Geometry of river networks. I. Scaling, fluctuations, and deviations

Peter Sheridan Dodds[1,2,*] and Daniel H. Rothman[2,†]
[1]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*
[2]*Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*
(Received 18 May 2000; published 27 December 2000)

This paper is the first in a series of three papers investigating the detailed geometry of river networks. Branching networks are a universal structure employed in the distribution and collection of material. Large-scale river networks mark an important class of two-dimensional branching networks, being not only of intrinsic interest but also a pervasive natural phenomenon. In the description of river network structure, scaling laws are uniformly observed. Reported values of scaling exponents vary, suggesting that no unique set of scaling exponents exists. To improve this current understanding of scaling in river networks and to provide a fuller description of branching network structure, here we report a theoretical and empirical study of fluctuations about and deviations from scaling. We examine data for continent-scale river networks such as the Mississippi and the Amazon and draw inspiration from a simple model of directed, random networks. We center our investigations on the scaling of the length of a subbasin's dominant stream with its area, a characterization of basin shape known as Hack's law. We generalize this relationship to a joint probability density, and provide observations and explanations of deviations from scaling. We show that fluctuations about scaling are substantial, and grow with system size. We find strong deviations from scaling at small scales which can be explained by the existence of a linear network structure. At intermediate scales, we find slow drifts in exponent values, indicating that scaling is only approximately obeyed and that universality remains indeterminate. At large scales, we observe a breakdown in scaling due to decreasing sample space and correlations with overall basin shape. The extent of approximate scaling is significantly restricted by these deviations, and will not be improved by increases in network resolution.

DOI: 10.1103/PhysRevE.63.016115 PACS number(s): 64.60.Ht, 92.40.Fb, 92.40.Gc, 68.70.+w

I. INTRODUCTION

Networks are intrinsic to a vast number of complex forms observed in the natural and manmade world. Networks repeatedly arise in the distribution and sharing of information, stresses, and materials. Complex networks give rise to interesting mathematical and physical properties as observed in the Internet [1], the ''small-world'' phenomenon [2], the cardiovascular system [3], force chains in granular media [4], and the wiring of the brain [5].

Branching, hierarchical geometries make up an important subclass of all networks. Our present investigations concern the paradigmatic example of river networks. The study of river networks, though more general in application, is an integral part of geomorphology, the theory of earth surface processes and form. Furthermore, river networks are held to be natural exemplars of allometry, i.e., how the dimensions of different parts of a structure scale or grow with respect to each other [6–10]. The shapes of drainage basins, for example, are reported to elongate with increasing basin size [11–13].

At present, there is no generally accepted theory explaining the origin of this allometric scaling. The fundamental problem is that an equation of motion for erosion, formulated from first principles, is lacking. The situation is somewhat analogous to issues surrounding a description of the dynamics of granular media [14,15], noting that erosion is arguably far more complex. Nevertheless, a number of erosion equations have been proposed, ranging from deterministic [16–19] to stochastic theories [20–25]. Each of these models attempts to describe how eroding surfaces evolve dynamically. In addition, various heuristic models of both surface and network evolution also exist. Examples include simple lattice-based models of erosion [17,26–28], an analogy to invasion percolation [29], the use of optimality principles and self-organized criticality [7,30,31], and even uncorrelated random networks [10,32,33]. Since river networks are an essential feature of eroding landscapes, any appropriate theory of erosion must yield surfaces with network structures comparable to that of the real world. However, no model of eroding landscapes or even simply of network evolution unambiguously reproduces the wide range of scaling behavior reported for real river networks.

A considerable problem facing these theories and models is that the values of scaling exponents for river network scaling laws are not precisely known. One of the issues we address in this work is universality [10,12]. Do the scaling exponents of all river networks belong to a unique universality class, or are there a set of classes obtained for various geomorphological conditions? For example, theoretical models suggest a variety of exponent values for networks that are directed versus nondirected, created on landscapes with heterogeneous versus homogeneous erosivity, and so on [10,34,35]. Clearly, refined measurements of scaling exponents are imperative if we are to be sure of any network belonging to a particular universality class. Moreover, given

*Author to whom correspondence should be addressed; Electronic address: dodds@segovia.mit.edu; URL:http://segovia.mit.edu/
†Electronic address: dan@segovia.mit.edu

that there is no accepted theory derivable from simple physics, more detailed phenomenological studies are required.

Motivated by this situation, here we perform a detailed investigation of the scaling properties of river networks. We analytically characterize fluctuations about scaling, showing that they grow with system size. We also report significant and ubiquitous deviations from scaling in real river networks. This implies surprisingly strong restrictions on the parameter regimes where scaling holds, and cautions against measurements of exponents that ignore such limitations. In the case of the Mississippi basin, for example, we find that although our study region spans four orders of magnitude in length, scaling may be deemed valid over no more than 1.5 orders of magnitude. Furthermore, we repeatedly find the scaling within these bounds to be only approximate, and that no exact, single exponent can be deduced. We show that scaling breaks down at small scales due to the presence of linear basins and at large scales due to the inherent discreteness of network structure and correlations with overall basin shape. Significantly, this latter correlation imprints upon river network structure the effects and history of geology.

This paper is the first of a series of three on river network geometry. Having addressed scaling laws in the present work, we proceed, in the second and third papers [36,37], to consider river network structure at a more detailed level. In Ref. [36] we examine the statistics of the ''building blocks'' of river networks, i.e., segments of streams and subnetworks. In particular, we analytically connect distributions of various kinds of stream length. Part of this material is employed in the present paper, and is a direct generalization of Horton's laws [38,39]. In the third paper [37], we proceed from the findings of Ref. [36] to characterize how these building blocks fit together. Central to this last work is a study of the frequency and spatial distributions of tributary branches along the length of a stream, and is itself a generalization of the descriptive picture of Tokunaga [40–42].

II. BASIN ALLOMETRY

A. Hack's law

In addressing these broader issues of scaling in branching networks, we set as our goal to understand the river network scaling relationship between basin area a and the length l of a basin's main stream:

$$l \propto a^h. \tag{1}$$

Known as Hack's law [11], this relation is central to the study of scaling in river networks [12,39]. Hack's exponent h is empirically found to lie in the range from 0.5 to 0.7 [11–13,43–48]. Here we postulate a generalized form of Hack's law that shows good agreement with data from real world networks.

We focus on Hack's law because of its intrinsic interest, and also because many interrelationships between a large number of scaling laws are known, and only a small subset are understood to be independent [12,39]. Thus our results for Hack's law will be in principle extendable to other scaling laws. With this in mind, we will also discuss probability densities of stream length and drainage area.

Hack's law is stated rather loosely in Eq. (1), and implicitly involves some type of averaging which needs to be made explicit. It is most usually considered to be the relationship between *mean* main stream length and drainage area, i.e.,

$$\langle l \rangle \propto a^h. \tag{2}$$

Here $\langle \cdot \rangle$ denotes an ensemble average, and $\langle l \rangle = \langle l(a) \rangle$ is the mean main stream length of all basins of area a. Typically, one performs linear regression analysis on $\log\langle l \rangle$ against $\log a$ to obtain the exponent h.

B. Fluctuations and deviations

In seeking to understand Hack's law, we are naturally led to wonder about the underlying distribution that gives rise to this mean relationship. By considering fluctuations, we begin to see Hack's law as an expression of basin morphology. What shapes of basins characterized by (a,l) are possible, and with what probability do they occur?

An important point here is that Hack's law does not exactly specify basin shapes. An additional connection to the Euclidean dimensions of a basin is required. We may think of a basin's longitudinal length $L_\parallel$ and its width $L_\perp$. The main stream length l is reported to scale with $L_\parallel$ as

$$l \propto L_\parallel^d, \tag{3}$$

where typically $1.0 \lesssim d \lesssim 1.1$ [12,50]. Hence we have $a \propto l^{1/h} \propto L_\parallel^{d/h}$. All other relevant scaling law exponents can be related to the pair of exponents (d,h) which therefore characterize the universality class of a river network [10,39]. If $d/h = 2$, we have that basins are self-similar, whereas if $d/h < 2$, we have that basins are elongating. So, while Hack's law gives a sense of basin allometry, the scaling properties of main stream lengths also need to be known in order to properly quantify the scaling of basin shape.

In addition to fluctuations, complementary insights are provided by the observation and understanding of deviations from scaling. We are thus interested in discerning the regularities and quirks of the joint probability distribution $P(a,l)$. We will refer to $P(a,l)$ as the *Hack distribution*.

Hack distributions for the Kansas River basin and the Mississippi River basin are given in Figs. 1(a) and 1(b). Fluctuations about and deviations from scaling are immediately evident for the Kansas River and to a lesser extent for the Mississippi River. The first section of the paper will propose and derive analytic forms for the Hack distribution under the assumption of uniform scaling with no deviations. Here, as well as in the following two papers of this series [36,37], we will motivate our results with a random network model originally due to Scheidegger [33].

