River landscapes and Optimal Channel Networks

Paul Balister[a], **József Balogh**[b], **Enrico Bertuzzo**[c], **Béla Bollobás**[a,d,e], **Guido Caldarelli**[f,g,h], **Amos Maritan**[i], **Rossana Mastrandrea**[f], **Robert Morris**[l], **and Andrea Rinaldo**[m,n,1]

[a]Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, US; [b]Department of Mathematics, University of Illinois, 1409 W. Green Street, Urbana IL 61801, US; [c]Dipartimento di Scienze Ambientali e Informatica, Università Cà Foscari Venezia, Venice, IT; [d]Trinity College, Cambridge CB2 1TQ, UK; [e]London Institute for Mathematical Sciences, South Street, London, UK; [f]IMT School for Advanced Studies, Lucca, Piazza S. Francesco 19, 55100 Lucca, IT; [g]Institute of Complex Systems CNR, via dei Taurini 19, 00185 Rome, IT; [h]European Centre for Living Technology, University of Venice, IT; [i]Dipartimento di Fisica e Astronomia, and INFN, University of Padua, Padova IT; [l]IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ Brazil; [m]Laboratory of Ecohydrology, École Polytechnique Fédérale de Lausanne CH; [n]Dipartimento ICEA, University of Padua, Padova IT

This manuscript was compiled on March 14, 2018

We study tree structures termed Optimal Channel Networks (OCNs) that minimize the total gravitational energy loss in the system, an exact property of steady-state landscape configurations that prove dynamically accessible and strikingly similar to natural forms. Here, we show that every OCN is a so-called *natural river tree*, in the sense that there exists a height function such that the flow directions are always directed along steepest descent. We also study the natural river trees in an arbitrary graph in terms of forbidden sub-structures, which we call k*-path obstacles*, and OCNs on a d-dimensional lattice, improving earlier results by determining the minimum energy up to a constant factor for every $d \geqslant 2$. Results extend our capabilities in environmental statistical mechanics.

Spanning Trees | Graph Theory | Slope-Area Law | Landscape Evolution

River networks can be viewed as rooted trees that, when extracted from fluvial landscapes by topographic steepest descent directions, show deep similarities of their parts and the whole, often across several orders of magnitude, despite great diversities in their geology, exposed lithology, vegetation, and climate (1). The large related body of observational data provide quintessential examples of real physical phenomena that can be effectively modelled using graph theory (2, 3). River networks are in fact the loopless patterns formed by fluvial erosion over a drainage basin. Such patterns, referred to as *spanning trees*, are such that a directed graph route exists for flows from every site of the catchment to an outlet, and are strictly related to the topographical surface whose gradients determine the edges (the drainage directions). A specific class of spanning trees, called Optimal Channel Networks (OCNs), was obtained by minimising a specific functional (4–6), later shown to be an exact property of the stationary solutions of the general equation describing landscape evolution (7, 8). The static properties and the dynamic origins of the scale-invariant structures of OCNs proved remarkable (1). OCNs are suboptimal (that is, dynamically accessible given initial conditions and quenched randomness frustrating the optimum search) configurations of a spanning network mimicking landscape evolution and network selection (Supporting Information, SI). Empirical and theoretical works have generally focused on the two dimensional case, though recently (inspired by vascular systems) three dimensional settings have also been analysed for an OCN embedded in a lattice (9). Several exact results have been derived for OCNs (7–16).

A model for a river network is obtained by taking a reasonably dense set of points on a terrain, and then joining each point to a nearby point downhill. Experimental observations suggest that OCNs should satisfy the following properties: (i) each portion of the drainage basin has a single output; (ii) a geomorphological stability condition holds. OCNs are defined on a finite graph G, consisting of a set $V(G)$ of vertices (or nodes), which correspond to sites in a drainage river basin, and a set $E(G)$ of edges between nearby sites. Moreover, each node v is endowed with a height, h_v. Water from each site/node v is *entirely* directed along the steepest descent, that is, towards the neighbour u of v which maximizes $h_v - h_u$, and we write

$$\Delta h_v = \max_{u \in N(v) \cup \{v\}} \left\{ h_v - h_u \right\}$$

where $N(v) = \{u : uv \in E(G)\}$ denotes the set of neighbours of node v. It is easy to see that the set of edges along which water flows is acyclic; in a single river basin it will be an oriented spanning tree, with all water flowing to a unique root. Note that the notation $N(v) \cup \{v\}$ also includes the possibility of finding nodes with no outlet. The second empirical property provides a connection between the gradient of the descent, and the landscape-forming flux of water flowing along the path. More precisely, in the (geological) steady state, if the landscape-forming water flowrate m_v at a site v increases, then the heights of the sites are modified by water erosion in such a way as to keep the product $m_v^{1/2} \Delta h_v$ constant. The exponent $1/2$ is experimentally determined by local slope-area plots where one assumes that landscape-forming discharges are proportional to total contributing area, and refers to the fluvial domain thereby neglecting unchanneled parts of the landscape (1). Such an exponent also emerges as the leading term of a general nonlocal erosion term under reparametrizazion invariance and a small-gradient approximation (8). Thus, the first rule microscopically ensures that no

Significance Statement

Optimal Channel Networks (OCNs) are a well-studied static model of river network structures. We present new exact results showing that every OCN is a natural river tree where a landscape exist such that the flow directions are always directed along its steepest descent. We characterize the family of natural river trees in terms of certain forbidden structures, called k*-path obstacles*. We thus determine conditions for river landscapes to imply a rooted tree as its network of topographic gradients. Our results are significant in particular for applications where OCNs may be used to produce statistically identical replicas of realistic matrices for ecological interactions.

