Trade-off between Multiple Constraints Enables Simultaneous Formation of Modules and Hubs in Neural Systems

Yuhan Chen[1,2], Shengjun Wang[2,3], Claus C. Hilgetag[4,5,6], Changsong Zhou[1,2,7,8]*

1 Department of Physics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China, 2 Centre for Nonlinear Studies, and Beijing-Hong Kong-Singapore Joint Centre for Nonlinear and Complex Systems (Hong Kong), Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China, 3 Department of Physics, Shaanxi Normal University, Xi'An, Shaanxi Province, China, 4 Department of Computational Neuroscience, University Medical Center Eppendorf, Hamburg, Germany, 5 School of Engineering and Science, Jacobs University Bremen, Bremen, Germany, 6 Department of Health Sciences, Boston University, Boston, Massachusetts, United States of America, 7 Beijing Computational Science Research Center, Beijing, People's Republic of China, 8 Research Centre, HKBU Institute of Research and Continuing Education, Virtual University Park Building, South Area Hi-tech Industrial Park, Shenzhen, China

Abstract

The formation of the complex network architecture of neural systems is subject to multiple structural and functional constraints. Two obvious but apparently contradictory constraints are low wiring cost and high processing efficiency, characterized by short overall wiring length and a small average number of processing steps, respectively. Growing evidence shows that neural networks are results from a trade-off between physical cost and functional value of the topology. However, the relationship between these competing constraints and complex topology is not well understood quantitatively. We explored this relationship systematically by reconstructing two known neural networks, Macaque cortical connectivity and *C. elegans* neuronal connections, from combinatory optimization of wiring cost and processing efficiency constraints, using a control parameter α, and comparing the reconstructed networks to the real networks. We found that in both neural systems, the reconstructed networks derived from the two constraints can reveal some important relations between the spatial layout of nodes and the topological connectivity, and match several properties of the real networks. The reconstructed and real networks had a similar modular organization in a broad range of α, resulting from spatial clustering of network nodes. Hubs emerged due to the competition of the two constraints, and their positions were close to, and partly coincided, with the real hubs in a range of α values. The degree of nodes was correlated with the density of nodes in their spatial neighborhood in both reconstructed and real networks. Generally, the rebuilt network matched a significant portion of real links, especially short-distant ones. These findings provide clear evidence to support the hypothesis of trade-off between multiple constraints on brain networks. The two constraints of wiring cost and processing efficiency, however, cannot explain all salient features in the real networks. The discrepancy suggests that there are further relevant factors that are not yet captured here.

Introduction

It is widely appreciated that complex neuroanatomical networks are the physiological basis for brain dynamics, information processing and mental function [1–5]. In the last years, understanding the organization of neural systems and structural-functional relationship using graph theoretical approaches and methods of complex network research has attracted a great deal of interest [6–16]. Neural systems possess several pronounced network features, including the small-world property [17–19] characterized by short path lengths and high clustering [7,12,13], hub nodes with much larger degree than the average node degree of the network [20–22] and network modules broadly coinciding with functional subdivisions of the systems [14,23–27]. While these

features of neural networks are similar to those in many other real-world complex networks [28–30], the mechanisms underlying the formation of such complex network organization are still poorly understood.

The simultaneous existence of modules and hubs is an ubiquitous mesoscopic structural property in neural networks, and may play a significant role in the information processing and functioning of the systems. It was shown that cortical brain connectivity comprises dense communities, which are more densely linked internally than externally [24,25,27]. Such a modular organization was observed in structural networks obtained from tract-tracing studies [14], or diffusion spectrum imaging tractography [21], across various species, such as human [21,31–33], cat [34], rhesus Macaque monkey [35] and *C. elegans*

Author Summary

What are essential relationships between fundamental physical constraints and the architecture of neural systems? Most existing investigations have considered a single constraint, either wiring cost or processing path efficiency, and little is known about how characteristic neural network features, such as the simultaneous existence of modules and hubs, are related to the constraints from multiple requirements. Here we emphasized the competition between the global wiring cost and an important functional requirement, path efficiency, as factors in forming Macaque cortical connectivity and *C. elegans* neuronal connections. By comparing real to reconstructed networks using optimization under multiple constraints, we found that several network features are related to the competition of these two constraints, in particular the simultaneous formation of network modules and hubs. However, not all the properties of the real networks could be attributed to these two constraints, suggesting that, likely, there exists additional structural or functional requirements.

[36] as well as in functional networks derived from EEG/MEG, fMRI and MEA experiments [31,37–39]. The modules of the cortical network, in broad agreement with functional subdivisions of the cerebral cortex, are spatially segregated — as the areas within the same functional subsystems (visual, auditory etc.) are mostly spatially co-localized [22,36,40,41]. A recent study [42] showed that that modularity similar to that in human functional networks can be obtained based on an objective function combining the number of common nearest neighbors with a power-law decay of connections over distance, implying that modularity may be closely related to local connections.

Highly connected hub nodes have been shown in the structural network of the human brain, constructed by diffusion tensor approaches and based on 70–90 cortical gray matter areas [43–45] or in cortical networks based on 998 region-of-interest [21], as well as in functional networks using fMRI [46] or other imaging techniques (EEG, MEG, MEA) [39], also in other species, such as Macaque cortex [20] and *C. elegans* [47]. Hubs could effectively integrate information that is segregated due to the existence of modules [22]. In agreement with this idea, the identified high-degree hubs were mostly multimodal association regions [16,22]. Moreover, in a recent study of brain pathology, a MEG study of connectivity provided additional evidence that a degradation of the small-world property in patients with Alzheimer's disease was due to disease-related changes in hubs [48].

What factors influence the formation of modules and hubs is still an open question. Nonetheless, the organization of neural networks is frequently attributed to fundamental constraints, such as metabolic cost [49], signal propagation efficiency [50], evolutionary history [51] and others. It has been speculated that the network organization is the result of an economical trade-off between the physical cost and the functional values of the topology [52]. But it is still not well understood what these functional constraints are and what the relationship is between network properties and functional values, in spite of intensive research on complex brain networks over the last decade.

One of the most extensively discussed aspects is the constraint of wiring cost [53–68], which is related to the possible minimization of neural wiring volume [54,62] or wiring length [55,58,59,61,63,65,69] in the nervous system. Most previous studies investigated whether the actual component placement

layout of neural systems has been optimized for wiring minimization, by comparing the actual wiring cost to the perturbed component placement while keeping the network connectivity as in the real systems [55,59,63–65,69]. In coarse-grained data sets, it was found that the wiring length of the Macaque prefrontal areas [59] and *C. elegans* ganglia is optimized [55]. However, in other networks, such as those linking Macaque or cat visual cortical areas, the wiring was found not to be fully optimized, but relatively more optimal than other subsets of the cortex [64]. Moreover the wiring of the whole neural network of the Macaque cortex and *C. elegans* neuronal network appeared not to be optimized under the single wiring cost constraint– the total wiring could be decreased to 64% of the original length in Macaque and to 52% in *C. elegans* [63] when applying the component placement optimization (CPO) to minimize the total wiring length while preserving the specific network connectivity. Alternatively, it has been suggested that constraints such as signal propagation efficiency, measured by the global minimization of processing steps across the network, may shape the organization of neural systems [63,68,70,71].

Generally none of the single constraints is likely able to account for all the functional values of the network. However, the processing efficiency perhaps is the most established network measure shown to correlate with functional performance in normal subjects and dysfunction in various brain diseases. As large-scale information processing and communication systems, neural networks favor reducing the number of intermediate transmission steps in order to respond quickly, with the tendency to minimize the average shortest path length (i.e., graph distance) [50,72]. Minimizing the graph distance has some important functional advantages. First, a small number of intermediate transmission steps might reduce energy consumption during signal processing. In fact, about 50% of the brain's energy is used to drive signal processing, suggested by halved brain energy consumption in deep anesthesia that blocks neural signaling [73]. Second, minimizing the graph distance would increase the speed of signal processing, ultimately leading to faster behavior for decisions and actions [50]. Third, minimizing the graph distance enables neighboring and distant brain areas (or neurons) to receive signals nearly simultaneously to allow synchronous functional processing [74–77]. Fourth, since there exists abundant noise in signal transmission (e.g., ionic channel noise or synaptic noise) [50,78] as well as a high failure rate for signal transmission (between 50% and 90% in individual synapses [50]), minimizing the path length would limit the interference by noise and increase the robustness of neural systems [72]. Indeed, the global efficiency (the inverse of the average shortest paths) of the resting-state brain network has been found to be strongly associated with the intelligence quotient(IQ) [79–82]. In disease, it was found that the efficiency of the human cortical network was disrupted in a manner proportional to the extent of white matter lesions [83–87]. Therefore, the graph distance/processing efficiency could be taken as a representative measure conferring functional value of brain networks.

To date, however, most studies either considered the influence of just a single constraint (mostly wiring minimization) [53–55,57–62,64–66], or evaluated two constraints (such as the metabolic cost constraint and the propagation efficiency constraint) separately [63]. However, this approach does not mean that the two constraints are independent. In fact, they may have partly opposite impact on network organization. The processing efficiency constraint favors network shortcuts that link topologically distant parts of the network, which may take the form of long-distance connections, in which case they would act against the metabolic cost constraint of wiring length minimization. Conversely, wiring

minimization favors the creation of links among spatially adjacent network nodes, which may also be topological neighbors (cf. Fig. S3 in [63]). Such networks with mainly local connections typically possess low path efficiency due to a large average number of processing steps. In conclusion, the two constraints need to be considered in combination. The coexistence of modular organization and hubs in networks could be, at least partially, understood by a balance of these two constraints. Indeed, there is growing evidence to support the idea that neural network connectivity is not optimized either to minimize connection costs or to maximize advantageous topological properties, but rather is an economical trade-off between the physical cost and the adaptive value of its topology [88], see [52] for a recent extensive review. It has been proposed that the cost-efficiency balance of the human functional network [89] may be related to the behavioral performance in cognitive tasks [79], but the anatomical mechanism underlying such desirable functional connectivity is not clear.

In the present work we explored the relationship between multiple constraints and complex network architecture by systematically testing the effect of the competition of multiple constraints. We considered the neuronal network of *C. elegans* [63,65,90,91] and the cortico-cortical network of the non-human primate (Macaque) brain based on tract-tracing studies [63,92], for which information is available for both the spatial positions of the nodes and the network connectivity. In a previous analysis of the Macaque cortical network [27,63], the division of the motor areas was not very highly resolved, which feature might induce biased results when analyzing network modules and spatial clustering. Therefore, in the present work we first improved the data set by a more detailed division of the motor areas based on the CoCoMac database [92], extending the former 6 motor areas to 15 areas with an additional 128 connections (see *Materials and Method* and Fig. S1 for the adjacency matrix). Different from the CPO method, in our scheme we compared the real network connectivity to reconstructed networks derived from multiple constraints by fixing the spatial position of each network node and the total number of (directed) connections as in the real networks. The reconstructed networks were obtained under various balancing conditions of wiring cost and processing efficiency constraints. As in previous studies [53–55,57–66], we used the total physical distance of the wiring L_p to represent the effect of the wiring cost constraint, and the total graph distance of the shortest paths L_g to represent the influence of the processing efficiency constraint, and defined an objective function L as a combination of both constraints using a weight parameter α, namely, $L = (1-\alpha)L_g + \alpha L_p$, with L_g and L_p appropriately normalized. So $\alpha = 0$, or $\alpha = 1$ corresponds to a single constraint of path efficiency or wiring cost, respectively. Then we reconstructed the connections of the network with the help of a simulated annealing approach to minimize the objective function L at different values of α, starting from 50 random configurations (see *Materials and Methods*). We studied the general properties of the competition between the two constraints in a 1D model with one-dimensional uniform spatial layout of nodes and directed connections. For the real neural networks, we investigated the modularity and hub properties of the reconstructed networks and studied the relationship to the spatial layout and compared them to those of the real networks. We found that for certain intervals of balancing these two constraints, the reconstructed networks showed a very similar modular structure and similar spatial positions of the hubs as the real networks. These results are also related to the nonuniform layout and clustering of the network nodes (neurons for *C. elegans* and areas for Macaque cortical brain) in space. Despite the observed agreements, there still exists significant discrepancies between model and real networks,

suggesting that there are additional functional requirements to be considered in the future.

Results

Competition between cost and efficiency constraints leads to the coexistence of hubs and local connections

The qualitative properties of the competition between the constraints at different values of α were found to be quite common for the 1D model (Fig. 1), the real Macaque cortical network (Fig. 2) and *C. elegans* neuronal network (Fig. 3), the latter two having highly non-uniform spatial layout of nodes. These properties can be most clearly seen in the 1D model. In addition to L_p and L_g, we used several parameters such as the number of hubs N_{hub}, the average degree of the hubs $\langle K_{hub} \rangle$ and the fraction of spatially local connections $P_{nearest}$ to characterize the reconstructed optimal networks at various α values (see *Materials and Methods*). The results are summarized below.

At $\alpha = 0$, the network is only optimized by the processing efficiency constraint and achieves minimization of the topological distance. The optimal network configuration depends on the density of connections. If the initial network is sparse (roughly, connection probability $p < 1/\sqrt{N}$), as in the 1D model ($p = 5\% < 1/\sqrt{N} = 7\%$, Fig. 1A) or *C. elegans* neuronal network ($p = 3.82\% < 1/\sqrt{N} = 5.8\%$, Fig. 3C)), the reconstructed network is characterized by a single hub connecting to all other nodes while the rest of connections appears to be random. The position of the hub node is arbitrary. A single hub is very effective for reducing the graph distance when the network is sparse enough, because the path length between any pair of nodes is either 1 (direct link) or 2 (connected through the hub). When the network becomes denser ($p > 1/\sqrt{N}$), such as the Macaque cortical network ($p = 23.7\% > 1/\sqrt{N} = 10.2\%$), the pathlength between any pair of nodes in a random network cannot be larger than 2. It is then very unlikely to obtain a single hub when optimizing from an initially random configuration. Thus, the reconstructed network connections appear to be random (Fig. 2C). The dependence of the network configuration on the connection density is also more systematically shown in the 1D model (Fig. S2).

