}[p(1-p)^n]^l, \tag{2}$$

independently on the particular form of the distribution for the scalars, $\eta(h)$. The binomial random graph (also coined in the physics literature as the Erdős–Rényi random graph) is constructed by taking all pairs (i, j) of N nodes and connecting them with probability p, independently from other connections. The exact form (2) is obtained after averaging over randomness both in the scalar field and the substrate graph. As one can see from figure 3, form (2) is approached relatively fast by the simulations. Figure 3 presents both the curves in (2) and the result of averaging over 10^4 independent numerical runs.

We will also show that the gradient network ∇G becomes a *scale-free network* with respect to the in-degree distribution, in the scaling limit $N \to \infty$ and $p \to 0$, such that $Np = z = \text{const} \gg 1$, up to the cut-off degree $l_c = z$. The in-degree distribution in this limit is described by the law:

$$R(l) \simeq \frac{1}{zl}, \qquad 0 < l \leqslant z. \tag{3}$$

a behavior which is also apparent from figure 3. This power law is a rather surprising result, since the substrate graph is a random graph which is *not* scale-free, its degree distribution (in the same limit) being Poisson, with a well-defined average degree z (setting the scale) and faster than exponential decaying tails [6].

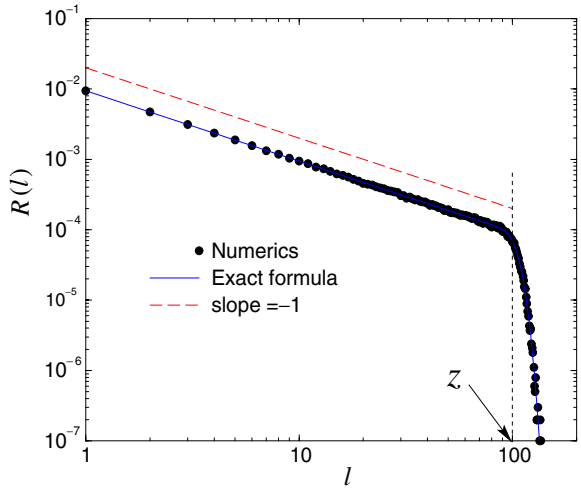

Figure 3. Comparison between the exact formula (2) and convergence of numerics [1]. Here $N = 1000$, $p = 0.1$ ($z = 100$). The numerical values are obtained after averaging over 10^4 sample runs.

Gradient networks are not the only scale-free networks that can be induced on $G_{N,p}$ random graphs. The paper [23] by Lakhina *et al* reports finding scale-free graphs relative to the trace-route measurements that are used to sample the structure of the internet. Namely, they find that the spanning trees obtained this way on $G_{N,p}$'s have a degree distribution that obeys the $1/k$ law. In [24], Clauset and Moore present an analytical approach to this law. Another method [25], which is based on pruning edges from $G_{N,p}$, also generates scale-free structures as a result.

3.2. A combinatorial derivation of the exact expression

In this subsection we give a combinatorial derivation for formula (2) first, for it is the shortest one. Another approach based on integral transforms is presented in the following subsection, which actually was our original method, and it had inspired the combinatorial one.

In order to calculate the in-degree distribution $R(l)$, we first distribute the scalars on the node set V, then find those link configurations which contribute to $R(l)$ when building the random graph $G_{N,p}$ over these nodes.

Without restricting the generality we will calculate the distribution of in-links for node 0. Let us consider a set of n nodes from V, that does not contain node 0, and it has the property that the scalar values at these nodes h_i are all larger than h_0. We will denote this set by $[\tau]_n$. The complementary set of $[\tau]_n$ in $V\backslash\{0\}$ will be denoted by $C_{[\tau]_n}$, see figure 4. Note that every link that connects to node 0 from a node of $C_{[\tau]_n}$ which is not connected to $[\tau]_n$ is a gradient link.