We then expand our discussion to consider deviations from exact scaling. In the case of the Kansas River, a striking example of deviations from scaling is the linear branch separated from the body of the main distribution shown in Fig. 1(a). This feature is less prominent in the lower resolution Mississippi data. Note that this linear branch is not an artifact

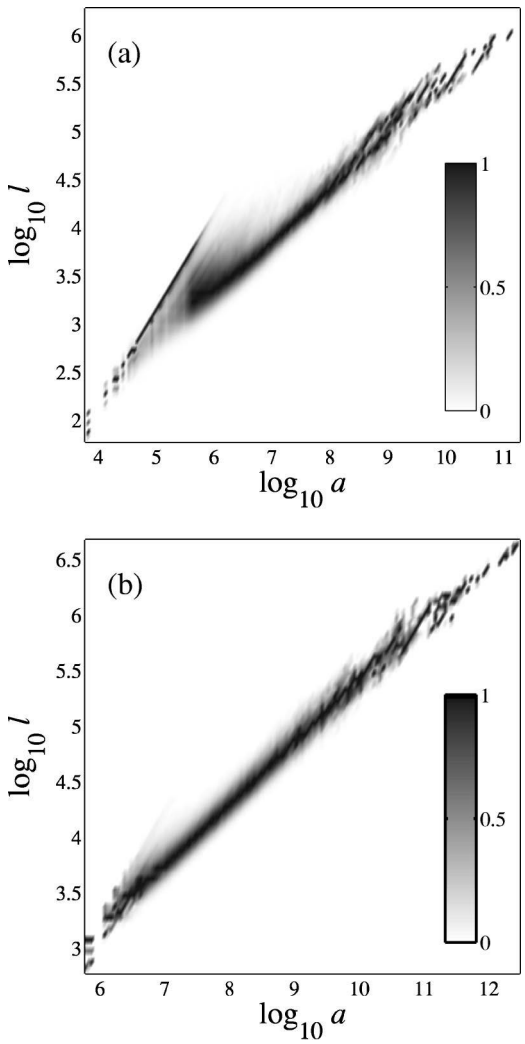

FIG. 1. The Full Hack distribution for the Kansas, (a) and Mississippi (b) River basins. Area a and length l are in m^2 and m, respectively. For each value of a, the distribution has been normalized along the l direction by $\max_l P(a,l)$ [49].

of the measurement technique or data set used. This will be explained in our discussion of deviations at small scales in Sec. VII.

We then consider the more subtle deviations associated with intermediate scales. At first inspection, the scaling appears to be robust. However, we find gradual drifts in ''exponents'' that prevent us from identifying a precise value of h, and hence a corresponding universality class.

Both distributions also show breakdowns in scaling for large areas and stream lengths, and this is addressed in the final part of Sec. VIII. The reason for such deviations is partly due to the decrease in number of samples and hence self-averaging, as area and stream lengths are increased. However, we will show that the direction of the deviations depends on the overall basin shape. We will quantify the extent to which such deviations can occur, and the effect that they have on measurements of Hack's exponent h. Throughout the paper, we will return to the Hack distributions for the Kansas and Mississippi Rivers, as well as data obtained for the Amazon, Nile and Congo Rivers.

III. FLUCTUATIONS: AN ANALYTIC FORM FOR THE HACK DISTRIBUTION

To provide some insight into the nature of the underlying Hack distribution, we present a line of reasoning that will build up from Hack's law to a scaling form of $P(a,l)$. First let us assume for the present discussion of fluctuations that an exact form of Hack's law holds,

$$\langle l \rangle = \theta a^h, \tag{4}$$

where we have introduced the coefficient θ which we discuss fully later on. Now, since Hack's law is a power law, it is reasonable to postulate a generalization of the form

$$P(l|a) = \frac{1}{a^h} F_l\left(\frac{l}{a^h}\right). \tag{5}$$

The prefactor $1/a^h$ provides the correct normalization, and F_l is the ''scaling function'' we hope to understand. The above will be our notation for all conditional probabilities. Implicit in Eq. (5) is the assumption that all moments and the distribution itself also scale. For example, the qth moment of $P(l|a)$ is

$$\langle l^q(a) \rangle \propto a^{qh}, \tag{6}$$

which implies

$$\langle l^q(a) \rangle = k^{-qh} \langle l^q(ak) \rangle, \tag{7}$$

where $k \in \mathbb{R}$. Also, for the distribution $P(l|a)$, it follows from Eq. (5) that

$$k^h P(lk^h|ak) = \frac{k^h}{a^h k^h} F_l\left(\frac{lk^h}{a^h k^h}\right) = P(l|a). \tag{8}$$

We note that previous investigations of Hack's law [12,13] consider the generalization in Eq. (5). The authors of Ref. [48] also examined the behavior of the moments of the distribution $P(l|a)$ for real networks. Here we will go further to characterize the full distribution $P(a,l)$ as well as both $P(l|a)$ and $P(a|l)$. Along these lines, Rigon *et al.* [13] suggested that the function $F_l(x)$ is a ''finite-size'' scaling function analogous to those found in statistical mechanics, i.e., $F_l(x) \to 0$ as $x \to \infty$ and $F_l(x) \to c$ as $x \to 0$. As we will detail below, the restrictions on $F_l(x)$ can be made stronger, and we will postulate a simple Gaussian form. More generally, $F_l(x)$ should be a unimodal distribution that is nonzero for an interval $[x_1, x_2]$, where $x_1 > 0$. This is so because for any given fixed basin area a, there is a minimum and maximum l beyond which no basin exists. This is also clear upon inspection of Figs. 1(a) and 1(b).

We observe that neither drainage area nor main stream length possess any obvious features so as to be deemed the independent variable. Hence we can also view Hack's law as its inversion $\langle a \rangle \propto l^{1/h}$. Note that the constant of proportionality is not necessarily $\theta^{1/h}$, and is dependent on the nature of the full Hack distribution. We thus have another scaling ansatz as per Eq. (5):

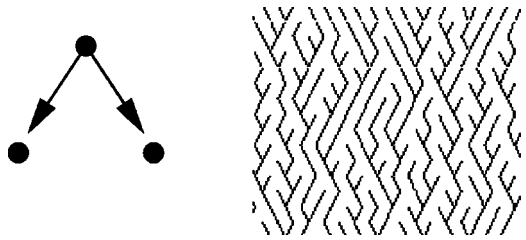

FIG. 2. Scheidegger's model of random, directed networks. Flow is down the page, and, at each site, stream flow is randomly chosen to be in one of the two downward diagonals. Stream paths and basin boundaries are thus discrete random walks.

$$P(a|l) = 1/l^{1/h} F_a(a/l^{1/h}). \tag{9}$$

The conditional probabilities $P(l|a)$ and $P(a|l)$ are related to the joint probability distribution as

$$P(a,l) = P(a)P(l|a) = P(l)P(a|l), \tag{10}$$

where $P(l)$ and $P(a)$ are the probability densities of the main stream length and the area. These distributions are in turn observed to be power laws both in real world networks and models [7,51,52],

$$P(a) \sim N_a a^{-\tau} \quad \text{and} \quad P(l) \sim N_l l^{-\gamma}. \tag{11}$$

where N_a and N_l are appropriate prefactors, and the tilde indicates asymptotic agreement between both sides for large values of the argument. Furthermore, the exponents τ and γ are related to Hack's exponent h via the scaling relations [12,39]

$$\tau = 2 - h \quad \text{and} \quad \gamma = 1/h. \tag{12}$$

Equations (5), (9), (10), (11), and (12) combine to give us two forms for $P(a,l)$,

$$P(a,l) = 1/a^2 F(l/a^h) = 1/l^{2/h} G(a/l^{1/h}), \tag{13}$$

where $x^{-2}F(x^{-h}) = G(x)$ and, equivalently, $F(y) = y^{-2}G(y^{-1/h})$.

IV. RANDOM DIRECTED NETWORKS

We will use results from the Scheidegger model [33] to motivate the forms of these distributions. In doing so, we will also connect with some problems in the theory of random walks.

Scheidegger's model of river networks is defined on a triangular lattice, as indicated by Fig. 2. Flow in the figure is directed down the page. At each site, the stream flow direction is randomly chosen between the two diagonal directions shown. Periodic boundary conditions are applied in all of our simulations. We take the triangular lattice to be a square one tilted by 45°, so each site locally drains an area of α^2, where the lattice unit α is the distance between neighboring sites, and each segment of stream has a length α. For simplicity, we will take α to be unity. We note that connections exist between the Scheidegger model and models of particle aggregation [51,53], Abelian sandpiles [54–56], and limiting cases of force chain models in granular media [4].

Since Scheidegger's model is based on random flow directions, its Hack distributions has a simple interpretation. The boundaries of drainage basins in the model are random walks. Understanding Hack's law therefore amounts to understanding the first collision time of two random walks that share the same origin in one dimension. If we subtract the graph of one walk from the other, we see that the latter problem is itself equivalent to the first return problem of random walks [57].

