INVITED PAPER

Impact of Changing Technology on the Evolution of Complex Informational Networks

This paper proposes a general framework to understand the fundamental advantages of connectivity in complex informational networks.

By Luís M. A. Bettencourt

ABSTRACT | We live in an era of increasing connectivity in human societies and in technology. These structural changes in the ways we interact with each other and with increasingly ubiquitous computational and communication devices have been formalized in research across several disciplines through the dynamics of complex informational networks. Complex networks are (mathematical) graphs, connecting nodes (people, computers) via edges (relationships, wires). While much progress in methods for network analysis has been achieved, the fundamental principles that drive network growth in human societies and in worldwide computer networks remain rather obscure. Mechanistic models for the origin of certain structural graph elements have now become common, but the formal connection between large empirical studies of network evolution and fundamental concepts of information, learning, and social theory remains only latent. To address these issues, I argue here that the most interesting aspect of the dynamics of informational networks in complex systems is that they are the physical manifestations of processes of evolution, inference, and learning, from natural ecosystems, to cities and to online environments. I formalize the general problem of learning and computation in network environments in terms of average structural network changes and propose a conceptual framework to explain the transition from initially static, undifferentiated, and information-poor environments to dynamical, richly diverse, and interconnected systems. I illustrate these ideas empirically by providing examples from cities, and from global computer networks and webs of documents. I finish with an overview of expected changes to urban form and function and to computational hardware under likely technological scenarios.

KEYWORDS | Collaborative work; communication networks; complex networks; distributed computing; intelligent systems; learning systems; urban areas; Wikipedia; World Wide Web (WWW)

I. INTRODUCTION

We live at a time of increasing connectivity. This is true of many of our most important technologies as well as of human societies themselves [1]. The rise of telecommunications and of globally networked information technologies, such as the World Wide Web (WWW), is clearly changing human societies in many ways, from our ability to create and store information and run civic institutions [2], [3] to scientific research [4], and from technological innovation [5] to human development [6], [7]. Not only are people and human organizations increasingly connected worldwide but also so are devices and engineering systems, through developments in information and communication technologies (ICTs), such as the Internet of Things [8].

Why do these trends toward greater global connectivity seem so irresistible? Why now? To attempt to answer these questions, I propose that we must build a general framework to understand the fundamental advantages of connectivity in complex informational networks, as well as their associated costs. Through the analysis of these tradeoffs and their change over time we will gain new perspectives from which we can take a new unified look at the history of many technologies—from cities, to transportation and telecommunications—and assess the future

Manuscript received May 10, 2014; revised October 15, 2014; accepted October 27, 2014. Date of current version November 18, 2014. This work was supported in part by the Bill and Melinda Gates Foundation under Grant OPP1076282; the MacArthur Foundation under Grant 13-105749-000-USP; the James S. McDonnell Foundation under Grant 220020195; the John Templeton Foundation under Grant 15705; the Army Research Office Minerva program under Grant W911NF1210097; and by a gift from Jeannie and Michael L. Klein.
The author is with Santa Fe Institute, Santa Fe, NM 87505 USA (e-mail: bettencourt@santafe.edu).

Digital Object Identifier: 10.1109/JPROC.2014.2367132

0018-9219 © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

of a more connected global human society and its relationships to Earth's other complex adaptive systems [9].

The ideas developed below are about informational networks, which are a subset of what we now refer to as complex networks [10]. Informational networks are made of nodes that are capable of changing their internal states in ways that can expand their information content. This happens in tandem with changes of connectivity as determined by their relative benefits and costs. For example, people can acquire new knowledge and expertise and exchange it through an expanded set of socioeconomic relations with others. In this sense, I will not have much to say about complex networks such as power grids, metabolic graphs, or transportation networks, which do not share these properties. Examples of informational networks may include human socioeconomic systems at different scales, networks of evolving documents, as well as, possibly, ecological networks and neuronal networks capable of learning (natural and artificial), though the latter will not be discussed here.

Because they deal with learning, informational networks pose a set of difficult problems, tied to the dynamics of innovation and productivity in distributed systems. How then may we start to understand these systems in general but simple ways? I propose here that the place to start is the realization that general computation and network connectivity are useful, because information, its discovery, and aggregation are necessary conditions for development in human societies [6], [7], [11] and for learning processes more generally. This point is not new: It is the basis for our best current ideas about economic growth [12] and human development, as well as, in different forms, about natural and technological "evolution" [5], [13], [14].

The interplay between the structure of various complex networks and their embedded information runs in both directions. While technological change that enables larger and more connected individuals provides the conditions for differentiation and learning [15]–[17], it is also true that the acquisition and management of information will remain limited when not embedded in dynamical network structures [18].

Specifically, the creation and management of information typically requires greater human social connectivity through the specialization and interdependence of knowledge in individuals and organizations [19]. These processes of individual and social change are much older than technological progress in modern computing. Although some precedents existed in simple human societies [20], they became manifest in earnest with the advent of the first urban civilizations [21]. With the industrial revolution and the development of new transportation and communication technologies, these processes gained new speed and scope, culminating with the worldwide information revolution currently under way [1], [22]. Thus, it is important to understand the basic conditions necessary for networks of intelligent nodes (such as people) to acquire information in open-ended ways, and the nature of situations when such dynamics may stall. Providing a simple (mean-field) framework to understand these dynamical changes is the main objective of this paper.

II. INFORMATIONAL BASIS FOR INCREASING CONNECTIVITY

Unlike what happens in simple physical systems where interactions between a system's elements constrain its overall structure [23], increases in connectivity in informational networks can, in some specific circumstances, lead to greater individual freedom. This statement may appear paradoxical and requires further explanation. In this section, I formalize these ideas and show when a dynamics of diversification and learning in networks can develop and how it benefits from technological change.

Before I introduce a more formal description, I would like to tap on some of our common intuition for this familiar but perhaps unexpected phenomenon. Many human social environments encourage, and indeed require, that individuals pursue different interests and vocations [24]–[26]. In modern human societies, these processes are partly formalized in educational and professional organizations. In this sense, individuals are encouraged to learn and create new information and new expertise and should expect to be rewarded for such efforts. Much of the modern explosion in technology, science, and the arts depends on the properties of these environments [24]–[26] as does entrepreneurship and human development [6], [7]. Urban [21], [24]–[26] and online environments [27] are general examples of networks where such dynamics of personal expression, learning, and sharing are not only possible on vast scales, but are, in fact, in some sense necessary.

Thus, the type of freedom that networked systems open up is intimately tied to their ability to allow individuals and groups to acquire new information, new roles, and new relationships in fluid ways. In other words, this freedom resides in the various forms of knowledge and relationships that nodes can create, and not so much on the reduction of the number of such interactions [26], [28].

Crucially, this increase in diversity and (implicitly) in individual learning and social expression in networked systems has important material consequences to human societies as it typically leads to greater (economic) productivity as a function of the size of the system, a concept know in economics as increasing returns to scale [29]. Below I discuss in more detail how this can happen as the result of network effects and, in fact, how these effects not only follow, but are necessary, to support a process of increasing connectivity and learning as part of a dynamical virtuous cycle of development. Important examples of these effects are that the economies of bigger cities tend to be larger on a per capita basis [30]–[32] than those of smaller places. Similarly, online systems such as the WWW or Wikipedia become more productive per capita (in ways that I will specify in Sections III-B and III-C) as a growing function of their size.

The essential property for this type of individual differentiation to occur is that nodes can increase their information through learning. Note that this is a necessary but not sufficient condition, as I demonstrate below. This is a natural property of human social networks, but it is currently only incipient in many technological networks, where learning at the node level (webpage, article, computer program) still requires human intervention [33].

The rest of this section is dedicated to formalizing these ideas in general terms, first by grounding the benefits of connectivity in foundational concepts from economics and social sciences, then by discussing the origin and costs of connectivity and finally by bringing it all together in a general quantitative framework for open-ended network dynamics.

