Supporting Information

Supporting information (SI) contains two major parts. In the first part, we describe the measurements used for statistical analysis of dendritic arbor shapes, and present data to further support the universality and self-similarity. In the second part, we present our theory that accounts for statistical properties of basal dendrites of pyramidal cells from maximizing the repertoire of possible connectivity patterns between dendrites and surrounding axons for a given dendritic cost.

Our theoretical calculation relies on the analogy with maximization of the entropy for a given energy in statistical physics. In particular, calculating the connectivity repertoire by enumerating different arbor shapes of given dimensions is analogous to counting conformations of a branched polymer in statistical physics. We show that the sparseness of basal dendrites results from avoiding multiple potential synapses with the same axon, which is mathematically analogous to the Flory-type self-repulsive energy of a polymer chain in 2D (1).

I. Experimental procedures and statistical analysis of dendritic arbors

I.1 Reconstruction of dendritic arbors

The “2D” data set consists of 2161 basal dendritic arbors from layer 3 pyramidal cells of primate neocortex (see Supporting Table). Experimental procedures were described previously (2). Briefly, cells were labeled *in vitro* with Lucifer yellow, projected on a 2D plane parallel to the cortical surface and reconstructed with a Camera Lucida. All reconstructions were corrected for tissue shrinkage.

Supporting table

Summary of pyramidal dendritic arbors in the "2D" data set					
Species	Case	Gender	Hemisphere	Cortical areas	Number of cells
Baboon(3-5)	B2	Male	Right	V1, V2, V4, 3b, 4, 5, 6, 7, 23, 24, 46	446
Macaque monkey(2, 6-10)		Male	Right	V1, V2, TEO, TE, STP, MT, LIP, 7a, FEF	395
Vervet monkey(11-13)		Male	Right	V1, V2, V4, IT, 3b,4, 5, 6, 7a, 9d, 9m, 23, 24, 32	880
Marmoset monkey(14)			Right	V1, V2, ITc, ITr, DA, DM, DL, PP, MT	330
Galago(15, 16)			Right	V1, V2, DL, IT	110

The "3D" data set is consists of ten basal dendritic arbors of pyramidal cells from cat primary visual cortex (V1) (17). Reconstructions are publicly accessible from www.neuromorpho.org [see also(18)]. Cells were labeled with biocytin *in vivo* and reconstructed with Neurolucida (MicroBrightField, Cochester, VT) from multiple tissue sections. All reconstructions were corrected for tissue shrinkage in three dimensions.

The "Purkinje" data set consists of ten reconstructions of Purkinje dendrites available at www.neuromorpho.org [see also (18), (19-21)].

I.2 Dendritic Arbor Radius

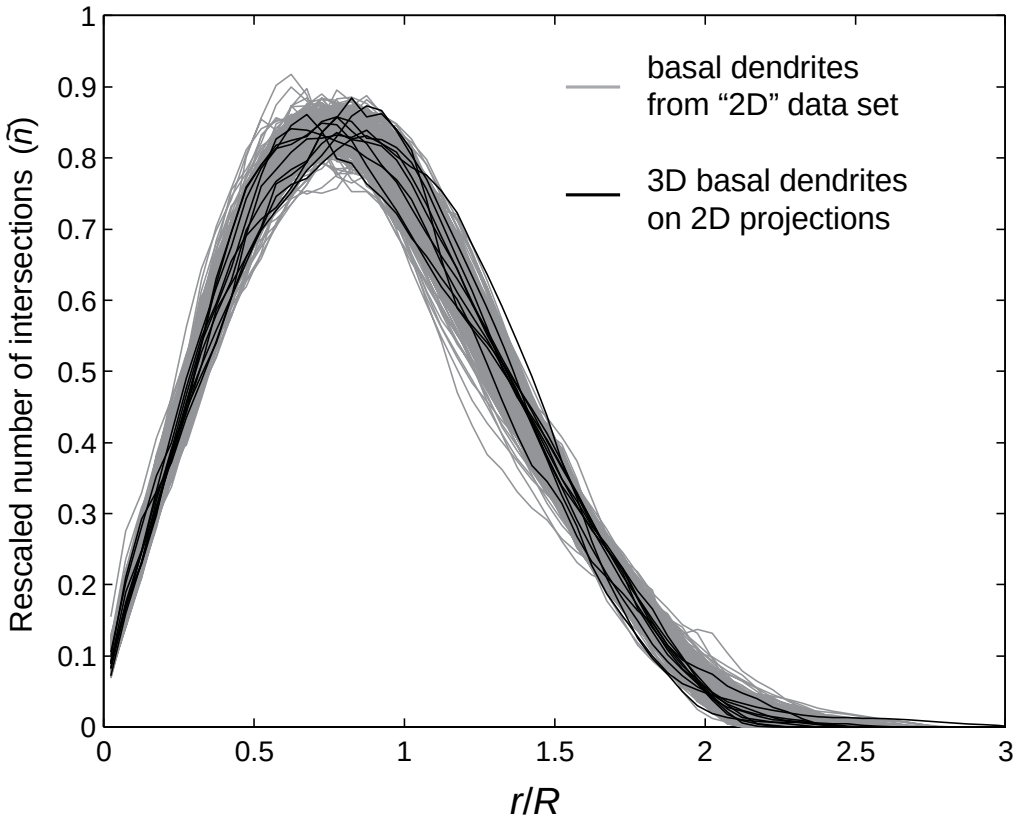
To quantify the linear size of an arbor, we define the arbor radius R of a dendritic arbor, — similar to the gyration radius in polymer physics — as

$$R^2 = \frac{1}{L^2}\sum_{i=1}^{K}\sum_{\substack{j=1 \\ j\neq i}}^{K}\delta l_i \delta l_j (\mathbf{r_i} - \mathbf{r_j})^2 \,, \qquad \text{(S1)}$$

where $\mathbf{r_i}$ is the position vector of a dendritic segment i, δl_i is the length of the segment, K is the total number of segments, and L is the total dendritic length.

As dendrites in “3D” and “Purkinje” data sets were reconstructed in vector format, we can use Eq. (S1) to calculate R.

Scarcity of 3D arbor reconstructions corrected for tissue shrinkage from homogeneous populations of neurons compelled us to use a collection of 2D reconstructions. This choice is supported by the fact that measurements on existing tissue shrinkage corrected 3D reconstructions are consistent with those on 2D reconstructions (Figure 3 and Supporting Figure 1).



Supporting Figure 1 | Spatial pairwise correlation between dendritic segments of a neuron from 2D projections of the “3D” data set and “2D” data set overlap on the rescaled coordinates. The 2D projections for arbors in the “3D” and “2D” data sets are parallel to the cortical surface. These data suggests that the measurements from pyramidal dendrites in the “3D” and the “2D” data sets are consistent with each other.

For the “2D” data set, we first skeletonized black and white images so that dendritic segments are represented by lines one pixel wide. Then we used Eq. (S1) to calculate R, where δl_i is given by the pixel size. In this case, L should be calculated as the product of the number of pixels in the arbor skeleton and the pixel size. We note that due to isotropic distribution of dendritic branches, our estimate of L would be similar to other more accurate length estimator algorithms (22).

I.3 Spatial correlations among dendritic branches

The statistical analysis of the distribution of branches within an arbor has a long history. The famous Sholl analysis (23-33), developed more than fifty years ago, measures the number of branches that intersect concentric spheres (or circles for 2D projections) centered at the soma as a function of sphere radius. The Sholl plot can be fit by a Gamma distribution (34), by a polynomial function (35), or, in the rising part, by a power law (36). However, the significance of the functional form of the Sholl plot remains unclear. Another tool, called cumulative-mass method, measures the total length of branches within a sphere as a function of the sphere radius (37, 38). Unlike the Sholl analysis, in which all spheres are centered at the soma, the cumulative-mass method involves averaging the dendritic length over identical spheres centered at different locations on the arbor. When the sphere radius is smaller than R, the average dendritic length within a

sphere scales with the sphere radius to some power called the fractal dimension (39). Nevertheless, the above methods have not revealed universal properties of dendritic arbor shape.

Here, we generalized the traditional Sholl plot (23) by repeating the branch-sphere intersections for spheres centered on randomly chosen dendritic segments, just like in the cumulative mass method (37). To avoid boundary effects due to regions with low density, we restricted the choice of the sphere center to the interior of the arbor, which is a distance less than $R/\sqrt{2}$ from the center of mass of the arbor. For each sphere radius r, we averaged the number, n, of branch-sphere intersections over center locations (Supporting Figure 2a). In the language of statistics, the generalized Sholl plot measures pairwise correlation in the locations of dendritic segments.

As most very small Sholl spheres intersect only one branch, they are not informative about the branch distribution and can be safely omitted. For this reason, our pairwise correlation measure start from sphere radius of 10 μm on the regular scale and from 0.1 R in rescaled units.