All authors determined the problem, discussed the results and wrote the paper. PB, JB, BB and RM obtained the mathematical proofs.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: andrea.rinaldo@epfl.ch

cycles are created, while the second rule introduces a feedback between river structure and landscape erosion.

Here, we present new analytical results about the properties of the structure that minimizes the total gravitational energy loss. In particular, we will give a theoretical characterization of river networks in terms of forbidden sub-structures, which provides a simple way of determining which of the spanning trees of a graph are natural river trees. We will also study OCNs embedded in a d-dimensional lattice, proving, for each dimension $d \geqslant 2$, upper and lower bounds on the energy that differ by only a constant factor.

We begin with a formal description of OCNs, then we present the analytical results concerning the characterization of river trees, and the energy of an OCN embedded in lattices of any dimension d.

For each constant $\gamma \in (0, 1)$, we can associate to the (rooted) spanning trees of a graph G an *energy* function of the following form:

$$\mathcal{E}_\gamma(T) := \sum_{v \in V(G)} A_v(T)^\gamma, \qquad [1]$$

where $A_v(T)$ is the number of vertices $u \in V(G)$ such that the path in T from u to the root contains the vertex v (in other words, the number of vertices in the sub-tree of T 'rooted' at v). A spanning tree of G which minimizes the energy in Eq. (1) is called an *Optimal Channel Network* of G. In the case of river networks, if we imagine an open system where injection occurs at each site (vertex) at rate 1, and flowing along the edges of T until it reaches the root, then $A_v(T)$ represents the total flow of water out of v. The energy function defined in Eq. (1) has its origins in the exact result that steady-state configurations of river networks minimize the total loss of gravitational energy. This corresponds to:

$$\mathcal{E}_\gamma(T) = \sum_{v \in V(G)} m_v g \Delta h_v, \qquad [2]$$

where m_v is the mass of water leaving vertex v, so m_v is proportional to $A_v(T)$. Furthermore, experimental evidence suggests the following relationship between $A_v(T)$ and the difference of heights of two adjacent nodes: $\Delta h_v \propto A_v(T)^{-1/2}$ (1). We therefore obtain:

$$\mathcal{E}_\gamma(T) \propto \sum_{v \in V(G)} A_v(T)^{1/2}. \qquad [3]$$

Replacing the exponent $1/2$ with the parameter $\gamma \in (0, 1)$, and ignoring the (unimportant) constant factor, gives Eq. (1). When $\gamma = 1$ the class of directed networks that minimize the mean distance to the outlet is obtained, whose energy function is the same (7, 17). This is true if the functional in Eq. (1) is replaced by a more general one where T is substituted with the graph G itself including all edges. For such functional, every tree is a local minimum (8) and minimization is carried out under the constraint that locally (for each node) and globally a conservation law holds matching injection rates and inflows/outflows (7, 8, 17).

Given a rooted graph G (with one or more roots), we say that T is a *rooted spanning forest* of G if T is an acyclic subgraph of G, with one root of G in each component of T, and all edges oriented towards the corresponding root. The following natural definition will play an important role in our investigation of the optimal channel network(s) of a graph.

Definition. Let G be a rooted graph, and T be a rooted spanning forest of G. We say that T is a *natural river spanning forest* of G if there is an injective *height* function $f\colon V(G) \to \mathbb{R}$ such that if $v \in V(G)$ is a non-root, and $u \in N(v)$, then $v \to u$ is an edge of T if and only if

$$f(u) = \min\big\{ f(x) : x \in N(v) \cup \{v\} \big\}. \qquad [4]$$

We remark that we will usually be interested in the case when G has exactly one root, in which case we say that T is a *natural river spanning tree* of G.

Note that (i) every non-root vertex of T has out-degree exactly 1, (ii) $f(v)$ strictly decreases along every directed path in T, and (iii) the function f attains its minimum at one of the roots. In Theorem 1, below, we will provide a simple characterization of the natural river spanning trees of a graph in terms of forbidden sub-structures. We will also show (see Theorem 3) that every optimal channel network in a graph is also a natural river spanning tree of that graph. As a consequence, in order to study (either analytically or numerically) the tree structures that minimize the energy function $\mathcal{E}_\gamma(T)$ defined in Eq. (1), it suffices to consider natural river trees.

Which networks are river trees?. The definition of a natural river spanning tree creates various constraints on its possible shape and structure; in this section we shall consider these constraints.

Definition. Let G be a rooted graph, and let T a spanning tree of G. A *k-path obstacle* for T consists of a sequence of directed paths $(\eta_1, \ldots, \eta_k)$ in T, with the following property: for each $i \in \mathbb{Z}_k$, the last vertex of η_i is connected to the first vertex of η_{i+1} by an edge of G that is not in T.

In other words, if η_i is the path $a_i \to c_i(1) \to \cdots \to c_i(\ell_i) \to b_i$ in T, and

$$b_1 a_2, b_2 a_3, \ldots, b_k a_1 \in E(G) \setminus E(T),$$

then we say that T contains a k-path obstacle (see Figure 1). We emphasize that the paths η_i do not need to be vertex-disjoint (and it will be convenient in the proofs below to allow them to intersect); however, if k is minimal such that T contains a k-path obstacle, then the paths in any k-path obstacle will be disjoint. Observe that we will always consider the indices modulo k, so that (for example) $b_0 = b_k$ and $b_1 = b_{k+1}$, and that if $k = 1$ then necessarily $\ell_1 > 0$.