A. General Advantages of Connectivity in Informational Networks

Evidence for the advantages of connectivity in informational networks is everywhere, from global trade to living in cities and from the achievements of science to the many uses of the WWW. But it has not always been like this and for a good reason: as I show below, connectivity is very costly and requires system scale and density to pay off. Here, I start from the very beginning by revisiting the basic concepts by which large-scale social connectivity in human societies has been justified and understood. In this way, I attempt to create a picture of how connectivity gradually develops in networks of certain types and how it can become more and more pervasive, under certain conditions that rely on technological change.

The general advantages of connectivity in networked informational systems are perhaps best introduced through the foundational concept of economics [34] and sociology [35], [36]: The division and coordination of labor. In the book that created modern economics, Adam Smith dedicated the three opening chapters of the *Wealth of Nations* [34] to the spectacular increases in productivity achieved by the division of labor. Smith illustrated this point through the increases in productivity of workers at a pin factory: By specializing in different small tasks and coordinating their labor as a whole to produce the final product, Smith estimated that each worker was able to produce about 480 pins/day, while a person working alone may master just a few. Thus, we obtain an increase in average labor productivity of about 100.

What are the basic ingredients of this spectacular gain in labor productivity? Adam Smith originally identified three types of effect, each of which remains important in modern complex networks though they may not be sufficient.

First, he considered the effect of learning to perform a task better, that is, the process of acquiring knowledge and expertise through accumulated experience [37]. This sort of effect has since been extensively studied in manufacturing at the organizational level and in cognitive science, at the individual level [38], [39]. A second source of productivity gains arises from the time savings resulting from avoiding switching between tasks. Finally, a third source of gains relies on the possibility that a task that has been rendered sufficiently simple through the successive division of labor can be made automatic, and, in that sense, be performed by a machine, thus saving human labor. Many technologies started this way, by the observation from a specialized worker of a solution that can save his/her labor and time.

These different sources of productivity gains are very general and clearly transcend the context of economic production in manufacturing. Thus, we should not think of the process of the division of labor, in terms of vertical integration of minutely specialized jobs in manufacturing firms (though it is that as well) but rather of the distribution of tasks in networks that are generally not hierarchies and the necessary creation of knowledge entailed by the specialized task and its integration (recombination) in many products and services [5], [17]. In this form, the ancient concept of the division and coordination of labor gains new life as a modern process, at play around us everywhere. In this modern form, it emphasizes information and communication in evolving complex networks. Many of the most modern socioeconomic phenomena— from online collective intelligence to the share economy—depend in fundamental ways on these processes.

In this light, the creation and interdependence of knowledge requires the development of complex and dynamical network structures as an evolving process that is sketched in Fig. 1. I start, for simplicity, by imagining a situation where a set of nodes (people, or other informational objects, books, computers) each approximately replicates the same functions [denoted by different colors in Fig. 1(a)]. An example is that of a subsistence human society [20], where despite some specialization of labor by sex and age, all households perform essentially the same tasks of hunting, gathering, or small-scale farming. The information content of such societies is replicated in each nuclear family [Fig. 1(a)] as they survive as individual units in interaction with their natural environment. Thus, the total information content in this situation is that of the typical unit, because each node is redundant (nondifferentiated) with all the others. This is why, in this type of disconnected network phase, information does not accumulate with increases in system size (nodes).

The situation changes radically as large-scale connectivity becomes common [Fig. 1(b)] [17]. It is then possible for nodes to differentiate and specialize on different tasks, relying on their functional complementarities to preserve overall function at the network level. So, for example, in modern urban societies most of us do not grow our own food or harvest energy and instead devote our time to extremely specialized tasks, often in services and in learning and organization. We rely on a vast number of different people (most of them strangers) for our survival in terms of

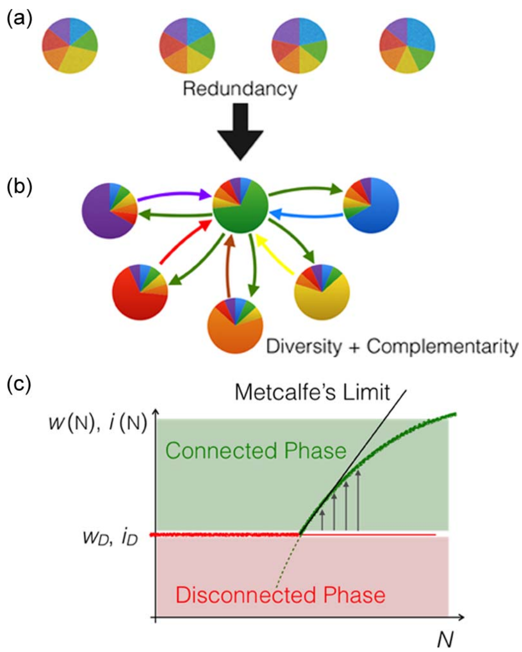

Fig. 1. ***Structural transformation in informational networks resulting from interconnection and knowledge specialization. The disconnected phase (a) is characterized by low levels of connectivity, functional redundancy (duller colors), low productivity, and slow learning; see Section II-B. As connectivity becomes less costly, a transition to a new phase (b) is possible, characterized by increasing connectivity with scale, accompanied by greater complementarity of functions (brighter colors), growing productivity, and fast overall learning at the individual and network levels. In this phase, nodes become functionally interdependent and exchange information, goods, and services (curved arrows). (Only the arrows in which the green node participates are shown, for simplicity.) (c) Nature of the transition (see Section II-D): Small systems where communication and exchange are costly will tend to be in the disconnected phase, while larger systems with inexpensive connectivity will tend to be in the connected phase and under continuous growth of information and productivity with network size N. Thus, as size or technological circumstances change, disconnected systems may become susceptible to entering the connected phase and vice versa. As this happens, the transition may be gradual, along the thick line, or sudden (arrows). Metcalfe's limit refers to the situation when productivity and information increase linearly with system size; see Table 1.***

the most critical products and services we need daily (food, water, etc.). In this situation, the information content of a network can scale up with the size of the society N as individual differentiation becomes the norm. Thus, even if the size of a network N were to remain constant, its information content will now be much larger, roughly proportional to the number of its nodes, naturally conferring economic and technological advantages to large connected systems. This effect is observable in some specific networks, such as associate memory models like Hopfield networks where it can be derived formally [40]. In real human societies, differentiation is typically not fully extensive (many people perform the same professions, for example). To deal with this issue, Adam Smith originally posited that (economic) specialization is in fact proportional to the "extent of the market" [16], [17], [34]. The factors that may limit functional differentiation of nodes in a network and consequently its potential information capacity, productivity, and diversity will be discussed further below together with a simple quantitative framework that makes more precise the qualitative picture introduced in this section.

B. Quantifying the Benefits of Connectivity

Having stated the general benefits of increasing network connectivity in terms of gains in information and productivity, I now quantify these effects in general terms, as functions of network size. I develop a simple "mean-field" model of these processes, where only average properties are taken into consideration. The development of a full statistical model, capable of accounting for the effect of fluctuations on the transition between network phases, requires additional technical development and will be pursued elsewhere.

Let us begin with the simpler disconnected phase; see Fig. 1(a). In this regime, the system is very simple and can be characterized by quantities that are independent of network size, because the nodes lack large-scale connectivity. As such we denote the constant connectivity per node (degree) $k = k_D$, the constant information content $i = i_D$, and the constant productivity per person $w = w_D$; see Table 1. Two related quantities are also worth specifying: the number of functions per person (a measure of individual specialization, e.g., the number of professions per person) $d = d_D$ and the average time spent on each task $t = t_D = T/d$, where T is the total activity time for an individual. As discussed above, I assume that in this phase, each node is a "subsistence generalist," defined by low connectivity and a large number of tasks it needs to perform to survive. As a consequence, the time per task is small, leading to low productivity as each task is associated with relatively small amounts of learning. Though the total knowledge (information) of each node may be large, it is redundant with other nodes performing similar tasks, and the information content of the system is low and of order i_D. In other words, because of low connectivity, individual functional differentiation is minimal and learning (information acquisition) is very slow, as a result of a very small amount of time spent on each task, even if network size N is large. Fig. 1(c) summarizes this situation (constant i_D, w_D on N) by the horizontal red line.