By combining the dendrites in the 3D data set, we found that the rescaled pairwise correlation function can be well fit by a function of the form,

$$\tilde{n}(x) = g x^{\mu} \exp(-h x^{\gamma}) \tag{S2}$$

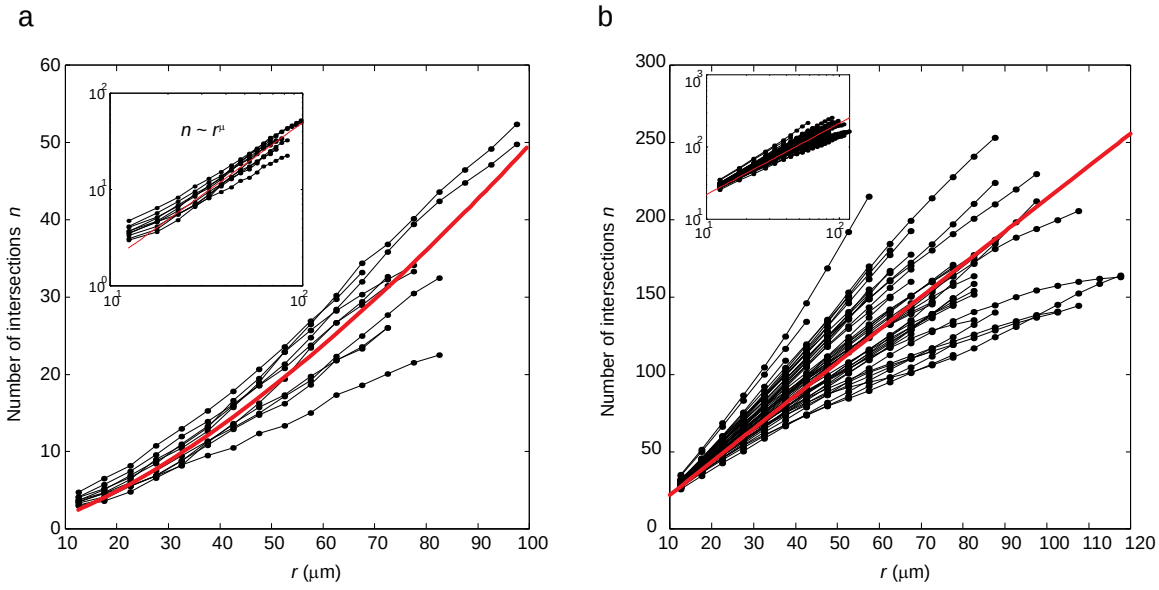
where $x = r/R$. Note that the fit only contains two free parameters, which we choose to be μ and γ, whereas the coefficients g and h were determined from the following normalization conditions:

$$\int_0^{\infty} \tilde{n}(x) dx = 1, \tag{S3}$$

which means that L in rescaled units is equal to one. The second moment of the distribution is equal to the R^2 and should be set to unity in rescaled units:

$$\int_0^\infty \tilde{n}(x)x^2 dx = 1. \qquad \text{(S4)}$$

In addition, the rising part of the pairwise correlation function in rescaled units (with x ranging from ~ 0.1 to 0.5) and on the regular scale (with r ranging from ~ 10 μm to $R/2$) were fit separately by power law functions (Figure 3b inset and Supporting Figure 2a). The goodness of the fit is not different from that in a canonical universal behavior, such as the 3D random walk (Supporting Figure 2b). Note that although the exponents μ obtained from the power law fit are slightly different from that obtained by fitting the entire curve with Eq. (S2) (Figure 3b), all exponents are well within the uncertainty of the measurement, compelling us to keep the same notation.



Supporting Figure 2 | The spread in the pairwise correlation function for pyramidal dendritic arbors is similar to that for 3D random walks. a, Rising part of the pairwise correlation function on regular scale for pyramidal dendritic arbors

and the same data on log-log scale (inset) show a power law relationship supporting self-similarity. Red line is the best fit and the adjusted coefficient of determination $r^2 = 0.91$, which gives the proportion of variability in the data set that is accounted by our statistical model. b, Rising part of the pairwise correlation function for a canonical universal behavior — 3D random walk. A simulated 3D random walk contains 10^5 steps and each step is 5 μm in length. Each of the 50 curves represents a different trajectory. The rising part of the pairwise correlation function for random walks should be characterized by a power law with exponent $\mu = 1$. Red line is the best fit with the exponent $\mu = 0.989 \pm 0.015$, which is not significantly different from the theoretical value. For this simulated data set, $r^2 = 0.83$. Thus, power law accounts as much variability in the rising part of the pairwise correlation function for pyramidal dendritic arbors as it does in random walks.

I.4 Relation between scaling exponents

If the statistical properties of different dendritic arbors were universal and dendritic arbors were indeed self-similar, the scaling exponent, ν, from the power law between R and L, would be related to the exponent from the rising part of spatial pairwise correlation function between dendritic segments, μ, as shown below.

For self-similar dendritic arbors built by common rules, the rising part of the pairwise correlation function on the regular scale, $n(r)$, would be described by the power law function (Supporting Figure 2a),

$$n \sim r^{\mu}, \; r < R \, . \tag{S5}$$

In addition, the pairwise correlation function would be characterized by a single length scale, R. Specifically, if $n(r)$ decays fast for $r > R$, and the rising part of $n(r)$ remains a

power law until r is on the order of R (as in agreement with our data), then the normalization condition $\int_0^\infty n(r)dr = L$ can be well approximated by introducing a cutoff R

$$\int_0^R n(r)dr \approx L\,. \tag{S6}$$

By using Eq. (S5) for $n(r)$ in Eq. (S6), we obtain

$$L \sim R^{\mu+1}. \tag{S7}$$

Recall that our measurement for arbor dimensions shows $R \sim L^{\nu}$ (Figure 3a). Therefore, we arrive at a relation between exponents μ and ν

$$\mu = \frac{1}{\nu} - 1. \tag{S8}$$

In addition, given Eqs. (S5) and (S8), the rising part of the rescaled pairwise correlation function should follow a similar power law $\tilde{n} \sim (r / R)^{\mu}$. This can be derived straightforwardly from the relation $\tilde{n} = nR / L$, which is a consequence of the normalization conditions.

I.5 Tortuosity of pyramidal dendrites

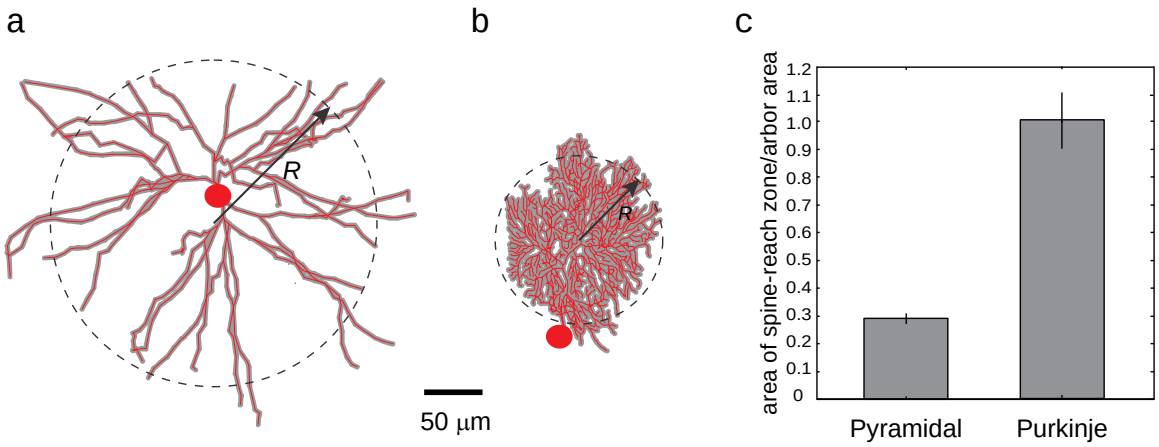
We measured the tortuosity of dendrites in 345 pyramidal cells from various cortical areas of one vervet and one baboon monkey. These cells come from the most complete subsets of neurons that one could obtain from only two brains. We vectorized cell drawings by manually tracing dendrites using NeuronJ (a plug-in to ImageJ). The retraced and digitally reconstructed data set is accessible from http://research.janelia.org/Chklovskii/Data.htm.