Many facets of the first return problem are well understood. In particular, the probability of n, the number of steps taken by a random walk until it first returns to the origin, is asymptotically given by

$$P(n) \sim \frac{1}{2\sqrt{\pi}} n^{-3/2}. \tag{14}$$

But this number of steps is also the length of the basin l. Therefore, we have

$$P(l) \sim \frac{2}{\sqrt{\pi}} l^{-3/2}, \tag{15}$$

where, because we are considering the difference of two walks, we use $l = 2n$. Also, we have found the prefactor $N_l = 2/\sqrt{\pi}$.

We thus have that $\gamma = 3/2$ for the Scheidegger model. The scaling relations of Eq. (12) then give $h = 2/3$ and $\tau = 4/3$. The value of h is also readily obtained by noting that the typical area of a basin of length l is $a \propto l \times l^{1/2} = l^{3/2} = l^{1/h}$, since the boundaries are random walks.

V. AREA-LENGTH DISTRIBUTION FOR RANDOM, DIRECTED NETWORKS

Something that is less well studied is the joint distribution of the area enclosed by a random walk and the number of steps to its first return. In terms of the Scheidegger model, this is precisely the Hack distribution.

We motivate some general results based on observations of the Scheidegger model. Figure 3 shows the normalized distributions $P(al^{-3/2})$ and $P(la^{-2/3})$ as derived from simulations of the model. Given the scaling ansatzes for $P(l|a)$ and $P(a|l)$ in Eqs. (5) and (9), we see that $P(y = la^{-2/3}) = F_l(y)$ and $P(x = al^{-3/2}) = F_a(x)$.

Note that we have already used Hack's law for the Scheidegger model with $h = \frac{2}{3}$ to obtain these distributions. The results are for ten realizations of the model on a $10^4 \times 10^4$ lattice, taking 10^7 samples from each of the ten instances. For $P(a|l)$, only sites where $l \geqslant 100$ were taken, and similarly, for $P(l|a)$, only sites where $a \geqslant 500$ were included in the histogram.

We postulate that the distribution $P(y = la^{-2/3})$ is a Gaussian having the form

$$P(l|a) = \frac{1}{\sqrt{2\pi} a^{2/3} \eta} \exp\{-(la^{-2/3} - \theta)^2 / 2\eta^2\}. \tag{16}$$

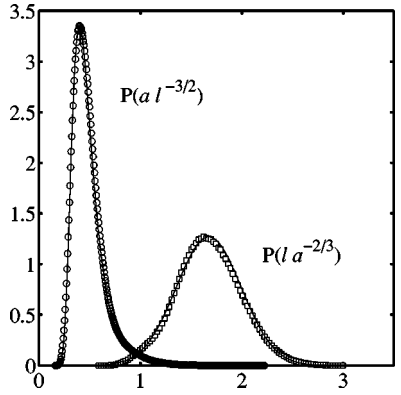

FIG. 3. Cross-sectional scaling functions of the Hack distribution for the Scheidegger model, with the lattice constant equal to unity. Both distributions are normalized, and all quantities are dimensionless. The right distribution is for l fixed and a varying, and is postulated to be a normal distribution. The left distribution for a fixed and l varying and is a form of an inverse Gaussian [58]. The data used were obtained for all sites with with $l \geqslant 100$ and $a \geqslant 500$ respectively. Each distribution was obtained from ten realizations of the Scheidegger model on a $10^4 \times 10^4$ lattice.

We estimate the mean of F_l to be $\theta \simeq 1.675$ [this is the same θ as found in Eq. (4)], and the standard deviation to be $\eta \simeq 0.321$. The fit is shown in Fig. 3 as a solid line. The above equation agrees with the form of the scaling ansatz of Eq. (5), and we now have the assertion that F_l is a Gaussian defined by the two parameters θ and η.

Note that the θ and η are coefficients for the actual mean and standard deviation. In other words, for fixed a, the mean of $P(l|a)$ is $\theta a^{2/3}$ and its standard deviation is $\eta a^{2/3}$. Having observed their context, we will refer to θ and η as the Hack mean coefficient and the Hack standard deviation coefficient.

From this starting point we can create $P(a,l)$ and $P(a|l)$, the latter providing a useful test. Since $P(a) \sim N_a a^{-\tau} = N_a a^{-4/3}$, as per Eq. (11), we have

$$P(a,l) = \frac{N_a}{a^{4/3}} \frac{1}{\sqrt{2\pi} a^{2/3} \eta} \exp\{-(l a^{-2/3} - \theta)^2 / 2\eta^2\},$$

$$= \frac{N_a}{\sqrt{2\pi} a^2 \eta} \exp\{-(l a^{-2/3} - \theta)^2 / 2\eta^2\}. \tag{17}$$

As expected, we observe the form of Eq. (17) to be in accordance with that of Eq. (13). Note that the scaling function F (and equivalently G) is defined by the three parameters θ, η, and N_a, the latter of which may be determined in terms of the former, as we will show below. Also, since we expect all scaling functions to be only asymptotically correct, we cannot use Eq. (17) to find an expression for the normalization N_a. Equation (17) ceases to be valid for small a and l. However, we will be able to use it once we have $P(a|l)$, since we are able to presume l is large, and therefore that the scaling form is exact. Using Eq. (15) and the fact that $P(a|l) = P(a,l)/P(l)$ from Eq. (10), we then have

$$P(a|l) = \frac{\sqrt{\pi} l^{3/2}}{2} \frac{N_a}{\sqrt{2\pi} a^2 \eta} \exp\{-(l/a^{2/3} - \theta)^2 / 2\eta^2\},$$

$$= \frac{N_a l^{3/2}}{2^{3/2} a^2 \eta} \exp\{-(l/a^{2/3} - \theta)^2 / 2\eta^2\},$$

$$= \frac{1}{l^{3/2}} \frac{N_a}{2^{3/2} \eta} (a/l^{3/2})^{-2} \times \exp\{-[(a/l^{3/2})^{-2/3} - \theta]^2 / 2\eta^2\}. \tag{18}$$

In rearranging the expression of $P(a|l)$, we have made clear that its form matches that of Eq. (5).

A closed form expression for the normalization factor N_a may now be determined by employing the fact that $\int_{a=0}^{\infty} da\, P(a|l) = 1$,

$$1 = \int_{a=0}^{\infty} da\, P(a|l),$$

$$= \int_{a=0}^{\infty} \frac{da}{l^{3/2}} \frac{N_a}{2^{3/2} \eta} (a/l^{3/2})^{-2} \times \exp\{-[(a/l^{3/2})^{-2/3} - \theta]^2 / 2\eta^2\},$$

$$= \frac{N_a \sqrt{3}}{2^{5/2} \eta} \int_{u=0}^{\infty} du\, u^{1/2} \exp\{-(u-\theta)^2 / 2\eta^2\}, \tag{19}$$

where we have used the substitution $a/l^{3/2} = u^{-3/2}$, and hence also $l^{-3/2} da = (-3/2) u^{-5/2} du$. We therefore have

$$N_a = \frac{2^{5/2} \eta}{\sqrt{3}} \left[\int_{u=0}^{\infty} du\, u^{1/2} \exp\{-(u-\theta)^2 / 2\eta^2\} \right]^{-1}. \tag{20}$$

We may thus write down all of the scaling functions F_l, F_a, F, and G for the Scheidegger model:

$$F_l(z) = \frac{1}{\sqrt{2\pi} \eta} \exp\{-(z-\theta)^2 / 2\eta^2\}, \tag{21}$$

$$F_a(z) = \frac{N_a}{2^{3/2} \eta} \exp\{-(z^{-2/3} - \theta)^2 / 2\eta^2\}, \tag{22}$$

$$F(z) = \frac{N_a}{\sqrt{2\pi} \eta} \exp\{-(z-\theta)^2 / 2\eta^2\}, \text{ and} \tag{23}$$

$$G(z) = \frac{N_a}{\sqrt{2\pi} \eta} z^{-2} \exp\{-(z^{-2/3} - \theta)^2 / 2\eta^2\}. \tag{24}$$

Recall that all of these forms rest on the assumption that $F_l(z)$ is a Gaussian. In order to check this assumption, we return to Fig. 3. The empirical distribution $P(z = a/l^{3/2})$ is shown on the left, marked with circles. The solid line through these points is $F_a(z)$, as given above in Eq. (22). There is an excellent match so we may be confident about our proposed form for $F_a(z)$. We note that the function $F_a(z)$ may be thought of as a fractional inverse Gaussian

distribution, the inverse Gaussian being a well known distribution arising in the study of first passage times for random walks [57]. It is worth contemplating the peculiar form of $P(a|l)$ in terms of first return random walks. Here we have been able to postulate the functional form of the distribution of areas bound by random walks that first return after n steps. If one could understand the origin of the Gaussian and find analytic expressions for θ and η, then the problem would be fully solved.