When connectivity and interdependence become possible, a network can express very different properties for the same size N. Let us for a moment ignore the costs of creating and maintaining connectivity, which are discussed in Section II-C. Then, let us suppose further that connectivity per capita increases with N. For illustration, I assume that it varies according to a scale invariant function (a power law) of the form $k(N) = k_C N^{\delta}$, where the amplitude k_C depends on technology and time (and cost)

Table 1 Characteristics of Complex Informational Network Phases

Symbol	Node Property	Disconnected Phase		Connected Phase (general)		Connected Phase (example)	
k	connectivity (degree)	k_D	(small)	$k(N)$	(increasing)	$k(N) = k_C(t)N^d$	(increasing)
d	number of functions	d_D	(large)	$d(N) = Ak^{-1}(N)$	(decreasing)	$d(N) = d_C(t)N^{-d}$	(decreasing)
i	information	i_D	(small)	$i(N) \sim t(N) \sim k(N)$	(increasing)	$i(N) = i_C(t)N^d$	(increasing)
w	productivity	w_D	(low)	$w(N)$ µ $i(N) \sim k(N)$	(increasing)	$w(N) = w_C(t)N^d$	(increasing)
t	time per function	t_D	(small)	$t(N) = A / d(N) \sim k(N)$	(increasing)	$t(N) = t_C(t)N^d$	(increasing)
c	cost per connection	c_D	(large)	$c(N) = Rj^2k(N)$	(increasing)	$c(N) = c_C(t)N^d$	(increasing)

Metcalfe's limit of the connected phase is obtained as $\delta \to 1$, see Figure 1. If information learned for the same functions across the network is redundant the total information in the connected phase is $I_C(N)=N\, i(N)d(N) \sim N$, which increases linearly with network size.

and the exponent (elasticity) δ is the rate of increase in connectivity with the size of the network and is assumed to be independent of N. This scale invariant form is predicted by urban scaling theory [41] and is observed in urban cell phone networks [42].

Then, I will assume that the number of functions accessible to each individual remains constant but becomes increasingly available to him/her through network connections [Fig. 1(b)], such that $k(N)d(N) = A$, which is independent of N. This is necessary because specializing individuals require access to functions they once maintained, e.g., a car mechanic needs access to a food producer and *vice versa*. This means specifically that the rate of increase in average individual specialization with system size equals in magnitude that of the increase in connectivity: $\dot{k}/k = -\dot{d}/d$, where the dots denote derivatives with respect to N. This quantitative behavior is observed, for example, in patterns of professional specialization and social connectivity in U.S. cities [17], and may already be present in simpler human societies [43]. Then, we conclude that $d(N) = A/k(N) = d_C N^{-\delta}$, so that each individual on average specializes in a smaller number of tasks. As a result he/she spends on average an amount of time on each task, $t(N) = T/d(N) = t_C N^{\delta}$, which increases proportionally to connectivity. Finally, we should expect that the total new information acquired (human capital) is proportional to the time on task, that is, $i(N) \propto t(N) = i_C N^{\delta}$, and that productivity is proportional to such information, $w(N) \propto i(N) = w_C N^{\delta}$, and thus, ultimately to connectivity.

These patterns are summarized in Fig. 1(c), as the rising green line, and in Table 1. They hold for any other dependence of connectivity on size, not just the illustrative power law, and express how a dynamical phase of network growth can take hold and lead to associated increases in overall information content, functional diversity, and individual productivity.

I can now also show how this reasoning maps to network effects and give an estimate of the value of connectivity. According to the so-called Metcalfe's law [44]: the *value* of connectivity is proportional to the square N^2 of the number of connected nodes N in a network. Up to a multiplicative constant, this is the maximum number of connections that can be realized in a network of N nodes. This result is readily obtained from the reasoning given above in the particular case when the parameter $\delta \to 1$, implying that each individual is connected to all others, and that, as a consequence, $k = N - 1$, and the total number of connections K is $K = Nk/2 \propto N(N-1)/2 \sim N^2$. I will refer to the regime $\delta \to 1$ as Metcalfe's limit; see Fig. 1(c).

Though Metcalfe's argument captures the potential maximum number of connections in a network, it is neither a direct measure of value nor a realistic assignment of the number of nodes that can actually be connected in a large system [41]. In order to derive the extent of connectivity in large networks, we must consider its costs, to which we now turn.

C. Costs of Connectivity

Connectivity is generally costly. A sense of the problem can be obtained by considering connectivity as a physical act of exchange. This exchange may involve the motion of physical goods, the transportation of people, or the transmission of information. Thus, each process of connectivity is mediated by a current j. In all macroscopic networks, there are dissipative energy losses associated with such exchanges that depend on the current as $C = Rj^2$. Here R is a resistance, set by whatever dissipative processes are relevant for the given exchange (e.g., friction in transportation or resistive losses for electricity) [41]. A well-known example is energy dissipation in electrical circuits (Joule's law). This reasoning also shows that energy costs are inevitable in irreversible exchanges (as a consequence of the second law of thermodynamics), whereas the translation of these costs into other units, such as money, may vary more widely, for example, as a result of the price of energy and choices of technology.

Thus, in our networked model, we expect a dissipative cost associated with each connection (as an independent current) proportional to the square of the intensity of the exchange times the relevant resistance parameter. The result in our network setting is that the cost per node $c(N) = Rj^2k(N) = c_C N^\delta$. (This becomes $c(N) = c_C N$ in Metcalfe's limit.) Thus, the ratio c_C/k_C measures the average cost per connection: It is independent of the size of the system, but is in general a function of time through technological and organizational innovations.

D. Cost-Benefit Analysis and Network Transitions

Finally, we can assemble the general picture of benefits and costs of connectivity in informational networks to derive average expectations about when the connected network phase and its associated dynamics of learning and increasing productivity may take place.

First, let us consider the net gains $w_n(N)$ from connectivity as $w_n(N) = w(N) - c(N)$. Using the expressions derived above, I can write $w_n(N)$ as

$$w_n(N) = (w_c - c_c)N^\delta = \left(\frac{w_C}{k_C} - \frac{c_C}{k_C}\right)k(N).$$

From this expression, we immediately see that the connected network phase does not always pay off; see Fig. 1(c). In particular, if the costs exceed the productivity per connection, $c_C \geq w_C$, the behavior typical of the connected phase cannot develop at all. Only in the opposite regime, when connectivity becomes inexpensive in units of productivity, does this network become able to probe its dynamical learning regime and explore the advantages of the division and interdependence of labor and information. To see the nature of this transformation more clearly, I note that the transition should occur when net productivity in the connected phase can be larger than in the disconnected phase, or, mathematically

$$w_n(N) > w_D \rightarrow k(N) > \frac{w_D k_C}{w_C - c_C} \Leftrightarrow N > \left(\frac{w_D}{w_C - c_C}\right)^{\frac{1}{\delta}}$$

where I used the power-law parameterization of $k(N)$ in the last expression. Otherwise, a different finite threshold for N would result in similar qualitative behavior. This condition shows that, everything else being equal, the transition to the connected phase is inexorable as network size N increases. Although this phenomenon may be related to ideas of development through population pressure [45] and circumscription theory [46], clearly the key to these structural changes in complex networks can be more general and the underlying necessary conditions likely more subtle. However, this transition can also be produced at fixed size N as the tradeoff between advantages and costs of connectivity in the connected phase shift. Thus, this transition may be smooth or sudden (as in a tipping point) depending on whether the system is able to immediately capitalize on the new available dynamics of connectivity, or remains temporarily stuck in the disconnected phase even as favorable circumstances for the shift develop [47]. It should also be clear that, while the mean-field model introduced here allows us to anticipate a transition between the two network phases, it does not reveal its detailed nature. For example, we cannot tell whether this transition is smooth, or first or second order, in the language of physics. Such questions will require a full statistical approach to the two network phases.

The role of technology in complex informational networks now starts to come into focus: by creating a positive benefit-to-cost tradeoff for connectivity across the largest possible realm of interactions, technological change can place networked systems on a path of collective learning and of gains in terms of diversity and productivity. Technologies here should be understood in the broadest possible sense, from cultural and political institutions that help realize the benefits of social interdependence to fast computing, or large memory and bandwidth. Most often, transformative technologies must operate both in the purely technological realm and on extant social conditions.