I.6 Expected number of potential synapses

To characterize the ability of a dendritic arbor to establish synaptic connections, we calculated the expected number of potential synapses, *P*, between a dendritic arbor and a straight axon passing through the arbor. First, we considered dendrites of a Purkinje cell, which are confined to a plane orthogonal to their major synaptic inputs – Parallel fibers. We define the dendritic arbor territory as the interior of the circle with radius *R* centered on the centroid of an arbor (Supporting Figure 3b). Then, *P* is given by the ratio of the spine-reach zone area, which is twice the average spine length *s* times the total arbor length in the arbor territory ($\approx L$), to the area of the arbor territory, πR^2,

$$P \approx \frac{2sL}{\pi R^2}. \qquad \text{(S9)}$$

We used $s = 2$ μm estimated from the center of a dendritic segment to the tip of a spine head (40-42).



Supporting Figure 3 | Pyramidal and Purkinje dendritic arbors differ in sparseness and, consequently, in the expected number of potential synapses per axon passing through the arbor territory. a, Basal dendritic arbor (red) of a pyramidal cell from rat cortex (43) and its spine-reach zone (grey). The arbor territory is approximated by the dashed circle. b, Dendritic arbor of a Purkinje cell from mouse cerebellum and its spine-reach zone. c, Ratio of the area of the spine-reach zone to the area of the arbor territory on a 2D projection [n = 10 for both pyramidal and Purkinje cells (17-21)]. Pyramidal basal dendrites are sparser than Purkinje dendrites ($p < 10^{-5}$, t-test with unequal variances) and, therefore, have lower probability to form a potential synapse with an axon passing through its territory. See Supporting information I.6 for details. Error bars are s.e.m. These digitally reconstructed arbors are publically accessible from http://NeuroMorpho.Org [see also (18)].

Second, we calculated the expected number of potential synapses between the basal dendritic arbor of a pyramidal cell and a straight axon passing through the arbor (Supporting Figure 3a). The expected number of potential synapses can be estimated on a projection as the ratio of the area of spine-reach zone to the arbor area. The area of the

spine-reach zone seen by the axon is given by the length of the projection of the 3D arbor onto a plane orthogonal to the axon L_{proj} times twice the spine length s. As axons run in various directions, we averaged over projections onto different planes:

$$P \approx \left\langle \frac{2sL_{proj}}{\pi R_{proj}^2} \right\rangle_{proj}, \quad \text{(S10)}$$

where R_{proj} is calculated by applying Eq. (S1) to the 2D projection of the arbor.

In the above we approximated the dendritic territory by a circle with radius R. This choice is consistent with statistical physics conventions and convenient for the following exposition. We verified that the results of this sub-section stand for other choices of the dendritic arbor territory, such as the convex hull.

I.7 Curve fitting and error analysis

Plots were fit by using the nonlinear least squares method from the MATLAB curve-fitting toolbox (MathWorks, Natick, MA). The errors of the fitting coefficients correspond to the 95 percent confidence bounds. For power law fit, we used linear regression method in the log-log coordinates as well, and the results are not significantly different from those obtained from the nonlinear method.

II. Theory and comparison with experimental data

II.1 Formulation of the theoretical principle

Our theory is based on the hypothesis that a dendritic arbor maximizes its functionality for a given dendritic cost to the organism. Consider a neuron that needs to make synapses with N out of M ($M \gg N$) axons in a given territory. We assume that each axon can be characterized by its appropriateness index i, so that the arbor functionality is proportional to the minimum appropriateness among connected axons. In the course of development, the arbor searches for the combination of axonal inputs that has maximum appropriateness. Below, we show that maximizing the arbor functionality reduces to maximizing the connectivity repertoire, which is defined as the natural *logarithm* of the total number of different combinations of axons that could synapse on a dendritic arbor of given dimensions.

We assume that $C(i)$ axons with appropriateness index i or greater are distributed randomly and uniformly among M axons passing through the dendritic arbor territory (see next section for experimental support of this assumption). Then the probability that the neuron would encounter N inputs with appropriateness index i or greater is given by the fraction:

$$f(i) = \frac{\binom{C(i)}{N}}{\binom{M}{N}}. \qquad \text{(S11)}$$

If *all* $\binom{M}{N}$ combinations of axonal inputs are available to a neuron, it could access all $\binom{C(i)}{N}$ different input combinations with appropriateness index i or greater. In reality,

the neuron can access only a small subset Ω of all $\binom{M}{N}$ combinations. One or more combinations of N inputs with appropriateness index i or greater are accessible to the neuron only if

$$f(i)\Omega \simeq \Lambda \ (\Lambda \geq 1 \text{ and is constant}), \tag{S12}$$

which is a *necessary* condition for finding combinations of N inputs with appropriateness index i or greater. Furthermore, if the searching algorithm is sufficiently good during development, then the neuron is able to find a combination of N inputs with appropriateness index i or greater.

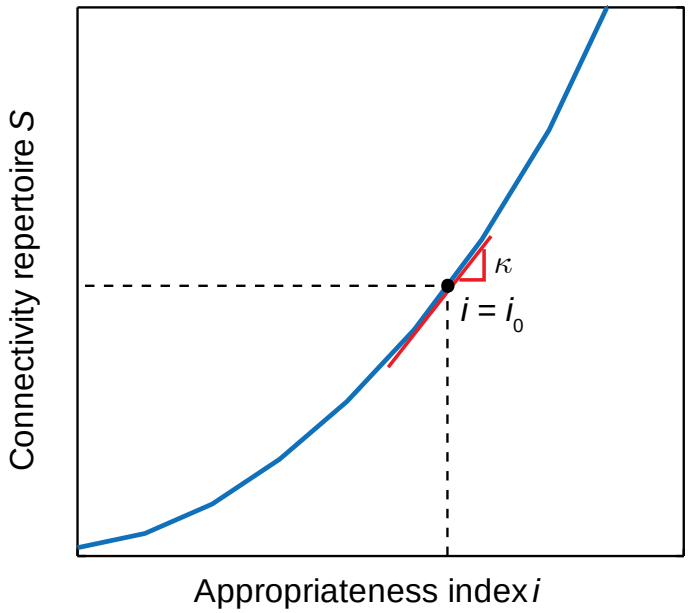
Now by substituting Eq. (S11) into (S12), we find that the connectivity repertoire, defined as $S = \ln\Omega$, is given by

$$S = \ln\Lambda + \ln\binom{M}{N} - \ln\binom{C(i)}{N}. \tag{S13}$$

Since $M \gg N$, and $C(i) \gg N$, Eq. (S13) can be well approximated as

$$S \approx \ln\Lambda - N\ln\frac{C(i)}{M}. \tag{S14}$$

As appropriate inputs are randomly and uniformly distributed in space, the fraction C/M is independent of arbor dimensions such as arbor radius R. Moreover, C/M decreases monotonically with i, and, hence, S increases monotonically with i (Supporting Figure 4). Thus, maximizing the arbor functionality, which we equate to finding axons with the highest minimum appropriateness index value, reduces to maximizing S.



Supporting Figure 4 | Connectivity repertoire is a monotonic increasing function of the minimum appropriateness index value i among axons connected with a dendritic arbor. The functionality of a real arbor operates in the vicinity of $i=i_0$, at which the slope of the curve is given by κ.

Maximizing arbor functionality for a given dendritic cost E reduces to maximizing their combination $i - \alpha E$, where α is a Lagrange multiplier reflecting the weight of E relative to functionality. As a consequence, the arbor functionality would operate at i_0 so that $\left.\frac{\partial i}{\partial E}\right|_{i=i_0} = \alpha$. Moreover, we can reduce the functionality-cost optimization to the minimization of the combination of S and E:

$$\mathcal{F} = -S + \beta E , \tag{S15}$$

where $\beta = \frac{\partial S}{\partial E} = \left.\frac{\partial i}{\partial E}\right|_{i=i_0} \left.\frac{\partial S}{\partial i}\right|_{i=i_0} = \alpha\kappa$, and κ is the slope of the S-i curve at $i = i_0$ (Supporting Figure 4). Minimizing $\mathcal{F}$ is analogous to minimizing free energy in statistical physics: S may be viewed as entropy and E as energy; β is analogous to inverse

temperature in statistical physics. If the solution of this optimization problem is consistent with observed properties of dendritic arbors, S and β can be determined from experimental data. Then from the S-i curve, κ, α and i_0 can be determined as well.

The use of the logarithm in defining S is motivated by the following: (1) The number of different combinations of axons synapsing on the whole arbor is a multiplicative function of the numbers of combinations calculated for the parts of the arbor. Then the logarithm of the number of combinations is an additive and extensive quantity. (It doubles when the total dendritic length is doubled, in agreement with our intuition.) (2) Dendritic cost is also an additive and extensive quantity. Therefore, combining cost with the connectivity repertoire in a reasonable optimization framework requires addition, an operation that calls for a logarithm of different connectivity patterns.