In order to have *exactly* l nodes pointing their gradient edges into node 0, we must fulfil the following conditions: first, they have to be connected to node 0 and, second, they *must not* be connected to the set $[\tau]_n$ (otherwise, they would be connected to a node with a scalar value larger than h_0, according to the definition of $[\tau]_n$). The probability for one node to fulfil these two conditions is $p(1-p)^n$, and since the links are drawn independently, for l nodes this probability is $[p(1-p)^n]^l$. We must also require that no other nodes will have their gradient links pointing into node 0. Obviously, by definition, nodes from $[\tau]_n$ will not be

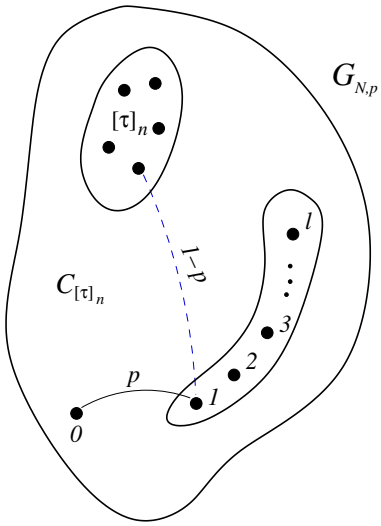

Figure 4. Schematic of the construction given in the main text.

pointing gradients into node 0. Therefore, we have to make sure, that none of the remaining $N-1-l-n$ nodes from $C_{[\tau]_n}$ will be pointing into 0. For one such node this will happen with probability $1-p(1-p)^n$. For all the $N-1-l-n$ such nodes this probability will be $[1-p(1-p)^n]^{N-1-l-n}$. Thus, given a specific set $[\tau]_n$, the probability of exactly l in-links to node 0 is

$$\binom{N-1-n}{l}[p(1-p)^n]^l[1-p(1-p)^n]^{N-1-l-n}. \tag{4}$$

The combinatorial factor in (4) counts the number of ways the set of l nodes which point their gradient edges to node 0, can be chosen from $C_{[\tau]_n}$.

The probability in (4) was computed by fixing h_0 and the set $[\tau]_n$. Next, we compute the probability Q_n of such an event for a given n, while letting the field $\mathbf{h}$ vary according to its distribution. The probability for a node to have its scalar value larger than h_0 is

$$\gamma(h_0) = \int_{h_0} \mathrm{d}h\, \eta(h). \tag{5}$$

The probability to have exactly n nodes with this property is given by

$$[\gamma(h_0)]^n[1-\gamma(h_0)]^{N-1-n}. \tag{6}$$

The number of ways the n nodes can be chosen from $V\backslash\{0\}$ is just the binomial $\binom{N-1}{n}$. Thus, the total probability Q_n will be given by

$$Q_n = \binom{N-1}{n}\int \mathrm{d}h_0\, \eta(h_0)[\gamma(h_0)]^n[1-\gamma(h_0)]^{N-1-n} = \frac{1}{N}, \tag{7}$$

where the last equality in the above equation is obtained after performing the change of variables $\mathrm{d}u \equiv \mathrm{d}\gamma(h_0) = \mathrm{d}h_0\, \eta(h_0)$.

As a final step, by combining (7) with (4), and summing over all possible n values, we arrive at (2).

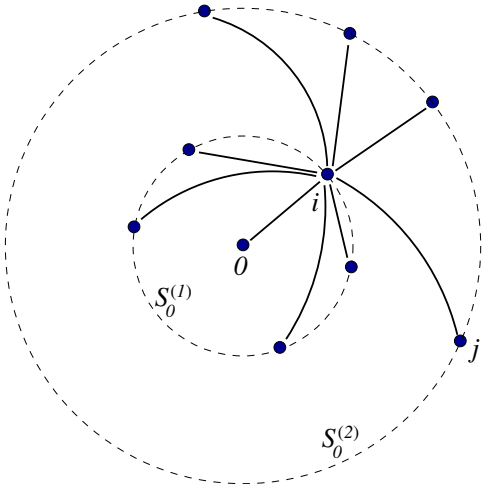

Figure 5. Node 0 has a gradient edge from i, if its scalar value is larger than the scalars of all its neighbors $j \neq 0$.