III. EXAMPLES: CITIES, GLOBAL ONLINE NETWORKS, AND WIKIPEDIA

To anchor some of these concepts, I now discuss some of the network properties of a few important sociotechnical systems, from cities to online informational networks.

A. Cities

Cities are first and foremost social networks of people. Throughout most of human history humans lived in small, self-sufficient groups and assumed stereotypical roles within these groups, adapted to their natural environments, e.g., as hunter-gatherers or subsistence farmers [Fig. 1(a)] [20]. This small socioeconomic connectivity and lack of strong interdependence is a general characteristic of simple human societies, from those early in history to those that remain rural and "underdeveloped" today. Thus, the (more) disconnected state of human societies is visible both cross-sectionally from large cities to the smallest towns, across places characterized by very different levels of socioeconomic development and over time. It is an important question whether the change from this immemorial way of life to modern interconnected societies can be understood as a true network transition of the type introduced above. The apparent stability of simple subsistence societies suggests the hypothesis that maybe it is.

In contrast to disconnected networks, urbanizing societies are characterized by growing settlements where frequent social interactions with many different people become possible [48], [49]. Through co-location and faster internal transportation, cities reduce the cost of social

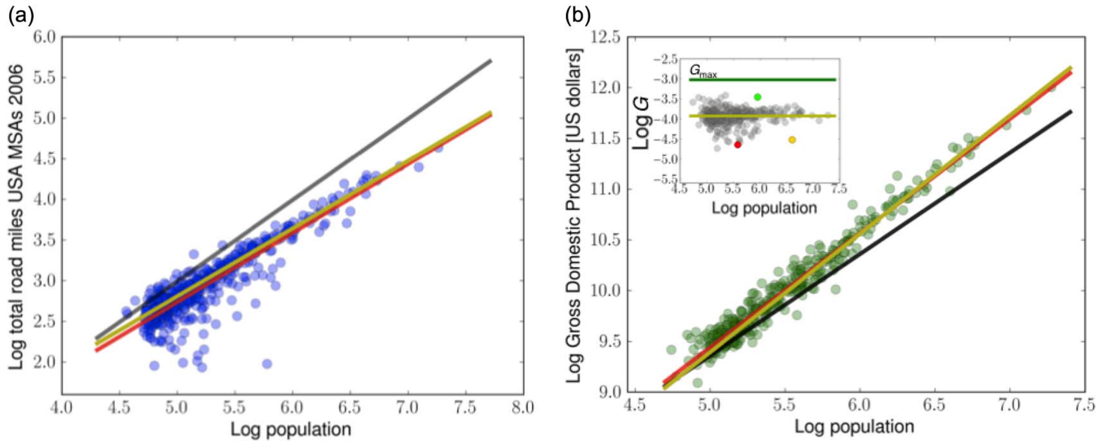

Fig. 2. ***General scaling properties of urban networks (adapted from [41]). (a) Total lane miles (volume) of roads in U.S. metropolitan areas (MSAs) in 2006 (blue dots). Data for 415 urban areas was obtained from the Office of Highway Policy Information from the Federal Highway Administration. Lines show a best fit to a power-law scaling relation with $b = 0.85$ (95% CI $= [0.81, 0.89]$, $R^2 = 0.65$). (b) Gross metropolitan product of MSAs in 2006 (green dots). Data obtained for 363 MSAs from U.S. Bureau of Economic Analysis. Lines describe best fit (red) to data, $b = 1.13$ (95% CI $= [1.10, 1.16]$, $R^2 = 0.96$). The black line shows a linear relation with unit slope. The yellow line shows the theory's prediction [41]. The inset shows the estimate of G for 313 U.S. MSAs, measured as the product of GDP and road volume, both per capita. Observed values of G for different cities are city size independent and cluster around a mean value expressing maximum net productivity, bounded by a maximum (green line) as predicted by the theory.***

connectivity and set in motion profound structural transformations in human societies [49], [50]. Thus, although the character of particular cities at specific times in history may appear more specific, urban centers are at heart the general means to the open-ended processes of human social development possible in the networked phase of social systems. In this light, as I emphasized elsewhere [41], [49], cities are the ultimate general-purpose "social reactors" [41], [49].

Because of the growing availability of data on urban quantities in cities all over the world, we are now able to measure quantitatively the expression of these processes and write them in terms of the framework developed in Section II. To set the stage, consider the results shown in Fig. 2, for the scaling of total road surface and gross domestic product (GDP) of U.S. metropolitan statistical areas (a definition of cities as interacting networks); see [41] and [50]. Both these quantities show average scaling behavior according to an expression of the form

$$Y(N,t) = Y_0(t)N^b$$

with $b = 1 + \delta$, where δ is the exponent introduced in the previous section. This follows from the fact that per capita quantities (denoted by lower case letters) and total quantities are proportional up to a factor of N, e.g., $Y = Ny$. Y, expressed as GDP, is a measure of (economic) output, and its main component, wages, which measures labor productivity, scales in the same way [41]. The plots of Fig. 2 show that the value of the exponent $\delta \simeq 1/6$, as predicted by theory [41], is based on the integrated modeling of cities as social networks embedded in infrastructural space and subject to both connectivity gains and costs as a result of social interactions in space.

In this framework, the exponents and prefactors can be computed explicitly from models of the general geometry of urban infrastructure; see [41] for details. Briefly, in this context, the exponent δ measures the rate of densification of people in public spaces (proxied by the infrastructure network volume V_n), $n = N/V_n = n_0 N^{\delta}$, where n is density. V_n, shown in Fig. 2(a), is in turn the result of building an infrastructure network with the general properties that 1) it connects all spatial parts of the city together; and 2) remains open-ended in the sense that it can be expanded gradually as the city grows. While some aspects of this calculation are specific to cities as spatial structured systems, I want to emphasize that it is the increase in density in some space that promotes greater average connectivity per unit time, $k(N) \sim n(N) \sim N^{\delta}$, as we observe directly in cities via cell phone connectivity [42]. I show in Sections II-B and II-C that this effect persists in online networks, with different exponent values, where no explicit reference to physical space is necessary.

Second, the value of prefactors such as $Y_0(t)$ is the result of the optimization of the cost benefit structure of gains minus costs following from social connectivity across all dimensions of life, that is, of the factor $w_C - c_C$. In cities, it can be shown that both benefits and costs of interaction depend on certain parameters associated with mobility (of people, goods, and information), and that these in turn are a function of the adaptation of human behavior to the characteristics of infrastructure and *vice versa* [41]. As a result it can be shown that the analog quantity to $w_C - c_C$, G in Fig. 2(b) (inset), is city size independent and has a well-defined maximum value across cities. Different cities manifest a slightly different value of their G, around this value (dots in the inset). In an analogous way, we should expect that different instantiations of networks in the connected phase should manifest values of their $w_C - c_C$ around a mean value.

Finally, while it remains difficult to find direct proxies for learning and the specialization of knowledge in cities, one can measure $d(N)$ through the consideration of the statistics of employment across a large number of different of professional occupations [17]. This analysis, applied to U.S. metropolitan areas, confirms the expectation for $d(N) \sim N^{\delta}$, with $\delta \approx 1/6$, and of labor productivity (measured through wages) exhibiting the property that $w(N)d(N) = A$, a constant independent of N, and thus that $w(N) \sim N^{\delta}$, as hypothesized above. Similarly, the social connectivity of urban social networks can be estimated at the individual level using cell phone networks [42], resulting on the scaling behavior of connectivity consistent with these observations, as predicted by theory [41].

The consistency of these ideas must continue to be explored empirically, especially in different nations and through more microscopic studies. These results do, however, provide an important illustration of the dynamics of the connected network phase proposed above and of its development over space and time in circumstances that are fundamental to understanding human sustainable development, technological change, and economic growth.