It is not difficult to see that a natural river tree cannot contain a k-path obstacle for any $k \in \mathbb{N}$; our first theorem states that this simple necessary condition is also sufficient, and hence characterizes the natural river trees.

Theorem 1. *Let G be a finite graph with a single root. A spanning tree T of G is a natural river spanning tree of G if, and only if, it has no path obstacle.*

Looking at fig. 1 one can easily deduce that a natural river tree by definition cannot contain any 3-path obstacle (i.e., cannot contain the red dashed edges). The same reasoning can be intuitively extended to any natural river tree and value of k. We will give here a formal proof of the necessary and sufficient conditions for T to be a natural river network.

Proof. We will show first that, for each $k \geqslant 1$, a natural river tree T cannot contain a k-path obstacle. Indeed, let

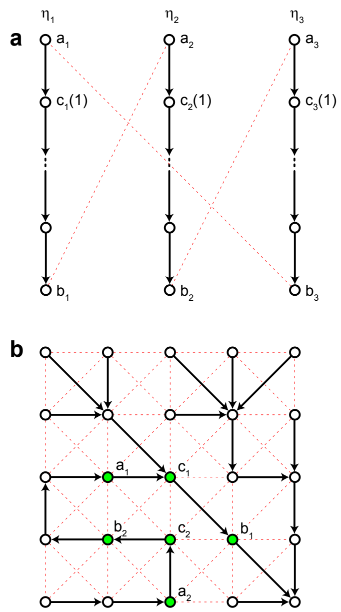

Fig. 1. k-path obstacles. (**a**) A 3-path obstacle. For $i \in \{1, 2, 3\}$, η_i is the directed path $a_i \to c_i(1) \to \cdots \to c_i(\ell_i) \to b_i$ in T, where ℓ_1, ℓ_2, ℓ_3 are arbitrary non-negative integers. Dashed red lines represent the edges $b_1 a_2$, $b_2 a_3$ and $b_3 a_1$ belonging to G but not to T. (**b**) An example of a 2-path obstacle ($a_1 \to c_1 \to b_1$ and $a_2 \to c_2 \to b_2$) where G (dashed red lines) is a regular two-dimensional lattice and T (black arrow) a spanning tree.

$f\colon V(G) \to \mathbb{R}$ be the height function associated to T, and suppose that there exist paths $\eta_1, \ldots, \eta_k$ in T, where η_i is the path $a_i \to c_i(1) \to \cdots \to c_i(\ell_i) \to b_i$ in T, and $b_1a_2, b_2a_3, \ldots, b_ka_1 \in E(G) \setminus E(T)$. Since $a_i \to c_i(1)$ (or $a_i \to b_i$ if $\ell_i = 0$) is in T, and $a_ib_{i-1} \in E(G)$, it follows from the definition that $f(b_i) \leqslant f(c_i(1)) < f(b_{i-1})$ for each $i \in \mathbb{Z}_k$, where $b_0 := b_k$. But then,

$$f(b_1) > f(b_2) > \cdots > f(b_k) > f(b_1), \qquad [5]$$

and this contradiction proves the claim.

We now turn to our main task: showing that if T has no path obstacle then it is a natural river tree. Given such a rooted spanning tree T in G, we introduce a directed graph D with vertex set $V(G)$ and

$$E(D) = \big\{x \to y \,:\, \text{either } x \to y \text{ is an edge of } T\text{, or} \\ \exists z \text{ such that } xz \in E(G) \setminus E(T) \text{ and } (z \to y) \in E(T)\big\}.$$

Claim. *The directed graph D is acyclic.*

Proof. Suppose that D contains a directed cycle with exactly k edges that are not edges of T, and note that $k \geqslant 1$, since T is a tree. Let us denote the vertices of the cycle as follows:

$$c_1(1) \to \cdots \to c_1(\ell_1) \to b_1 \to c_2(1) \to \cdots \to c_2(\ell_2) \to \\ b_2 \to \cdots \to b_{k-1} \to c_k(1) \to \cdots \to c_k(\ell_k) \to b_k \to c_1(1),$$

where $b_1, \ldots, b_k$ are the vertices whose out-neighbour in the cycle is not an edge of T, and $\ell_1, \ldots, \ell_k \geqslant 0$. Thus there exist vertices $a_1, \ldots, a_k$ such that, for each $i \in \mathbb{Z}_k$, we have $b_{i-1}a_i \in E(G) \setminus E(T)$ (where $b_0 := b_k$), and $a_i \to c_i(1)$ (or $a_i \to b_i$ if $\ell_i = 0$) is an edge of T. This gives a k-path obstacle, contradicting our assumption. □

It follows from the claim that the directed graph D can be extended to a linear ordering $<$ on $V(G)$ (i.e., a linear ordering with $x > y$ for every edge $x \to y$ of D). Choose a height function f such that $f(x) < f(y)$ if and only if $x < y$; we claim that T is a natural river spanning tree with this height function.