3.3. An integral transform based method

Although this method is more involved, it can be used to calculate the in-degree distribution of gradient networks for substrate graphs other than $G_{N,p}$, and for some non-i.i.d. scalar distributions. Calculations for such cases will be presented elsewhere.

When calculating the degree distribution, we have to perform two averages: one corresponding to the scalar field disorder

$$\langle \bullet \rangle_h = \int \mathrm{d}h_0 \cdots \mathrm{d}h_{N-1}\, \eta(h_0) \cdots \eta(h_{N-1}) \bullet, \tag{8}$$

and the other to an average over the network (graph ensemble):

$$\langle \bullet \rangle_{nw} = \sum_{a_{01}} \cdots \sum_{a_{N-2N-1}} v(a_{01}) \cdots v(a_{N-2N-1}) \bullet, \tag{9}$$

where $v(a) = p^a(1-p)^{1-a}$, $a \in \{0, 1\}$ and $\sum_a \equiv \sum_{a=0}^{1}$. Here G is the binomial random graph $G_{N,p}$ with N nodes and link-probability p. The integrals in (8) are computed over the range of the scalar field and the summation in (9) is over all $N(N-1)/2$ pairs (i, j) with $i < j$.

In order to calculate the in-degree distribution, we define first a counter operator for the in-links. Without restricting the generality we calculate the in-degree of the gradient network for node 0 namely, $k_0^{(in)}$. Let us introduce $\mathbf{B} = \mathbf{I} + \mathbf{A}$, where $\mathbf{I}$ is the $N \times N$ identity matrix so $b_{ij} = \delta_{i,j} + a_{ij}$, and the quantities $H_i(j) = 1 - b_{ij} + b_{ij}\theta(h_0 - h_j)$ for $i, j \in V$, and $i \in S_0^{(1)}$. Thus, the in-link counter can be written as

$$k_0^{(in)} = \sum_{i=1}^{N-1} a_{0i} \prod_{j=1}^{N-1} H_i(j). \tag{10}$$

With the aid of figure 5 we see that indeed this expression will count the number of gradient edges into node 0: $H_i(j)$ is zero only if the *neighbor* j of i (except node 0) has a larger scalar value than node 0, i.e., $h_0 < h_j$, otherwise $H_i(j)$ is equal to unity. Therefore a term under the sum in (10) will be non-zero if and only if for *all* neighbors j of i (i.e., $b_{ij} = 1$) $h_j < h_0$ holds, making the edge $(i, 0)$ to be the gradient edge for node i.

The probability that a node will have l in-degree on the gradient network $\mathbf{F}$, is

$$R(l) = \left\langle \left\langle \delta_{l,k_0^{(in)}} \right\rangle_h \right\rangle_{nw} = \int_{-\pi}^{\pi} \frac{\mathrm{d}q}{2\pi}\, \mathrm{e}^{\mathbf{i}ql} \left\langle \left\langle \mathrm{e}^{-\mathbf{i}qk_0^{(in)}} \right\rangle_h \right\rangle_{nw}. \tag{11}$$

First, we compute the average over the scalar field. (The order of the averages does not matter, however, it is formally easier this way.) Let us denote

$$L_G(q) = \left\langle \mathrm{e}^{-\mathbf{i}qk_0^{(in)}} \right\rangle_h \tag{12}$$

We have

$$L_G(q) = \int \mathrm{d}h_0 \cdots \int \mathrm{d}h_{N-1}\, \eta(h_0) \cdots \eta(h_{N-1})\, \mathrm{e}^{-\mathbf{i}q \sum_{i=1}^{N-1} a_{0i} \prod_{j=1}^{N-1} [1 - b_{ij} + b_{ij}\theta(h_0 - h_j)]}$$