II.2 Calculation of the connectivity repertoire — assumptions and approximations

Prior to minimizing $\mathcal{F}$, we should subsequently introduce each term in Eq. (S15). To calculate S, we would make the following assumptions and approximations:

i. We ignore the differences in synaptic weights among connectivity patterns because this contribution (44) is independent of the arbor shapes.

ii. We ignore the contribution from both the variation in dendritic locations and the number of synaptic contacts between a pair of neurons to the connectivity repertoire. The latter scenario could be represented by a single contact with a combined synaptic weight, and, hence, is reduced to assumption i. [The effect of dendritic location on the number of input combinations has been considered by Poirazi and Mel (45).]

iii. Axonal inputs thought appropriate to a particular basal dendritic arbor, for example, through the action of learning rules, are randomly and uniformly distributed in space.

iv. Axons are uniformly and isotropically distributed in space.

v. Axons are straight and non-branching within the dendritic arbor territory.

vi. A dendritic arbor is sparse so that when an axon passes through the dendritic arbor field, there is a low probability of establishing a potential synapse (Supporting Figure 3a).

vii. Dendritic arbors are characterized as flexible chains jointed by rigid segments. The length of a segment, called persistence length in polymer physics, is much longer than the inter-synapse interval on a dendrite.

viii. Dendritic branches are nearly straight (Figure 4e). In the language of polymer physics, dendrites are “strongly stretched”.

Assumption iii is supported by the following experimental observations. First, the numerically dominant inputs to pyramidal dendrites are pyramidal axons, and a pyramidal axonal arbor consists of sparsely distributed straight branches and the arbor size is much larger than the basal dendritic arbor radius (17, 26, 46-48). Second, most excitatory synapses of a pyramidal cell come from distant neurons. For example, in cat primary visual cortex, 83 percent of synapses on a dendrite originate from neurons located farther than 200 μm in the cortical plane (49). For these pre-synaptic neurons, typically only one axonal branch courses through the dendritic arbor territory. Thus, on the scale of a basal dendritic arbor, pyramidal axons are rarely branching straight lines running in different directions (50). As a result, basal dendrites located at different depths in the cortex have no way of predicting where they would meet appropriate axonal inputs. Third, there is no correlation in the location of pyramidal axonal and dendritic branches (51).

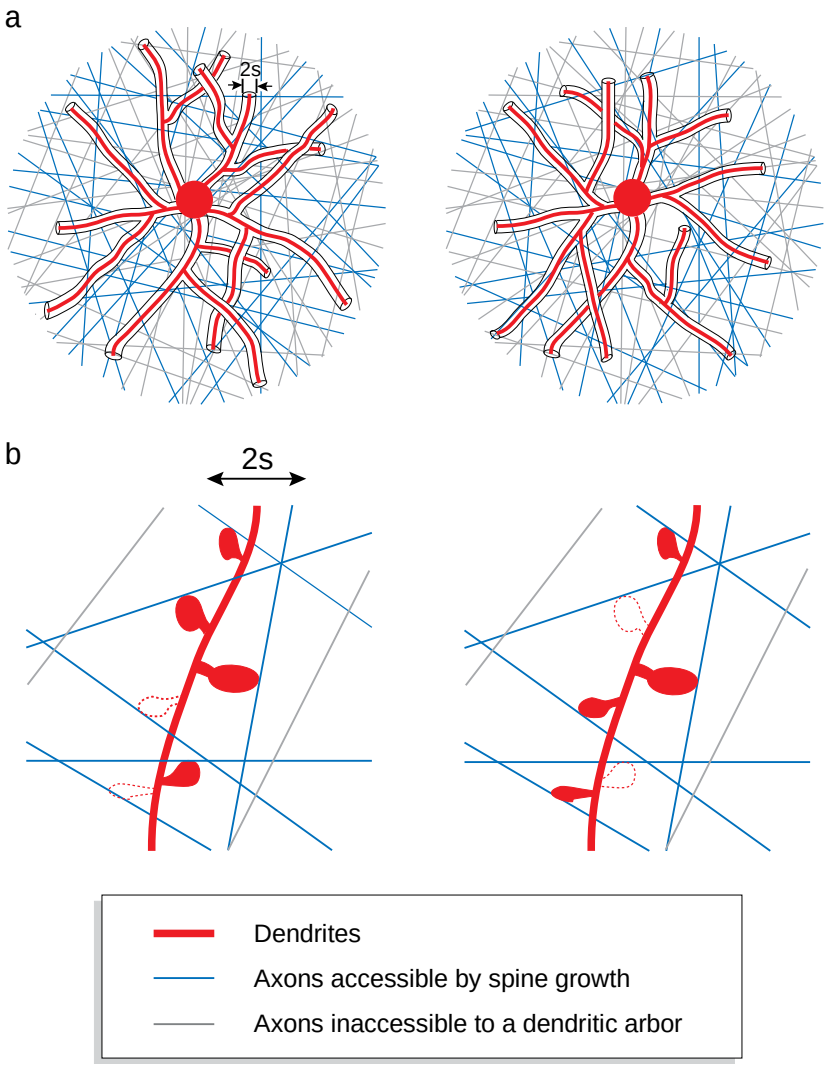
Note that the random spatial distribution of axons does not imply random formation of actual connections. The former scenario could exist even if the exact neuronal circuit is specified genetically, that is, the circuits in different animals could be wired up identically despite the variation in the axonal locations. If the circuit is not predetermined, then there is no preferred spatial distribution, and random locations of synaptic inputs would be expected.

We will recover assumptions vi-viii from the solution of minimizing $\mathcal{F}$ in Eq. (S15) , thus making our theory self-consistent. The theory can be extended to the situation when axons are tortuous and branching, thus relaxing assumption v.

As shown in the main text, the connectivity repertoire could be expressed as the sum of two contributions

$$S = \ln \Omega_p + \ln \Omega_a \,. \qquad \text{(S16)}$$

First, Ω_p is the total number of distinct arbor shapes (Supporting Figure 5a), and each arbor shape selects a different subset of surrounding axons as potential synapses. Because dendritic arbor is thin (~ 1 μm in diameter) and long ($L \sim 10^3 - 10^4$ μm), calculating the first contribution is analogous to counting the number of different conformations of a branched polymer chain and $\ln\Omega_p$ is called the conformational entropy. In this analogy, a piece of dendrite shorter than the persistence length a behaves like a stiff rod, due to the mechanical properties of cytoskeletal filament assembly. However, on the scale much longer than the persistence length, the dendrite can behave quite flexibly. Such a picture is consistent with the observation that dendrites can make sharp turns by substantially changing their orientations (≥ 90 degree, see Figure 2). Second, Ω_a is the total number of ways of choosing actual connections out of potential synapses for given arbor shape (Supporting Figure 5b). Note that Eq. (S16) relies on the assumption that different arbor shapes could have very few identical combinations of inputs. In other words, we may ignore the possibility that all the actual connections on distinct arbors could be from the same axons. Below, we show that our approximation is justified under assumptions vi and vii.



Supporting Figure 5 | Two steps of calculating the connectivity repertoire. a, Distinct conformations of dendritic arbors result in different sets of potential synapses between dendrites and surrounding axons. b, For a given arbor shape, different input patterns can be achieved by choosing actual connections out of the set of potential synapses.

We start by estimating the probability that the connectivity pattern established by one arbor could be achieved by another arbor. If P is the probability of an axon to establish a potential synapse with a dendritic arbor, the probability of having N actual

connections in common is P^N, if the locations of neighboring synapses are independent. As the number of different arbor shapes is Ω_p, a given input pattern can be achieved by $P^N\Omega_p$ arbor shapes. To guarantee that the total number of common input patterns is small, we must have

$$P^N\Omega_p \ll 1. \tag{S17}$$

By taking the logarithm on both sides, Eq. (S17) is equivalent to

$$N \ln P + \ln \Omega_p < 0. \tag{S18}$$

According to results from polymer physics, the total conformation entropy $\ln\Omega_p \simeq L/a$ (1), where a is the persistence length. Substituting the entropy expression into Eq. (S18), we find that for a reasonably sparse dendritic arbor (say $P \leq 0.3$) with the persistence length longer than the inter-synapse interval ($a > L/N$), the number of common input patterns achievable by different arbors is small and can be neglected.

II.3 Multiple potential synapses reduce the connectivity repertoire

Next, we elaborate upon each term in Eq. (S16). First, for a given arbor shape, we calculate the number of different ways to choose actual connections out of potential axonal targets. As individual connections can be selected independently, the number of choices Ω_a is simply the binomial coefficient $\binom{N_p}{N}$, where N_p is the total number of potential synapses between the dendritic arbor and its surrounding axons, and N is the number of actual connections per dendritic arbor.