VI. AREA-LENGTH DISTRIBUTION EXTENDED TO REAL NETWORKS

We now seek to extend these results for Scheidegger's model to real world networks. We will look for the same functional forms for the Hack distributions that we have found above. The conditional probability distributions pertaining to Hack's law take the forms

$$P(l|a) = 1/a^{-h} F_l(la^{-h}) = \frac{a^{-h}}{\sqrt{2\pi}\eta} \exp\{-(la^{-h}-\theta)^2/2\eta^2\} \tag{25}$$

and

$$P(a|l) = l^{-1/h} F_a(al^{-1/h}) = l^{-1/h} \frac{N_a}{\sqrt{2\pi} N_l \eta} (al^{-1/h})^{-2} \times \exp\{-[(al^{-1/h})^{-h}-\theta]^2/2\eta^2\}, \tag{26}$$

and the full Hack distribution is given by

$$P(a,l) = a^2 F(la^{-h}) = l^{-2/h} G(al^{-1/h}) = \frac{N_a}{\sqrt{2\pi}\eta a^2} \times \exp\{-(la^{-h}-\theta)^2/2\eta^2\}, \tag{27}$$

with N_a determined by Eq. (20). The three parameters—the Hack exponent h, the Hack mean coefficient θ, and the Hack standard deviation coefficient η—are in principle landscape dependent. Furthermore, in η we have a basic measure of fluctuations in the morphology of basins.

Figures 4 and 5 present Hack scaling functions for the Mississippi and Nile River basins. These figures are to be compared with the results for the Scheidegger model in Fig. 3.

For both rivers, Hack's exponent h was determined first from a stream ordering analysis (we discuss stream ordering later in Sec. X). Estimates of the parameters θ and η were then made using the scaling function F_l presuming a Gaussian form.

We observe that the Gaussian fit for the Mississippi is more satisfactory than that for the Nile. These fits are not rigorously made because even though we have chosen data ranges where deviations (which we address in Sec. VII) are minimal, deviations from scaling do still skew the distributions. The specific ranges used to obtain F_l and F_a, respectively, are $8.5 < \log_{10} a < 9.5$ and $4.75 < \log_{10} l < 6$, for the Mississippi and $9 < \log_{10} a < 11$ and $5.5 < \log_{10} l < 6$ for the Nile (areas are in km^2, and lengths in km). Furthermore, we observe that the estimate of h has an effect on the resulting forms of F_l and F_a. Nevertheless, here we are attempting to capture the essence of the generalized form of Hack's law in real networks.

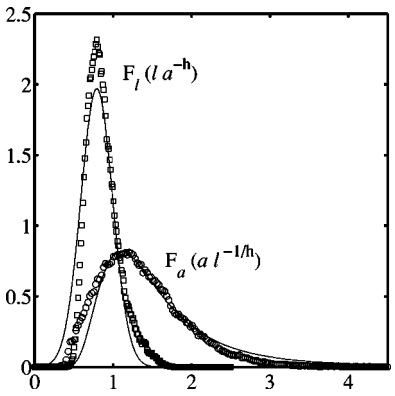

FIG. 4. Cross-sectional scaling functions of the Hack distribution for the Mississippi. Area a and length l are in m^2 and m, respectively. The estimate used for Hack's exponent is $h = 0.55$, the determination of which is discussed in Sec. VII. The fits indicated by the smooth curves to the data are made as per the Scheidegger model in Fig. 3, and according to Eqs. (25) and (26). The values of the Hack mean coefficient and standard deviation coefficient are estimated to be $\theta \simeq 0.80$ and $\eta \simeq 0.20$.

We then use the parameters h, θ, and η, and Eq. (26) to construct our theoretical F_a, the smooth curves in Figs. 4 and 5. As for the Scheidegger model data in Fig. 3, we see in both examples approximate agreement between the measured F_a and the one predicted from the form of F_l. Table I shows estimates of h, θ, and η for the five major river basins studied.

Given our reservations about the precision of these values of θ and η, we are nevertheless able to make qualitative distinctions. Recalling that $\langle l \rangle = \theta a^h$, we see that, for fixed h, higher values of θ indicate relatively longer stream lengths

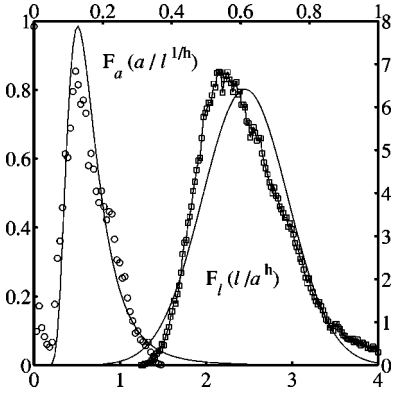

FIG. 5. Cross-sectional scaling functions of the Hack distribution for the Nile. Dimensions are as per Fig. 4. The top and right axes correspond to F_a and the bottom and left to F_l. The Hack exponent used is $h = 0.50$, and the values of the Hack mean coefficient and standard deviation coefficient are estimated to be $\theta \simeq 2.45$ and $\eta \simeq 0.50$ [59].

TABLE I. Estimates of Hack distribution parameters for real river networks. The scaling exponent h is Hack's exponent. The parameters θ and η are coefficients of the mean and standard deviations of the conditional probability density function $P(l|a)$, and are fully discussed in the text [60].

River network	θ	η	h
Mississippi	0.80	0.20	0.55
Amazon	1.90	0.35	0.52
Nile	2.45	0.50	0.50
Kansas	0.70	0.15	0.57
Congo	0.89	0.18	0.54

for a given area, and hence longer and thinner basins. The results therefore suggest the Nile, and to a lesser degree the Amazon, have basins with thinner profiles than the Congo and, in particular, the Mississippi and Kansas. This seems not unreasonable since the Nile is a strongly directed network constrained within a relatively narrow overall shape. This is somewhat in spite of the fact that the shape of an overall river basin is not necessarily related to its internal basin morphology, an observation we will address later in Sec. X.

The importance of θ is tempered by the value of h. Hack's exponent affects not only the absolute measure of stream length for a given area but also how basin shapes change with increasing area. So, in the case of the Kansas, the higher value of h suggests that basin profiles thin with increasing size. This is in keeping with overall directedness of the network. Note that our measurements of the fractal dimension d of stream lengths for the Kansas place it to be $d=1.04 \pm 0.02$. Therefore, $d/h \simeq 1.9 < 2$, and elongation is still expected when we factor in the scaling of l with $L_{\|}$.

The Nile and Amazon also all have relatively high η, indicating greater fluctuations in basin shape. In comparison, the Mississippi, Kansas, and Congo appear to have less variation. Note that the variability of the Kansas is in reasonable agreement with that of the whole Mississippi River network for which it is a sub-basin.

Finally, regardless of the actual form of the distribution underlying Hack's law, fluctuations are always present, and an estimate of their extent is an important measurement. Thus, the Hack mean and standard deviation coefficients θ and η are suggested to be of sufficient worth so as to be included with any measurement of the Hack exponent h.

VII. DEVIATIONS FROM SCALING

In generalizing Hack's law, we have sought out regions of robust scaling, discarding ranges where deviations become prominent. We now turn our attention to the nature of the deviations themselves.

We observe three major classes of deviations which we will define by the scales at which they occur: small, intermediate, and large. Throughout the following sections we primarily consider deviations from the mean version of Hack's law, $\langle l \rangle = \theta a^h$, given in Eq. (4). Much of the understanding we gain from this will be extendable to deviations for higher moments.

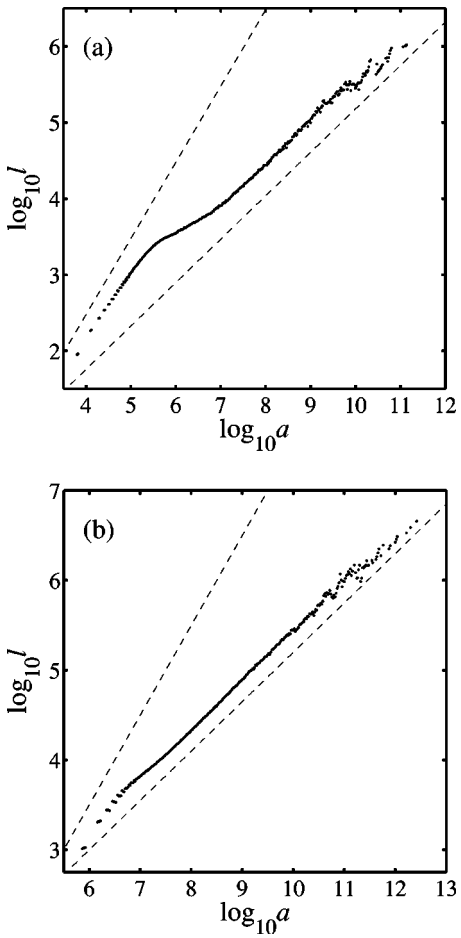

FIG. 6. The mean version of Hack's law for the Kansas (a) and the Mississippi (b). The units of lengths and areas are meters and square meters. These are calculated from the full Hack distributions shown in Figs. 1(a) and 1(b) by finding $\langle l \rangle$ for each value of basin area a. Area samples are taken every 0.02 orders of magnitude in logarithmic space. The upper dashed lines represent a slope of unity in both plots, and the lower lines the Hack exponents 0.57 and 0.55 for the Kansas and Mississippi, respectively. For the Kansas, there is a clear deviation for small area which rolls over into a region of very slowly changing derivative before breaking up at large scales. Deviations from scaling are present for the lower resolution dataset of the Mississippi, but to a lesser extent.