Let $M_i(m) = \prod_{j=1}^{m} H_i(j)$. So

$$L_G(q) = \int \mathrm{d}h_0 \cdots \int \mathrm{d}h_{N-1}\, \eta(h_0) \cdots \eta(h_{N-1})\, \mathrm{e}^{-\mathbf{i}q \sum_{i=1}^{N-1} a_{0i} M_i(N-1)}. \tag{13}$$

Using the recursion

$$M_i(m) = [1 - b_{im} + b_{im}\theta(h_0 - h_m)]\, M_i(m-1), \tag{14}$$

the integral over h_{N-1} can be performed

$$L_G(q) = \int \mathrm{d}h_0 \cdots \int \mathrm{d}h_{N-2}\, \eta(h_0) \cdots \eta(h_{N-2}) \Big\{ \gamma(h_0)\, \mathrm{e}^{-\mathbf{i}q \sum_{i=1}^{N-1} a_{0i} M_i(N-2)} + [1 - \gamma(h_0)]\, \mathrm{e}^{-\mathbf{i}q \sum_{i=1}^{N-1} a_{0i}[1 - b_{iN-1}] M_i(N-2)} \Big\}$$

where $\gamma(x) = \int^x \mathrm{d}y\, \eta(y)$. Performing all the integrals recursively, except for h_0, we obtain

$$L_G(q) = \sum_{n=0}^{N-1} J(N,n) \sum_{[\tau]_n \in \mathcal{P}_n(N-1)} \mathrm{e}^{-\mathbf{i}q \sum_{i=1}^{N-1} a_{0i} \prod_{j=1}^{n} (1 - b_{i\tau(j)})} \tag{15}$$

where $J(N,n) = \int \mathrm{d}h_0\, \eta(h_0) [\gamma(h_0)]^{N-1-n} [1 - \gamma(h_0)]^n$. Here $[\tau]_n = \{\tau(1), \tau(2), \ldots, \tau(n)\}$ is an n-subset of the set $\{1, 2, \ldots, N-1\}$ and $\mathcal{P}_n(N-1)$ denotes the set of all n-subsets of $\{1, 2, \ldots, N-1\}$. We have $|\mathcal{P}_n(N-1)| = \binom{N-1}{n}$. After a change of variables $u = \gamma(h_0)$ and using $\mathrm{d}u = \mathrm{d}\gamma(h_0) = \eta(h_0)h_0$ the integral $J(N,n)$ yields $J(N,n) = \frac{1}{N}\binom{N-1}{n}^{-1}$, i.e., the in-degree distribution is independent of the choice of the $\eta(h)$ distribution!

In the following, we perform the network average $\langle L_G(q) \rangle_{nw}$. For a fixed n-subset $[\tau]_n$, let us denote

$$Z_{[\tau]_n}(q) \equiv \left\langle \mathrm{e}^{-\mathbf{i}q \sum_{i=1}^{N-1} a_{0i} \prod_{j=1}^{n} (1 - b_{i\tau(j)})} \right\rangle_{nw}. \tag{16}$$

Thus,

$$\langle L_G(q) \rangle_{nw} = \frac{1}{N} \sum_{n=0}^{N-1} \binom{N-1}{n}^{-1} \sum_{[\tau]_n \in \mathcal{P}_n(N-1)} Z_{[\tau]_n}(q). \tag{17}$$

Let

$$T_n = [\tau]_n \cup \bigcup_{j=1}^{n} S^{(1)}_{\tau(j)} \tag{18}$$

be the set of vertices $[\tau]_n$ and its neighbors on G.