However, an essential correction needs to be made. Even with the assumption of sparse dendritic arbors, a small number of axons establish multiple potential synapses with different dendritic segments in an arbor. These multiple potential synapses from the same axon do not expand the connectivity repertoire and should be subtracted. Thus, the number of choices for different actual connections reduces to

$$\Omega_a = \binom{N_p - N_2}{N}, \tag{S19}$$

where N_2 is the number of axons establishing two potential synapses, which we call double hits. Higher numbers of potential synapses occur much less frequently than double-hits and can be ignored in the calculation.

If $N_2 \ll N_p$, the logarithm of Ω_a can be approximated as

$$\ln \Omega_a \approx \ln \binom{N_p}{N} + N_2 \ln(1 - N / N_p). \tag{S20}$$

For cortical neurons, the connectivity fraction N/N_p is much smaller than one (42). Then, using the relation $\ln(1- N/N_p) \approx -N/N_p$, we can simplify Eq. (S20) further as

$$\ln \Omega_a \approx \ln \binom{N_p}{N} - N \frac{N_2}{N_p}. \quad \text{(S21)}$$

II.4 Avoiding multiple potential synapses generates effective repulsion among dendrites

In this section, we calculate the dependence of the connectivity repertoire, $\ln\Omega_a$, on arbor dimensions. In particular, we compute a contribution of multiple hits to the repertoire. This calculation is mathematically analogous to calculating a self-exclusion energy of a polymer chain in 2D (52). First, let us calculate N_p in Eq. (S21). It was shown previously (42) that

$$N_p = \frac{\pi}{4} 2sL\rho_a, \quad \text{(S22)}$$

where s is the average spine length and ρ_a is the length density of axons with units of μm^{-2}, that is the total length of axons enclosed in a given volume divided by the volume.

The expression for N_p can be understood by considering a 2D projection of a dendritic arbor. This idea will also turn out to be useful in calculating the number of double hits, N_2. Each axon establishing potential synapses has to pass through the spine-reach zone of the dendritic arbor (i.e., gray area surrounding the dendrites in Fig. 3). If axons are oriented perpendicular to dendritic segments, N_p is simply the product of the spine-reach zone area $2sL$ and the axonal density.

In a more realistic scenario, axons run in different directions and we may assume that the directions of axon are distributed isotropically relative to the dendrite. Then, in Eq. (S22), one should substitute $2sL$ by the mean area of the spine-reach zone on a random 2D projection, which is reduced by a numerical factor $\pi/4$. We refer to the previous work (42) for a general derivation of the result. Below, we consider a simpler version of the calculation.

The area of the spine-reach zone on a projection is $2s\sum_i a\sin\theta_i$, where a is the persistence length and θ_i is the angle between a segment and the normal vector of the 2D plane. Then the area of the spine-reach zone averaged over random 2D projections is $\langle\sin\theta_i\rangle 2s\sum_i a = \langle\sin\theta_i\rangle 2sL$ and the mean $\sin\theta_i$ is given by

$$\langle\sin\theta_i\rangle = \frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi}|\sin\theta|\sin\theta d\theta d\varphi = \frac{\pi}{4}, \tag{S23}$$

where the solid angle 4π is a normalization factor.

Next we estimate N_2 in Eq. (S21). If axons are straight, a double hit takes place when an axon pass through the sites where the spine-reach zones of two dendritic segments overlap on a 2D projection (Figure 4a), and N_2 should be proportional to the number of double hit sites,

$$N_2 \approx \#\text{ double hit sites}\times(2s)^2\rho_a, \tag{S24}$$

where $(2s)^2$ approximates the area of one double hit site.

To calculate the number of double hit sites, we adopt a mean-field approach developed by Flory (52) in polymer physics. Consider a dendritic arbor projecting onto a plane with arbor territory area πR_{proj}^2. We may divide the area into small pixels $2s$ in size. The probability that a pixel contains a dendritic segment is

$$P \approx \frac{2sL_{proj}}{\pi R_{proj}^2}, \tag{S25}$$

where L_{proj} is the total dendritic length on a 2D projection.

If we ignore the correlations among dendritic segments, which is the essence of Flory mean-field approximation, the probability that two dendritic segments stay in the same pixel is P^2. Multiplying P^2 by the total number of pixels $\pi R_{proj}^2/4s^2$, we have

$$\text{\#double hit sites} \approx \left\langle \frac{L_{proj}^2}{\pi R_{proj}^2} \right\rangle . \quad \text{(S26)}$$

where $\langle\ \rangle$ denotes an average over different 2D projections of an arbor.

Eq. (S26) is in agreement with measurements from 3D basal dendritic arbors of pyramidal cells (Supporting Figure 6). To count the actual number of double hit sites, each 3D arbor (n = 10) was first projected randomly onto a 2D plane. On the 2D arbor, a disk with diameter 2*s* (*s* = 2 μm) was centered on each segment 2*s* in length. The total overlapping area between different pairs of disks dividing πs^2 gives the number of double hit sites on one projection. The actual number of double hit sites for a dendritic arbor is the average over different 2D projections.

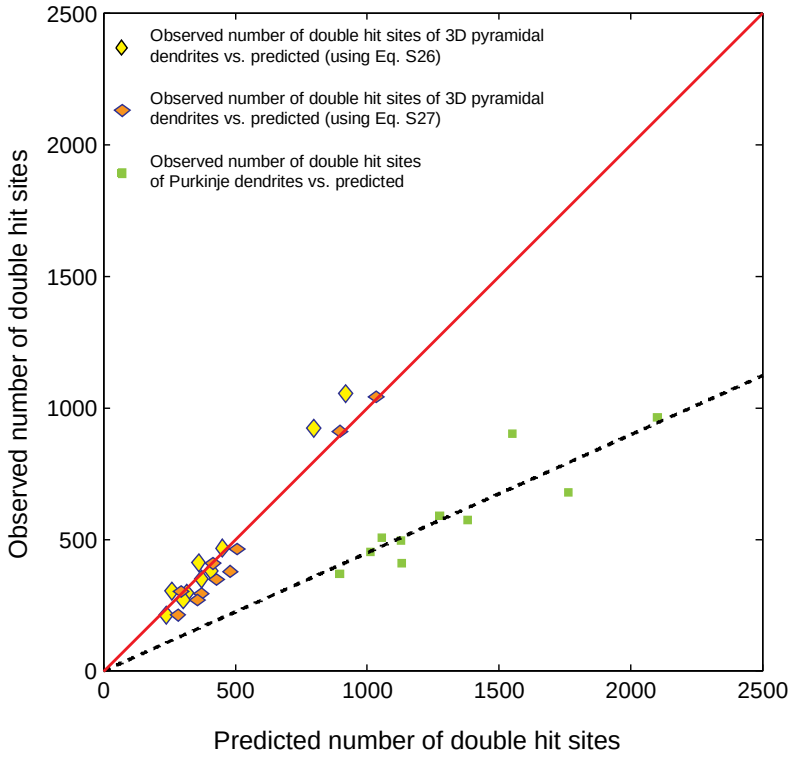
In addition, Eq. (S26) can be re-written in terms of *L* and *R* defined in 3D. As $L \simeq L_{proj}$, $R \simeq R_{proj}$, we may write

$$\text{\#double hit sites} \approx \frac{L^2}{\pi R^2}. \quad \text{(S27)}$$

The agreement between Eqs. (S26)-(S27) and the measurements (Supporting Figure 6) strongly support the applicability of the mean-field approximation. By substituting Eq. (S27) into Eq. (S24), we obtain

$$N_2 \simeq \frac{L^2}{R^2} s^2 \rho_a , \quad \text{(S28)}$$

which shows that the number of double hits decreases with *R*.



Supporting Figure 6 | Number of double-hit sites measured from arbor reconstructions versus that predicted from the Flory approximation. Perfect predictions fall on the diagonal. The number of double hit sites for 3D pyramidal cells is in agreement with the theory whereas Purkinje dendrites have fewer double hit sites than expected.

By substituting the expressions for N_2 and N_p [Eqs. (S28) and (S22)] into Eq. (S21), we have

$$\ln \Omega_a \simeq \ln \binom{sL\rho_a}{N} - \frac{sL}{R^2} N \,. \tag{S29}$$

This formula is essential to understand the sparseness of a dendritic arbor. The first term, which is independent of R, is irrelevant for optimization. But the second term, arising from double hit, should be minimized with respect to R to maximize the connectivity

repertoire. As this term is inversely proportional to the arbor area, it can be viewed as an effective repulsion among dendrites that favors sparse arbors.

However, we note that real physical force due to volume exclusion between branches cannot account for the sparseness of a pyramidal dendritic arbor. This is because the neuropil where a neuron is located is densely packed by other cells (50), and the volume density is about constant everywhere so that dendritic branches experience no net force. This effect is well known in the physics of polymer melts (1), an analogous system containing densely packed polymer chains.