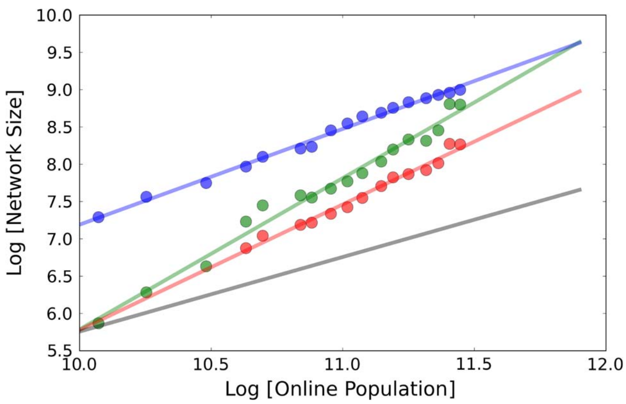

Fig. 3. ***Scaling of global computer networks with online population size. The size of the Internet, measured in terms of DNS hosts (blue) is characterized by an exponent 1.28 (95% CI = [1.22, 1.34], $R^2 = 0.99$), while the growth of the WWW, in terms of an estimate of total webpages (green), is characterized by an exponent 2.03 (95% CI = [1.88, 2.17], $R^2 = 0.98$) and of active pages (red) by an exponent 1.68 (95% CI = [1.55, 1.82] $R^2 = 0.98$). In all cases, the size of online networks has been growing superlinearly with the number of Internet users, indicating that more pages and more computation is effectively used per capita as the network grows, much like in other open-ended social systems (e.g., cities). Exponents are manifestly different from those observed for cities.***

B. The Internet and the World Wide Web

Over the last 20 years, progress in computing and telecommunication technologies has enabled unprecedented growth in connectivity between distant people. These technologies are also creating networks of knowledge that are, in specific senses, external to individual humans and their social networks and where information is instead encoded in webs of interlinked "documents," without explicit spatial location. The Internet and the World Wide Web (WWW) embody these global changes and continue to evolve from more specific and smaller networks to new and more pervasive realms.

It is, therefore, interesting to study the evolution of these networks in light of the general concepts developed above: To what extent are the Internet and the WWW examples of the connected phase dynamics of informational networks? What sort of productivity and learning are they creating?

Such enterprise is fraught with conceptual caveats and empirical limitations, some of them discussed in here below and in Section VI. The main difference from the analysis of urban data is that we do not have the ability to perform cross-sectional analysis with the Internet or the WWW, and must hope instead that their time evolution can give us a sense of the size dependence of their informational properties. If, in addition to growth in their size, the prefactors change exponentially in time, however, they will contribute to the magnitude of the exponents and are likely to result in their overestimation.

Perhaps surprisingly, the actual size and connectivity of these networks remain largely unknown both because they have become immensely large, but also because of their decentralized and bottom-up dynamics (see data sources in Section VI). Mapping them requires in practice that the entire network is visited, node by node, in order to estimate their global structure. Nevertheless, several surveys give us a sense at least of the broad structural dynamics of these networks as their user-base has increased over time. Fig. 3 shows the number of DNS hosts and two estimates of the number of webpages as functions of the total worldwide online population, in analogy to Fig. 2 for cities. The first interesting feature of Fig. 3 is that, in both cases, we observe clear superlinear scaling ($b > 1, \delta > 0$). While this sort of behavior for DNS hosts suggests an increase in task load (cost) on servers with each additional person online, the growth in number of pages is especially interesting as it suggests a (more strongly superlinear) increase in content, and thus perhaps in the productivity of the system. The number of total webpages (which we should think functionally as links between users), in particular, exhibits scaling with an exponent consistent with Metcalfe's law. However, this webpage count is plagued by certain spurious effects related to incentives to artificially create pages (see Section VI) so that the number of active sites, which scale with a smaller superlinear exponent, may be a more accurate measure.

In any case, it is interesting that we infer from these results that the number of webpages has grown with online population size at a rate much faster than social connections with the population size of urban agglomerations. In this sense, each individual may typically have access to more pieces of information online and be able to specialize on his/her own production to a larger extent.

To my knowledge, this is the first demonstration of pervasive superlinear scaling of the Internet and the WWW with online population size. However, these measures remain very rough estimates of the growth of these networks, and it would be very interesting to revisit the present results with better data. In addition, it would be desirable to obtain other measures more directly related to online connectivity, information, and individual attention and of their evolution over time.

C. Wikipedia

Another, more particular online network example, where more thorough measures of network properties are available, is Wikipedia, the online encyclopedia. Wikipedia started in January 2001 and has grown spectacularly even since, comprising currently of over 30 million articles across its large set of different languages.

Wikipedia is not a general-purpose network aimed at increasing general productivity or connectivity. Its goal is to create encyclopedia articles collaboratively, through the contributions of anyone who wishes to participate. In this sense, nodes, treated as articles, do increase their information content over time through the intervention (edits) of human contributors. Thus, even though nodes do not learn *per se*, we can treat them in analogy to the scheme developed above, with humans being a part of the connectivity structure (and bearing some of the costs) of creating and growing the network and its nodes.

The growth of the body of cross-referenced articles hence created does then provide us with a picture of how information as a whole increases and how its productivity in terms of impact may change in tandem. This happens in two ways: 1) through the iterative process of improvement of each document (which is a process of collective learning encoded as the article); and 2) through the linkages (connectivity) that an entry establishes to others, both internal and external to Wikipedia. Thus, it is this network of documents that encodes information and it is its change that represents the process of learning. Although readers of Wikipedia may also benefit (and learn) from this encoding of knowledge, contributors to Wikipedia may not individually possess all the knowledge that a single page reflects (that is the point of the collaborative model). This turns the process of learning in cities (and the parallel suggestive structure of scaling in the WWW) upside down and suggests that the best measure for the size of Wikipedia are articles and that their connectivity is supplied by human contributors as well as document links, not *vice versa*.

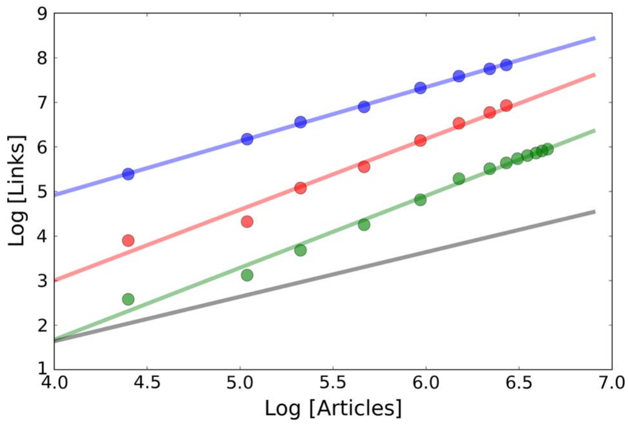

Fig. 4. ***Contributors and external and internal links to Wikipedia articles scale superlinearly. The number of contributors (green) scales with the number of Wikipedia articles with exponent $b = 1.61$ (95% CI = $[1.51, 1.72]$, $R^2 = 0.99$). The number of external links (red) scales approximately in the same way with exponent $b = 1.59$ (95% CI = $[1.40, 1.79]$, $R^2 = 0.98$). Finally, the larger number of internal links (blue) scales more slowly with an exponent $b = 1.21$ (95% CI = $[1.18, 1.24]$, $R^2 = 0.99$). If we interpret these quantities as different measures of connectivity between articles, we see that they all scale with exponents larger than those observed for social connectivity in cities.***

If we adopt this perspective, we find scaling results that broadly agree with those we invoked for cities and the WWW, but with different exponents. Fig. 4 shows how the network connectivity, measured in terms of human contributors as well as internal and external links, grows superlinearly with the number of articles. Fig. 5 helps justify the identification of Wikipedia contributors with page links, by showing that these scale linearly (proportionally) to each other. Fig. 5 also shows that the number of edits in Wikipedia is proportional to the number of contributors, supporting the assertion that the cost of connectivity per link is constant in N, as is human effort in cities [41]. Finally, Fig. 6 shows that a proxy for the productivity of an average article (and of each contributor) increases superlinearly, at least in terms of audience reach. This also establishes that the benefit of creating a connection (the effort of a contributor) is outpaced by its benefits in terms of audience reach, suggesting indeed that Wikipedia is an informational network expressing the connected phase dynamics.