Indeed, let $v \in V(G)$ be a non-root, and note that there is a unique $u \in N(v)$ such that $v \to u$ is an edge of T, since all edges of T are directed towards the root. Note that $v \to u$ is an edge of D, by definition, so $f(u) < f(v)$. Now let $u \neq w \in N(v)$, and observe that $w \to u$ is an edge of D (since $wv \in E(G) \setminus E(T)$ and $v \to u$ is an edge of T), so $f(u) < f(w)$. Hence

$$f(u) = \min\big\{f(x) : x \in N(v) \cup \{v\}\big\},$$

as required. It follows that a spanning tree of G with no path obstacle is a natural river spanning tree of G, completing the proof of Theorem 1. □

Recall that an Optimal Channel Network (with parameter γ) in a graph G is a spanning tree of G that minimizes the energy $\mathcal{E}_\gamma(T)$. Our next theorem states that such a tree does not contain a path obstacle.

Theorem 2. *Let G be a finite graph with a single root, let T be a spanning tree of G, and let $k \in \mathbb{N}$ and $\gamma \in (0, 1)$. If T is an Optimal Channel Network in G with parameter γ, then T does not contain a k-path obstacle.*

The proof of this theorem is presented in section Material and Methods. Theorems 1 and 2 immediately imply the following fundamental fact.

Theorem 3. *Every Optimal Channel Network is a natural river spanning tree.*

In other words, if G is a finite graph with a single root, and T is an OCN in G with parameter $\gamma \in (0, 1)$, then T contains no path obstacle, by Theorem 2, and hence, by Theorem 1, is a natural river spanning tree of G.

Energy of Optimal Channel Networks in d-dimensional grids. In this section we will study in more detail the energy of an OCN in a d-dimensional lattice. In particular, we will focus on the case of the d-dimensional grid $[n]^d$, with root $(1, \ldots, 1)$, and edges between vertices at ℓ_∞-distance 1, though it will be clear from the proofs below that our results can be easily extended to other lattices. For each $n, d \in \mathbb{N}$ and $\gamma \in (0, 1)$, define

$$T(n, d, \gamma) := \min\big\{\mathcal{E}_\gamma(T) \,:\, T \text{ is a spanning tree of } [n]^d\big\},$$

the energy of an optimal channel network in the d-dimensional grid. The following theorem determines $T(n, d, \gamma)$ up to a constant factor (depending on d and γ) for all values of the parameters. It extends results of Colaiori et al. (13), who proved the lower bounds for $d = 2$ and $\gamma \neq 1/2$ (in fact, on an orthogonal grid), and improves a result of Briggs and Krishnamoorthy (9), who obtained the lower bound $3n^2/2 - O(n)$ in the case $d = 2$ and $\gamma = 1/2$.

Theorem 4. *For every integer $d \geqslant 2$, we have*

$$T(n, d, \gamma) = \begin{cases} \Theta(n^d), & \text{if } 0 < \gamma < 1 - 1/d, \\ \Theta(n^d \log n), & \text{if } \gamma = 1 - 1/d, \\ \Theta(n^{1+d\gamma}), & \text{if } 1 - 1/d < \gamma < 1, \end{cases} \qquad [6]$$

as $n \to \infty$.

Proof. Let us first prove the upper bounds, which follow from a straightforward recursive construction. We may assume for simplicity that $n = 2^k$, and consider 2^d copies of the construction $T_{n/2}$ for $[n/2]^d$, oriented so that $2^d - 1$ of them have their root near the center of $[n]^d$, and the last one has its root at $(1, \dots, 1)$, see Fig. 2. Join the vertex $(n/2, \dots, n/2)$ in this last subcube to each of the roots of the other subcubes, which are each of the form $(n/2, \dots, n/2) + \mathbf{x}$ for some $\mathbf{x} \in \{0,1\}^d$, so as to form a spanning tree T_n in $[n]^d$.

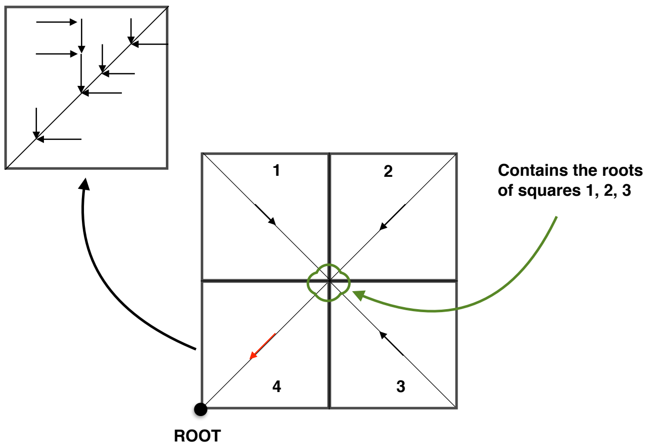

Fig. 2. The upper bound construction in the proof of Theorem 4. The tree is formed (in two dimensions) by 4 copies of the construction for $[n/2]^2$, placed so that squares 1, 2 and 3 have their root adjacent to the center of $[n]^2$, and the last one has its root at (1,1) (the black dot in the left bottom corner).