To provide an overview of what follows, examples of mean Hack distributions for the Kansas and Mississippi Rivers are shown in Figs. 6(a) and 6(b). Hack's law for the Kansas River exhibits a marked deviation for small areas, starting with a near linear relationship between stream length and area. A long crossover region of several orders of magnitude in area then leads to an intermediate scaling regime, wherein we attempt to determine the Hack exponent h. In doing so, we show that such regions of robust scaling are surprisingly limited for river network quantities. Moreover, we observe that, where present, scaling is only approximate, and that no exact exponents can be ascribed to the networks we study here. It follows that the identification of universality classes based on empirical evidence is a hazardous step.

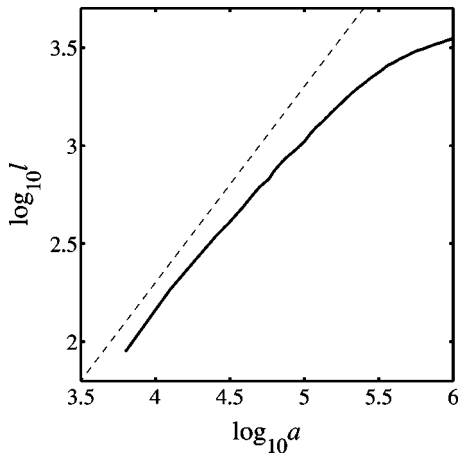

FIG. 7. The linearity of Hack's law at small scales for the Kansas River. Dimensions are as per Fig. 6. The linear regime enters a crossover region after almost 1.5 orders of magnitude in area. This is an expanded detail of the mean Hack's law given in Fig. 6(a), and areas and lengths are in square meters and meters.

Finally, the approximate scaling of this intermediate region then gives way to a breakdown in scaling at larger scales due to low sampling and correlations with basin shape. The same deviations are present in the relatively coarse-grained Mississippi data, but are less pronounced.

VIII. DEVIATIONS AT SMALL SCALES

At small scales, we find that the mean Hack distribution follows a linear relationship, i.e., $l \propto a^1$. This feature is most evident for the Kansas River, as shown in Fig. 6(a), and in more detail in Fig. 7. The linear regime persists for nearly 1.5 orders of magnitude in basin area. To a lesser extent, the same trend is apparent in the Mississippi data, [Fig. 6(b)].

Returning to the full Hack distribution of Figs. 1(a) and 1(b), we begin to see the origin of this linear regime. In both instances, a linear branch separates from the body of the main distribution. Since l cannot grow faster than a^1, the linear branch marks an upper bound on the extent of the distribution in (a,l) coordinates. When averaged to give the mean Hack distribution, this linear data dominates the result for small scales.

We find this branch evident in all Hack distributions. It is not an artifact of resolution, and in fact becomes more pronounced with increased map precision. The origin of this linear branch is simple: data points along the branch correspond to positions in narrow subnetworks, i.e., long, thin ''valleys.'' To understand the separation of this linear branch from the main body of the distribution, consider Fig. 8, which depicts a stream draining such a valley with length and area (l_1,a_1) that meets a stream from a basin with characteristics (l_2,a_2). The area and length of the basin formed at this junction is thus $[(a_3,l_3)=(a_1+a_2,\max(l_1,l_2)]$ [in the figure, $\max(l_1,l_2)=l_2$]. The greater jump in area moves the point across into the main body of the Hack distribution, creating a separation of the linear branch.

In fact, the full Hack distribution is itself comprised of many such linear segments. As in the above example, until a

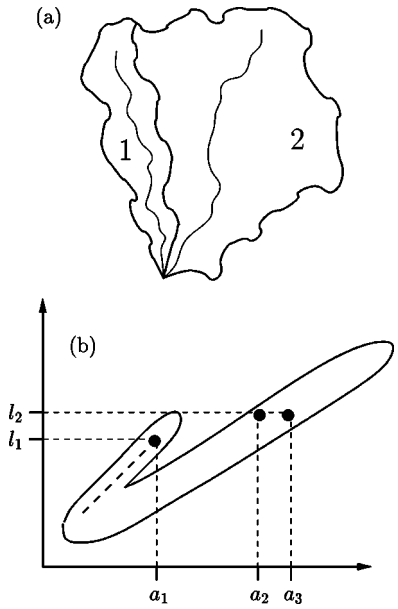

FIG. 8. Origin of the linear branch in the Hack distribution. The subbasins labeled 1 and 2 depicted in (a) have areas a_1 and a_2 and lengths l_1 and l_2. These subbasins combine to form a basin of area $a_3=a_1+a_2$ and main stream length $l_3=\max l_1,\ l_2=l_2$. Since subbasin 1 is a linear subnetwork (a valley) the pair (a_1,l_1) lie along the linear branch of the Hack distribution as shown in (b). Points along the main stream of subbasin 1 lie along the dashed line, leading to the coordinate (a_1,l_1). On combining with the second basin, the jump in the resultant area creates a jump from (a_1,l_1) to (a_3,l_2) in the main body of the Hack distribution.

stream does not meet any streams of comparable size, then its area and length will roughly increase in a linear fashion. When it does meet such a stream, there is a jump in area, and the trace of a new linear segment is started in the distribution. We will see this most clearly later on, when we study deviations at large scales.

For very fine scale maps, on the order of meters, we might expect to pick up the scale of the unchannelized, convex regions of a landscape, i.e., ''hillslopes'' [61]. This length scale represents the typical separation of branches at a network's finest scale. The computation of stream networks for these hillslope regions would result in largely nonconvergent (divergent or parallel) flow. Therefore, we would have linear ''basins'' that would in theory contribute to the linear branch we observe [10]. Potentially, the crossover in Hack's law could be used as a determinant of hillslope scale, a crucial parameter in geomorphology [61,62]. However, when long, thin network structures are present in a network, this hillslope scale is masked by their contribution.

Whether because of the hillslope scale or linear network structure, we see that at small scales, Hack's law will show a crossover in scaling from $h=1$ to a lower exponent. The crossover's position depends on the extent of linear basins in the network. For example, in the Kansas River basin, the crossover occurs when $(l,a)=(4\times10^3\ \mathrm{m},\ 10^7\ \mathrm{m}^2)$. Since increased map resolution can only increase measures of length, the crossover's position must occur at least at such a

length scale which may be many orders of magnitude greater than the scale of the map.

However, the measurement of the area of such linear basins will potentially grow with coarse-graining. Note that for the Mississippi mean Hack distribution, the crossover begins around $a=10^{6.5}\,\mathrm{m}^2$ whereas for the Kansas the crossover initiates near $a=10^{5.5}\,\mathrm{m}^2$, but the ends of the crossovers in both cases appear to agree, occurring at around $a=10^{7}\,\mathrm{m}^2$. Continued coarse graining will of course eventually destroy all statistics, and introduce spurious deviations. Nevertheless, we see here that the deviation which is only suggested in the Mississippi data is well confirmed in the finer-grained Kansas data.

IX. DEVIATIONS AT INTERMEDIATE SCALES

As the basin area increases, we move out of the linear regime, observing a crossover to what would be considered the normal scaling region of Hack's law. We detail our attempts to measure the Hack exponents for the Kansas and Mississippi examples. Rather than relying solely on a single regression on a mean Hack distribution, we employ a more precise technique that examines the distribution's derivative.

To determine h, we consider Hack's law [Eq. (4)] explicitly in logarithmic coordinates,

$$\log_{10}\langle l(a)\rangle=\log_{10}\theta+h\log_{10}a. \tag{28}$$

The derivative of this equation with respect to $\log_{10}a$ then gives Hack's exponent as a function of area:

$$h(a)=\frac{d}{d\log_{10}a}\log_{10}\langle l(a)\rangle. \tag{29}$$

We may think of $h(a)$ as a ''local Hack exponent.'' We calculate the discrete derivative as above for the Kansas and Mississippi. We smooth the data by taking running averages with varying window sizes of n samples, the results for $n=32$ being shown in Figs. 9(a) and 9(b), where the spacing of $\log a$ is 0.02 orders of magnitude. Thus the running averages for the figures are taken over area ranges corresponding to 0.64 orders of magnitude.

Now, if the scaling law in question is truly a scaling law, the above type of derivative will fluctuate around a constant value of exponent over several orders of magnitude. With increasing n, these fluctuations will necessarily decrease, and we should see the derivative holding steady around the exponent's value.

At first glance, we note a considerable variation in $h(a)$ for both data sets, with the Kansas standing out. Fluctuations are reduced with increasing n, but we observe continuous variation of the local Hack exponent with area. For the example of the Kansas, the linear regime and ensuing crossover appear as a steep rise followed by a drop and then another rise during all of which the local Hack exponent moves well below 1/2.