At $\alpha = 1$, the processing efficiency constraint plays no role while the wiring cost constraint is fully dominant, and the network has the minimal total wiring length. Most of the connections are local and there is no hub ($N_{hub} = 0$) in any of the three systems (Figs. 1A, 2F, 3F). For the 1D model or the reconstructed networks of the Macaque, all the connections are local ($P_{nearest} = 1.0$, Fig. 4J,K); L_p is reduced to 53.5% of that in the real Macaque network (Fig. 4B). As for the *C. elegans* network, many neurons in the head and tail are very densely distributed with very small distance among them, while the distance between neighboring neurons in the ventral cord is much larger. Therefore, significant numbers of non-nearest but short-distant connections within the head or the tail let the physical distance L_p become very small, only 1.25% of the real *C. elegans* (Fig. 4C), but $P_{nearest}$ is clearly smaller than 1.0 (Fig. 4L)). And in the real networks, the spatial layout of the network nodes is non-uniform, forming spatial clusters; these local connections make the adjacency matrix become non-uniform, showing some clustered pattern similar to the original network connectivity (Figs. 2F, 3F). Spatial clustering and module organization in the reconstructed and real networks are studied below.

For $0 < \alpha < 1$, the processing efficiency and wiring cost constraints combine their impact, resulting in two distinct regimes, depending on the networks. When α is positive, but not very large,
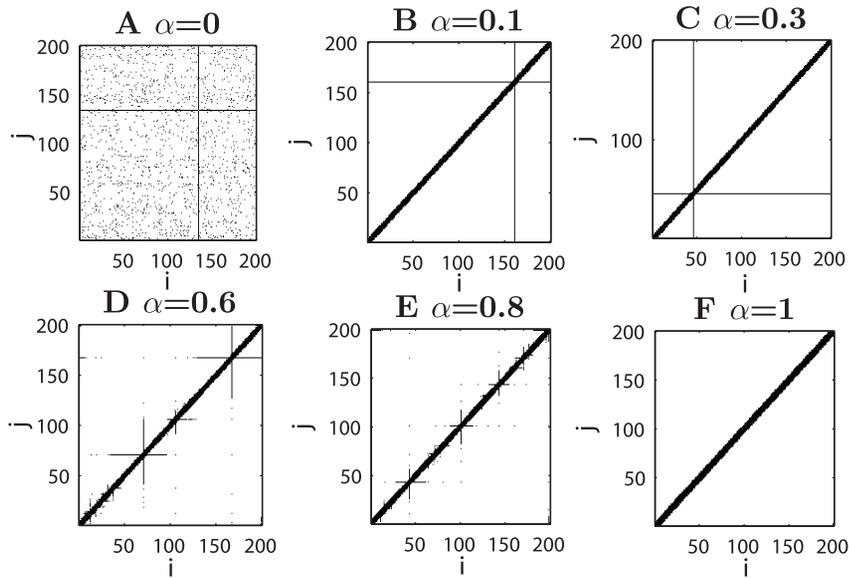
Figure 1. Constructed networks in a 1D model. There are 200 nodes uniformly placed on a one-dimensional circle linked by a total of 2000 directed connections. Shown are the adjacency matrices obtained at various α values, as indicated on the top of the plots. The nodes are indexed by their locations on the circle, common for all panels. (A) $\alpha=0$, (B) $\alpha=0.1$, (C) $\alpha=0.3$, (D) $\alpha=0.6$, (E) $\alpha=0.8$ and (F) $\alpha=1$.
doi:10.1371/journal.pcbi.1002937.g001

such as $\alpha=0\sim0.3$ for the 1D model (Fig. 1B,C), $\alpha=0\sim0.8$ for Macaque (Fig. 2D) and $\alpha=0\sim0.3$ for *C. elegans* (Fig. 3D), there is a single hub connecting to all the other nodes in all the three systems ($k_{in}=k_{out}=N-1$). Most of the other connections are linked to the nearest neighbors, and $P_{nearest}$ is close to 1.0 in the 1 D model (Fig. 4J) and the reconstructed networks of Macaque (Fig. 4K), but is clearly smaller in *C. elegans* for the reason stated above

($P_{nearest}\sim0.65$, Fig. 4L). This single, global hub is a very effective configuration to provide high efficiency, when the other connections are short-distance due to the cost constraint. While the placement of the hub is arbitrary in the 1D model due to a symmetrical spatial layout (Fig. 1B,C), it is unique in the reconstructed networks of Macaque and *C. elegans*, located close to the global geographical center of the whole network, as will be
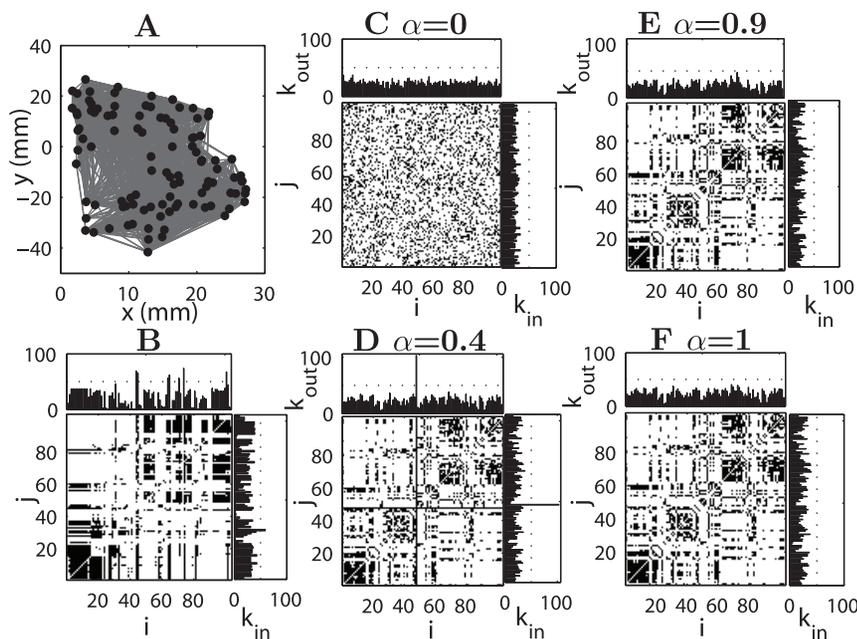
Figure 2. Comparison of reconstructed and original connectivity of Macaque cortical network. The left two plots (A and B) are for the original network. (A) Layout placement of 103 areas and connections between them. (B) Adjacency matrix, the output (k_{out}) and input (k_{in}) degrees of the areas. The right four plots (C–F) show adjacency matrices and the degrees of areas in the reconstructed networks at various values of α. (C) $\alpha=0$, (D) $\alpha=0.4$, (E) $\alpha=0.9$ and (F) $\alpha=1$. The index of the cortical areas is the same for (C–F) and the names of the areas are listed in Table S1 of Supporting Information (SI).
doi:10.1371/journal.pcbi.1002937.g002
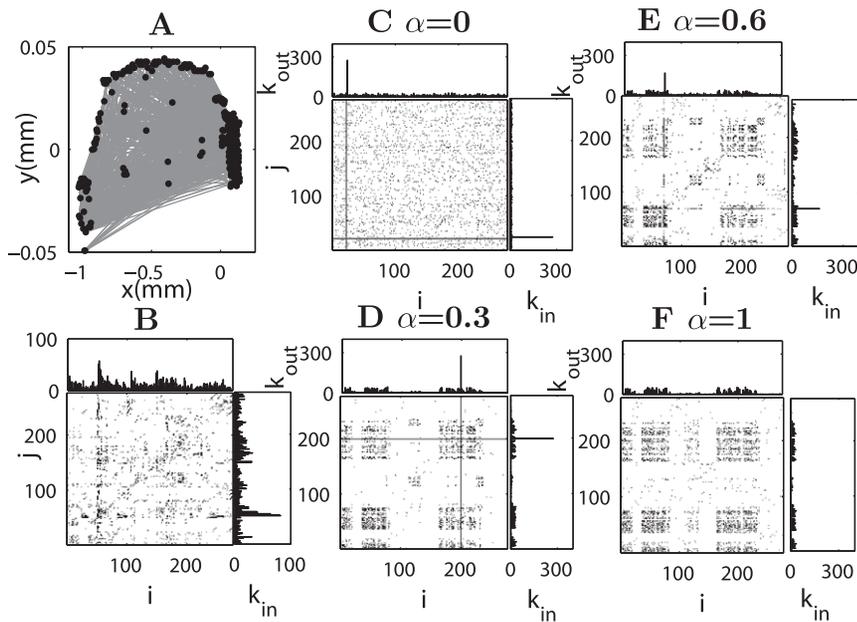
Figure 3. Comparison of reconstructed and original connectivity of *C. elegans* neuronal network. The left two plots (A and B) are for the original network. (A) Layout placement of 276 neurons and connections between them. (B) Adjacency matrix and the output (k_{out}) and input (k_{in}) degrees of the areas. The right four plots (C–F) show adjacency matrices and the degrees of areas in the reconstructed networks at various values of α. (C) $\alpha=0$, (D) $\alpha=0.3$, (E) $\alpha=0.6$ and (F) $\alpha=1$. The index of the neurons is the same for (C–F) and the names of the neurons are listed in Table S2 of SI.
doi:10.1371/journal.pcbi.1002937.g003
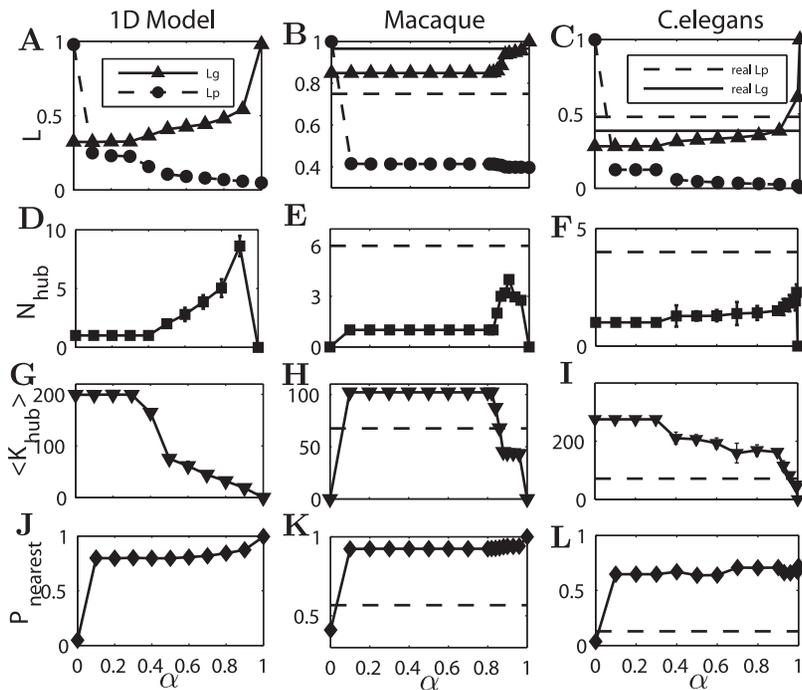
Figure 4. Properties of the reconstructed networks as functions of α. Results are shown for the 1D model (left panel), Macaque cortical network (middle panel) and the *C. elegans* neuronal network (right panel). (A, B, C) L_g and L_p (normalized) of the reconstructed networks (with triangles or dots respectively), which are compared to those in the original networks in B and C. (D, E, F) N_{hub}, the number of hubs in the reconstructed networks vs. α. Here a node with z-score of total degrees (input and output) larger than 2 is considered as a hub. (G, H, I) The average degree of the hubs in the reconstructed networks. (J, K, L) The probability $P_{nearest}$ of connections to the spatial nearest-neighbors. The results are obtained by averaging over 50 realizations of the reconstructed networks for each α. In most cases the error bars are smaller than the symbol size. The dashed lines are the corresponding results for the real network.
doi:10.1371/journal.pcbi.1002937.g004

discussed in detail below. In this regime, L_p, L_g, N_{hub} and $P_{nearest}$ are all constant with respect to α, since the optimal network configuration does not change with α, though the speed of convergence in the optimization process does. While both L_g and L_p in the reconstructed networks are smaller than those in the real networks in this region, L_p is much more smaller (reduced to 55.2% in Macaque and to 25.8% in *C. elegans*, see Fig. 4B,C).

The insensitivity of the optimal configuration to α in this regime can be understood from the objective function $L = (1-\alpha)L_g + \alpha L_p$. The variation is $\Delta L = (1-\alpha)\Delta L_g + \alpha \Delta L_p$, where $\Delta L_g > 0$ and $\Delta L_p < 0$ with increasing α. The network configuration would change with α only if $\Delta L < 0$, i.e., $|\frac{\Delta l_g}{\Delta l_p}| < \frac{\alpha}{1-\alpha}$. However, for the Macaque network, we found that $|\frac{\Delta l_g}{\Delta l_p}| > 4$ for any perturbation to the configuration obtained at $\alpha = 0.1$ (a single global hub and all other local links). When $\alpha < 0.8$, $\frac{\alpha}{1-\alpha} < 4$, thus $|\frac{\Delta l_g}{\Delta l_p}| > \frac{\alpha}{1-\alpha}$, then $\Delta L > 0$, and the configuration change cannot be accepted by the optimization. Only when $\frac{\alpha}{1-\alpha} > 4$ ($\alpha > 0.8$), it becomes possible to have $|\frac{\Delta l_g}{\Delta l_p}| < \frac{\alpha}{1-\alpha}$ for certain perturbations, and a configuration change will happen. For this reason, in all the three networks, the optimal solution is the same in a range of α values, depending on the spatial layout of the network.