We claim that the subtree T_n of $[n]^d$ constructed in this way contains the diagonal edge $(x+1, \dots, x+1) \to (x, \dots, x)$ for each $1 \leqslant x < n$. Indeed, this holds (trivially) when $n = 1$, and by induction it follows for all $n = 2^k$, since we add the diagonal edge $(n/2+1, \dots, n/2+1) \to (n/2, \dots, n/2)$ in the induction step, and all other diagonal edges are in T_n by the induction hypothesis. It follows that all the the mass from $2^d - 1$ copies of $[n/2]^d$ flows down the diagonal of the final copy, and hence we have the following recursive formula:

$$\begin{aligned}\mathcal{E}_\gamma(T_n) \;&=\; 2^d \cdot \mathcal{E}_\gamma(T_{n/2}) + \\ &\quad + \sum_{x=1}^{n/2} \Big(\big(A_{(x,\dots,x)}(T_{n/2}) + n^d - (n/2)^d \big)^\gamma + \\ &\quad - A_{(x,\dots,x)}(T_{n/2})^\gamma \Big). \qquad [7]\end{aligned}$$

Since $(y+z)^\gamma \leqslant y^\gamma + z^\gamma$, and noting that $(n/2)\big(n^d - (n/2)^d\big)^\gamma \leqslant n^{1+d\gamma}$, it follows that

$$\mathcal{E}_\gamma(T_n) \leqslant 2^d \cdot \mathcal{E}_\gamma(T_{n/2}) + n^{1+d\gamma},$$

and hence, noting that $\mathcal{E}_\gamma(T_1) = 1$, and recalling that $n = 2^k$, we obtain

$$T(n,d,\gamma) \leqslant \mathcal{E}_\gamma(T_n) \leqslant n^{1+d\gamma} + \cdots + 2^{\ell d} \left(\frac{n}{2^\ell} \right)^{1+d\gamma} + \cdots + n^d. \qquad [8]$$

This inequality immediately implies the claimed upper bounds. Indeed, if $1 + d\gamma < d$ then the right-hand side of Eq. (8) is an increasing geometric series, so

$$T(n,d,\gamma) \;\leqslant\; \frac{n^d}{1 - 2^{1+d\gamma-d}} \;=\; O(n^d).$$

If $1 + d\gamma > d$ then it is a decreasing geometric series, so

$$T(n,d,\gamma) \;\leqslant\; \frac{n^{1+d\gamma}}{1 - 2^{d-1-d\gamma}} \;=\; O(n^{1+d\gamma}).$$

Finally, if $1 + d\gamma = d$ then all terms are equal, so

$$T(n,d,\gamma) \leqslant n^d \big(\log_2 n + 1 \big) = O(n^d \log n),$$

as required.

We next turn to the lower bounds. Observe first that the lower bound

$$\mathcal{E}_\gamma(T) \;=\; \sum_{v \in V([n]^d)} A_v(T)^\gamma \;\geqslant\; \sum_{v \in V([n]^d)} 1 \;=\; n^d$$

holds trivially for every spanning tree T of $[n]^d$, and every $\gamma \in (0,1)$, since we have $A_v(T) \geqslant 1$ for every vertex v. Hence $T(n,d,\gamma) \geqslant n^d$ holds always.

For the other lower bounds we will find it useful to work in a more general setting, assuming for simplicity that $n = 3^k$. Let $[n]^d_\circ$ denote a copy of the d-dimensional grid in which every vertex on the boundary is a root, and define

$$F(n,d,\gamma) := \min \big\{ \mathcal{E}_\gamma(F) \,:\, F \text{ is a rooted spanning forest of } [n]^d_\circ \big\},$$

where each component of F is assumed to have a single root (which lies on the boundary of $[n]^d$), and the energy $\mathcal{E}_\gamma(F)$ is defined to be the sum of the energies of the components. Note that $F(n,d,\gamma) \leqslant T(n,d,\gamma)$, since a spanning tree in $[n]^d$ contains a rooted spanning forest of $[n]^d_\circ$. We claim that

$$F(3n,d,\gamma) \;\geqslant\; 3^d \cdot F(n,d,\gamma) + \frac{\gamma}{3^{d(1-\gamma)}} \cdot n^{1+d\gamma}. \qquad [9]$$

To prove Eq. (9), we partition $[3n]^d$ into 3^d subcubes of size $[n]^d$. In each subcube we must send the mass to the boundary of this subcube, and this gives a contribution of at least $3^d \cdot F(n,d,\gamma)$ to the energy. Additionally, the middle subcube has to send its mass from its boundary to the boundary of the large cube $[3n]^d$, and therefore each drop of water must pass through at least n additional vertices on its way to a root. Note that, since $A_v(T) \leqslant (3n)^d$ and $\gamma < 1$, we have

$$A_v(T)^\gamma - \big(A_v(T) - x \big)^\gamma \;\geqslant\; \gamma \cdot A_v(T)^{\gamma - 1} \cdot x \;\geqslant\; \gamma \cdot (3n)^{d\gamma - d} \cdot x$$

for any $x > 0$. Therefore, moving the mass of the centre subcube to the boundary increases the energy by at least

$$n^d \cdot n \cdot \gamma \cdot (3n)^{d\gamma - d} = \frac{\gamma}{3^{d(1-\gamma)}} \cdot n^{1+d\gamma},$$

and hence we obtain Eq. (9) (Fig. 3).