It is after these small scale fluctuations that we would expect to find Hack's exponent. For the Kansas River data, we see the derivative gradually climbs for all values of n before reaching the end of the intermediate regime where the putative scaling breaks down altogether.

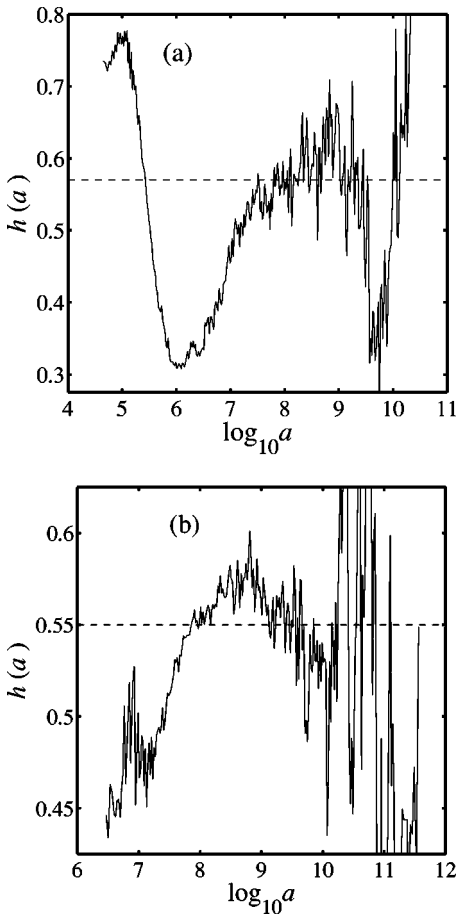

FIG. 9. Variation in Hack's exponent for the Kansas (a) and the Mississippi (b). The area a is in m^2. The plots are derivatives of the mean Hack distributions given in Figs. 6(a) and 6(b). Both derivatives have been smoothed by taking running averages over 0.64 orders of magnitude in a. For the Kansas, the dashed line is set at $h=0.57$, Hack's exponent estimated via simple regression analysis for points with $10^{7}<a<10^{10}\,\mathrm{m}^2$. The local exponent is seen to gradually rise through the $h=0.57$ level, indicating that scaling is not robust. For the Mississippi in (b), the dashed line is a Hack exponent of 0.55 calculated from regression on data in the interval $10^{9}<a<10^{11.5}\,\mathrm{m}^2$. The local Hack exponent is seen to gradually rise and fall about this value.

For the Kansas River data shown in Fig. 9(a), the dashed line represents $h=0.57$, our estimate of Hack's exponent from simple regression on the mean Hack distribution of Fig. 6(a). For the regression calculation, the intermediate region was identified from the figure to be $10^{7}<a<10^{10}\,\mathrm{m}^2$. We see from the smoothed derivative in Fig. 9(a) that the value $h=0.57$ is not precise. After the crossover from the linear region has been completed, we observe a slow rise from $h\simeq 0.54$ to $h\simeq 0.63$. Thus the local Hack exponent $h(a)$ gradually climbs above $h=0.57$ rather than fluctuating around it.

A similar slow change in $h(a)$ is observed for the Mississippi data. In Fig. 9(b) we see a gradual rise and then fall in $h(a)$. The dashed line here represents $h=0.55$, the value

of which was determined from Fig. 6(b) using regression on the range $10^9 < a < 10^{11.5}\,\mathrm{m}^2$. The range of $h(a)$ is roughly [0.52, 0.58]. Again, while $h = 0.55$ approximates the derivative throughout this intermediate range of Hack's law, we cannot claim it to be a precise value.

We observe the same drifts in $h(a)$ in other datasets and for varying window size n of the running average. The results suggest that we cannot assign specific Hack exponents to these river networks, and are therefore unable to even consider what might be an appropriate universality class. The value of h obtained by regression analysis is clearly sensitive to the the range of a used. Furthermore, these results indicate that we should maintain healthy reservations about the exact values of other reported exponents.

X. DEVIATIONS AT LARGE SCALES

We turn now to deviations from Hack's law at large scales. As we move beyond the intermediate region of approximate scaling, fluctuations in $h(a)$ begin to grow rapidly. This is clear on inspection of the derivatives of Hack's law in Figs. 9(a) and 9(b). There are two main factors conspiring to drive these fluctuations up. The first is that the number of samples of subbasins with area a decays algebraically in a. This is just the observation that $P(a) \propto a^{-\tau}$ as per Eq. (11). The second factor is that fluctuations in l and a are on the order of the parameters themselves. This follows from our generalization of Hack's law, which shows, for example, that the moments $\langle l^q \rangle$ of $P(l|a)$ grow like a^{qh}. Thus, the standard deviation grows like the mean: $\sigma(l) = (\langle l^2 \rangle - \langle l \rangle^2)^{1/2} \propto a^h \propto \langle l \rangle$.

A. Stream ordering and Horton's laws

So as to understand these large scale deviations from Hack's law, we need to examine network structure in depth. One way to do this is by using Horton-Strahler stream ordering [38,63] and a generalization of the well-known Horton's laws [36,38,39,64,65]. This will naturally allow us to deal with the discrete nature of a network that is most apparent at large scales.

Stream ordering discretizes a network into a set of stream segments (or, equivalently, a set of nested basins) by an iterative pruning. Source streams (i.e., those without tributaries) are designated as stream segments of order $\omega = 1$. These are removed from the network, and the new source streams are then labeled as order $\omega = 2$ stream segments. The process is repeated until a single stream segment of order $\omega = \Omega$ is left, and the basin itself is defined to be of order Ω.

Natural metrics for an ordered river network are n_ω, the number of order ω stream segments (or basins); $\overline{a}_\omega$, the average area of order ω basins; $\overline{l}_\omega$, the average main stream length of order ω basins; and $\overline{l}_\omega^{(s)}$, the average length of order ω stream segments. Horton's laws state that these quantities change regularly from order to order, i.e.,

$$\frac{n_\omega}{n_{\omega+1}} = R_n \quad \text{and} \quad \frac{\overline{X}_{\omega+1}}{\overline{X}_\omega} = R_X, \tag{30}$$

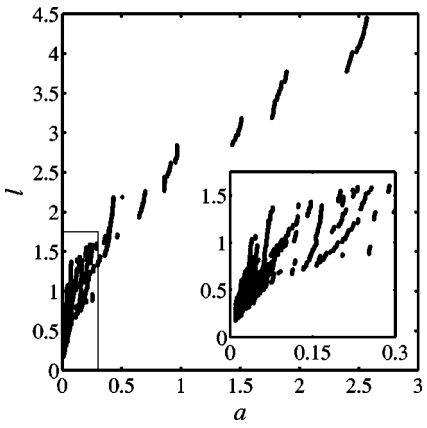

FIG. 10. Hack distribution for the Mississippi plotted in linear space. Area a is given in units of $10^{12}\,\mathrm{m}^2$, and length l in 10^6 m. The discreteness of the basin structure is clearly indicated by the isolated, linear fragments. The inset is a blowup of the box on the main graph.

where $X = a$, l, or $l^{(s)}$. Note that all ratios are defined to be greater than unity since areas and lengths increase but number decreases. Also, there are only two independent ratios since $R_a \equiv R_n$ and $R_l \equiv R_{l^{(s)}}$ [39]. Horton's laws mean that stream-order quantities change exponentially with order. For example, Eq. (30) gives that $l_\omega \propto (R_l)^\omega$.

B. Discrete version of Hack's law

Returning to Hack's law, we examine its large scale fluctuations with the help of stream ordering. We are interested in the size of these fluctuations, and also how they might correlate with the overall shape of a basin. First, we note that the structure of the network at large scales is explicitly discrete. Figure 10 demonstrates this by plotting the distribution of (a,l) without the usual logarithmic transformation. Hack's law is seen to be composed of linear fragments. As explained above in Fig. 8, areas and length increase in proportion to each other along streams where no major tributaries enter. As soon as a stream does combine with a comparable one, a jump in drainage area occurs. Thus, in Fig. 10 we see isolated linear segments which, upon ending at a point (a_1, l_1), begin again at $(a_1 + a_2, l_1)$, i.e., the main stream length stays the same but the area is shifted.

We consider a stream ordering version of Hack's law given by the points $(\overline{a}_\omega, \overline{l}_\omega)$. The scaling of these data points is equivalent to scaling in the usual Hack's law. Also, given Horton's laws, it follows that $h = \ln R_l / \ln R_n$ (using $R_a \equiv R_n$). Along the lines of the derivative we introduced to study intermediate scale fluctuations in Eq. (29), we have here an order-based difference:

$$h_{\omega,\omega-1} = \frac{\log \overline{l}_\omega / \overline{l}_{\omega-1}}{\log \overline{a}_\omega / \overline{a}_{\omega-1}}. \tag{31}$$

We can further extend this definition to differences between nonadjacent orders:

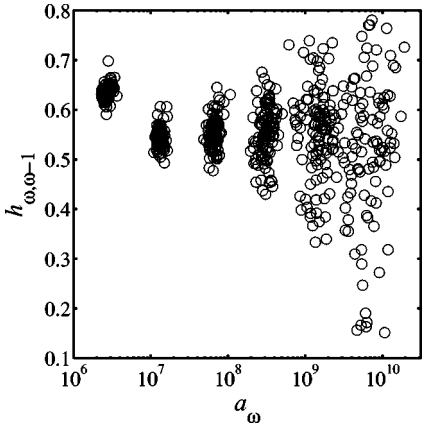

FIG. 11. Differences of the stream order-based version of Hack's law for 104 order $\Omega=7$ basins of the Mississippi (compare the continuous versions given in Fig. 9). The units of area are square meters, and the exponent $h_{\omega,\omega-1}$ is dimensionless. The plots are overlaid to give a sense of the increase in fluctuations of the local Hack exponent $h_{\omega,\omega-1}$ with increasing order ω. The clusters correspond to $\omega=1,2,\ldots,7$, moving from left to right.