With further increase of α, the influence of the wiring cost constraint becomes stronger. A single hub is no longer found, because very long-distant connections are prohibited and the systems move into a different regime ($0.4 < \alpha < 0.9$ for the 1D model, $0.8 < \alpha < 0.96$ for Macaque and for *C. elegans* $0.3 < \alpha < 0.99$). Several smaller hubs emerge, with connections extending to nodes in their spatial neighborhood, and the connection range of such regional hubs becomes smaller at larger α, as can be most clearly seen in the 1D model (Figs. 1D,E), and is also true in the real networks (Figs. 2E and 3E). Consequently, L_p is further reduced slightly, but L_g increases and is close to that of the real network (Fig. 4B,C). As shown below, the spatial positions of the hubs in the reconstructed networks are close to the real hubs in the two original networks.

When α is very close to 1.0, where the efficiency constraint is weak while the wiring cost constraint is almost fully dominant, most of the connections are local and there are no pronounced hubs ($N_{hub} = 0$). L_g increases quickly and becomes larger than that in real networks.

We also found that for all α values, the input and output degrees of the nodes in the reconstructed networks are largely symmetrical (Fig. S3A and Fig. S3C). While the input and output degrees were found to be significantly correlated in the real networks (Fig. S3B and Fig. S3D), the discrepancy between the optimization model and real data is quite large, because this model does not include possible requirements that could generate the asymmetry, for instance, input-output information flowing as in real networks (e.g., from sensory neurons to motor neurons in *C. elegans*).

The above results show that the coexistence of local connections and hubs in the cortical networks could be a solution to the multiple constraints of wiring cost and processing efficiency. There is a regime ($\alpha = 0.8 \sim 0.96$ for Macaque and $\alpha = 0.3 \sim 0.99$ for *C. elegans*) in which the competition of the constraints can allow the formation of several hubs connecting to many of the nodes in the neighborhood. Here, L_g in the reconstructed networks is very close that of the real networks, but L_p can be much smaller. In the

following sections, we show that both the modular structure and the positions of hubs are quite similar to the real networks in this regime.

Modules in network connectivity and spatial clustering of nodes

The above results (Figs. 2D–F and 3D–F) showed the emergence of a modular organization in the reconstructed networks similar to the real ones, which is derived from the non-uniform spatial distribution of the nodes and the local connections due to the constraints. In order to further explore these relations, we examined the relation between the spatial clustering of nodes and the modules in network connectivity both in the real and reconstructed networks, and compared the similarity between the modules in the real and reconstructed networks.

Macaque cortical network. In the the real Macaque cortical network, the nodes (which represent the centers of mass of the cortical areas) are distributed non-uniformly in the three-dimensional space due to size variations of the areas. Visual inspection of a two-dimensional layout in Fig. 5A suggested that the nodes could be divided into two spatially separated clusters. This was confirmed by clustering analysis (see *Materials and Methods*), shown by different colors (blue and red) of the nodes in Fig. 5A. Such spatial clustering was found to be quite robust in data sets with different parcellations of the areas. The clustering of this more highly resolved data set with finer division of the motor areas ($N = 176$) was compared with a parcellation of $N = 103$ areas [93] (see *Materials and Methods*) and was found to be very similar. The clustering boundaries of the two data sets differed only by 5 areas (referring to the $N = 176$ dataset) (see Fig. S4 and S5). Unfortunately, network connectivity is not available for this finest parcellation of $N = 176$ areas. We also examined the spatial density of the nodes within a radius around a given node, which we called the *neighborhood density* (see *Materials and Methods*), and found that it was also quite non-uniform due to heterogeneous area size. As shown in the following sections, many properties of the reconstructed and original networks of the Macaque cortex are related to these properties of the spatial layout of the network nodes.

The connectivity in the original Macaque cortical network shown in Fig. 5A also suggests that the connections are denser within spatial clusters than between clusters, indicating that the spatial clustering of the areas and densely connected network modules are related to each other. We analyzed modules in the original Macaque cortical network and the reconstructed networks (see *Materials and Methods*). The optimal modularity is $Q = 0.395$ in the original Macaque network (Fig. 5A), corresponding to two modules shown by open and filled circles in Fig. 5B. The modularity did not show uncertainty with respect to the modular boundaries and is quite significant when compared to the corresponding random networks with the same number of nodes and connections ($Q = 0.08 \pm 0.004$). We can see that the modules and spatial clusters have strong overlap, and only 14.6% of the total nodes are mismatched. An example of a reconstructed network ($\alpha = 0.9$) is shown in Fig. 5C, which has much fewer connections between the two spatial clusters than the real network in Fig. 5A, therefore has a larger optimal modularity $Q = 0.50$ (also two modules, without uncertainty). The two modules from the reconstructed networks overlap almost completely with the two spatial clusters, respectively (Fig. 5D), with only one mismatched areas (0.97%). We examined the mismatched areas between the module partitions of the original and the reconstructed network at various $\alpha > 0$ and found that they were consistent. The number of the mismatched areas remains small (around 15% of all the areas,
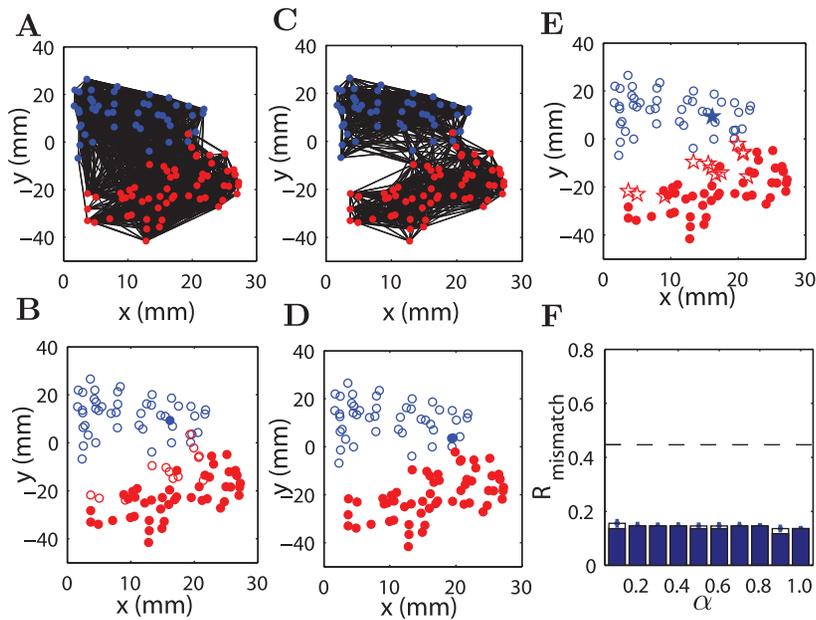
Figure 5. Modularity of original Macaque network and reconstructed networks. (A) Layout placement of 103 areas and the connections of the Macaque cortical network. (B) The two modules of the real network (open and filled circles) are compared to the two spatial clusters (blue and red). The corresponding modularity is $Q=0.395$. (C) As in (A), but for a reconstructed network at $\alpha=0.9$. The blue and red colors of the nodes represent the two spatial clusters. (D) The same as (B), but the modules are from the reconstructed network in (C), with $Q=0.50$. (E) Mismatch between the module partitions of the reconstructed and real networks. The mismatching areas are indicated by the pentagrams. (F) $R_{mismatch}$ shows the fraction of mismatching areas between reconstructed and original networks in module partition, with respect to α. The result did not include $\alpha=0$ where the reconstructed networks did not show strong modularity. Blue bars are for the mismatching areas appearing in more than 60% of all 50 realizations at each α, while white bars are for the mismatching areas appearing at least once in the 50 realizations at the given α. The dashed line represents the mismatching rate between the real networks and shuffled modules (see section "Matching between partitions" in *Materials and Methods*).
doi:10.1371/journal.pcbi.1002937.g005

much smaller than the mismatching rate for real networks and shuffled modules (see *Materials and Methods*) for all $\alpha>0$, but is clearly smaller for $\alpha=0.8\sim0.9$ (blue bars in Fig. 5F). The modularity Q of the reconstructed networks was larger than that in the real network for $\alpha>0$ (see supplementary Fig. S6A). Note when $\alpha=0$, the reconstructed solution under the single efficiency constraint is a random-like network without clear modular structure (modularity Q$=0.096\pm0.005$).

C. elegans neuronal network. The spatial layout and connectivity of the real C. elegans neural network are shown in Fig. 6A. There are four network modules, with modularity $Q=0.423$ (Fig. 6B, different symbols). Again, there was no uncertainty in the module partition, and the modularity is significant compared to the corresponding random networks ($Q=0.175\pm0.003$). Using the spatial clustering method, the neurons were grouped into three clusters (shown by different colors). Interestingly, the relationship between connectivity modules and spatial clusters is similar to the Macaque cortical network. While not perfect, there exists some overlap between modules and spatial clusters (Fig. 6B), about 30.8% of the neurons are mismatched. For the reconstructed network of *C. elegans*, quite similar to the Macaque cortical network, there are fewer connections between different spatial clusters (Fig. 6C, $\alpha=0.6$) when compared to the real network, consistent with a larger modularity $Q=0.560\pm0.012$ for four modules. The overlap between the modules and the spatial clusters is stronger than in the real network (Fig. 6D), and the ratio of mismatched neurons is reduced to 11.96%. The number of mismatched neurons between the modular partitions of the original and the reconstructed networks remains small (about 28% of all neurons) for $0<\alpha<1$,

approaching the smallest value (nearly 23%) when $\alpha=0.8\sim0.9$, but increasing clearly for α very close to 1.0 ($\alpha=0.995$) (blue bars in Fig. 6F). Again the mismatching rate is much smaller than that between the real network and shuffled modules. The modularity Q of the reconstructed networks is larger than that of the real network for $\alpha>0$ (see supplementary Fig. S6B). Note that for the singular constraint of path efficiency at $\alpha=0$ or of wiring cost at $\alpha=1$, the reconstructed networks do not exhibit evident modularity ($Q=0.1651\pm0.001$ and $Q=0.154\pm0.009$, respectively).

We would like to emphasize that the results reported above have already taken the uncertainty of module partition and spatial clustering into consideration (see *Materials and Methods*). First, the uncertainty in the module partition is quite small. The Macaque cortex did not show degeneracy in module partition when (the toolbox of brain connectivity) BCT method was applied for many times, although degeneracy was indeed observed in a corresponding random network with the same number of nodes and connections ($Q=0.08\pm0.00004$, here the error-bar is from 50 applications of the BCT method). The real network in *C. elegans* did not have degeneracy either, but there was a small uncertainty in the reconstructed networks (with a typical fluctuation of 1–5 neurons from run to run). Second, we used 8 different distance measures to detect the hierarchical clustering. When comparing the clustering partition from these 8 methods to the modular division, the partition based on inner squared distance generated the minimal set that matched the module partition. The rest of the nodes formed the largest possible mismatched group that covered all the mismatched groups from different clustering methods (see *Materials and Methods*). In particular, there were 89 cortical areas in the Macaque which match between the modules and the clusters
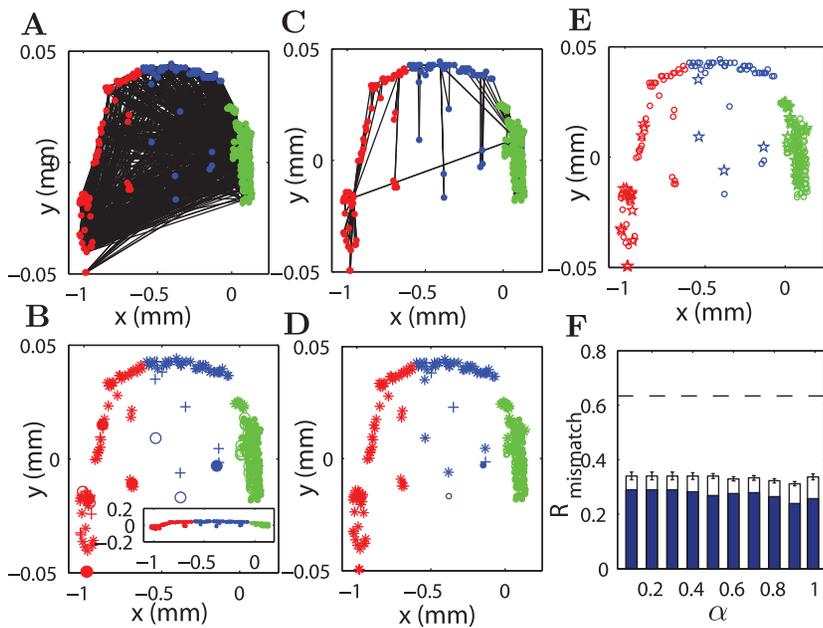
Figure 6. Modularity of original *C. elegans* neuronal network and reconstructed networks. (A) Layout placement of 276 neurons and the connections of the network. (B) The four modules of the real network (open, filled circles, plus and asterisk) are compared to the three spatial clusters (blue, green and red). The corresponding modularity is $Q = 0.423$. (C) As in (A), but for a reconstructed network at $\alpha = 0.6$. The red, blue, and green colors of the nodes represent the three spatial clusters. In (A) and (C), we used different scales for the x and y axis for clear presentation. The connections among dense nodes cannot be seen. The inset of (B) shows the positions and clusters of neurons in identical scales. (D) The same as (B), but the modules are from the reconstructed network in (C), with $Q = 0.560 \pm 0.012$. (E) Mismatch between the module partition of the reconstructed and real networks. The mismatched areas are indicated by the pentagrams. (F) $R_{mismatch}$ shows the ratio of mismatched areas between reconstructed and original networks in module partition, with respect to α. The result did not include $\alpha = 0$ and $\alpha = 1$ where the reconstructed networks did not show strong modularity. Blue bars are for mismatched areas appearing in more than 60% of all 50 realizations, while white bars are for mismatched areas appearing at least once in the 50 realizations for a given α. The dashed line represents the mismatching rate between the real network and the shuffled modules.
doi:10.1371/journal.pcbi.1002937.g006

from all the 8 methods; the remaining set of 14 areas is thus the largest possible group covering all the unmatched groups across different methods of clustering and was reported in Fig. 5B (among these 14 unmatched areas, 5 areas appear to be the unmatched areas in all these 8 methods, 1 area in 7 methods, 6 areas in 5 methods, and the 2 areas in 3 methods). In the real network of *C. elegans*, this set of most stable nodes contains 191 neurons and the remaining 85 neurons are considered as the the largest group of mismatched neurons (Fig. 6B). In the same vein, the reported mismatching between the module partitions of real and reconstructed networks is the largest possible group when taking the uncertainty into consideration (Figs. 5E,F, and 6E,F).