Note, that $\prod_{j=1}^{n} (1 - b_{i\tau(j)}) = 1$ if and only if $i \notin T_n$ otherwise it is zero. Therefore,

$$\sum_{i=1}^{N-1} a_{0i} \prod_{j=1}^{n} (1 - b_{i\tau(j)}) = \text{the number of neighbors of 0 } \textit{which do not belong to } T_n. \tag{19}$$

From (9)

$$Z_{[\tau]_n}(q) = \sum_{a_{01}} \cdots \sum_{a_{N-2N-1}} v(a_{01}) \cdots v(a_{N-2N-1}) \prod_{i=1}^{N-1} \mathrm{e}^{-\mathbf{i}q a_{0i} \prod_{j=1}^{n}(1-b_{i\tau(j)})} \tag{20}$$

Since $\tau(j) \neq 0$, ($[\tau]_n \in \mathcal{P}_n(N-1)$), the sums over the matrix variables a_{0i} can be performed

$$\sum_{a_{0i}} v(a_{0i})\, \mathrm{e}^{-\mathbf{i}q a_{0i} \prod_{j=1}^{n}(1-b_{i\tau(j)})} = 1 - p + p\, \mathrm{e}^{-\mathbf{i}q \prod_{j=1}^{n}(1-b_{i\tau(j)})}, \tag{21}$$

and therefore

$$Z_{[\tau]_n}(q) = \sum_{a_{12}} \cdots \sum_{a_{N-2N-1}} v(a_{12}) \cdots v(a_{N-2N-1}) \prod_{i=1}^{N-1} \left[1 - p + p\, \mathrm{e}^{-\mathbf{i}q \prod_{j=1}^{n}(1-b_{i\tau(j)})}\right]. \tag{22}$$

The set of vertices $\{1, 2, \ldots, N-1\}$ is split into two groups: $[\tau]_n$ and its complementary in $\{1, 2, \ldots, N-1\}$. Without changing anything, we can rename the vertices, such that $\{1, 2, \ldots, n\} = [\tau]_n$ and $C_{[\tau]_n} = \{n+1, n+2, \ldots, N-1\}$ be the complementary set of $[\tau]_n$. It is easy to see that only cross-terms (a_{ij} involving one node i from $[\tau]_n$ and one node j from $C_{[\tau]_n}$) give non-trivial contribution (i.e., different from unity) in (22). Thus

$$Z_{[\tau]_n}(q) = \prod_{i=n+1}^{N-1} \sum_{a_{1i}} \cdots \sum_{a_{ni}} v(a_{1i}) \cdots v(a_{ni}) \left[1 - p + p\, \mathrm{e}^{-\mathbf{i}q \prod_{j=1}^{n}(1-a_{ji})}\right] \tag{23}$$

Let $\alpha_1 = 1 - p$ and $\beta_1 = p$. Then

$$\sum_{a_{1i}} v(a_{1i}) \left[\alpha_1 + \beta_1\, \mathrm{e}^{-\mathbf{i}q(1-a_{1i})\cdots(1-a_{ni})}\right] = \alpha_2 + \beta_2\, \mathrm{e}^{-\mathbf{i}q(1-a_{2i})\cdots(1-a_{ni})} \tag{24}$$

where $\alpha_2 = (1-p)\alpha_1 + p$ and $\beta_2 = (1-p)\beta_1$. The summation over the rest of the matrix elements can be similarly performed to give (for a fixed node i)

$$\alpha_{n+1} + \beta_{n+1}\, \mathrm{e}^{-\mathbf{i}q}. \tag{25}$$

The coefficients are determined from the recursion

$$\begin{cases} \alpha_k = (1-p)\alpha_{k-1} + p, & \alpha_1 = 1 - p \\ \beta_k = (1-p)\beta_{k-1}, & \beta_1 = p \end{cases} \tag{26}$$

which obeys $\alpha_k + \beta_k = 1$ for all k. These recursions are easily solved:

$$\alpha_{n+1} = 1 - p(1-p)^n, \qquad \beta_{n+1} = p(1-p)^n. \tag{27}$$

Thus (25) becomes $1 - p(1-p)^n + p(1-p)^n\, \mathrm{e}^{-\mathbf{i}q}$. Since for all indices i in (23) the result of the summations is the same, one finally obtains

$$Z_{[\tau]_n}(q) = [1 - p(1-p)^n(1 - \mathrm{e}^{-\mathbf{i}q})]^{N-1-n}. \tag{28}$$