II.5 Conformational entropy of dendritic arbors

Next, we switch our attention to the first term in Eq. (S16) and estimate the number of different arbor shapes Ω_p for given dimensions. By using results from statistical physics of branched polymers (53, 54), the conformational entropy may be written as

$$\ln \Omega_p(R,\ell) = \ln \Omega_0 + \frac{L}{a}\ln(1-R/\ell) - \frac{\ell^2}{La}, \tag{S30}$$

where ℓ is the average path length from the tip of a branch to the soma.

The first term $\ln\Omega_0 \simeq L/a$ is the total conformational entropy of dendritic arbors for given L and different arbor radii. The second term is a correction due to the fact that straighter branches have fewer arbor shapes. This logarithmic term is applicable only if dendritic branches are nearly straight (53), i.e., $R \approx \ell$, an assumption that will be recovered later by minimizing $\mathcal{F}$ in Eq. (S15), thus making our theory self-consistent. The last term is a correction due to the fact that branchy arbors have more arbor shapes, which is applicable only if $\ell \ll L$ (54).

In calculating the conformational entropy, we have adopted the freely-jointed chain model: dendritic segments are completely rigid on the scale shorter than the persistence

length, but could be highly flexible on the scale longer than the persistence length. If a segment is stiff but can still fluctuate and bend, dendrites would be better characterized by the worm-like chain model (55). In the latter model, the conformational entropy may be given by

$$\ln \Omega_p(R,\ell) = \ln \Omega_0 - \frac{L}{a(1-R/\ell)} - \frac{\ell^2}{La}. \tag{S31}$$

Because we do not possess enough experimental data to distinguish between the two chain models, we primarily consider the freely jointed chain model. Results from the worm-like chain model are similar and are briefly compared below.

Now, by substituting Eqs. (S29) and (S30) into Eq. (S16), the expression for the connectivity repertoire is given by

$$S \simeq \log\binom{sL\rho_a}{N} - \frac{sL}{R^2}N + \ln \Omega_0 + \frac{L}{a}\ln(1-R/\ell) - \frac{\ell^2}{La}. \tag{S32}$$

By introducing S_0 to absorb the terms independent of R and ℓ and assuming $N \sim L$ (56), we obtain the expression for S introduced in the main text

$$S \simeq S_0 + \frac{L}{a}\ln(1-R/\ell) - \frac{\ell^2}{La} - \frac{L^2}{R^2}. \tag{S33}$$

II.6 Maximizing the connectivity repertoire alone predicts tortuous dendritic branches

In this section, we will show that maximizing the connectivity repertoire alone would predict more tortuous dendritic branches than those in a pyramidal cell, a discrepancy that motivates us to introduce the dendritic path length cost in the following sections.

To calculate the tortuosity of dendritic branches, we need to find optimal ℓ and R. However, maximizing Eq. (S33) with respect to ℓ and R would not lead to a self-consistent result satisfying $R \approx \ell$ and $\ell \ll L$, two conditions that justify the applicability of Eq. (S33). To find a self-consistent solution, we would assume that $R \ll \ell$. In this case, the logarithmic term in Eq. (S33) should be replaced by $R^2/(\ell a)$ (54) and the connectivity repertoire reads

$$S \simeq S_0 - \frac{R^2}{\ell a} - \frac{\ell^2}{La} - \frac{L^2}{R^2}. \tag{S34}$$

By taking $\partial S / \partial \ell = 0$ and $\partial S / \partial R = 0$, we find that $\ell \simeq a^{2/10} L^{8/10}$, $R \simeq a^{3/10} L^{7/10}$, and hence $R \simeq a^{1/8} \ell^{7/8}$ (54). Given that a is only a few microns, $R \approx \ell^{7/8} < \ell$, thus making our calculation self-consistent. In Figure 4e, the relation $R = \ell^{7/8}$ was plotted (yellow line) and compared with experimental measurements for pyramidal dendrites.

II.7 A cost model for dendritic arbors

As maximizing the connectivity repertoire alone leads to more tortuous branches than observed, we must include the dendritic cost in the optimization framework. The cost must grow with the path length ℓ (34, 57, 58), because longer path length from synapses to soma invokes greater metabolic energy for intracellular transport and for counteraction

of attenuation of synaptic currents. In addition, the cost is likely to increase with L as well (59-61), for longer dendrites occupy more space and require more metabolic energy.

After considering several simplified models based on the above reasoning (57), we propose that the cost scales with L and ℓ as:

$$E \sim L\ell^{\delta}, \tag{S35}$$

The value of exponent δ is model-dependent. Below, we consider one simplified model that yields a specific value of δ.

We assume that E is proportional to the total area of dendritic surface for the following reason. To maintain the resting potential of a neuron, or low [Na+] inside a cell, sodium ions permeating into a neuron should be actively extruded by Na+/K+ pumps (62). The amount of ATP expended is proportional to the membrane capacitance for action potential propagation or membrane input conductance in the absence of spikes (62). Hence, in both considerations, E is proportional to the membrane surface area,

$$E \sim Ld, \tag{S36}$$

where d is the mean dendritic diameter.

Although E decreases as the mean dendritic diameter d is reduced, a small diameter would detrimentally affect dendritic function. One possible reason is that proteins must be transported from the cell body to synapses and the rate of transport is proportional to the number of microtubules at a given cross-section of a dendrite (63). As the rate of protein transport through a cross-section must be proportional to the number of synapses downstream of that cross-section. If the density of microtubules (per cross-sectional area) is roughly invariant, the cross-sectional area must be proportional to the number of microtubules and hence to the number of synapses downstream (63-65). If the typical

branch has length ℓ, the number of synapses on it is $N\ell/L$. Therefore, a typical dendritic diameter is:

$$d^2 \simeq AN\ell / L\,, \tag{S37}$$

where A is a cross-sectional area needed to support one synapse. By substituting Eq. (S37) into Eq. (S36), we find that

$$E \sim N^{1/2} L^{1/2} \ell^{1/2}\,. \tag{S38}$$

Given that $N \sim L$ (56), we obtain

$$E \sim L\ell^{1/2}\,. \tag{S39}$$

II.8 Minimizing the dendritic cost stretches dendrites

With all the parts of the theory given in Eq. (S35) and Eq. (S33), now we can combine them into $\mathcal{F}$ in Eq. (S15), which will be minimized with respect to R and ℓ. Note that we would only consider the regime $\beta E \gg \ell^2/(La)$, and, thus, ignore the conformational entropy term due to branching. The opposite regime has already been discussed in Section II.7, and will not be considered any more. Ignoring all the terms independent of R and ℓ, we have

$$\mathcal{F} \simeq \frac{L^2}{R^2} - \frac{L}{a}\ln(1 - R/\ell) + \beta L \ell^{\delta}\,, \tag{S40}$$

where the weight of E relative to S, β, has absorbed all the constant pre-factors. Eq. (S40) will be minimized below.

Let us start by considering two terms in $\mathcal{F}$ that depend on ℓ. One term,

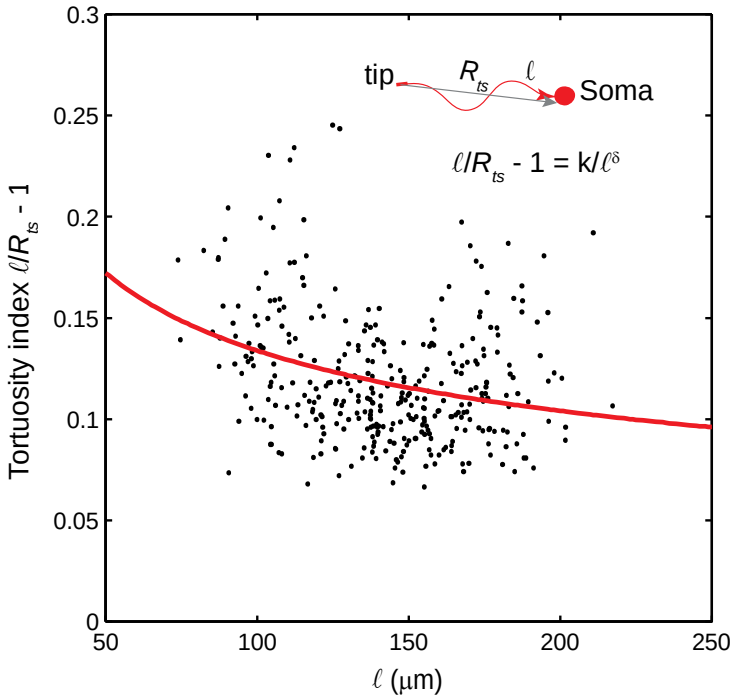
$$\mathcal{F}_{el} = -\frac{L}{a}\ln(1 - R/\ell)\,, \tag{S41}$$

is analogous to the elastic energy in polymer physics. Minimizing this term favors $\ell \gg R$, and, hence, tortuous branches. The dendritic cost $L\ell^{\delta}$ favors shorter path length, and hence, stretches the branches. When we add them, the competition leads to a minimum, at which the tortuosity of branches can be calculated. By taking $\partial \mathcal{F} / \partial \ell = 0$, we find that

$$\frac{\ell}{R} - 1 = \frac{1}{\delta \beta \ell^{\delta} a}. \qquad \text{(S42)}$$

The worm-like chain model would give a similar result but with a different exponent for ℓ: ℓ/R -1 ~ $1/\ell^{\delta/2}$.