These examples supply evidence that informational networks typically enable payoffs that are superlinear on the number of learner elements and that they are limited in their growth primarily by the cost of establishing and maintaining this connectivity. Whenever benefits outstrip costs, these networks can connect (and in some cases grow) explosively [Fig. 1(c)]. Eventually, they may equilibrate to a scale invariant regime where costs and benefits scale in the same (superlinear) way.

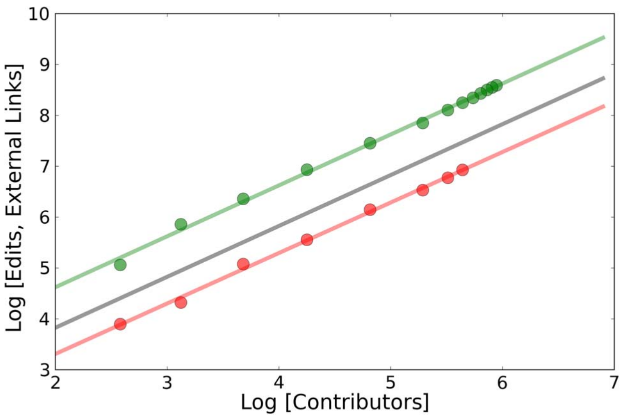

Fig. 5. *Total number of edits (green) and external links (red) to Wikipedia articles is proportional to the number of individual contributors. The green line shows the best fit to the number of edits versus contributors with exponent $b = 1.00$ (95% CI = $[0.97, 1.03]$, $R^2 = 0.99$). The red line shows the best fit to the number of external links versus contributors with exponent $b = 0.99$ (95% CI = $[0.95, 1.03]$, $R^2 = 0.99$). The black line shows exact proportionality $b = 1$, for comparison.*

IV. TECHNOLOGICAL TRENDS IN COMPLEX INFORMATIONAL NETWORKS

Finally, I would like to discuss several scenarios for the interplay between physical and informational networks and the impact of changing technology on their evolution.

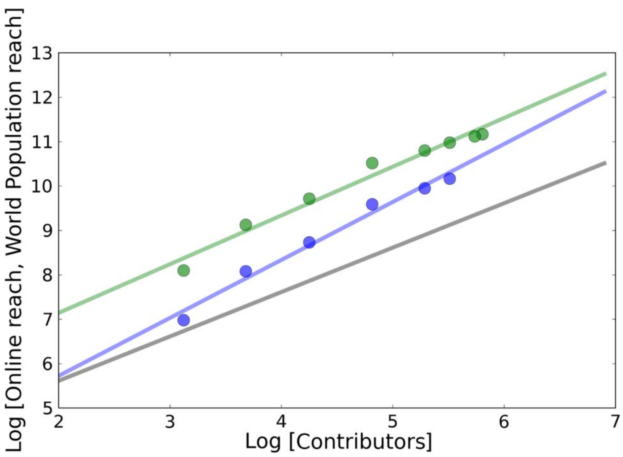

Fig. 6. *Audience reach of contributors to Wikipedia increases superlinearly. The green line shows the best fit to Alexa online reach surveys, which estimates the number of Internet users who read Wikipedia, with $b = 1.10$ (95% CI = $[0.95, 1.25]$, $R^2 = 0.97$). The blue line shows the best fit to the reach in terms of total worldwide online population with $b = 1.31$ (95% CI = $[1.23, 1.38]$, $R^2 = 0.98$). The blue line is steeper than the green line because it accounts for the growth of the online population versus the total world population, which in 2013 is estimated at 39%.*

A. The Death of Distance? Can Cyberspace Replace Physical Space?

I argued above that the dynamics characteristic of the connected network phase are primary to its "means" or processes. This means that whether this dynamics can be realized in cities, where space and infrastructure play an essential role, or online, where attention and time are apparently more relevant "spaces," is a secondary consideration.

Except, of course, to the extent that one of these modes overcomes disadvantages of another and may hence substitute it altogether. In this light, a common question is whether the Internet and information and telecommunication technologies (ICTs) can eliminate the need for cities [51], or, instead, whether these very different networks play synergetic and mutually supportive roles. The hypothesis that physical proximity becomes unnecessary is often described by the concept of the "death of distance" [51]–[56].

Research over the last decade has pursued answers to these questions with mixed results. Two general findings seem to stand out and are worth noting: 1) online and ICT networks are local; and 2) the uses of the Internet and of local social networks tend to be integrated and tend to complement, rather than substitute one another. On the first point, it has been found empirically that more online content is available in larger cities [52], [53], so that these new technologies tend to reinforce rather than replace the connectivity dynamics of larger places. This means in particular that maps, services, etc., are disproportionally available online if they stem from larger cities. In this vein, it is probably interesting to remark that previous introductions of informational and telecommunications technologies, from the newspaper and the postal service to the telegraph and the telephone, were always skewed toward larger environments, and not simply for economic reasons related to their cost [25]. On the second point, most findings are both intuitive and obvious: whether for shopping and commerce [53]–[55] or for telecommuting [56], new ICTs are extremely useful in helping organize the complex life patterns typical of larger cities, including the fine temporal coordination involved in meetings. In this sense, new informational technologies are most useful in the most intensely connected network phases, which typically are to be found in large cities [41], precisely because they lower the cost of their pervasive connectivity.

Despite these findings, the question remains whether vastly improved telecommunication technologies, capable of reproducing the nuances of sharing space and meeting face to face, can one day replace personal travel [56]. While there is no reason to exclude such possibility, my guess is that all connected networks will tend to mesh together and reinforce one another and that substitution is only possible when new modes fully include and transcend the advantages of previous modes of interaction and learning.

B. Technological Change, Connectivity, and Information in Constant-Size Networks

In the framework of Section II, I entangled the issue of increases in the information content and productivity of a networked information system together with changes in its size N. Here, I briefly discuss how these processes can take place independently, that is, even when N is constant. The key issue is the variation in the remaining parameters, prefactors, and exponents. This can generally be captured by their time dependence.

Consider then changes in the baseline productivity per social connection w_C and in the costs of connectivity c_C. The first can be affected by the adoption or improvement of new production methods or increased learning rates, while the latter may change due to improvements in transportation or through the development of social and political institutions that handle conflict resolution more effectively [41]. Thus, the crucial consideration is how baseline net productivity $w_C - c_C$ changes over time.

An increase in this quantity moves the green curve in Fig. 1(c) up, shifting its intersection with the horizontal line to the left and consequently makes the connected network phase more advantageous sooner; that is, at smaller network sizes N. Conversely, a decrease in the baseline net productivity moves the green line down and reduces the attractiveness of connectivity, delaying its onset to larger N.

Thus, we see that certain temporal shifts in these baseline parameters are sufficient, at fixed N, to produce the transition to large-scale learning and growth in complex networks. This is analogous to ideas of intensive economic growth [12] through technological change in theories of endogenous economic growth [14]. Endogenizing growth, in turn, would require additional models for how N changes over time and its associated costs, and whether a partial allocation of productivity to such costs can generate a virtuous cycle of growth and learning, or will instead fizzle out. Such elaborations are left to future work.

C. Resilience of Connected Informational Networks

Finally, it is important to discuss some of the pitfalls of the simple dynamics of differentiation, learning, and growth described above. A more complete consideration of other important factors, not treated in detail in this paper, is addressed in Section V.

First, the path of increasing individual specialization may, in some circumstances, lead to static arrangements that stall processes of open-ended learning and productivity increases. In real circumstances in human societies, extreme labor and knowledge specialization are sometimes only possible inside vertically integrated organizations (hierarchies), such as those of large firms and of universities and research laboratories [57]. Such environments can promote the stability and continuity necessary for the pursuit of more speculative R&D or for extreme specialization, say in an assembly line, in ways that economic markets often do not support. The danger of this internal specialization is that knowledge hence created remains tied to its very specific context and cannot be used in new generative ways in large networks [58]. In science and technology, specific communication channels, such as the publication of scientific manuscripts and patents, help bridge this gap, but much knowledge still remains tacit and local. This difficulty may prove more severe in large-scale manufacturing where factory-floor workers are typically at once very specifically matched to their tasks and redundant with each other and with automated solutions. In these circumstances, labor is not free to specialize further, or to learn in ways that may benefit the individual over the long term or the networked system as a whole.