These results show that the modular structure of the original networks is strongly related to the fact that the network nodes are spatially clustered, and many of the connections are contained within the clusters due to the effect of the wiring cost constraint. However, the best match of the modular partition with the real network involves the processing efficiency constraint ($\alpha < 1.0$) that enforces some long-range connections between the clusters.

Location of hubs

The results in Section I. showed that the emergence of hubs could be the result of the combination of the wiring cost and path efficiency constraints, since hubs connecting to many nodes are very effective for improving path efficiency, while allowing most of the other connections to remain local to satisfy the cost constraint. If the hub is a global one, connecting to all the nodes, then it is reasonable to expect that the position of the hub is not arbitrary, but is near the geographical center (the node with minimal total

distance to all the other nodes) of the whole network in order to maximally reduce the total wiring length of the connections from the hub. If the hub is a regional one, connecting to most of the nodes within the local region (e.g., within one of the spatial clusters), then the position should be close to the geographical center of this region. Thus, a node with many other nodes densely distributed in the neighborhood (having a high neighborhood density, see *Materials and Methods*), could be an candidate for a hub under the two constraints.

In the following section, we present findings regarding the location of the hubs in the reconstructed as well as the real networks. In our analysis, network nodes with z-score of total degrees (input and output) larger than 2 (see *Material and Methods*) are considered as hubs. The locations of the hubs in all 50 realizations of the reconstructed networks at each α were identified. We would like to point out that identifying hubs in this way is heuristic and may introduce some ambiguity when comparing different networks (real or reconstructed at different α). However, hubs that were defined in this way indeed provided a plausible way to describe the variation of the nodes with the largest degrees in the network as α changes.

Macaque cortical network. For the Macaque cortical network, there was only one global hub connecting to all the other areas in all optimally reconstructed networks for $0 < \alpha < 0.8$. Either the primary auditory area A1 (the area index No. 41) or the area Ri (No. 73) was selected as the single hub, with larger than 80% and smaller than 20% probabilities, respectively (Fig. 7A). When checking the global geographical centrality (see *Materials and Methods*), we found that A1 and Ri were ranked No. 1 and No. 2,
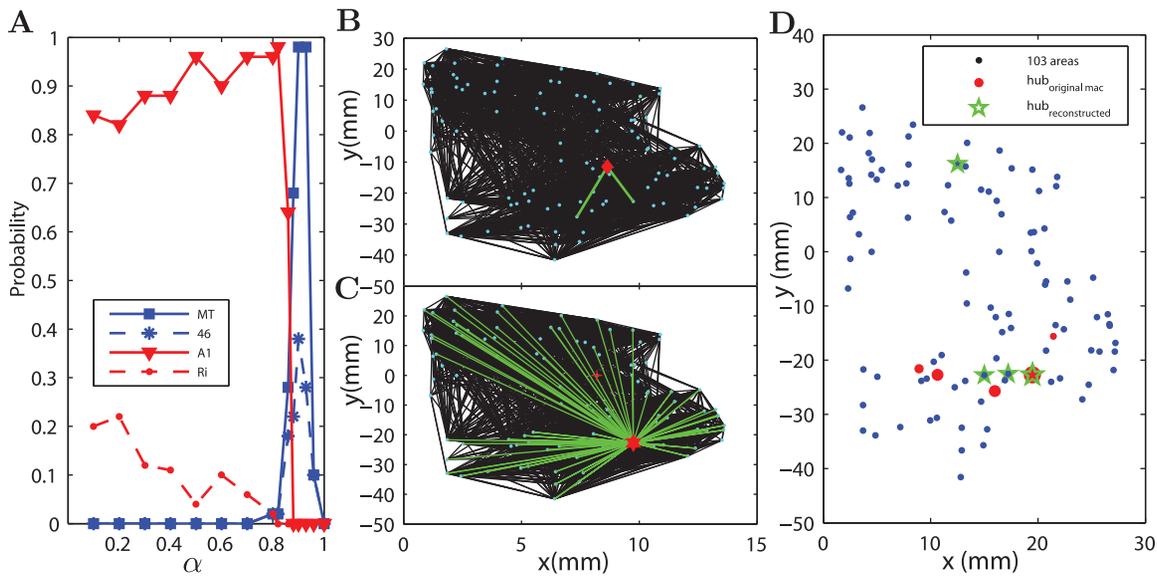
Figure 7. Hubs in the reconstructed and original Macaque networks. (A) Four areas A1, Ri and V5/MT and area 46 had the large probabilities to appear as one of the hubs in the reconstructed networks at different α. Note that V5/MT is also the biggest hub and area 46 is the biggest out-degree hub in the original network. (B) The position and connections (green lines) of the area A1 (global geographical center) in the real network. (C) The position and connections of the biggest hub V5/MT in the real network. (D) 6 hubs in the real network (red bullets, with the size of the symbol indicating the total degree) and 4 hubs (V5/MT, area 46, MSTm, MSTd) in the reconstructed network at $\alpha \sim 0.9$ (green stars). The positions of the reconstructed or real hubs coincide or are close to each other.
doi:10.1371/journal.pcbi.1002937.g007

respectively, which is consistent with our expectation. However, in the real network these two areas are non-hub nodes, with only 2 connections for A1 and 21 for Ri. The position and connections of A1 in the real network are shown in Fig. 7B.

When $0.8 < \alpha < 0.96$, the number of hubs increases while their average degree decreases, and the hubs become regional. Further increasing of α to be very close to $\alpha = 1$ prevents the emergence of hubs, because the impact of the wiring cost constraint becomes too strong relative to the influence of the processing efficiency constraint. For α around 0.9, four areas, V5/MT, area 46, MSTm and MSTd appeared as hubs in many realizations of the reconstructed networks (Fig. 7D, green stars). It is important to note that area V5/MT and area 46 are respectively the biggest hubs in the original network in terms of total degree and output degree (these two areas are also identified as hubs in [20]). V5/MT has a total of 112 input and output links, with 87 of them from the red spatial cluster (connecting to 87.9% of nodes in this cluster) and 25 from the blue cluster. The location of area V5/MT and its all connections in the real network are shown in Fig. 7 C. Although V5/MT is only ranked No. 27 in terms of the global geographical centrality, it is ranked No. 3 of the regional geographical centrality of the red cluster. It is also ranked No. 1 in terms of high neighborhood density in a range of radius (specifically, $r/r_{max} \in (0.24, 0.28)$, see *Material and Methods*). The biggest output hub area 46 has 59 efferent links in the real network. It is ranked No. 2 of the regional geographical centrality of the blue cluster and No. 3 by neighborhood density in a range of radius $r/r_{max} \in (0.24, 0.28)$.

In the reconstructed networks with α around 0.9, V5/MT and area 46 were respectively selected as hubs in 94% and about 40% of the realizations (Fig. 7A). The other two areas MSTm and MSTd were selected as hubs in the reconstructed networks, because they are respectively ranked No. 1 and No. 2 in terms of the geographical centrality of the red cluster and No. 2 and No. 4

by neighborhood density. But these two areas are not hubs in the real network.

Apart from the biggest hub area V5/MT, there are five more hubs (z-score above 2.0 either in total, input or output degree) in the real Macaque cortical network, areas 7a, VIPl, 7b, LIPv and area 46. All the hub areas are shown by the red bullets in Fig. 7D, with the size of the symbol indicating the total degree. Except for area 46 located close to the geographical center of the blue cluster, the other five hubs are all located within the red cluster and are among the top 18.5% in terms of the regional geographical centrality of this cluster. The ranking of a few nodes with the largest total degrees in terms of neighborhood density and geographical centrality is also shown in Fig. S7. The positions of the real hubs thus coincide or are close to those hubs in the reconstructed networks which are located in the regional geographical centers.

***C. elegans* neuronal network.** For the real *C. elegans* neuronal network, four neurons were identified as hubs (z-score above 2.0), which are AVAL(left)/AVAR(right) and AVBL/R ($k = 133, 137, 101, 104$), all located in the head (red bullets in Fig. 8). These 4 hubs are among the top 27.2% in terms of the regional geographical centrality of the head. Note that the neuron FLPL having the highest global geographical centrality of the whole system (red square Fig. 8) is also in the head, as most of the neurons (166 out of 276) are in the head.

In the reconstructed network, the global geographical center FLPL was indeed taken as the single hub connecting to all the other neurons for $\alpha \in (0, 0.3)$. With stronger impact of the wiring cost constraint for $0.3 < a < 0.99$, the number of hubs increases while their average degree decreases. In total, 14 neurons were chosen as hubs (with a probability above 30% to appear in the reconstructed networks at certain $\alpha = 0.3 \sim 0.99$), with degree ranging from 140 to 200, connecting 50.7% to 72.5% of the neurons. 10 of them were from 5 groups with left and right symmetry, ADFL/R, ASEL/R, ASHL/R, AWCL/R and
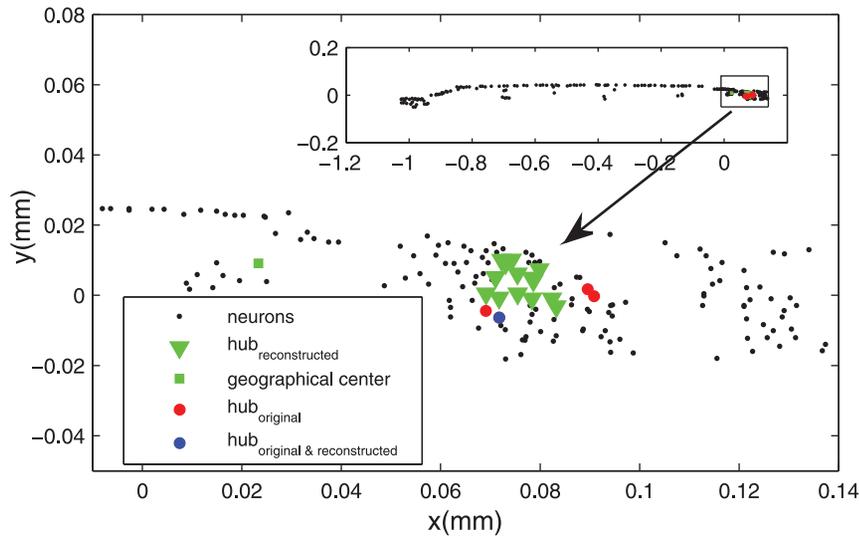
Figure 8. Hubs in the reconstructed and original *C. elegans* networks. The inset shows the two-dimensional layout of all 276 neurons of *C. elegans*. The square region in the inset is zoomed out, showing the four hubs in the real network (red and blue bullets), the global geographical center (green square) and the hubs in the reconstructed networks (green triangles) and the overlap with the real hub AVBR (blue bullets).
doi:10.1371/journal.pcbi.1002937.g008

RMHL/R, and the other 4 neurons were AIBR, RIBR, SIBDR and SIBVL. Interestingly, these hubs in the reconstructed networks (green triangles, Fig. 8) were close in location to the 4 real hubs (red bullets, Fig. 8). In fact, these 14 neurons are among the top 8.33% in terms of the regional geographical centrality in the head. Notably, we also found an overlap of reconstructed with real hubs. When $\alpha \sim 0.9-0.93$, the real hub AVBR has a relatively small probability (24%) to be chosen as the hub in the reconstructed network (marked as the blue bullet in Fig. 8).

One may wonder whether it is meaningful to say that the hubs were spatially close when most neurons are located in the tiny head of the worm. Therefore, a more quantitative comparison is desirable to support the visual inspection. We calculated the sum distance for each of the 166 neurons in the head to the 4 real hubs and sorted the neurons in the order of increasing distance (Fig.S8). It is clear that most of the neurons taken as hubs in the reconstructed networks are among the closest neurons to the real hubs.

While generally there is ambiguity in defining hubs in the networks, our observation is clear: with stronger influence from the wiring cost constraint, the reconstructed networks shift from a configuration with a single global hub to configurations having several nodes with relatively large degrees, and their spatial locations move from the global geographical center to the regional geographical centers accordingly. In real networks, the nodes with the largest degrees are close to the regional geographical centers, showing that both the processing efficiency and cost constraints are at work. This finding is strong evidence to support the hypothesis of a trade-off, though it appears there are more nodes with larger degrees in the real networks than the reconstructed networks.

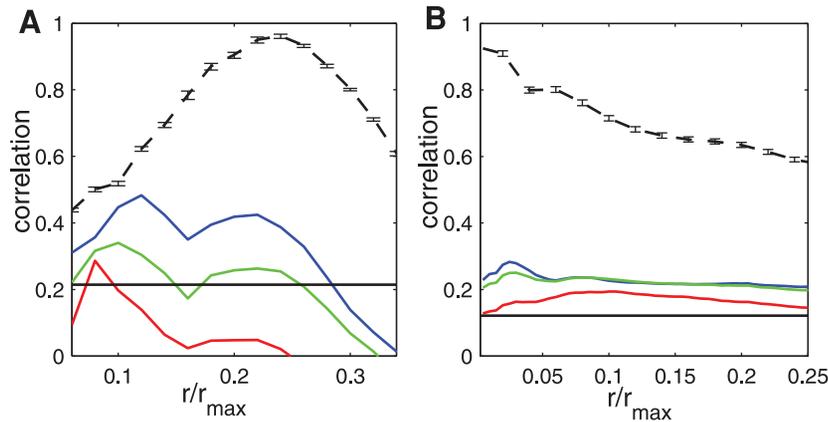## Degrees of nodes and correlation with neighborhood density of nodes

The results in the above sections suggested that the mesoscopic properties of networks, the simultaneous formation of modules and hubs, can be partially explained by the combination of the wiring cost and processing efficiency constraints. Now we examine the degree of nodes in the original and reconstructed networks.