Because the result in (28) is not specific of the $[\tau]_n$ set, for all realizations of $[\tau]_n$, $Z_{[\tau]_n}(q)$ is the same expression, and thus the sum over all realizations of $[\tau]_n$ in (17) will generate the factor $|\mathcal{P}_n(N-1)| = \binom{N-1}{n}$ which cancels the combinatorial factor in (17). Thus $\langle L_G(q)\rangle_{nw} = \frac{1}{N}\sum_{n=0}^{N-1} Z_{[\tau]_n}(q)$. Plugging this into (11), and performing the integral over the q variable we obtain

$$R(l) = \frac{1}{N} \sum_{n=0}^{N-1} \binom{N-1-n}{l} [1 - p(1-p)^n]^{N-1-n-l} [p(1-p)^n]^l \tag{29}$$

with the usual convention $\binom{M}{m} = 0$ for $M < m$. Equation (29) is the exact expression for the in-degree distribution of the gradient network ∇G.

3.4. Derivation of the $1/l$ *scaling for the in-degree distribution*

In order to obtain the scaling $1/l$ valid in the limit $N \to \infty$, $p \to 0$, such that $z = pN = \text{const} \gg 1$, for $0 \leqslant l \leqslant z$, we will use the saddle-point method. We write equation (2) first in the form $R(l) = \frac{1}{N}\sum_{n=1}^{N} r_N(n,l)$ and then exponentiate the argument. Using Stirling's formula to the first order ($\ln(x!) \approx x(\ln x - 1)$), one obtains that $r_N(n,l) \approx \mathrm{e}^{q_N(n,l)}$, where

$$\begin{aligned} q_N(n,l) = {} & (N-n)\ln[(N-n)/e] - l\ln(l/e) \\ & -(N-n-l)\{\ln[(N-n-l)/e] - \ln(1-p(1-p)^{n-1})\} \\ & +l[\ln p + (n-1)\ln(1-p)]. \end{aligned} \tag{30}$$

To calculate the largest contributor under the sum in (2) we use the saddle-point method: $\int \mathrm{d}x\, \mathrm{e}^{f(x)} \approx \sqrt{2\pi}\, \mathrm{e}^{f(x_0)}/\sqrt{-f''(x_0)}$ where $f'(x_0) = 0$. In our case, we thus need to consider

$$\left.\frac{\partial q_N(n,l)}{\partial n}\right|_{n^*(l)} = 0 \tag{31}$$

where $n^*(l)$ denotes the index of the maximal contributor for a given l. The difficulty we get into by trying to find $n^*(l)$ from (31) is that the equation cannot be solved explicitly for $n^*(l)$. To get around this, let us consider instead the derivative

$$\left.\frac{\partial q_N(n,l)}{\partial l}\right|_{\hat{l}(n)} = 0 \tag{32}$$

defining $\hat{l}(n)$. Performing the derivation the solution is easily found as

$$\hat{l}(n) = (N-n)p(1-p)^{n-1}. \tag{33}$$

Since $\hat{l}(n)$ is a monotonic function of n, it is invertible ($\hat{l}'(n) < 0$). The inverse of (33) will be denoted by $\hat{n}(l)$. This means that

$$\left.\frac{\partial q_N(n,l)}{\partial l}\right|_{\hat{n}(l)} = 0. \tag{34}$$

Next, we observe that $\hat{l}(n)$ satisfies (31) when inserting it into its explicit expression. Accordingly, it will also be satisfied by $\hat{n}(l)$. Assuming that there is only one solution to (31) it thus follows that

$$n^*(l) = \hat{n}(l). \tag{35}$$

If we now calculate $q_N(n,l)$, at the saddle point, we find that $q_N(n^*(l),l) = 0$ (using the fact that the parametric curve of the maximum can be written as either $(n^*(l),l)$ or $(n,\hat{l}(n))$ and thus calculating $q_N(n,\hat{l}(n))$). This means that we need to go one step further in the Stirling series, in order to calculate the leading piece of $\mathrm{e}^{\ln r_N(n,l)}$ at the saddle point. For the saddle point itself, we use the same expression as previously (obtained with the first-order Stirling approximation) because as can be shown, the corrections introduced by the next term in the Stirling approximation are vanishing as $N \to \infty$ and therefore they will be neglected. Thus using the next order term in the Stirling series ($\ln(x!) \approx x(\ln x - 1) - \ln(\sqrt{x}) + \ln(\sqrt{2\pi})$) and writing