Setting $C = \gamma \cdot 3^{-(d+1)}$, and noting that $F(1,d,\gamma) = 1$, we obtain

$$T(n,d,\gamma) \;\geqslant\; F(n,d,\gamma) \;\geqslant\; C \cdot n^{1+d\gamma} + \cdots \\ \cdots + C \cdot 3^{\ell d} \left(\frac{n}{3^\ell} \right)^{1+d\gamma} + \cdots + n^d,$$

where in the penultimate term $\ell = k - 2$. If $1 + d\gamma < d$ then this gives only a minor improvement over the trivial lower bound n^d, but for $1 + d\gamma > d$ the improvement is more substantial. Indeed, it implies that

$$T(n,d,\gamma) \;\geqslant\; \frac{\gamma + o(1)}{3^{d+1}} \cdot \frac{n^{1+d\gamma}}{1 - 3^{d-1-d\gamma}} \;=\; \Omega\big(n^{1+d\gamma}\big)$$

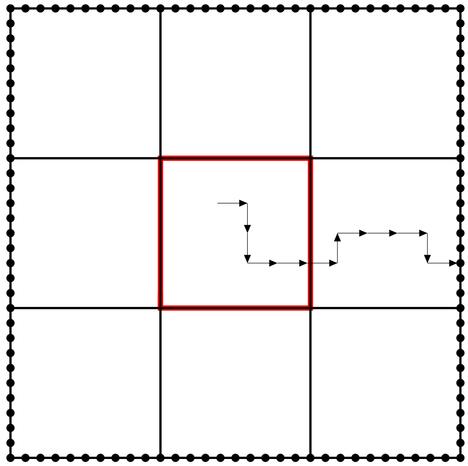

Fig. 3. **The proof of the lower bounds in Theorem 4.** Black dots represent the root on the boundary of $[3n]^d$ for each component of the spanning forest. The red square highlights the central subcube; the water from this subcube must travel to a root on the boundary, and to do so it must first cross its own boundary, and then pass through at least n additional vertices.

as $n \to \infty$. Finally, if $1 + d\gamma = d$, then we obtain

$$T(n, d, \gamma) \geqslant \frac{\gamma}{3^{d+1}} \cdot n^d \big(\log_3 n - 2\big) = \Omega\big(n^d \log n\big),$$

as required. □

An interesting (and likely difficult) challenge would be to decrease the implicit constant factors between the upper and lower bounds in Theorem 4 to factors of $1 + o(1)$, and hence to determine $T(n, d, \gamma)$ asymptotically.

Discussion

When one considers only the structure of the network (without imposing a height function on the vertices) the requirement that a river network at stationarity minimizes total energy dissipation takes a simple and elegant form: we need to minimize the function

$$\mathcal{E}_\gamma(T) \propto \sum_{v \in V(G)} A_v(T)^\gamma. \qquad [10]$$

We have studied the energy and the structure of the OCN (the tree that minimizes Eq. (10)), first on a general graph, and then (in more detail) for the particular case of a discrete lattice (Fig. 4, SI). Every OCN is a natural river tree, meaning that there exists a height function which determines the tree, and we characterized the natural river trees of a graph in terms of forbidden substructures. In the case of a d-dimensional lattice, we determined the energy of the OCN up to a constant factor for every $d \geqslant 2$ and $\gamma \in (0, 1)$, extending and improving previous work (9, 13).

We have also verified that, while it is natural to expect that any OCN on a discrete lattice should take the highly symmetric form of a Peano structure (1, 18), this is in fact not the case: for example, on the square lattice several suboptimal minimum energy structures are obtained by breaking the Peano structure symmetry (1) (SI). Much efforts have been devoted to reconciling the features of the ground state (known exactly since (11)) with those of feasible configurations. The resulting exponents associated with the global minimum did not match either the observational data or the numerical simulations (1). Because every local minimum of the OCN functional is a stationary solution of the general landscape evolution equation, any self-organization of the fluvial landscape corresponds to the dynamical settling of optimal structures into suboptimal niches of their fitness landscape. Thus feasible optimality, i.e., the search for optima that are accessible to the dynamics given the initial conditions and path obstacles, is the rule in the fluvial landscape and this might apply to a broad spectrum of interfaces in nature (17). Different network shapes give different values of $\mathcal{E}_\gamma(T)$ on a discrete lattice. Such case imposes sharp conditions on the minimising structures; in particular a specific subset of all the possible trees was selected as the natural rivers restricting the quest of minimum in this set.

Among the consequences of the results derived here, one notes that the constructability of elevation fields and topographies compatible with the planar imprinting of OCNs (Fig. 4 and SI), implies the possibility of building replicas of statistically identical matrices for ecological interactions. This is due to the flexibility of random search algorithms in a 'greedy' mode (SI), that cannot access the absolute minimum but rather suboptimal configurations that are known to reproduce accurately, differently from the ground state, those of natural fluvial landforms forms (17). This echoes other scaling properties of suboptimal interfaces (like e.g. in (19)). Interestingly, an application of limit scaling properties as a test of optimality has been been proposed for foodwebs (20) and a direct use of OCN landscapes has been made to explain elevational gradients of biodiversity (21).

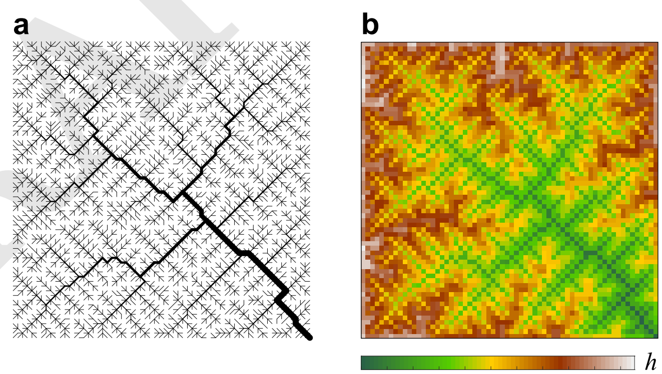

Fig. 4. **OCN and landscapes**. (**a**) An OCN derived within a 64× 64 regular two-dimensional lattice by means of a simulated annealing approach (see SI) assuming $\gamma = 1/2$. The algorithm ensures that the state is a local, dynamically accessible minimum, not necessarily the absolute minimum of the function $\mathcal{E}_\gamma(T)$. (**b**) The corresponding landscape (i.e. the elevation h_v) computed assuming $\Delta h_v \propto A_v(T)^{-1/2}$. We have verified numerically that the obtained OCN is a natural river spanning tree.