When constructing networks under the wiring cost constraint, the nodes tend to connect to their nearest spatial neighbors, which is confirmed by a high value of $P_{neareast}$. Therefore, the number of connections of a node (degree) is expected to be related to the neighborhood density of the node in a certain spatial range. We calculated the density of nodes for various radii r and evaluated its correlation with the degree of nodes. The correlation between the degrees and density in reconstructed network with strong enough wiring cost constraint (α close to 1) is quite large in a range of r for both the Macaque and *C. elegans*, as shown by the black dashed curve in Fig. 9A,B.

Interestingly, the correlation between degrees and density is also significantly present in the real network (colored curves in Fig. 9A,B). For example, the correlation can reach 0.49 at $r/r_{max}=0.14$ for Macaque and 0.29 at $r/r_{max}=0.04$ for *C. elegans*, much larger than the 95% of the significance level in the corresponding surrogate data. The results are consistent with the observation above that hubs in the real network are ranked high in term of the neighborhood density.

Although the correlation between degree and neighborhood density is significant in both real and reconstructed networks, the discrepancy is also quite large. In particular, the degree distribution in the real network was not well reproduced in the reconstructed network (see Fig. S9 for a comparison). In both neural systems, the real networks have higher probabilities to have large-degree nodes, but the reconstructed networks have higher probabilities to generate intermediate-degree nodes, because the connections in the reconstructed networks are much more strongly determined by the neighborhood density. It has been shown that both the in- and out- degree of the *C. elegans* neuronal network obey a power-law distribution [47]. Our test of the significance of the power-law fitting to the distribution (see *Materials and Methods*) confirmed this statement ($p=0.807$, $\gamma=3.5$, $x_{min}=14$ for in-degree and $p=0.675$, $\gamma=3.2$, $x_{min}=13$ for the out-degree, Fig. S9). The real Macaque cortical network and the reconstructed networks of both systems do not show scale-free features. For all

Figure 9. Degrees of nodes in the reconstructed and real networks. (A) and (B) Correlation between degrees and neighborhood density vs. the normalized radius r/r_{max} for reconstructed ($\alpha = 0.9$) and real networks. (A) for Macaque and (B) for *C. elegans*. The results differ for the output (blue line), input (red line) and total degree (green line) in the real network, but are almost the same in the reconstructed network (black dashed line). The horizontal black line represents 95% of significance level. The error-bar is from 50 realizations of the reconstructed networks.
doi:10.1371/journal.pcbi.1002937.g009

these networks, the largest p value is 0.02 ± 0.02 for the out-degree distribution of *C. elegans* obtained at $\alpha = 0.9$.

These results indicate that degrees of the individual nodes reflect the impact of the wiring cost constraint, however there are still some other unknown factors in addition to the wiring cost and processing efficiency constraints that may strongly influence the node degrees in the original network.

Link recovery

In the above sections we explored the similarity between reconstructed and real networks in global network measures (e.g., physical and graphical distance), and in mesoscopic properties (e.g., modules, hubs, degree and degree distribution). Next we analyzed how well the reconstruct networks recovered the connections in the real network. We calculated the recovery rate R by comparing the overlap of the adjacency matrices of the original and reconstructed networks entry by entry, taking both asymmetry of 0's and 1's entries and the directionality of connections into consideration (see *Material and Methods*). R is maximal at 1.0 when two networks are identical in all connected and un-connected pairs.

The overall recovery rate R of the whole network as a function of α is shown in Fig. 10A,B for the two neural systems. For the Macaque cortical network, when $\alpha > 0$, the recovery rate is nearly 60% (Fig. 10 A), which is clearly larger than the recovery rate of random benchmark networks obtained by rewiring the original network while retaining the input and output degrees. For the *C. elegans* neuronal network, when $\alpha > 0$, the recovery rate is more than 30%, which is significantly larger than 19.14% of random benchmark networks (Fig. 10 B). The recovery rate is much smaller than that of the Macaque networks, partially due to much sparser connectivity in *C. elegans*.

Importantly, the recovery rate is not uniform in the networks. It is higher within the spatial clusters, but lower between the clusters, as shown as green bars in Figs. 10C, D for Macaque and *C. elegans* respectively (both for $\alpha = 0.9$). Notably, for *C. elegans*, the recovery rate of the connections within clusters could approach 40% (Fig. 10 D), almost doubling that of the random networks (blue bars). For the connections between the clusters, the recovery rate is even smaller than that in the random benchmark networks in both systems, especially significant for *C. elegans*, because the connections between clusters become too sparse in the reconstructed

networks (see Figs. 5C, 6C). The real neural network has more long-range connections between spatial clusters, likely due to additional functional requirements.

Although the recovery rate within modules (red bars) is also larger than that of the random networks, particularly within spatial clusters, the recovery rate between modules is much higher than for random benchmarks for both systems, in contrast to the case between spatial clusters. The different results between clusters and between modules originate from the mismatching of modules and clusters.

In the combination of wiring cost and efficiency constraints, most of the connections are local. Therefore, the recovery rate is expected to be higher for the local connections. Our more detailed analysis of the recovery rate for the connections of different physical distance confirmed this expectation (Fig. 10E,F). In both the reconstructed and real networks, the probability of two nodes to be connected decreases with distance (insets, Fig. 10E,F). However, in real networks, the connection probability for very short distance is much smaller than in the reconstructed network, but it decays more slowly with the distance; therefore there is much higher probability of long-range connections, especially in *C. elegans*, likely due to other functional requirements. Nevertheless, the recovery rate is much higher for short-distant links, around 50% in both systems.

There are some differences between the *C. elegans* neuronal network and Macaque cortical network. (1) The *C. elegans* neuronal network is quite sparse (connection probability $p = 5.8\%$), while the Macaque cortical network is rather dense ($p = 23.7\%$). Therefore the recovery rate of the random benchmark networks is smaller in *C. elegans* (19.14%) than Macaque (44.5%). Furthermore, when $\alpha = 0$ (efficiency constraint only), the reconstructed networks of *C. elegans* possess a global hub, but for the Macaque network they are quite similar to random networks (Fig. 2C and Fig. 3 C). Thus for *C. elegans*, $R = 18.24\%$ at $\alpha = 0$ is even smaller than that of the random networks (19.14%). (2) About 60% of the neurons in *C. elegans* gather in the tiny head, and the others scatter in the ventral cord and tail. This highly heterogeneous spatial distribution of neurons requires more long-distant connections in the real network. Indeed, *C. elegans* has a smaller fraction of local connections (12.9%) when compared to Macaque (56.7%) (Fig. 4K, L, dashed lines). These properties are perhaps the main reasons that the recovery rate in *C. elegans* is low, because too many short-
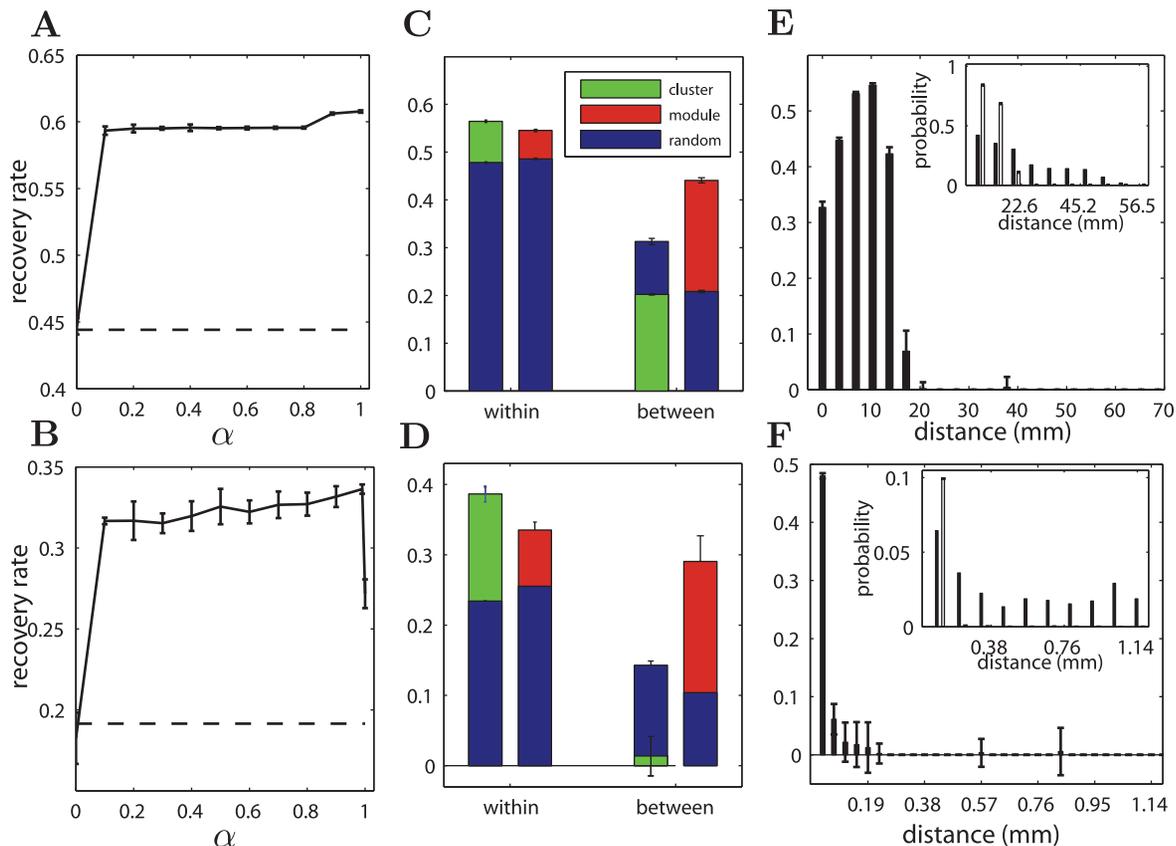
Figure 10. The rate of connection recovery in the reconstructed networks. The upper panel is for the Macaque cortical network and lower panel for the *C. elegans* neuronal network. (A, B) The recovery rate R as function of α. (C, D) Recovery rates among the nodes within the two spatial clusters and between the clusters (green bars for the reconstructed networks), as well as within and between modules (red bars for the reconstructed networks). The blue bars are the recovery rate of random networks with the same input and output degrees as the real networks. (E, F) The recovery rate as function of the spatial distance of the connections in real networks. The insets show the probability of connections as function of the spatial distance between two nodes, in real networks (dark bars) and reconstructed networks (white bars). The results in (C,D,E,F) are for $\alpha = 0.9$ for both Macaque and *C. elegans*. The error-bars are from 50 realizations of the reconstructed networks.
doi:10.1371/journal.pcbi.1002937.g010

distant links are put into the small head in reconstructed networks when compared to the real networks (Fig. 6 and Fig. 10E,F).

Discussion

Summary

The formation of complex neural networks is subject to many structural and functional constraints. In particular, it has been speculated that the network organization is the result of an economical trade-off between the physical cost and the functional value of the topology [52]. Among various graph theoretical measures, the average shortest pathlength representing the processing efficiency can be taken as an essential representative of functional constraints. In this study, we showed in a systematic and quantitative manner that the competition between these two straightforward constraints, the overall wiring cost and signal propagation efficiency, plays an important role in the network organization of Macaque cortical connectivity and *C. elegans* neuronal connections. By reconstructing network connections using optimization under multiple constraints, while fixing the component layout as in the real network and comparing the properties the reconstructed to the real networks at different topological scales, we revealed that the connectivity in both neural

networks is closely related to the spatial arrangement of the nodes. The main findings are as follows. *i*) The combination of the wiring cost and processing efficiency constraints can lead to the simultaneous formation of local connections and hubs. *ii*) When the spatial layout of the nodes is not uniform, but clustered as seen in the real networks, this combination will lead to the formation of modules and hubs. The modules have strong overlap with the spatial clusters and the hubs are located at global or regional geographical centers. *iii*) In certain regimes of competition between the two constraints, the reconstructed networks have a modular organization quite similar to that in real networks, and the positions of the hubs coincide or are close to the actual hubs. *iv*) The analysis also revealed that the degrees of nodes in the real networks are correlated with their neighborhood density, and the reconstructed networks can recover a significant portion of the individual links of the real networks, especially for short-distant connections. These observations support the idea of a trade-off between cost and functional values. The two constraints, however, cannot fully explain all the properties in the real networks. There are discrepancies in several important aspects: (1) There are significantly more long-range connections in both real systems when compared to the model. (2) The correlation between the individual degree and the neighborhood density

in the real networks is much lower than in the optimization model. (3) The degree distribution in the real network is different from that in the reconstructed network. (4) The model cannot generate asymmetry in the input and output degrees as in the real networks. All these differences suggest that there are additional important factors influencing the formation of real neural networks. In the following sections we discuss these points in more detail.

Real neural networks are not optimized solely to minimize wiring length or to maximize processing efficiency

It is clear that real networks are not solely optimized by topological efficiency, due to a large number of local connections. Neither are they optimized only for minimal wiring cost, even though this cost constraint is playing a significant role. Previous studies investigated whether the component placement layouts in the real neural networks satisfy the concept of wiring length minimization, by comparing the wiring cost of real networks to perturbed component placement while keeping the network connectivity fixed. The accumulated evidence showed that the wiring cost constraint indeed plays an important role [55,58,59,61,63,65,69]. However, for the Macaque cortical network and *C. elegans* neuronal network, previous work using CPO while fixing the network connectivity showed that the wiring length is not minimized [63]. Here we reconstructed the network connectivity while keeping the component placement layout as in the real network. Under only the wiring cost constraint at $\alpha = 1$, the reconstructed network differed significantly from the real networks: all the connections were short-distant and there were no hubs and no long-range connections as in the real networks.