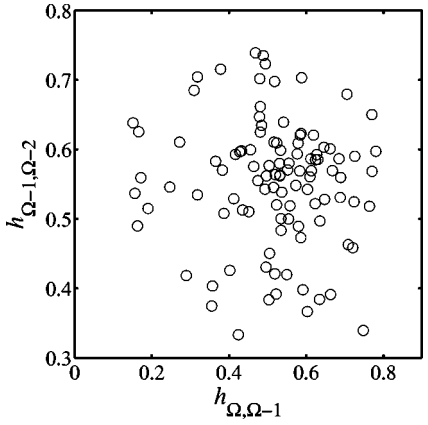

FIG. 12. A comparison of the stream order Hack derivatives $h_{\Omega,\Omega-1}$ and $h_{\Omega-1,\Omega-2}$ for each of the 104 order $\Omega=7$ basins of the Mississippi. Both axes are dimensionless. The linear correlation coefficient is $r=-0.06$, and the Spearman correlation coefficient is $r_s=-0.08$. The latter has probability $p_s=0.43$, indicating there are no significant correlations.

$$h_{\omega,\omega'}=\frac{\log \overline{l}_\omega/\overline{l}_{\omega'}}{\log \overline{a}_\omega/\overline{a}_{\omega'}}. \tag{32}$$

This type of difference, where $\omega'<\omega$, may be best thought of as a measure of trends rather than an approximate discrete derivative.

Using these discrete differences, we examine two features of the order-based versions of Hack's law. First we consider correlations between large scale deviations within an individual basin and, second, correlations between overall deviations and basin shape. For the latter, we will also consider deviations as they move back into the intermediate scale. This will help to explain the gradual deviations from scaling we have observed at intermediate scales.

Since deviations at large scales are reflective of only a few basins, we require an ensemble of basins to provide sufficient statistics. As an example of such an ensemble, we take the set of order $\Omega=7$ basins of the Mississippi basin. For the dataset used here, where the overall basin itself is of order $\Omega=11$, we have 104 order $\Omega=7$ subbasins. The Horton averages for these basins are $\overline{a}_7\simeq 16600\,\mathrm{km}^2$, $\overline{l}_7\simeq 350\,\mathrm{km}$, and $\overline{L}_7\simeq 210\,\mathrm{km}$.

For each basin, we first calculate the Horton averages $(\overline{a}_\omega,\overline{l}_\omega)$. We then compute $h_{\omega,\omega-1}$, the Hack difference given in Eq. (31). To give a rough picture of what is observed, Figure 11 shows a scatter plot of $h_{\omega,\omega-1}$ for all order $\Omega=7$ basins. Note the increase in fluctuations with increasing ω. This increase is qualitatively consistent with the smooth versions found in the single basin examples of Figs. 9(a) and 9(b). In part, less self-averaging for larger ω results in a greater spread in this discrete derivative. However, as we will show, these fluctuations are also correlated with fluctuations in basin shape.

C. Effect of basin shape on Hack's law

In what follows, we extract two statistical measures of correlations between deviations in Hack's law and overall basin shape. These are r, the standard linear correlation coefficient and r_s, the Spearman rank-order correlation coefficient [66–68]. For N observations of data pairs (u_i,v_i), r is defined to be

$$r=\frac{\sum_{i=1}^{N}(u_i-\mu_u)(v_i-\mu_v)}{\sum_{i=1}^{N}(x_i-\mu_u)^2\Sigma_{i=1}^{N}(y_i-\mu_v)^2}=\frac{C(u,v)}{\sigma_u\sigma_v}, \tag{33}$$

where $C(u,v)$ is the covariance of the u_i's and v_i's, μ_u and μ_v their means, and σ_u and σ_v their standard deviations. The value of Spearman's r_s is determined in the same way, but for the u_i and v_i replaced by their ranks. From r_s, we determine a two-sided significance p_s via Student's t distribution [66].

We define κ, a measure of basin aspect ratio, as

$$\kappa=L^2/a. \tag{34}$$

Long and narrow basins correspond to $\kappa\gg 1$, while for short and wide basins we have $\kappa\ll 1$

We now examine the discrete derivatives of Hack's law in more detail. In order to discern correlations between large scale fluctuations within individual basins, we specifically look at the last two differences in a basin: $h_{\Omega,\Omega-1}$ and $h_{\Omega-1,\Omega-2}$. For each of the Mississippi's 104 order $\Omega=7$ basins, these values are plotted against each other in Fig. 12. Both our correlation measurements strongly suggest these differences are uncorrelated. The linear correlation coefficient is $r=-0.06\simeq 0$ and, similarly, we have $r_s=-0.08\simeq 0$. The significance $p_s=0.43$ implies that the null hypothesis of uncorrelated data cannot be rejected.

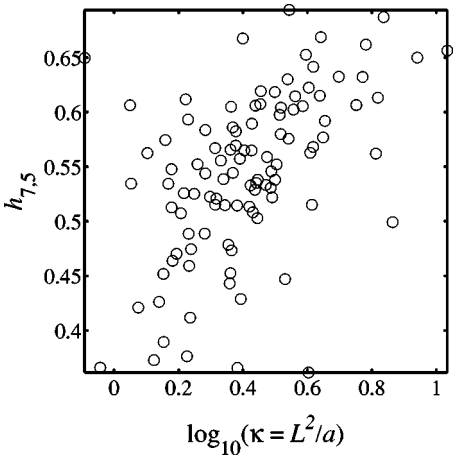

FIG. 13. Correlation between trends in Hack's law and the aspect ratio of a basin as estimated by the dimensionless quantity $\kappa = L^2/a$. The data is for the order $\Omega = 7$ basins of the Mississippi, and the specific trend is $h_{7,5}$. The correlation measurements give $r = 0.50$, $r_s = 0.53$, and $p_s < 10^{-8}$.

Thus, for Hack's law in an individual basin, large scale fluctuations are seen to be uncorrelated. However, correlations between these fluctuations and other factors may still exist. This leads us to our second test which concerns the relationship between trends in Hack's law and overall basin shape.

Figure 13 shows a comparison of the aspect ratio κ and $h_{7,5}$ for the order $\Omega = 7$ basins of the Mississippi. The measured correlation coefficients are $r = 0.50$ and $r_s = 0.53$, giving a significance of $p_s < 10^{-8}$. Furthermore, we find the differences $h_{7,6}$ ($r = 0.34$, $r_s = 0.39$ and $p_s < 10^{-4}$) and $h_{6,5}$ ($r = 0.35$, $r_s = 0.34$ and $p_s < 10^{-3}$) are individually correlated with basin shape. We observe this correlation between basin shape and trends in Hack's law at large scales, namely, $h_{\Omega,\Omega-1}$, $h_{\Omega-1,\Omega-2}$, and $h_{\Omega,\Omega-2}$, repeatedly in our other data sets. In some cases, correlations extend further to $h_{\Omega-2,\Omega-3}$.

Since the area ratio R_a is typically in the range 4–5, Hack's law is affected by boundary conditions set by the geometry of the overall basin down to subbasins 1–2 orders of magnitude smaller in area than the overall basin. These deviations are present regardless of the absolute size of the overall basin. Furthermore, the origin of the basin boundaries being geologic or chance or both is irrelevant—large scale deviations will still occur. However, it is reasonable to suggest that particularly strong deviations are more likely the result of geologic structure rather than simple fluctuations.

XI. CONCLUSION

Hack's law is a central relation in the study of river networks and branching networks in general. We have shown Hack's law to have a more complicated structure than is typically given attention. The starting generalization is to consider fluctuations around scaling. Using the directed, random network model, a form for the Hack distribution underlying Hack's law may be postulated and reasonable agreement with real networks is observed. Questions of the validity of the distribution aside, the Hack mean coefficient θ and the Hack standard deviation coefficient η should be standard measurements because they provide further points of comparison between theory and other basins.

With the idealized Hack distribution proposed, we may begin to understand deviations from its form. As with any scaling law pertaining to a physical system, cutoffs in scaling must exist and need to be understood. For small scales, we have identified the presence of linear subbasins as the source of an initial linear relation between area and stream length. At large scales, statistical fluctuations and geologic boundaries give rise to basins whose overall shape produces deviations in Hack's laws. Both deviations extend over a considerable range of areas, as do the crossovers which link them to the region of intermediate scales, particularly the crossover from small scales.