Material and methods

Proof of Theorem 2. Let T be an optimal channel network in G with parameter γ, and suppose (for a contradiction) that T contains a path obstacle. Let k be minimal such that there is a k-path obstacle in T, so there exist paths $\eta_1, \ldots, \eta_k$ in T, where η_i is $a_i \to c_i(1) \to \cdots \to c_i(\ell_i) \to b_i$, and

$$b_1 a_2, b_2 a_3, \ldots, b_k a_1 \in E(G) \setminus E(T).$$

Now, for each $i \in \mathbb{Z}_k$, define a new tree T_i by removing the edge $a_i \to c_i(1)$ (or $a_i \to b_i$ if $\ell_i = 0$) from T, and adding the edge $a_i \to b_{i-1}$ (where $b_0 := b_k$). If $k = 1$, then T_1 is a spanning tree of G with lower energy than T (since $\ell_1 > 0$), contradicting our assumption that T is an optimal channel network. Let us therefore assume from now on that $k \geqslant 2$.

To show that T_i is a tree, we need to use the minimality of k. Indeed, if there is a cycle in T_i then it must contain the edge $a_i \to b_{i-1}$, and therefore there exists a path

$$b_{i-1} \to d(1) \to \cdots \to d(\ell) \to a_i$$

in T. But now we can construct a $(k-1)$-path obstacle in T by replacing the paths η_{i-1} and η_i by the path

$$\begin{aligned} a_{i-1} \to c_{i-1}(1) \to \cdots \to c_{i-1}(\ell_i) \to b_{i-1} \to d(1) \to \cdots \\ \cdots \to d(\ell) \to a_i \to c_i(1) \to \cdots \to c_i(\ell_i) \to b_i, \end{aligned}$$

contradicting the minimality of k. Hence T_i is a tree, and therefore, since T is an optimal channel network, we have $\mathcal{E}_\gamma(T) \leqslant \mathcal{E}_\gamma(T_i)$ for each $i \in \mathbb{Z}_k$.

We next bound $\mathcal{E}_\gamma(T_i)$ from above by introducing a new tree, T_i', by removing the edge $a_i \to b_{i-1}$ from T_i, and adding the edge $a_i \to c_{i-1}(1)$. Note that this may not be a subgraph of G; however, it is a tree (since, as above, there is no path in T from $c_{i-1}(1)$ to a_i, by the minimality of k), and therefore its energy $\mathcal{E}_\gamma(T_i')$ may still be defined via Eq. (1). Moreover, we have $\mathcal{E}_\gamma(T_i) \leqslant \mathcal{E}_\gamma(T_i')$, since $A_v(T_i) \leqslant A_v(T_i')$ for every vertex $v \in V(G)$, and hence, by the observations above,

$$\mathcal{E}_\gamma(T) \leqslant \mathcal{E}_\gamma(T_i) \leqslant \mathcal{E}_\gamma(T_i').$$

For each $i \in \mathbb{Z}_k$, let us write ξ_i for the (unique) path in T from $c_i(1)$ to the root of T. Observe that $A_v(T_i') = A_v(T) - w_i$ for each $v \in V(\xi_i) \setminus V(\xi_{i-1})$, where $w_i := A_{a_i}$, and that $A_v(T_i') = A_v(T) + w_i$ for each $v \in V(\xi_{i-1}) \setminus V(\xi_i)$, so

$$\begin{aligned} \mathcal{E}_\gamma(T_i') = \mathcal{E}_\gamma(T) &+ \sum_{v \in V(\xi_i) \setminus V(\xi_{i-1})} \Big((A_v(T) - w_i)^\gamma - A_v(T)^\gamma \Big) \\ &+ \sum_{v \in V(\xi_{i-1}) \setminus V(\xi_i)} \Big((A_v(T) + w_i)^\gamma - A_v(T)^\gamma \Big). \quad [11] \end{aligned}$$

Now, let us write $S_v := \{ i \in \mathbb{Z}_k : v \in V(\xi_i) \}$ for each vertex $v \in V(G)$, and define $V_R := \{ v \in V(G) : S_v = R \}$ for each set $R \subset \mathbb{Z}_k$. Note that the vertices in V_R form a subpath of each ξ_i such that $i \in R$, and set

$$f_R(x) := \sum_{v \in V_R} \Big((A_v(T) + x)^\gamma - A_v(T)^\gamma \Big).$$

Since $\mathcal{E}_\gamma(T) \leqslant \mathcal{E}_\gamma(T_i')$, it follows from Eq. (11) that

$$\sum_{\substack{R \,:\, i \in R, \\ i-1 \notin R}} f_R(-w_i) + \sum_{\substack{R \,:\, i \notin R, \\ i-1 \in R}} f_R(w_i) \geqslant 0. \quad [12]$$