The following comparison between real networks and reconstructed networks at $\alpha = 1$ showed more clearly that the wiring cost constraint is at work, but the real neural networks are not solely optimized by minimal wiring length. (1) The wiring cost L_p in the reconstructed networks is clearly lower than that of the real networks (53.5% for Macaque but only 1.25% for *C. elegans*), although the link recovery rate is still significant (60.8% for Macaque and 27% for *C. elegans*), since the wiring cost constraint does impose many short-range connections in the real networks. In the CPO solutions [63], by fixing the network connections as in the real networks, the wiring lengths were 68% and 52% of the real values of the networks in the Macaque and *C. elegans*, respectively. This finding shows that allowing hubs as in the real network is costly in wiring, especially for *C. elegans*. (2) The processing efficiency is much lower than in the real networks. As seen in Fig. 4B,C, L_g (the reverse of the processing efficiency) in the real networks is much smaller than the reconstructed networks at $\alpha = 1$. When α is slightly smaller than 1.0, the regional hubs in the reconstructed networks reduce the path length significantly and L_g is close to the values in real networks; interestingly these values are not much larger than the optimal L_g at $\alpha = 0$. This finding is consistent with the previous observation that biological neural networks appear to be optimized for processing efficiency [63]. (3) The recovery rate is quite high within the clusters. In particular, the recovery rate of the connections within Macaque visual system (covering 65.5% areas of the red cluster of Fig. 5) is 63.3%, and it is 47.7% for the rest of the connections of the network, which is consistent with previous observations that the wiring among the Macaque visual cortex is relatively more wiring cost optimal than other cortical subsets [64]. (4) In *C. elegans*, solely minimizing the wiring cost at $\alpha = 1$ cannot produce strong modular structure in reconstructed networks ($Q = 0.154 \pm 0.009$,

Fig. S6B) (as explained in the next section). However, similar modular organization as in the real network can be obtained under the balancing of two constraints (Fig. 3).

Overlap between network modules and spatial clusters reflects the impact of the wiring cost constraint

We showed that the wiring cost constraint can predominantly determine the modular organization in real neural networks, and the partition of the modules is not very sensitive to the effect of the processing efficiency constraint at different α. More precisely, the spatial clustering governs the connectivity modules in the reconstructed networks where most of the connections are extended to the spatial neighbors under the wiring cost constraint. Modules in the reconstructed network overlap largely with the spatial clusters. In the Macaque cortical network, the mismatched areas (mainly of the somatosensory system) are near the boundary of two spatial clusters (see Fig. 5E). This arrangement reduces the wiring cost when these areas make more connections to the other cluster to form a module, likely due to some further functional requirements.

Notably, the modular organization in the *C. elegans* neuronal network is most likely determined by the combination of the wiring cost and path efficiency constraints. With only the wiring cost constraint at $\alpha = 1$, the reconstructed network of *C. elegans* did not show strong modularity as seen in the real network (Fig. S6B, $Q = 0.154 \pm 0.009$ at $\alpha = 1.0$). The main reason is as follows. About 60% of the neurons gather in the tiny head, while other neurons scatter in the ventral cord and the tail with much longer physical distance between them. At $\alpha = 1$, almost all the links in the reconstructed networks were put in the head to form a highly connected core. The rest of the nodes in the body and tail form an approximately one-dimensional array with a minimal number of necessary links in order to avoid disconnection from the main core, but do not form dense modules. However, when the path efficiency constraint becomes effective for α slightly smaller than 1.0, some long-range connections are forced between the hubs in the head and the other parts of the network. These long-range connections from the hubs now can also take the role of avoiding disconnection of the nodes into subsets. Now the remainder of the connections is allowed to be more short-ranged in the tail and in the ventral cord, which form modules that coincide with the spatial clusters. With suitable combinatorial influence of the two constraints at $\alpha \approx 0.95$, the fraction of mismatched neurons is reduced to about 23% (Fig. 6F).

In both systems, the best overlap between the modules in the reconstructed networks and real networks appears for a region $0.8 < \alpha < 1$, suggesting that the combination of the two constraints plays an important role in the formation of real neural networks which possess both modules and hubs.

The processing efficiency constraint creates hubs by competition with the wiring cost constraint

The emergence of hubs can be attributed to the combination of the wiring cost and path efficiency constraints, because a single constraint on its own, either efficiency ($\alpha = 0$) or wiring cost ($\alpha = 1$), does not support hubs in networks as dense as the Macaque cortical network. For sparse networks, such as the one of *C. elegans*, the efficiency constraint on its own can generate a global hub, but the hub position is arbitrary. The wiring cost constraint is effective even for small α values. In both systems, there is a broad range of the competing parameter ($\alpha = 0 \sim 0.8$ for Macaque and $\alpha = 0 \sim 0.3$ for *C. elegans*) where the reconstructed networks are composed of local connections and just one dominant hub. For larger α, the

wiring cost constraint becomes stronger and several regional hubs emerge, similar to the organization in real networks. However, the wiring cost constraint on its own ($\alpha = 1.0$) cannot enforce hubs in the reconstructed networks.

The nodes chosen as the hubs in the reconstructed networks occupy the highest ranks of the regional geographical centrality and neighborhood density in order to be wiring economic. Interestingly, real hubs are found close to these nodes with high centrality. Quite strikingly, in the Macaque cortical network, the location of the biggest hub at area V5/MT and the output hub of area 46 in the real network of the Macaque cortical network can be reproduced in the reconstructed networks by the combination of the wiring cost and path efficiency constraint at $\alpha \sim 0.9$. The strong impact from the wiring cost constraint in real networks is further reflected by the economic wiring of these hubs nodes: (1) the areas V5/MT and area 46 stay near the geographical centers of the red and blue clusters, ranked No. 3 and No. 2 of the regional geographical centrality of the red and blue clusters, respectively; these areas have most of their connections within their respective cluster. (2) V5/MT and area 46 are ranked No. 1 and No. 3 in terms of high density of nodes in the neighborhood within a certain range of radius. In *C. elegans*, the locations of hubs in the reconstructed network were also found to be very close to or overlapping with the position of real hubs, and there is also a regime ($\alpha = 0.9 \sim 1$) which best matches modules and hubs to those of the real network. In this regime, L_g is close to the real networks, but L_p is much smaller, especially in the *C. elegans* neural network.

It appears in both systems that there are more nodes with larger degrees in the real networks than the reconstructed networks. However, importantly, the nodes with the largest degrees are close to the regional geographical centers similar to the model. Although one cannot expect that only two constraints can recover all the features of the real networks, the observation of the overlapping and close spatial locations of the reconstructed and real hubs (i.e., economic wiring of the hubs) provided clear evidence that probably both cost and efficiency constraints are at work in the real neural systems, supporting the idea of a trade-off between cost and functional values of the networks.

Robustness against node failure might be an additional constraint

In a broad range of the balance parameter ($\alpha = 0 \sim 0.8$ for Macaque and $\alpha = 0 \sim 0.3$ for *C. elegans*), the reconstructed networks have a single dominant hub linking to most of the nodes. Such a configuration is due to a strong impact of the path efficiency constraint, however, it is not functionally robust. Failure of the single dominant hub node greatly degrades the efficiency in information processing, because without the central hub, the network has mainly local connections. We performed a systematic analysis of the impact of removing the node with the largest degree in the reconstructed networks and measured the increase of L_g [94]. The results are shown in Fig. S10. It is seen that in the regime with a single, global hub, L_g in the reconstructed networks increases significantly when the hub is removed. The degrading effect is much more serious in *C. elegans*, because the reconstructed networks without the hub contain an almost one dimensional array for the neurons in the body that separates head and tail. In the next regime with several regional hubs, ($\alpha = 0.8 \sim 0.96$) for Macaque and $\alpha = (0.9 \sim 99)$ for *C. elegans*), the reconstructed networks are much more robust against the removal of the node with the largest degree (Fig. S10), very close to that in the real network. Thus, while hubs enhance the processing efficiency significantly, they are also the points of vulnerability to

>pathological damage [. The high energy consumption of the brain puts it under high vulnerability for energy undersupply, and the metabolically most expensive node is particularly vulnerable in pathological circumstances. There is evidence that metabolic costs of a node are proportional to its degree [; thus some brain disorders may be closely related to hubs' abnormalities. For example, this effect has been speculated to be a network mechanism for Alzheimer's disease [97,98]. In particular, laser ablation experiments have demonstrated that AVA neurons are required for normal spontaneous and evoked backward locomotion [100,102]. Therefore, avoiding the configuration of a single dominant hub could indeed be functionally significant. In conclusion, the robustness requirement might be an additional important factor that generates the evolutionary pressure for the real networks to have more hubs, so that both in Macaque cortical network and neuronal networks, the competition between the two constraints is settled down to a regime with \alpha <?mathtype % MathType!MTEF!2!0!+% feaagaart1ev2aqatCvAUfeBSjuy-ZL2yd9gzLbvyNv2CaerbuLwBLn% hiov2DGi1BTfMBaeXatLx-BI9gBaerbd9wDYLwzYbItLDharqqtubsr% 4rNCHbGeaGqi-Vu0Je9sqqrpepC0xbbL8F4rqqrFfpeea0xe9Lq-c9% vqaq-pepm0xbba9pwe9Q8fs0-qaqpepae9pg0FirpepeKkFr0xfrr9adba-qaaeGaciGaaiaabeqaamaabaabaaGcbaGaeqySdegaaa!378D!\alpha $ smaller, but very close to 1.0.

Differences between reconstructed and real networks imply additional constraints

We have shown that there is a regime ($\alpha \sim 0.9$) of the competition between the two constraints of wiring cost and processing efficiency where the reconstructed networks can reproduce some major properties of the real networks, including the simultaneous formation of hubs and modules, locating hubs or other large-degree notes close to the regional centers of the spatial clusters and similar resilience to node failure. However, there are still significant differences. The reconstructed networks have much smaller wiring length L_p (about 54% of the real wiring length for Macaque and nearly 5% for *C. elegans*). This difference stems from the fact that the real networks have more long-range connections: there are more large-degree nodes or hubs (Figs. 4E,F, Fig. S3), the fraction of spatially local connections $P_{nearest}$ is much smaller and the probability of long-range connections is much larger (insets in Figs. 10E,F). Moreover, the number of connections is not strongly determined by the neighborhood density (Fig. 9).

Another significant difference is the asymmetry in the in- and out-degrees in real networks. Although directed links are used in our model influencing the calculation of processing efficiency, the objective function did not contain possible constraints that can reflect the functional role of the asymmetry in the input and output links; the reconstructed networks are thus highly symmetric. In fact, the asymmetry in the biological neural networks may be closely related to functional requirements of specific signal processing flow. For instance, in *C. elegans*, there are more connections from the sensory neurons to motor neurons than vice versa. Also from the connection matrix of the Macaque cortical network, there are more connections from some visual areas to the motor system, and fewer along the opposite direction. Therefore, in future work, one may include additional constraints to enforce information flow between different types of nodes, such as input nodes (the sensory neurons or primary sensory areas) and output nodes (the motor neurons or motor areas).

In our view, the differences imply that most likely real neural networks are influenced by additional functional requirements and constraints. The trade-off between the wiring cost and the

processing efficiency with these new constraints could give a better account of the asymmetry and other properties of the real neural networks.

Conclusion and outlook

We studied the formation of complex network connectivity derived from multiple constraints, in particular the competition between wiring cost and path efficiency requirements. By reconstructing networks while preserving the spatial layout of the components, we obtained an understanding of the relationship between the spatial layout and network connectivity derived from the multiple constraints. This understanding guided us to investigate the relationship between the spatial layout and network connectivity of the real Macaque cortical network and *C. elegans* neuronal network. The results are consistent with previous observations that wiring cost and efficiency constraints are playing an important role in shaping the network organization and provide evidence to support the idea of a trade-off between them. While significant, the wiring cost and path efficiency constraints cannot completely explain all the features in the connectivity patterns of the real Macaque cortical network and *C. elegans* neuronal network. Other factors, such as robustness of networks against node failure, more long-range connections and asymmetric input and output information flow appear to be important. Moreover, with our combinatory optimization model, we have mainly discussed the question of which constraints are at work to influence the neural network properties, but have not yet addressed possible mechanisms underlying the biological implementation of these constraints. Exploring additional functional constraints and incorporating them into growth or generative models [42,103] may be worthwhile directions in the future research.

The present study on the impact of multiple constraints on the architecture of neural systems might also provide an anatomical foundation for the cost-efficiency balance in human functional networks [79]. A more detailed understanding of the relationship between the cost-efficiency balance in anatomical and functional neural networks will require the analysis and modeling of the dynamical interdependence in the reconstructed networks under different competition conditions relative to that in real networks.

Materials and Methods

Materials

Primate cortico-cortical network. We analyzed the connectivity of Macaque cortical network and its relationship with the three-dimensional spatial layout of the components and compared it to the reconstructed networks in order to understand the impact of multiple constraints. The Macaque network is largely based on a dataset with 94 cortical areas and 2,390 directed projections [63]. This data set is more suitable for our purpose compared to other connectivity data sets (e.g., cortical network of cat), since both the spatial positions of the components and the network connectivity are available. The connectivity data and three-dimensional spatial position (the average surface coordinate) of each cortical area were provided by M. Kaiser (http://www.biological-networks.org/) and amended with the help of the CoCoMac database (http://cocomac.org).

This dataset is extensive, but is still partly incomplete. Especially the divisions of motor regions are quite coarse with incomplete connections among several areas (e.g., the motor areas 4 and 6 cover a large territory, 6.5% of neocortex). Generally, it is a challenge to define anatomical parcellations that provide a high

spatial resolution as well as a high density of the associated connectivity data. In this study we first improved the data set with a more detailed parcellation of the motor areas based on CoCoMac [92], which consists of three primary databases: literature, mapping, and connectivity. For the motor areas, there were 39 unique brain maps, 84 unique literature sets, and 454 unique records interrelating brain maps to each other. However, there are many types of inconsistency, including the lack of connectional information of some brain maps in the database, the supra-structure relations not being the symmetric transpose of sub-structure relations, typographical errors, brain regions with different acronyms, but with the same full name, etc. We checked the inconsistent and conflicting data. For instance, those maps lacking of connectional information in the connectivity database or with different acronyms but representing the same regions would not be included in the updated Macaque network. Those connections from the source sites A to the target sites B, if not consistent with the connections of region B receiving from region A, were removed. And connections between a motor area and different sub-structure areas of an existing area in the former database were merged. This expanded collation of data extended the former 6 motor areas to 15 areas (4, 4C, 6, 6D,6DC, 6DR, 6DS, 6m, 6Va, 6Val, 6Vam, 6Vb, SMA, SMAr, SMA-proper) with an additional 128 projections. Spatial positions of these areas were taken as the average surface coordinate, as in a previous approach [63]. This improved dataset of the Macaque cortical network has 103 areas and 2518 connections in total. The labels of the areas are listed in Table S1 of SI, and the adjacency matrix is shown in Fig. S1, where the new areas and links are highlighted by blue color. Although we applied a detailed dataset for the Macaque cortical network, the connectivity data may still be incomplete. We noticed that the motor region owns relatively low connection density (21.3%), comparing to the connection density of the whole Macaque cortex (23.7%). However, strong fluctuations of node degree and regional connection density have also been observed in most known connectivity data sets [35,104,105].