$$\ln r_N(n,l) \approx q_N(n,l) + s_N(n,l), \tag{36}$$

where $s_N(n,l)$ is the correction generated this way, we obtain

$$\mathrm{e}^{s_N(n^*(l),l)} = \frac{1}{\sqrt{2\pi}}\sqrt{\frac{N-n^*(l)}{N-n^*(l)-l}}\frac{1}{\sqrt{l}} = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{l}} + \mathcal{O}\left(\frac{\ln z}{z}\right). \tag{37}$$

Calculating the second derivative $\partial^2 q_N(n,l)/\partial n^2$ at the point (33), one finally obtains

$$\frac{\partial^2 q_N(n,l)}{\partial n^2} = -l\frac{z^2}{N^2} - l\mathcal{O}\left(\frac{z}{N^2}\right) - l\mathcal{O}\left(\frac{z^3}{N^3}\right) \tag{38}$$

Combining (38) with (36), (37) in the saddle-point formula, one obtains that

$$R(l) \approx \frac{1}{zl}, \tag{39}$$

valid in the domain $1 \leqslant l \leqslant l_c$. The cutoff value l_c is determined by the validity range of the saddle-point method: since the function $n^*(l)$ is monotonically decreasing, at $l = l_c$ it will hit the lowest allowed value by the range of the integral (or sum), namely, at $n^*(l_c) = 1$. Since $\hat{l}$ is the inverse function of n^*, it follows that

$$l_c = \hat{l}(n^*(l_c)) = \hat{l}(1) = p(N-1) = z \tag{40}$$

meaning that the cutoff for the $1/l$ scaling law happens at z, which is indeed confirmed by the numerical simulations shown in figure 3.

4. Conclusions

If the substrate graph is a scale-free network (we used the Barabási–Albert (BA) process with parameter m to generate the scale-free network [7], but other uncorrelated scale-free graph models, such as the configuration model [26–28] will lead to similar conclusions as long as $\gamma > 2$), the gradient graph will still be scale-free as we show that via numerical simulations in [1]. The gradient network actually is a scale-free graph with the same exponent γ. We have performed (numerical) calculations for many other substrate graphs (both 'scaled' and scale-free) with the general conclusion that gradient networks tend to be power-law degree distributed graphs, thus offering a robust mechanism for producing such heterogeneous structures.

Note that our analysis did not assume anything about the properties of the edges. Certainly, if one wants to study actual throughput measures, the transport properties of the edges (conductance, or 'cost') must be specified. In this case, the overall performance of the network will depend not only on the properties of the gradient graphs but also on the distribution of the edge conductances. The notion of gradient networks has also been used to study congestion in transport networks from an optimization point of view, see [1, 30–32].

When deriving our results we made the assumption that the scalars were i.i.d. random variables. This is certainly not expected to hold in general. However, here we have only considered the simplest case, the 'zeroth'-order model. We also studied cases when there are correlations among the scalars (such as for protein folding landscapes [21]), which will be presented in future publications. In brief, those results also show that the power-law, or scale-free character of the gradient network is actually a robust feature. Typically, the correlations were found to affect the *value of the exponent* γ of these power laws.

In summary, we have shown that local gradients induced networks naturally form graph structures with power-law degree distributions. For a number of massive real-world networks, the network structure is sampled based on observation of flows on them. For such situations the flow-based observations will likely lead to scale-free structures.