Now, observe that f_R is a concave function of x for any $\gamma \in (0,1)$ and $R \subset \mathbb{Z}_k$, and note that $f_R(0) = 0$. It follows that there exists a constant $\alpha_R \in \mathbb{R}$ such that $f_R(x) \leqslant \alpha_R x$ for every $x \in \mathbb{R}$, and moreover $f_R(x) < \alpha_R x$ whenever $x \neq 0$ and $V_R \neq \emptyset$. Hence, it follows from Eq. (12), and the fact that $w_i > 0$, that

$$\sum_{\substack{R \,:\, i \in R, \\ i-1 \notin R}} \alpha_R \leqslant \sum_{\substack{R \,:\, i \notin R, \\ i-1 \in R}} \alpha_R, \quad [13]$$

and moreover the inequality is strict if $V_R \neq \emptyset$ for any set R included in either sum. Hence, adding the α_R with $\{i-1, i\} \subset R$ to both sides of Eq. (14), we obtain

$$\sum_{R \,:\, i \in R} \alpha_R < \sum_{R \,:\, i-1 \in R} \alpha_R, \quad [14]$$

since $c_i(1) \in V_{\{i\}}$, so $V_{\{i\}} \neq \emptyset$. Setting $\alpha_i := \sum_{R \,:\, i \in R} \alpha_R$, it follows that $\alpha_1 > \alpha_2 > \cdots > \alpha_k > \alpha_1$, which is the desired contradiction.

□

ACKNOWLEDGMENTS. PB and BB are partially supported by NSF grant DMS 1600742. JB is partially supported by NSF grant DMS-1500121 and an Arnold O. Beckman Research Award (UIUC Campus Research Board 15006). BB and GC are partially supported by MULTIPLEX grant 317532. RM is partially supported by CNPq (Proc. 303275/2013-8), by FAPERJ (Proc. 201.598/2014), and by ERC Starting Grant 680275 MALIG. AR and EB have been supported by ERC Advanced Grant RINEC 22761. Part of the research in this paper was carried out while JB was a Visiting Fellow Commoner at Trinity College, Cambridge

1. Rodríguez-Iturbe I, Rinaldo A (2001) *Fractal River Basins: Chance and Self-Organization.* (Cambridge University Press, Cambridge (UK)).
2. Bollobás B (1979) *Graph Theory, An Introductory course.* (Springer), 1 edition.
3. Bollobás B (1985) *Random Graphs.* (Academic Press, London).
4. Rodríguez-Iturbe I, Rinaldo A, Rigon R, Bras R, Ijjasz-Vasquez E (1992) Fractal Structures as least energy patterns: the case of river networks. *Geophysical Research Letters* 19:889–892.
5. Rinaldo A, et al. (1992) Minimum energy and fractal structure of drainage networks. *Water Resour Res* 28:2183 – 2190.
6. Rodríguez-Iturbe I, et al. (1992) Energy dissipation, runoff production, and the three-dimensional structures of river basins. *Water Resources Research* 28:1095–1103.
7. Banavar J, Colaiori F, Flammini A, Maritan A, Rinaldo A (2000) Topology of the fittest transportation network. *Phys Rev Lett* 84(20):4745–4748.
8. Banavar J, Colaiori F, Flammini A, Maritan A, Rinaldo A (2001) Scaling, optimality and landscape evolution. *J Stat Phys* 104(1):1–33.
9. Briggs LA, Krishnamoorthy M (2013) Exploring network scaling through variations on optimal channel networks. *Proc Natl Acad Sci USA* 110(48):19295–300.
10. Rinaldo A, Rodríguez-Iturbe I, Rigon R, Ijjasz-Vasquez E, Bras R (1993) Self-organized fractal river networks. *Phys Rev Lett* 70(6):822–825.
11. Maritan A, Colaiori F, Flammini A, Cieplak M, Banavar J (1996) Disorder, river patterns and universality. *Science* 272(5264):984–988.
12. Rinaldo A, Maritan A, Colaiori F, Flammini A, Rigon R (1996) Thermodynamics of fractal networks. *Phys Rev Lett* 76(18):3364–3367.
13. Colaiori F, Flammini A, Maritan A, Banavar JR (1997) Analytical and numerical study of optimal channel networks. *Phys Rev E* 55(2):1298.
14. Banavar JR, Maritan A, Rinaldo A (1999) Size and form in efficient transportation networks. *Nature* 399(6732):130–2.
15. Santambrogio F (2007) Optimal channel networks, landscape function and branched transport. *Interfaces and Free Boundaries* 9(1):149–169.
16. Armaghan A, Singh A, Yang Z (2017) Emergent spectral properties of river network topology: an optimal channel network approach. *Sci. Rep.* 7:11486.
17. Rinaldo A, Rigon R, Banavar JR, Maritan A, Rodriguez-Iturbe I (2014) Evolution and selection of river networks: statics, dynamics, and complexity. *Proc. Natl. Acad. Sci. USA* 111(7):2417–24.
18. Flammini A, Colaiori F (1996) Exact analysis of the Peano basin. *Journal of Physics A: Mathematical and General* 29(21):6701.
19. Swift M, Maritan A, Banavar J (1996) Scaling properties of suboptimal interfaces. *Phys Rev Lett* 77(26):5288–5291.
20. Garlaschelli D, Caldarelli G, Pietronero L (2003) Universal scaling relations in food webs. *Nature* 423(6936):165–8.
21. Bertuzzo E, et al. (2016) Geomorphic controls on elevational gradients of species richness. *Proc Natl Acad Sci USA* 113(7):1737–1742.