To examine the reliability of the spatial clustering (see "Spatial clustering" in *Materials and Methods*) of the components in this new dataset with 103 areas, we analyzed the clustering in another dataset -"PHT data" [93] with 176 areas from a finer parcellation of the subcortex and parts of the cortex in high magnification. The spatial positions of these areas were taken as the average surface coordinate estimated from surface parceling using the CARET software (http://sumsdb.wustl.edu/sums/index.jsp). However, connectivity information for this finer parcellation of areas is not available, and we only used the spatial positions of these areas for the spatial clustering analysis.

Spatial distance between the areas was calculated as Euclidean distance between the spatial positions of the areas, and the wiring length L_p is the sum of the distances between connected areas.

C. elegans neuronal network. We also analyzed the connectivity of the *C. elegans* neuronal network and its relationship with the two-dimensional spatial layout and compared to the reconstructed networks. The dataset of the spatial positions of 276 neurons was provided by M. Kaiser (http://www.biological-network.org). The connectivity in this dataset is not complete. Thus, we compared the databases by Varshney et al. [47] and by Oshio et al. [91] on online neuronal wiring of *C. elegans* (http://www.wormatlas.org/neuronalwiring.html), and added those links that were common in these two new databases but did not exist in the dataset of Kaiser. Thus the original database of Kaiser with 2105 connections was extended to 2902 connections. The links considered here are all chemical synapses. The

formation of gap junctions may require direct contact of cell membrane at certain stage of development, and they are not considered in this study of wiring cost and processing efficiency. The updated network has 276 neurons and 2902 links. The names of the neurons are listed in Table S2 of SI. Here the wiring length L_p is measured as the Euclidean length between all connected pairs of neurons.

Methods

Objective function and connectivity optimization. We reconstructed network connections for a given total number of connections while fixing the spatial position of the nodes. The reconstructed networks were to minimize an objective function combining the wiring cost and processing efficiency constraints,

$$L = (1 - \alpha)L_g + \alpha L_p,$$

where α is a parameter to represent the relative weight of the normalized physical wiring length $L_p = l_p / l_p^{\max}$ and the normalized graph length $L_g = l_g / l_g^{\max}$. Here l_p is the total wiring length of the links and l_g is the sum of the shortest pathlength between all pairs of nodes in the network. $l_p^{\max}$ is obtained at $\alpha = 0$ when minimizing l_g without considering any spatial constraint and $l_g^{\max}$ is obtained at $\alpha = 1$ when minimizing l_p without considering the efficiency constraint.

We applied a simulated annealing optimization algorithm [106] to search for network configurations that minimize the objective function L. In each step, we set the programme to avoid those choices which made the network disconnected. In the beginning, the simulated annealing approaches were employed to obtain the benchmark networks at $\alpha = 0$ and $\alpha = 1$ to obtain the physical distance $l_g^{\max}$ and graph distance $l_p^{\max}$, respectively. Using them for normalization, we made the value of the objective function L stay within (0, 1).

The simulated annealing optimization algorithm was implemented as follows. We started with a random network and a high temperature T_0, and then decreased temperature as $T_{n+1} = T_n / n$. At each temperature, we rewired the network for 1000 steps. At each step, we chose four random nodes to exchange the connections, and accepted the changes with the probability $\exp(-\Delta L / T)$. The program was terminated when $\Delta L \leq 10^{-5}$. For each α, the simulated annealing programme was performed to obtain 50 realizations of the optimized networks from different initial random networks.

We first applied the method to a 1D model with the nodes arranged uniformly on a 1 one-dimensional ring, then to the primate cortico-cortical network of Macaque and the neuronal network of *C. elegans*.

1D model. The nodes were arranged with equal distance on a one-dimensional circle. The links were directed. The wiring length of a connection is the Euclidean distance between two nodes $l_p(i,j)$ and $l_p = \sum l_p(i,j)$.

The optimal solution for $\alpha = 0$ (processing efficiency constraint only) is a network with apparently random connections and a single, global hub; and the optimal solution for $\alpha = 1$ (wiring cost constraint only) is a regular, local graph, while in-between the optimal solutions are networks with local connections and hubs (Fig. 1).

Modularity analysis. A modular community is defined as a subnetwork with a higher density of connections relative to the entire network, based on the work by Newman [28]. The modular structure is obtained through optimizing the partitioning of the network into several modules to maximize a quantitative measure of modularity, Q, defined as [107]

$$Q = \frac{1}{K} \sum_{i,j} [A_{ij} - \frac{k_{out}(i)k_{in}(j)}{K}] \delta_{c_i c_j}$$

where, $A_{ij} = 1$ when there is a directed projection from node i to node j, and 0, otherwise. K is the total number of links in this network, δ is the Kronecker delta ($\delta_{ij} = 1$ if $i = j$, and 0 otherwise), and c_i is the index of the community where node i is assigned. We used the modularity-dir.m [108] in (the toolbox of brain connectivity) BCT package for the module partition. Since there could be uncertainty and degeneracy in the algorithm [109,110], we applied the methods for 50 independent runs for each network considered (real or reconstructed networks). To check the significance of the modularity, we also measured modularity in the corresponding random networks with the same number of nodes and links as in the network examined. In the event, we found that the modular division had no degeneracy in both real networks of Macaque and *C. elegans* and in reconstructed networks of Macaque ($\alpha \neq 0$), but it had small degeneracy for the reconstructed networks of *C. elegans* (typically 1–5 neurons of mismatch from run to run).

Spatial clustering. As seen from two-dimensional projection of the spatial layout and the connectivity of the Macaque cortical network in Fig. 5A, the nodes can be roughly divided into two dense groups which are spatially distant from each other. To perform a quantitative analysis of the spatial clustering, we computed the pair-wise Euclidean distance of the nodes to obtain the distance matrix and applied the hierarchical clustering algorithm in MATLAB (linkage.m and cluster.m) to the distance matrix. With a suitable threshold, we could obtain the same number of clusters as network modules in order to compare them appropriately.

There are various ways to compute distance in agglomerating the hierarchical cluster tree, including *centroid distance, inner squared distance, the shortest distance, the furthest distance, average distance, weighted average distance, and weighted center of mass distance between clusters*. In addition to hierarchical clustering, we also applied the K-means method to minimize the within-cluster sums of point-to-cluster-centroid distances. While the results are deterministic for a given distance measure, they differ slightly for different distance measures. To take this uncertainty into consideration, in this work, we applied the different methods to the real networks of Macaque and *C. elegans*, and compared the clusters from different methods to the modules of the real network obtained by the BCT method (see sections "Modularity analysis" and "Matching between partitions" in *Method*). We found that in both systems, the clusters by inner squared distance has the minimal matching rate with the modules, and the group of mismatched nodes is the largest possible to cover all the mismatching sets in the other methods (centroid distance produced the same result as inner squared distance in Macaque). This means that the set of matching areas resulting from the inner squared distance is most stable when different possible clusterings are considered. Thus we used the clusters from the inner squared distance for further comparison with the modules in reconstructed networks.

The method was applied to the Macaque cortical network to obtain two clusters (Fig. 5) and to the *C. elegans* neuronal network to obtain four clusters (Fig. 6).

For the Macaque cortical network, the comparison of the clustering of two datasets (103 areas and 176 areas) was shown in Fig. S4 and S5 for the dendrograms (A,B) and for the boundaries between the division into two clusters (C). There are only five mismatched areas between the two datasets (areas 4, 6C, 3b, 2/1 and AI according to the dataset of 176 areas) (Fig. S5).

Matching between partitions. We compared the partition of spatial clusters with the partition of modules or the partitions of modules in real and reconstructed networks. Suppose the partition A gives m groups A_i and partition B generates n groups B_j with $m \leq n$. For each partition B_j, we found the corresponding group in partition A to which the group B_j should belong (the one covers the maximal number of nodes in B_j). This maximally covered number in B_j is added to the matched set and the rest of the nodes are counted in the mismatched set. Note that it is possible that two or more groups in B can belong to the same group in A. In this way, we identified all the nodes matching the two partitions. Taking the matching between the clusters and modules of the real *C. elegans* network as an example, there are 3 spatial clusters and 4 modules, and two modules belong to the same cluster at the head.

We also developed a way to judge the significance of matching between partitions by comparing the results to "shuffled modules". We kept the size of all subgroups B_j of partition B, but randomly shuffled two elements in different subgroups. Such shuffled data have the same n groups as in the real partition B, but the labels of nodes are randomized. We obtained an ensemble of such shuffled modules and calculated the average mismatching rate with the real partition A. In fact, this average value is independent of the group number n and the size of the groups when the ensemble is large enough. The results are shown as the dashed lines in Fig. 5 F and 6 F.

Identifying hubs in networks. There is no unique definition of hubs in networks. In general, the hub nodes should have a degree of about the same order of the network size, much larger than the mean degree of the whole network. Such a picture is usually reasonable when the degree distribution is not very narrow. Taking *C. elegans* for instance, the mean value of in-degree is 10.51, and the standard deviation in the reconstructed networks varies at different α, such as 18.2 ($\alpha = 0.3$), 13.2 ($\alpha = 0.9$), 10.2 ($\alpha = 0.95$) and 10.6 ($\alpha = 1.0$). The standard deviation is not small in our work, therefore we heuristically identified hubs in the real and reconstructed networks by the z-score of the degree k of the node, which is defined as $z = (k - \langle k \rangle)/\sigma_k$, where $\langle k \rangle$ is the average degree of the network and σ_k is the standard deviation of the degree distribution. In a previous analysis of hubs in cortical network [20], a node was regarded as a hub when $z \geq 1.0$. To focus on nodes with the degree significantly larger than the average degree, in our analysis, a node was regarded as a hub when $z \geq 2.0$. For example, in the reconstructed of *C. elegans* at $\alpha = 0.9$, the hubs identified in this way have an average degree of 42, which is reasonably large when compared to the mean degree 10.5 and the standard deviation 10.2 in the network.

Despite the ambiguity in defining hubs in networks, this method provides a pragmatic way to characterize the properties about the changes of the number of nodes with relatively large degrees, the average degree of them and their spatial positions with respect to the parameter α, which were quite clear and common across the different networks analyzed.

Locality measure $P_{nearest}$ of connections. We measured to what extent the links in the network were connected to spatial nearest neighbors. For each node i, we considered the total number of connections $k(i) = k_{in}(i) + k_{out}(i)$. The total wiring length $l_p(i)$ from this node can be computed as the sum of the wiring length of the k links. $l_p(i)$ for the node i will be minimal, $l_p^{min}(i)$, if all the k_{in} input links and k_{out} output links are connected to k_{in} and k_{out}

nearest neighbors of the node i, respectively. One can use the ratio $P_{nearest}(i) = l_p^{min}(i)/l_p(i)$ to quantify the locality of the connections of the nodes in the network. All the connections of a node are local if $P_{nearest}(i) = 1$. One can obtain a locality measure of the connections of the whole network as the average of $P_{nearest}(i)$, i.e.,

$$P_{nearest} = (1/N) \sum_i^N P_{nearest}(i).$$

$P_{nearest} \approx 1.0$ indicates that most of the connections in the network are to the spatial nearest neighbors. For random connections, $P_{nearest}$ is small. The connections with a distant hub reduce the locality measure of the network.

Geographical centrality of the nodes. Under the multiple constraints of wiring cost and processing efficiency, the centrality of a node in the spatial layout of the network, called *geographical centrality* here, may be closely related to its connectivity. To quantify the geographical centrality, we computed the total Euclidean distance $D_i = \sum_{j \neq i} d_{ij}$ from a node i to all the other nodes in the network, and ranked the nodes according to this distance D_i. The node with the smallest total distance has the highest geographical centrality.

We can also measure the geographical centrality with respect to a subset of nodes, for example, within one of the spatial clusters of nodes.

Neighborhood density of nodes in space. Unlike the 1D model, the spatial distribution of the nodes in the original Macaque cortical network is not uniform due to varying sizes and shapes of the cortical areas. The neurons in *C. elegans* are also distributed highly non-uniformly in space. To quantify the spatial heterogeneity, we evaluated the density of nodes in the neighborhood of a given node, called the neighborhood density $\rho_i(r)$, i.e.,

$$\rho_i(r) = N_i(r)/V_i(r),$$

where $N_i(r)$ represents the number of nodes lying within the sphere (for Macaque) or circle (for *C. elegant*) of radius r, centered at the given node i, and $V_i(r)$ is the volume of the sphere (area of the circle).

The radius r is referenced to the maximal distance r_{max} between two nodes in the network.

Comparing the connectivity of reconstructed and real networks and computing the recovery rate. The most direct way to compare the connectivity of the reconstructed network with the original network is to compare the adjacency matrices using measures such as the Hamming distance. One can measure the number of overlapping entries K_r of the two adjacency matrices and obtain the recovery rate as $R = K_r/(N(N-1))$ (without considering the diagonal entries). However, this result could be biased when the number of 1's (K_1) and the number of 0's (K_0) are strongly asymmetrical in the network. This is the case for both the Macaque cortical network ($K_1 = 2,518$ and $K_0 = 7,988$) and *C. elegans* neuronal network ($K_1 = 2,902$ and $K_0 = 72,998$). We followed the method proposed in [27] to compute the recovery rate as $R = \sqrt{R_0 R_1}$ where R_0 and R_1 are the recovery rate for 1's and 0's, respectively. In particular, $R_0 = K_{r1}/K_1$ and $R_1 = K_{r0}/K_0$, where K_{r1} and K_{r0} are the number of overlapping entries with value 1 and 0, respectively.