Eq. (S42) shows that dendrites with longer path length tend to have smaller tortuosity. This prediction agrees with measurements for pyramidal dendrites (Figure 4f and Supporting Figure 7). To compare the prediction with the experimental data, we have substituted R in the formula by the average Euclidean distance from the tip of a branch to the soma R_{ts} (Supporting Figure 7). The replacement will not affect the results, for R and R_{ts} are approximately equal to each other ($R \approx 1.01\ R_{ts}$ measured for basal pyramidal dendrites).



Supporting Figure 7 | Tortuosity index as a function of the average path length from branch tip to soma. R_{ts} is the average Euclidean distance from the tip of a branch to the soma. Each dot represents a different arbor. Red line is the power law fit with exponent δ = 0.36 +/- 0.12, which is not significantly different from the exponent extracted from Figure 4f. Note large cell to cell variability of tortuosity index. The coefficient in the power law function k = 0.71, which can be used to calculate the persistence length according to Eq. (S42).

As discussed in section II.6, Eq. (S41) is applicable only if $\ell \approx R$ (dendrites are "strongly stretched" in the language of polymer physics). Thus, one must have $\beta a \ell^{\delta} \gg 1$ so that our calculation is self-consistent. The numerical values of β and a will be computed later by fitting to the experimental data.

II.9 Scaling of the arbor radius with the total dendritic length

If we combine Eq. (S42) and Eq. (S41) for "strongly stretched" arbors, the elastic energy can be approximated as

$$\mathcal{F}_{el} \simeq \frac{L}{a}\ln(\beta R^{\delta} a)\,. \tag{S43}$$

One sees that $\mathcal{F}_{el} \ll (L/a)\beta R^{\delta} a \approx \beta E$ and can be ignored in $\mathcal{F}$. Then only two terms are left in Eq. (S40),

$$\mathcal{F} \simeq \frac{L^2}{R^2} + \beta L R^{\delta}\,. \tag{S44}$$

In Eq. (S44), the first term on the right side resembles the repulsive energy and favors a sparser arbor whereas the dendritic cost prefers a more compact arbor. The competition leads to a minimum for which the optimal arbor radius

$$R \simeq \left(\frac{L}{\beta}\right)^{\frac{1}{2+\delta}}, \tag{S45}$$

The same formula holds in the worm-like chain model as well. Note that to build a self-consistent theory, β should be sufficiently small so that the optimal arbor is sparse (assumption vi in section II.3).

In Eq. (S45), assuming all the coefficients are constant, we obtain the scaling law $R \sim L^{\nu}$ with

$$\nu = \frac{1}{2+\delta}\,. \tag{S46}$$

II.10 Comparison of the theory with experimental data

The theory leads to predictions that can be directly compared with experimental data. First, by using Eq. (S42) to fit the tortuosity of dendrites (Figure 4f), we found that the exponent in E (Eq. (S35)) $\delta = 0.309 \pm 0.005$. Second, substituting the measured δ into Eq. (S46), we predicted that $\nu = 0.43$. This value is not significantly different from the direct measurement of this exponent, $\nu = 0.44 \pm 0.01$.

We next determine the numerical values of β and the persistence length a by using the experimental data. By including all the numerical factors ignored in Eq. (S45), we find that β is given by

$$\beta = \frac{16sL}{\pi^2 \delta b R^{2+\delta}} \left(\frac{R}{\ell} \right)^{\delta}, \tag{S47}$$

where b is the average inter-spine interval. The values for $L/R^{2+\delta}$ and R/ℓ can be computed from the "2D" data set of pyramidal dendrites (see Figure 3a and Figure 4e). For other coefficients, we use b = 0.5 μm (50), s = 2 μm (42, 50) . We find $\beta = 1.2 \pm 0.2$ $\mu m^{-1.3}$.

To find the persistence length a, we use Eq. (S42) to fit the tortuosity of pyramidal dendrites (Supporting Figure 7) and compute the best-fit coefficient $\beta a \delta$ in the power law. Having determined β and δ, we find $a \approx 4$ μm. Since $a \gg b \approx 0.5$ μm, we recover assumption vii in section II.3, as promised previously. We note that this consistency also holds for the worm-like chain model.

In summary, we accounted for the scaling relation between R and L and the tortuosity of dendritic branches by minimizing $\mathcal{F}$ in Eq. (S15). In addition, by using two independent measurements, we showed that sparse dendritic arbors, "strongly stretched"

branches, and long persistence length, are all consistent with the solution of minimizing $\mathcal{F}$.

References

1. de Gennes P-G (1979) *Scaling concepts in polymer physics* (Cornell University Press, Ithaca, N.Y.) p 324.
2. Elston GN & Rosa MGP (1997) The occipitoparietal pathway of the macaque monkey: comparison of pyramidal cell morphology in layer III of functionally related cortical visual areas. *Cereb. Cortex* 7:432-452.
3. Elston GN, Benavides-Piccione R, Elston A, DeFelipe J, & Manger P (2005) Specialization in pyramidal cell structure in the cingulate cortex of the Chacma baboon (*Papio ursinus*): an intracellular injection study of the posterior and anterior cingulate gyrus with comparative notes on macaque and vervet monkeys. *Neuroscience Lett.* 387:130-135.
4. Elston GN, Benavides-Piccione R, Elston A, DeFelipe J, & Manger P (2005) Pyramidal cell specialization in the occipitotemporal cortex of the Chacma baboon (*Papio ursinus*). *Exp. Brain Res.* 167:496-503.
5. Elston GN, Benavides-Piccione R, Elston A, Manger P, & DeFelipe J (2005) Specialization in pyramidal cell structure in the sensory-motor cortex of the Chacma baboon (*papio ursinus*) with comparative notes on the macaque monkey. *Anat. Record* 286:854-865.
6. Elston GN & Rosa MGP (1998a) Morphological variation of layer III pyramidal neurones in the occipitotemporal pathway of the macaque monkey visual cortex. *Cereb. Cortex* 8:278-294.
7. Elston GN, Tweedale R, & Rosa MGP (1999a) Cortical integration in the visual system of the macaque monkey: large scale morphological differences of pyramidal neurones in the occipital, parietal and temporal lobes. *Proc. R. Soc. Lond. Ser. B* 266:1367-1374.
8. Elston GN (2001) Interlaminar differences in the pyramidal cell phenotype in cortical areas 7 m and STP (the superior temporal polysensory area) of the macaque monkey. *Exp Brain Res* 138(2):141-152.
9. Elston GN & Rockland K (2002) The pyramidal cell of the sensorimotor cortex of the macaque monkey: phenotypic variation. *Cerebral Cortex* 12:1071-1078.
10. Elston GN, Benavides-Piccione R, & DeFelipe J (2005) A study of pyramidal cell structure in the cingulate cortex of the macaque monkey with comparative notes on inferotemporal and primary visual cortex. *Cerebral Cortex* 15:64-73.
11. Elston GN, Benavides-Piccione R, Elston A, Manger P, & DeFelipe J (2005) Pyramidal cell specialization in the occipitotemporal cortex of the vervet monkey (*Cercopithecus pygerythrus*). *NeuroReport* 16:967-970.
12. Elston GN, Benavides-Piccione R, Elston A, DeFelipe J, & Manger P (2005) Specialization in pyramidal cell structure in the sensory-motor cortex of the vervet monkey (*Cercopithecus pygerythrus*). *Neuroscience* 134:1057-1068.