This creates an apparent contradiction: While the creation and full use of specialized knowledge often requires protective environments inside stable organizations, its value depends on broad openness and exchange at the network level. These two processes, taken together, suggest that dynamics of formation and dissolution of organizations (such as hierarchies) are likely necessary for new information to be created and for it to realize its full value. This concept is partly captured by the idea of quasi-decomposability of hierarchies [59]. This can be achieved through open and dynamical labor markets, entrepreneurship, and processes by which knowledge can be accumulated in stable but open ways, for example, through open platforms, such as in online Wikis and open-source repositories.

A second issue relates to the resilience of the connected network phase. It should be clear that the disconnected phase, though it is characterized by low productivity and information content, is generally very robust to the loss of nodes. This is a direct result of its informational redundancy at the node level, a mechanism that is often employed in engineering solutions to ensure against random local failures [60].

The source of resilience of the connected state emerges not so much from its structure, but rather from its dynamics: In this phase, the loss of nodes implies some loss of information, and the loss of connections may reverse the process of learning; but these processes are, to some extent, reversible. The idea is that being ultimately dynamical the system can adjust to a loss in system size by tracing its evolution backward. This implies some loss in knowledge as well as some degradation of productivity, but still maintains the system ready to bounce back and reevolve again.

Thus, the question of resilience in informational networks is whether, upon a shock, the connected system can degrade gracefully and bounce back quickly. Anecdotal evidence from recent disasters in cities suggests that people can take up many of the functions that are usually performed by infrastructure and services [61]. Examples are walking or bicycling as a substitute to mass transit. But, the possibility remains that fast and reversible adaptation may not always be possible and that sudden, hard to reverse transitions may occur, accompanied by the destruction of network connectivity and of critical information. The exploration of such important phenomena requires further development of the ideas discussed here in terms of their statistical dynamics.

V. DISCUSSION AND CONCLUSION

In this paper, I proposed general and intentionally very simple dynamics of network development and showed that under certain circumstances nodes capable of learning may remain disconnected, redundant, and relatively unproductive, while in other situations, they may develop pervasive connectivity and embark on an open-ended trajectory of growth and development.

These ideas were largely conceived in the context of cities, though I argued here that they likely reflect more fundamental concepts common across several disciplines and apply also, at a different scope and speed, to worldwide computer networks and webs of information, such as the Internet and Wikipedia.

While I hope that the framework developed here provides a simple integrated perspective on the open-ended evolution of complex informational networks, much remains to be done to further develop these ideas theoretically, to test them empirically, and to make them useful in guiding policy and technological development.

Regarding theory, a development of more formal definitions of information and learning processes and their relationship to network structure is clearly necessary. This issue is already reasonably well developed technically on some fronts, such as in information theory [40], [62], Bayesian networks [63], and related artificial neural networks [33]. Nevertheless, the concept of node differentiation and learning through changes in their internal states occurring in tandem with related changes in network connectivity may provide new formal developments in such models, which remain to date incapable of the kind of learning that is common for humans and their social networks [33].

On this vein, the consideration of network structures that go beyond changes in average connectivity will almost certainly be necessary. Connectivity and learning costs, as well as productivity increases, may be facilitated by the creation and dissolution of formal organizations, such as firms, civic associations, and other local and online communities. Such entities are characterized by complex internal organization, often in the form of emergent hierarchical structures. Consequently, I expect that, besides average increases in connectivity, local network structures should play an essential role in the development of any real system. It may be interesting to develop a version of the quantities considered here that is local within larger networks (such as firms inside markets) and thus gain access to the more explicit dynamics of spreading processes that may result from successful local adaptations as they gain a foothold across larger networks.

Finally, I purposefully avoided explicit considerations of the cost and benefits of learning and information. This is an old problem, bypassed in the original formulation of information theory [40], [62] but essential to the understanding of evolution and development in complex systems. Such costs and benefits can be measured in many different units, such as money or energy. However, it seems obvious that such measures fail to capture much of the dynamics of online networks. In these cases, different measures of value may be necessary, such as, for example, attention (i.e., human time) [64], but it remains unclear whether they are sufficient.

A crucial issue going forward is our ability to understand and predict quantitatively how these connected informational networks will evolve, including the detailed consequences of specific new technologies. At present, in society and in technology, we are dependent on human creativity and intervention to achieve most processes of learning and knowledge recombination. Understanding these processes in more integrated and detailed ways will ultimately create a new science of complexity that can explain massive evolving informational networks such as ecosystems, cities, and the World Wide Web. It will also create a world of more intelligent technologies, able to coevolve and coadapt in connection with humans and with other natural and artificial networked complex systems.

VI. MATERIALS AND METHODS

Data on the size of the WWW (the number of total websites and the number of active websites) shown in Fig. 2 were obtained through the Web Server Survey by Netcraft (news.netcraft.com). The methods for this survey are described at www.netcraft.com/active-sites, including the distinction between all sites and active sites. The need for this distinction, which is aimed at excluding some websites created automatically and others targeted at increasing search engine visibility, was less crucial in the early years of the WWW. The size of the Internet, measured by the number of domain name system (DNS) hosts, was obtained from the Internet domain survey by the Internet Systems Consortium (ftp.isc.org/www/survey/reports/2013/07). The number of worldwide people online was obtained from estimates in the International Telecommunications Union annual reports (www.itu.int) in percent and converted into a number using estimates of the world's total population obtained from geohive (www.geohive.com). I emphasize that the nature of most measurements of the size of the WWW, the Internet, and online population is obtained through surveys and may be subject to incompleteness and biases. Wikipedia size and usage data were obtained from online Wikipedia statistics (stats.wikimedia.org). ■

Acknowledgment

The author would like to thank J. Lobo and S. Ortman for discussions on theories of economic growth and innovation.

REFERENCES

[1] M. Castells, *Rise of the Network Society*. New York, NY, USA: Wiley-Blackwell, 1996.

[2] M. Hilbert and P. López, "The world's technological capacity to store, communicate, and compute information," *Science*, vol. 332, pp. 60–65, Apr. 2011.

[3] B. Goldstein, Ed., *Beyond Transparency: Open Data and the Future of Civic Innovation*. San Francisco, CA, USA: Code for America Press, 2013.

[4] J. Bollen *et al.*, "Clickstream data yields high-resolution maps of science," *PLoS ONE*, vol. 4, Mar. 2009, e4803.

[5] W. B. Arthur, *The Nature of Technology: What It Is and How It Evolves*. New York, NY, USA: Free Press, 2011.

[6] J. Holston, *Insurgent Citizenship*. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[7] W. R. Easterley, *The Elusive Quest for Growth*. Cambridge, MA, USA: MIT Press, 2002.

[8] J. Höller *et al.*, *From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence*. Amsterdam, The Netherlands: Elsevier, 2014.

[9] National Research Council, Policy Division, Board on Sustainable Development, *Our Common Journey: A Transition toward Sustainability*. Washington, DC: National Academy Press, 1999.

[10] M. Newman, *Networks: An Introduction*. Oxford, U.K.: Oxford Univ. Press, 2010.

[11] C. Hidalgo, B. Klinger, A.-L. Barabasi, and R. Hausmann, "The product space conditions the development of nations," *Science*, vol. 317, pp. 482–487, Jul. 2007.

[12] R. J. Barro and X. I. Sala-i-Martin, *Economic Growth*. Cambridge, MA, USA: MIT Press, 2003.

[13] S. A. Frank, "The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis," *Evolution*, vol. 51, pp. 1712–1729, Dec. 1997.

[14] P. M. Romer, "The origins of endogenous growth," *J. Econ. Perspect.*, vol. 8, pp. 3–22, Dec. 1994.

[15] F. Hayek, "The use of knowledge in society," *Amer. Econ. Rev.*, vol. 35, pp. 519–530, Sep. 1945.

[16] G. S. Becker and K. M. Murphy, "The division of labor, coordination costs, and knowledge," *Quart. J. Econ.*, vol. CVII, pp. 1137–1160, Nov. 1992.

[17] L. M. A. Bettencourt, H. Samaniego, and H. Youn, "Professional diversity and the productivity of cities," *Sci. Rep.*, vol. 4, p. 5393, Jun. 2014.