Finally, by focusing in detail on a few large scale examples networks, we have found evidence that river networks do not belong to well defined universality classes. The relationship between basin area and stream length may be approximately, and in some cases very well, described by scaling laws, but not exactly so. The gradual drift in exponents we observe suggests a more complicated picture, one where subtle correlations between basin shape and geologic features are intrinsic to river network structure.

ACKNOWLEDGMENTS

This work was supported in part by NSF Grant No. EAR-9706220, and Department of Energy Grant No. DE FG02-99ER 15004.

[1] R. Albert, H. Jeong, and A.-L. Barabasi, Nature (London) **406**, 378 (2000).
[2] D. J. Watts and S. J. Strogatz, Nature (London) **393**, 440 (1998).
[3] M. Zamir, J. Theor. Biol. **197**, 517 (1999).
[4] S. N. Coppersmith, C.-h. Liu, S. Majumdar, O. Narayan, and T. A. Witten, Phys. Rev. E **53**, 4673 (1996).
[5] C. Cherniak, M. Changizi, and D. W. Kang, Phys. Rev. E. **59**, 6001 (1999).
[6] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1983).
[7] I. Rodríguez-Iturbe and A. Rinaldo, *Fractal River Basins: Chance and Self-Organization* (Cambridge University Press, Cambridge, 1997).
[8] A. Rinaldo, I. Rodríguez-Iturbe, and R. Rigon, Annu. Rev. Earth Planet Sci. **26**, 289 (1998).
[9] P. Ball, *The Self-Made Tapestry* (Oxford University Press, Oxford, 1998).
[10] P. S. Dodds and D. H. Rothman, Annu. Rev. Earth Planet Sci. **28**, 571 (2000).

[11] J. T. Hack, U.S. Geol. Surv. Prof. Pap., **294-B**, 45 (1957).
[12] A. Maritan, A. Rinaldo, R. Rigon, A. Giacometti, and I. Rodríguez-Iturbe, Phys. Rev. E **53**, 1510 (1996).
[13] R. Rigon, I. Rodríguez-Iturbe, A. Maritan, A. Giacometti, D. G. Tarboton, and A. Rinaldo, Water Resour. Res. **32**, 3367 (1996).
[14] H. Jaeger, S. Nagel, and R. Behringer, Rev. Mod. Phys. **68**, 1259 (1996).
[15] L. Kadanoff, Rev. Mod. Phys. **71**, 435 (1999).
[16] T. R. Smith and F. P. Bretherton, Water Resour. Res. **3**, 1506 (1972).
[17] S. Kramer and M. Marder, Phys. Rev. Lett. **68**, 205 (1992).
[18] N. Izumi and G. Parker, J. Fluid Mech. **283**, 341 (1995).
[19] K. Sinclair and R. C. Ball, Phys. Rev. Lett. **76**, 3360 (1996).
[20] J. R. Banavar, F. Colaiori, A. Flammini, A. Giacometti, A. Maritan, and A. Rinaldo, Phys. Rev. Lett. **78**, 4522 (1997).
[21] E. Somfai and L. M. Sander, Phys. Rev. E **56**, R5 (1997).
[22] R. Pastor-Satorras and D. H. Rothman, Phys. Rev. Lett. **80**, 4349 (1998).
[23] R. Pastor-Satorras and D. H. Rothman, J. Stat. Phys. **93**, 477 (1998).
[24] M. Cieplak, A. Giacometti, A. Maritan, A. Rinaldo, I. Rodríguez-Iturbe, and J. R. Banavar, J. Stat. Phys. **91**, 1 (1998).
[25] A. Giacometti (unpublished).
[26] H. Takayasu and H. Inaoka, Phys. Rev. Lett. **68**, 966 (1992).
[27] R. L. Leheny, Phys. Rev. E **52**, 5610 (1995).
[28] G. Caldarelli, A. Giacometti, A. Maritan, I. Rodríguez-Iturbe, and A. Rinaldo, Phys. Rev. E **55**, 4865 (1997).
[29] C. P. Stark, Nature (London) **352**, 405 (1991).
[30] T. Sun, P. Meakin, and T. Jøssang, Phys. Rev. E **49**, 4865 (1994).
[31] T. Sun, P. Meakin, and T. Jøssang, Phys. Rev. E **51**, 5353 (1995).
[32] L. B. Leopold and W. B. Langbein, U.S. Geol. Surv. Prof. Pap. **500-A**, 1 (1962).
[33] A. E. Scheidegger, Bull. Int. Assoc. Sci. Hydrol. **12**, 15 (1967).
[34] S. S. Manna and B. Subramanian, Phys. Rev. Lett. **76**, 3460 (1996).
[35] A. Maritan, F. Colaiori, A. Flammini, M. Cieplak, and J. R. Banavar, Science **272**, 984 (1996).
[36] P. S. Dodds and D. H. Rothman, following paper, Phys. Rev. E **63**, 016116 (2000).
[37] P. S. Dodds and D. H. Rothman, third paper, Phys. Rev. E **63**, 016117 (2000).
[38] R. E. Horton, Bull. Geol. Soc. Am. **56**, 275 (1945).
[39] P. S. Dodds and D. H. Rothman, Phys. Rev. E **59**, 4865 (1999).
[40] E. Tokunaga, Geophys. Bull. Hokkaido Univ. **15**, 1 (1966).
[41] E. Tokunaga, Geogr. Rep., Tokyo Metrop. Univ. **13**, 1 (1978).
[42] E. Tokunaga, Trans. Jpn. Geomorphol. Union. **5**, 71 (1984).
[43] D. M. Gray, J. Geophys. Res. **66**, 1215 (1961).
[44] J. E. Mueller, Geol. Soc. Am. Bull. **83**, 3471 (1972).
[45] M. P. Mosley and R. S. Parker, Geol. Soc. Am. Bull. **84**, 3123 (1973).
[46] J. E. Mueller, Geol. Soc. Am. Bull. **84**, 3127 (1973).
[47] D. R. Montgomery and W. E. Dietrich, Science **255**, 826 (1992).
[48] R. Rigon, I. Rodríguez-Iturbe, and A. Rinalod, Water Resour. Res. **34**, 3181 (1998).
[49] The topography used to extract areas and stream lengths is a composite of United States Geological Survey 3-arcsec digital elevation models available on the Internet at www.usgs.gov. These datasets provide grids of elevation data with horizontal resolution on the order of 90 m. The Kansas river was analyzed directly from the data, while the Mississippi basin was studied on a coarse-grained version with a horizontal resolution of approximately 1000 m.
[50] D. G. Tarboton, R. L. Bras, and I. Rodríguez-Iturbe, Water Resour. Res. **26**, 2243 (1990).
[51] H. Takayasu, I. Nishikawa, and H. Tasaki, Phys. Rev. A **37**, 3110 (1988).
[52] P. Meakin, J. Feder, and T. Jøssang, Physica A **176**, 409 (1991).
[53] G. Huber, Physica A **170**, 463 (1991).
[54] D. Dhar and S. N. Majumdar, J. Phys. A **23**, 4333 (1990).
[55] D. Dhar, Physica A **186**, 82 (1992).
[56] D. Dhar, Physica A **263**, 4 (1999).
[57] W. Feller, *An Introduction to Probability Theory and Its Applicatioins*, 3rd ed. (Wiley, New York, 1968), Vol. I.
[58] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. (Wiley, New York, 1971), Vol II.
[59] The data for the Nile and Congo were obtained from the United States Geological Survey's 30-arcsec Hydrol K dataset, which may be accessed on the Internet at edcftp.cr.usgs.gov. Note that these Hydrol K datasets have undergone the extra processing of projection onto a uniform grid.
[60] The dataset used for the Amazon has a horizontal resolution of approximately 1000 m and comes from 30-arcsec terrain data provided by the National Imagery and Mapping Agency available on the Internet at www.nima.mil.
[61] W. Dietrich and D. Montgomery, in *Scale Dependence and Scale Invariance in Hydrology*, edited by G. Sposito (Cambridge University Press, Cambridge, 1998), pp. 30–60.
[62] W. E. Dietrich and T. Dunne, in *Channel Network Hydrology*, edited by K. Beven and M. Kirkby (Wiley, New York, 1993), Chap. 7, pp. 175–219.
[63] A. N. Strahler, EOS Trans. AGU **38**, 913 (1957).
[64] S. A. Schumm, Bull. Geol. Soc. Am. **67**, 597 (1956).
[65] S. Peckham and V. Gupta, Water Resour. Res. **35**, 2763 (1999).
[66] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* 2nd ed. (Cambridge University Press, Cambridge, 1992).
[67] E. L. Lehman, *Nonparametrics: Statistical Methods Based on Ranks* (Holden-Day, San Francisco, 1975).
[68] P. Sprent, *Applied Nonparametric Statistical Methods*, 2nd ed. (Chapman & Hall, New York, 1993).