13. Elston GN, Benavides-Piccione R, Elston A, Manger P, & DeFelipe J (2005) Regional specialization in pyramidal cell structure in the limbic cortex of the vervet monkey (*Cercopithecus pygerythrus*): an intracellular injection study of the anterior and posterior cingulate gyrus. *Exp. Brain Res.* 167:315-323.
14. Elston GN, Tweedale R, & Rosa MGP (1999b) Cellular heterogeneity in cerebral cortex. A study of the morphology of pyramidal neurones in visual areas of the marmoset monkey. *Journal of Comparative Neurology* 415:*33-51*.
15. Elston GN, Elston A, Kaas J, & Casagrande VA (2005) Regional specialization in pyramidal cell structure in the visual cortex of the galago. An intracellular injection study with comparative notes on New World and Old World monkeys. *Brain, Behav. Evol.* 66:10-21.
16. Elston GN, Elston A, Casagrande VA, & Kaas JH (2005) Pyramidal neurons of granular prefrontal cortex the galago: complexity in the evolution of the psychic cell. *Anat. Rec.* 285:610-618.
17. Kisvarday ZF & Eysel UT (1992) Cellular organization of reciprocal patchy networks in layer III of cat visual cortex (area 17). *Neuroscience* 46(2):275-286.
18. Ascoli GA (2006) Mobilizing the base of neuroscience data: the case of neuronal morphologies. *Nat Rev Neurosci* 7(4):318-324.
19. Martone ME, *et al.* (2003) The cell-centered database: a database for multiscale structural and protein localization data from light and electron microscopy. *Neuroinformatics* 1(4):379-395.
20. Rapp M, Segev I, & Yarom Y (1994) Physiology, morphology and detailed passive models of guinea-pig cerebellar Purkinje cells. *J Physiol* 474(1):101-118.
21. Vetter P, Roth A, & Hausser M (2001) Propagation of action potentials in dendrites depends on dendritic morphology. *J Neurophysiol* 85(2):926-937.
22. Dorst L & Smeulders AWN (1987) Length estimators for digitized contours. *Computer Vision, Graphics, and Image Processing* 40(3):311-333.
23. Sholl DA (1953) Dendritic organization in the neurons of the visual and motor cortices of the cat. *J Anat* 87(4):387-406.
24. Duan H, Wearne L, Morrison JH, & Hof PR (2002) Quantitative analysis of the dendritic morphology of corticocortical projection neurons in the macaque monkey association cortex. *Neuroscience* 114(2):349-359.
25. Jacobs B & Scheibel AB (2002) Regional dendritic variation in primate cortical pyramidal cells. *Cortical Areas: Unity And Diversity.*, eds Schüz A & Miller R (Taylor and Francis, London), pp 111-131.
26. Lund JS, Yoshioka T, & Levitt JB (1993) Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cereb. Cortex* 3:148-162.
27. Elston G, Rosa MGP, & Calford MB (1996) Comparison of dendritic fields of layer III pyramidal neurones in striate and extrastriate visual areas of the marmoset: a Lucifer Yellow intracellular injection study. *Cereb. Cortex* 6:807-813.
28. Duan HL, *et al.* (2003) Age-related dendritic and spine changes in corticocortically projecting neurons in macaque monkeys. *Cerebral Cortex* 13(9):950-961.
29. Li Y, Brewer D, Burke RE, & Ascoli GA (2005) Developmental changes in spinal motoneuron dendrites in neonatal mice. *Journal of Comparative Neurology* 483(3):304-317.

30. Elston GN, Benavides-Piccione R, & DeFelipe J (2001) The pyramidal cell in cognition: a comparative study in human and monkey. *J. Neurosci* 21:RC163(161-165).
31. Nakayama AY, Harms MB, & Luo LQ (2000) Small GTPases Rac and Rho in the maintenance of dendritic spines and branches in hippocampal pyramidal neurons. *Journal of Neuroscience* 20(14):5329-5338.
32. Horch HW, Kruttgen A, Portbury SD, & Katz LC (1999) Destabilization of cortical dendrites and spines by BDNF. *Neuron* 23(2):353-364.
33. Nedivi E, Wu GY, & Cline HT (1998) Promotion of dendritic growth by CPG15, an activity-induced signaling molecule. *Science* 281(5384):1863-1866.
34. Burke RE & Marks WB (2002) Some Approaches to Quantitative Dendritic Morphology. *Computational Neuroanatomy: Principles and Methods*, ed Ascoli GA (Humana Press), p 476.
35. Ristanovic D, Milosevic NT, & Stulic V (2006) Application of modified Sholl analysis to neuronal dendritic arborization of the cat spinal cord. *Journal of Neuroscience Methods* 158(2):212-218.
36. Rothnie P, Kabaso D, Hof PR, Henry BI, & Wearne SL (2006) Functionally relevant measures of spatial complexity in neuronal dendritic arbors. *J Theor Biol* 238(3):505-526.
37. Caserta F, *et al.* (1995) Determination of fractal dimension of physiologically characterized neurons in two and three dimensions. *J Neurosci Methods* 56(2):133-144.
38. Caserta F, *et al.* (1990) Physical mechanisms underlying neurite outgrowth: A quantitative analysis of neuronal shape. *Physical Review Letters* 64(1):95-98.
39. Mandelbrot BB (1982) *The fractal geometry of nature* (W.H. Freeman, San Francisco) p 460
40. Anderson JC, Douglas RJ, Martin KA, & Nelson JC (1994) Map of the synapses formed with the dendrites of spiny stellate neurons of cat visual cortex. *J Comp Neurol* 341(1):25-38.
41. Napper RM & Harvey RJ (1988) Quantitative study of the Purkinje cell dendritic spines in the rat cerebellum. *J Comp Neurol* 274(2):158-167.
42. Stepanyants A, Hof PR, & Chklovskii DB (2002) Geometry and structural plasticity of synaptic connectivity. *Neuron* 34(2):275-288.
43. Shepherd GM & Svoboda K (2005) Laminar and columnar organization of ascending excitatory projections to layer 2/3 pyramidal neurons in rat barrel cortex. *J Neurosci* 25(24):5670-5679.
44. Varshney LR, Sjostrom PJ, & Chklovskii DB (2006) Optimal information storage in noisy synapses under resource constraints. *Neuron* 52(3):409-423.
45. Poirazi P & Mel BW (2001) Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* 29(3):779-796.
46. Binzegger T, Douglas RJ, & Martin KA (2005) Axons in cat visual cortex are topologically self-similar. *Cereb Cortex* 15(2):152-165.
47. Fujita I & Fujita T (1996) Intrinsic connections in the macaque inferior temporal cortex. *Journal of Comparative Neurology* 368:467-486.

48. Anderson JC, Binzegger T, Douglas RJ, & Martin KA (2002) Chance or design? Some specific considerations concerning synaptic boutons in cat visual cortex. *J Neurocytol* 31(3-5):211-229.
49. Stepanyants A, Martinez LM, Ferecsko AS, & Kisvarday ZF (2009) The fractions of short- and long-range connections in the visual cortex. *Proc Natl Acad Sci U S A* 106(9):3555-3560.
50. Braitenberg V & Schüz A (1998) *Cortex : statistics and geometry of neuronal connectivity* (Springer, Berlin ; New York) 2nd thoroughly rev. Ed pp xiii, 249.
51. Stepanyants A, Tamas G, & Chklovskii DB (2004) Class-specific features of neuronal wiring. *Neuron* 43(2):251-259.
52. Flory PJ (1953) *Principles of polymer chemistry* (Cornell University Press, Ithaca,) p 672 p.
53. Grosberg AY & Khokhlov AR (1994) *Statistical physics of macromolecules* (AIP Press, New York) pp xxvii, 350.
54. Gutin AM, Grosberg AY, & Shakhnovich EI (1993) Polymers with Annealed and Quenched Branchings Belong to Different Universality Classes. *Macromolecules* 26(6):1293-1295.
55. Rubinstein M & Colby RH (2003) *Polymer physics* (Oxford University Press, Oxford ; New York) pp xi, 440 p.
56. Larkman AU (1991) Dendritic morphology of pyramidal neurones of the visual cortex of the rat: III. Spine distributions. *J Comp Neurol* 306(2):332-343.
57. Wen Q & Chklovskii DB (2008) A cost-benefit analysis of neuronal morphology. *J Neurophysiol* 99(5):2320-2328.
58. Cuntz H, Borst A, & Segev I (2007) Optimization principles of dendritic structure. *Theor Biol Med Model* 4:21.
59. Chklovskii DB (2004) Synaptic connectivity and neuronal morphology: two sides of the same coin. *Neuron* 43(5):609-617.
60. Mitchison G (1991) Neuronal branching patterns and the economy of cortical wiring. *Proc Biol Sci* 245(1313):151-158.
61. Cherniak C (1992) Local optimization of neuron arbors. *Biol Cybern* 66(6):503-510.
62. Attwell D & Laughlin SB (2001) An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab* 21(10):1133-1145.
63. Hillman D (1979) Neuronal shape parameters and substructures as a basis of neuronal form. *The Neurosciences : fourth study program*, (MIT Press, Cambridge, Mass.), pp vii, 1185 p.
64. Hsu A, Tsukamoto Y, Smith RG, & Sterling P (1998) Functional architecture of primate cone and rod axons. *Vision Res* 38(17):2539-2549.
65. Wittenberg GM & Wang SS-H (2007) Evolution and scaling of dendrites. *Dendrites*, eds Stuart G, Spruston N, & Häusser M (Oxford University Press, USA), 2rd Ed, pp 43–67.