Acknowledgments

The authors acknowledge useful discussions with M Anghel, B Bollobás, P L Erdős, E Lopez and D Sherrington. This work was supported by the US DOE under contract W-7405-ENG-36 (ZT, BK and NWH), KEB was supported by the NSF through DMR-0406323, DMR-0427538,

and the Alfred P Sloan Foundation. BK and GK were supported in part by NSF through DMR-0113049, DMR-0426488 and the Research Corp. grant no. RI0761.

References

[1] Toroczkai Z and Bassler K E 2004 *Nature* **428** 716
[2] Boccaletti S, Latora V, Moreno Y, Chaves M and Hwang D-U 2006 *Phys. Rep.* **424** 175
[3] Newman M E J, Barabási A-L and Watts D J 2006 *The Structure and Dynamics of Networks* (Princeton, NJ: Princeton University Press)
[4] Dorogovtsev S N and Mendes J F F 2003 *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford: Oxford University Press)
[5] Caldarelli G 2007 *Scale-Free Networks* (Oxford: Oxford University Press)
[6] Bollobás B 2001 *Random Graphs* 2nd edn (Cambridge: Cambridge University Press)
[7] Barabási A-L and Albert R 1999 *Science* **286** 509
[8] Bianconi G and Barabási A-L 2001 *Europhys. Lett.* **54** 436
[9] Krapivsky P L, Rodgers G J and Redner S 2001 *Phys. Rev. Lett.* **86** 5401
[10] Klemm K and Eguíluz V M 2002 *Phys. Rev.* E **65** 036123
[11] Aiello W, Chung F and Lu L 2000 *Proc. 32nd ACM Symp. on Theory of Computing* p 171
[12] Vázquez A, Flammini A, Maritan A and Vespignani A 2003 *ComPlexUs* **1** 38
[13] Goh K-I, Kahng B and Kim D 2001 *Phys. Rev. Lett.* **87** 278701
[14] Caldarelli G, Capocci A, de Los Rios P and Mũnoz M A 2002 *Phys. Rev. Lett.* **89** 258702
[15] Boguña M and Pastor-Satorras R 2003 *Phys. Rev.* E **68** 036112
[16] Ferrer i Cancho R and Solé R V 2003 *Lect. Notes Phys.* **625** 114
[17] Barthélemy M 2003 *Europhys. Lett.* **63** 915
[18] Nuccitelli R 1984 *Pattern Formation: A Primer in Developmental Biology* ed G M Malacinski (New York: MacMillan) p 23
[19] Rabani Y, Sinclair A and Wanka R 1998 *Proc. 39th Symp. on Foundations of Computer Science (FOCS)* p 694
[20] Anghel M, Toroczkai Z, Bassler K E and Korniss G 2004 *Phys. Rev. Lett.* **92** 058701
[21] Ravasz E, Gnanakaran S and Toroczkai Z 2007 submitted http://arxiv.org/abs/0705.0912
[22] da Fontoura Costa L and Travieso G 2005 *Int. J. Mod. Phys.* C **16** 1097
[23] Lakhina A, Byers J W, Crovella M and Xie P 2003 *Proc. IEEE Infocom*
[24] Clauset A and Moore C 2005 *Phys. Rev. Lett.* **94** 01870
[25] Petermann T and De Los Rios P 2004 *Eur. Phys. J.* B **38** 201
[26] Bekessy A, Bekessy P and Komlos J 1972 *Stud. Sci. Math. Hung.* **7** 343
[27] Bender E A and Caneld E R 1978 *J. Comb. Theory* A **24** 296
[28] Molloy M and Reed B 1995 *Comb. Probab. Comput.* **6** 161
[29] Sreenivasan S, Cohen R, Lopez E, Toroczkai Z and Stanley H E 2007 *Phys. Rev.* E **75** 036105
[30] Park K, Lai Y-C, Zhao L and Ye N 2005 *Phys. Rev.* E **71** 065105
[31] Gulbahce N 2007 *Chaos* **17** 026105
[32] Danila B, Yu Y, Earl S, Marsh J A, Toroczkai Z and Bassler K E 2006 *Phys. Rev.* E **74** 046114