[18] E. L. Glaeser and M. G. Resseger, "The complementarity between cities and skills," *J. Reg. Sci.*, vol. 50, pp. 221–244, Jan. 2010.

[19] S. E. Page, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton, NJ, USA: Princeton Univ. Press, 2007.

[20] A. W. Johnson and T. Earle, *The Evolution of Human Societies*. Stanford, CA, USA: Stanford Univ. Press, 2000.

[21] L. Mumford, *The City in History: Its Origins, Its Transformations, and Its Prospects*. New York, NY, USA: Harcourt, Brace & World, 1961.

[22] F. Braudel, *Civilization & Capitalism 15th–18th Century, Vol. 1, 2, 3*. New York, NY, USA: Harper, 1979.

[23] P. W. Anderson, "More is different," *Science*, vol. 177, pp. 393–396, Aug. 1972.

[24] R. Florida, *Cities and the Creative Class*. New York, NY, USA: Routledge, 2004.

[25] P. Hall, *Cities in Civilization*. New York, NY, USA: Pantheon Books, 1998.

[26] J. Jacobs, *The Economy of Cities*. New York, NY, USA: Vintage, 1970.

[27] D. Tapscott and A. D. Williams, *Wikinomics: How Mass Collaboration Changes Everything*. New York, NY, USA: Portfolio Trade, 2010.

[28] R. Botsman and R. Roo, *What's Mine Is Yours: The Rise of Collaborative Consumption*. New York, NY, USA: Harper Business, 2010.

[29] W. B. Arthur, *Increasing Returns and Path Dependence in the Economy*. Ann Arbor, MI, USA: Univ. Michigan Press, 1994.

[30] L. Sveikauskas, "The productivity of cities," *Quart. J. Econ.*, vol. 89, pp. 393–413, Aug. 1975.

[31] S. S. Rosenthal and W. C. Strange, "The determinants of agglomeration," *J. Urb. Econ.*, vol. 50, pp. 191–229, Sep. 2001.

[32] G. Duranton and D. Puga, "Diversity and specialisation in cities: Why, where and when does it matter?" *Urban Studies*, vol. 37, pp. 533–555, Aug. 2000.

[33] S. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach*. New York, NY, USA: Prentice-Hall, 2009.

[34] A. Smith, *An Inquiry Into the Nature and Causes of the Wealth of Nations*. London, U.K.: W. Strahan, 1776.

[35] E. Durkheim, *The Division of Labor in Society*. New York, NY, USA: Free Press.

[36] M. Weber, *The Theory of Social and Economic Organization*, T. Parsons, Ed. New York, NY, USA: Free Press, 1947.

[37] K. Arrow, "The economic implications of learning by doing," *Rev. Econ. Studies*, vol. 29, pp. 155–173, Jun. 1962.

[38] B. Nagy, J. D. Farmer, Q. M. Bui, and J. E. Trancik, "Statistical basis for predicting technological progress," *PLoS ONE*, vol. 8, Feb. 2013, e52669.

[39] K. A. Ericsson, N. Charness, P. J. Feltovich, and R. R. Hoffman, Eds., *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[40] D. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[41] L. M. A. Bettencourt, "The origin of scaling in cities," *Science*, vol. 340, pp. 1438–1441, Jun. 2013.

[42] M. Schläpfer *et al.*, "The scaling of human interactions with city size," *J. Roy. Soc. I.*, vol. 6, Sep. 2014, 20130789.

[43] M. A. Kline and R. Boyd, "Population size predicts technological complexity in Oceania," *Proc. Roy. Soc. Biol. Sci.*, vol. 277, pp. 2559–2564, Apr. 2010.

[44] Wikipedia, "Metcalfe's Law." [Online]. Available: http://en.wikipedia.org/wiki/Metcalfe's_law

[45] E. Boserup, *Population and Technological Change: A Study of Long-Term Trends*. Chicago, IL, USA: Univ. Chicago Press, 1981.

[46] R. L. Carneiro, "The transition from quantity to quality: A neglected causal mechanism in accounting for social evolution," *Proc. Nat. Acad. Sci. USA*, vol. 97, pp. 12 926–12 931, Nov. 2000.

[47] D. Achlioptas, R. D'Souza, and J. Spencer, "Explosive percolation in random networks," *Science*, vol. 323, pp. 1453–1455, Mar. 2009.

[48] L. Wirth, "Urbanism as a way of life," *Amer. J. Soc.*, vol. 44, pp. 1–24, Jul. 1938.

[49] L. M. A. Bettencourt, "The kind of problem a city is," SFI working paper 2013-03-008, 2013. [Online]. Available: http://www.santafe.edu/media/workingpapers/13-03-008.pdf

[50] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, "Growth, innovation, scaling, and the pace of life in cities," *Proc. Nat. Acad. Sci. USA*, vol. 104, pp. 7301–7306, Apr. 2007.

[51] S. Kirsch, "The incredible shrinking world? Technology and the production of space," *Environ. Plann. D*, vol. 13, pp. 529–555, Mar. 1995.

[52] J. Kolko, "The death of cities? The death of distance? Evidence from the geography of commercial Internet usage," in *The Cities in the Global Information Society: An International Perspective*. Cambridge, MA, USA: MIT Press, 1999, pp. 73–98.

[53] T. Sinai and J. Waldfogel, "Geography and the Internet: Is the Internet a substitute or a complement for cities?" *J. Urban Econ.*, vol. 56, pp. 1–24, Jul. 2004.

[54] E. Tranos and P. Nijkamp, "The death of distance revisited: Cyber-place, physical and relational proximities," *J. Reg. Sci.*, 2013, DOI: 10.1111/jors.12021.

[55] J. Pons-Novell and E. Viladecans-Marsal, "Cities and the Internet: The end of distance?" *J. Urban Tech.*, vol. 13, pp. 109–132, Dec. 2010.

[56] P. Zhu, "Are telecommuting and personal travel complements or substitutes?" *Ann. Reg. Sci.*, vol. 48, pp. 619–639, Apr. 2012.

[57] K. Arrow, *The Limits of Organizations*. New York, NY, USA: Norton, 1974.

[58] J. Jacobs, *The Nature of Economies*. New York, NY, USA: Random House, 2000.

[59] H. A. Simon, "The architecture of complexity," in *The Sciences of the Artificial*. Cambridge, MA, USA: MIT Press, 1969.

[60] S. Naeem, "Species redundancy and ecosystem reliability," *Conserv. Biol.*, vol. 12, pp. 39–45, Feb. 1998.

[61] E. G. T. Tompson, J. Benz, J. Agiesta, K. Cagney, and M. Meit, "Resilience in the wake of Superstorm Sandy," The Associated Press/NORC Center for Public Affairs Research. [Online]. Available: http://www.apnorc.org/projects/Pages/resilience-in-the-wake-of-superstorm-sandy.aspx

[62] E. T. Jaines, *Probability Theory: The Logic of Science*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[63] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[64] F. Wu and B. A. Huberman, "Novelty and collective attention," *Proc. Nat. Acad. Sci. USA*, vol. 104, pp. 17599–17601, Oct. 2007.

ABOUT THE AUTHOR

Luís M. A. Bettencourt was born in Lisbon, Portugal, in 1969. He received the Licenciatura degree (five years) in engineering physics from the Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal, in 1992 and the Ph.D. degree in theoretical physics from Imperial College, University of London, London, U.K., in 1996.

From 1997 to 2003, he was a Postdoctoral Fellow at The Institute for Theoretical Physics, Heidelberg, Germany; a Director's Fellow at Los Alamos National Laboratory, Los Alamos, NM; and a Postdoctoral Associate at the Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA, USA. In 2003, he became a Staff Member in the Computer and Computational Sciences at the Theoretical Division, Los Alamos National Laboratory. Since November 2011, he has been a Professor at the Santa Fe Institute, Santa Fe, NM, USA. Currently, he works on complex systems theory and different aspects of cities and urbanization. His research emphasizes the creation of new interdisciplinary synthesis to describe cities in quantitative and predictive ways, informed by the growing availability of empirical data worldwide. He is particularly interested in the interplay between information, structure, and scale in setting the properties of diverse complex systems.