The same method can be applied to obtain the recovery rate for subnetworks by considering the corresponding sub-matrices, for example, the connections within and between the spatial clusters

or the modules (Fig. 10C,D). It can also be extended to the connections between nodes separated by a distance with certain range ((Fig. 10E,F).

Testing whether degrees follow power-law distribution. We used the Kolmogorov-Smirnov statistic and maximum-likelihood fitting methods (MLEs) [111], to test whether the degree distribution in real or reconstructed networks can be fitted by a power-law function. First, the following power-law model

$$p(x) = \frac{\gamma - 1}{x_{\min}} \left(\frac{x}{x_{\min}}\right)^{-\gamma}$$

is fitted to the data to estimate the parameters $x_{\min}$ and γ. Then an ensemble of surrogate data (with the same number of data points as in the degree distribution) is generated from the above power-law function with the estimated x_{min} and γ. The distance $\Delta_s(i)$ between the ith surrogate data set and the model function can be obtained and compared to the distance Δ_d between the degree distribution and the model. A p value is obtained as the fraction of surrogate data sets in the ensemble with $\Delta_s(i) > \Delta_d$. Thus if p is large (close to 1), then the difference between the degree distribution and the power-law model can mainly be attributed to the statistical fluctuation. But if p is small ($p \leq 0.1$), then the power-law can be ruled out for this degree distribution.

Significance test of correlation. We studied several types of correlations between two data sets of small sizes, e.g., correlation between degrees and density in the reconstructed and real networks. To evaluate the significance of the correlation between two finite-size data sets, we used the method of surrogate data. We randomly reshuffled the order of one of the data sets and computed the correlation values for 100 realizations of the random reshuffling. From the distribution of the correlation of such surrogate data one can obtain the standard deviation σ. If the correlation between the two original data sets is outside 2σ of this distribution, it has about 95% of significance that the real correlation is not a coincidence due to finite size of the data sets.

Supporting Information

Figure S1 The adjacency matrix of the Macaque cortical networks with finer division of the motor areas. The links in the original data set of Kaiser are shown as gray, and the links from or to new, finer motor areas are shown by blue.
(EPS)

Figure S2 Adjacent matrices of the networks with various connection probabilities in 1D model at $\alpha = 0$. A single hub appears for sparse networks (A, $p = 2.5\%$) and (B, $p = 5\%$), but not in dense network (C, $p = 12.5\%$).
(EPS)

Figure S3 Input and output degrees in reconstructed and real networks. (A) Reconstructed network of the Macaque cortical network, with $\alpha = 0.9$. (B) Real network of the Macaque cortical network. The correlation between the input and output degrees is 0.589 which is larger than 95% significance level at 0.237 in the corresponding randomly reshuffled data (see *Materials and Methods*). (C) Reconstructed network of the *C. elegans* neuronal network, with $\alpha = 0.9$. (D) Real network of the *C. elegans* neuronal network. The correlation between the input and output degrees is 0.717, larger than 95% significance level at 0.158 in the corresponding shuffled data.
(EPS)

Figure S4 Comparison of the spatial clustering in two different databases of Macaque cortical network. (A) and (B) display the clustering dendrogram for datasets with $N = 103$ and $N = 176$ areas, respectively. In both datasets, the nodes can be divided into two pronounced clusters.
(EPS)

Figure S5 The spatial layout of the areas of Macaque cortex. It shows that the spatial layout of the areas(indicated by lines for 103 areas), with reference to parcellation of 176 areas (color). The boundaries between the two clusters are shown by bold lines (blue for for dataset of 103 areas and red line for 176 areas). These two boundaries are close to each other, indicating that spatial clustering is a robust property in the spatial layout of the cortical areas.
(TIF)

Figure S6 Modularity Q of reconstructed networks as a function of α, compared to Q of the real network. (A) Macaque cortical network. (B) *C. elegans* neuronal network. The corresponding random networks with the same number of nodes and connections have $Q = 0.08 \pm 0.004$ (Macaque) and 0.17 ± 0.037 (*C. elegans*). The error-bars are from 50 realizations of the reconstructed networks.
(EPS)

Figure S7 The ranking of a few nodes with the largest total degrees in real Macaque cortical networks in terms of neighborhood density and geographical centrality. (A) 7 nodes (cortical ares) with the largest total degree. (B) 20% of network nodes having top neighborhood density in a range of radius $r/r_{max} = 0.24 - 0.28$. (C) and (D) 40% of nodes from the red and blue clusters having top geographical centrality from the respective cluster (gray bars). The length of the bar indicates the corresponding values. The dark bars in (C) and (D) highlight the cortical areas listed in (A). The nodes with the largest total degrees are among the top list of neighborhood density as well as geographical centrality, which is a strong signature of wiring cost constraint.
(EPS)

Figure S8 Closeness between reconstructed and real hubs in *C. elegans*. The bars show the total distance of each of the 166 neurons in the head to the 4 real hubs (AVAL/R and AVBL/R, red and blue bullets), sorted from small to large values. Green triangles indicated the hubs in the reconstructed networks for $\alpha \sim 0.9 - 0.93$.
(EPS)

Figure S9 Degree distribution in the original and reconstructed networks. (A) shows the outdegree distribution for Macaque cortical network. (B) shows the outdegree distribution for *C. elegans* neuronal network. The degree distribution in reconstructed networks does not differ significantly for different α values, since the degrees are mainly determined by the neighborhood density of the node. In the insets, the dashed line with different color represents the fitting power-law distribution from the corresponding degree distribution with the same color (including distributions not obeying power-law). Inset of A: $p = 0.001$ (real network), $p = 0.002$ ($\alpha = 0.1$), $p = 0.0025$ ($\alpha = 0.9$). Inset of B: $p = 0.675, \gamma = 3.2$ (real network), $p = 0.01$ ($\alpha = 0.1$), $p = 0.02$ ($\alpha = 0.9$).
(EPS)

Figure S10 Degrading effect on processing efficiency when removing the node with the largest degree in reconstructed networks. Shown is the ratio between L_g after

and before the removal of the node. (A) Macaque cortical network. (B) *C. elegans* neuronal network. The error-bars are from 50 realizations of the reconstructed networks.
(EPS)

Table S1 The list of the labels of 103 cortical areas in Macaque is provided on-line.
(XLS)

Table S2 The list of the names of 276 neurons in *C. elegans* is provided on-line.
(XLS)

Author Contributions

Conceived and designed the experiments: CZ CCH. Performed the experiments: YC SW. Analyzed the data: YC SW CZ. Contributed reagents/materials/analysis tools: YC CCH CZ. Wrote the paper: YC CCH CZ.

References

1. Bressler S (1995) Large-scale cortical networks and cognition. Brain Research Reviews 20: 288–304.
2. Mesulam M (1998) From sensation to cognition. Brain 121: 1013.
3. McIntosh A (2000) Towards a network theory of cognition. Neural Networks 13: 861–870.
4. Friston K (2002) Beyond phrenology: what can neuroimaging tell us about distributed circuitry? Annual review of neuroscience 25: 221–250.
5. Buzsáki G (2006) Rhythms of the Brain. USA: Oxford University Press.
6. Sporns O, Chialvo D, Kaiser M, Hilgetag C (2004) Organization, development and function of complex brain networks. Trends in Cognitive Sciences 8: 418–425.
7. Bassett D, Bullmore E (2006) Small-world brain networks. The Neuroscientist 12: 512.
8. Zhou C, Zemanová L, Zamora G, Hilgetag C, Kurths J (2006) Hierarchical organization unveiled by functional connectivity in complex brain networks. Physical review letters 97: 238103.
9. Zemanová L, Zhou C, Kurths J (2006) Structural and functional clusters of complex brain networks. Physica D: Nonlinear Phenomena 224: 202–212.
10. Honey C, Kötter R, Breakspear M, Sporns O (2007) Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proceedings of the National Academy of Sciences 104: 10240.
11. Thivierge J, Marcus G (2007) The topographic brain: from neural connectivity to cognition. Trends in Neurosciences 30: 251–259.
12. Reijneveld J, Ponten S, Berendse H, Stam C (2007) The application of graph theoretical analysis to complex networks in the brain. Clinical Neurophysiology 118: 2317–2331.
13. Stam C, Reijneveld J (2007) Graph theoretical analysis of complex networks in the brain. Nonlinear Biomedical Physics 1: 3.
14. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10: 186–198.
15. Honey C, Thivierge J, Sporns O (2010) Can structure predict function in the human brain? Neuroimage 52: 766–776.
16. Zamora-López G, Zhou C, Kurths J (2011) Exploring brain function from anatomical connectivity. Front Neurosci 5: 83.
17. Watts D, Strogatz S (1998) Collective dynamics of small-worldnetworks. Nature 393: 440–442.
18. Stam C, Jones B, Nolte G, Breakspear M, Scheltens P (2007) Small-world networks and functional connectivity in alzheimer's disease. Cerebral Cortex 17: 92–99.
19. Liu Y, Liang M, Zhou Y, He Y, Hao Y, et al. (2008) Disrupted small-world networks in schizophrenia. Brain 131: 945–961.
20. Sporns O, Honey C, Kötter R (2007) Identification and classification of hubs in brain networks. PLoS One 2: e1049.
21. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey C, et al. (2008) Mapping the structural core of human cerebral cortex. PLoS Biology 6: e159.
22. Zamora-López G, Zhou C, Kurths J (2010) Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. Frontiers in Neuroinformatics 4: 1.
23. Young M (1992) Objective analysis of the topological organization of the primate cortical visual system. Nature 358: 152–155.
24. Hilgetag C, Burns G, O'Neill M, Scannell J, Young M (2000) Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat. Phil Trans R Soc Lond B 355: 91.
25. Hilgetag C, Kaiser M (2004) Clustered organization of cortical connectivity. Neuroinformatics 2: 353–360.
26. Zemanová L, Zhou C, Kurths J (2006) Structural and functional clusters of complex brain networks. Physica D: Nonlinear Phenomena 224: 202–212.
27. Costa L, Kaiser M, Hilgetag C (2007) Predicting the connectivity of primate cortical networks from topological and spatial node properties. BMC Systems Biology 1: 16.
28. Girvan M, Newman M (2002) Community structure in social and biological networks. Proceedings of the National Academy of Sciences of the United States of America 99: 7821.
29. Albert R, Barabási A (2002) Statistical mechanics of complex networks. Reviews of modern physics 74: 47.
30. Ravasz E, Somera A, Mongru D, Oltvai Z, Barabási A (2002) Hierarchical organization of modularity in metabolic networks. Science 297: 1551.
31. Schwarz A, Gozzi A, Bifone A (2008) Community structure and modularity in networks of correlated brain activity. Magnetic resonance imaging 26: 914–920.
32. Ferrarini L, Veer I, Baerends E, van Tol M, Renken R, et al. (2009) Hierarchical functional modularity in the resting-state human brain. Human brain mapping 30: 2220–2231.
33. Meunier D, Achard S, Morcom A, Bullmore E (2009) Age-related changes in modular organization of human brain functional networks. Neuroimage 44: 715–723.
34. Scannell J, Blakemore C, Young M (1995) Analysis of connectivity in the cat cerebral cortex. The Journal of Neuroscience 15: 1463.
35. Young M (1993) The organization of neural systems in the primate cerebral cortex. Proceedings of the Royal Society of London Series B: Biological Sciences 252: 13.
36. Pan R, Chatterjee N, Sinha S (2010) Mesoscopic organization reveals the constraints governing caenorhabditis elegans nervous system. PLoS one 5: e9240.
37. Schneidman E, Still S, Berry M, Bialek W(2003) Network information and connected correlations. Physical review letters 91: 238701.
38. Schneidman E, Berry M, II R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440: 1007.
39. Bettencourt L, Stephens G, Ham M, Gross G (2007) Functional structure of cortical neuronal networks grown in vitro. Physical Review E 75: 021915.
40. Salvador R, Suckling J, Schwarzbauer C, Bullmore E (2005) Undirected graphs of frequencydependent functional connectivity in whole brain networks. Philosophical Transactions of the Royal Society B: Biological Sciences 360: 937.
41. Henderson J, Robinson P (2011) Geometric effects on complex network structure in the cortex. Physical Review Letters 107: 18102.
42. Vértes P, Alexander-Bloch A, Gogtay N, Giedd J, Rapoport J, et al. (2012) Simple models of human brain functional networks. Proceedings of the National Academy of Sciences 109: 5868–5873.
43. Iturria-Medina Y, Canales-Rodriguez E, Melie-Garcia L, Valdes-Hernandez P, Martinez-Montes E, et al. (2007) Characterizing brain anatomical connections using diffusion weighted mri and graph theory. Neuroimage 36: 645–660.
44. Iturria-Medina Y, Sotero R, Canales-Rodríguez E, Alemán-Gómez Y, Melie-García L (2008) Studying the human brain anatomical network via diffusion-weighted mri and graph theory. Neuroimage 40: 1064–1076.
45. Gong G, He Y, Concha L, Lebel C, Gross D, et al. (2009) Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. Cerebral Cortex 19: 524.
46. Achard S, Salvador R, Whitcher B, Suckling J, Bullmore E (2006) A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. The Journal of Neuroscience 26: 63.
47. Varshney L, Chen B, Paniagua E, Hall D, Chklovskii D (2011) Structural properties of the caenorhabditis elegans neuronal network. PLoS computational biology 7: e1001066.
48. He Y, Chen Z, Evans A (2008) Structural insights into aberrant topological patterns of large-scale cortical networks in alzheimer's disease. The Journal of Neuroscience 28: 4756.
49. Laughlin S, van Steveninck R, Anderson J (1998) The metabolic cost of neural information. Nature neuroscience 1: 36–41.
50. Laughlin S, Sejnowski T (2003) Communication in neuronal networks. Science 301: 1870.
51. Striedter G (2006) Précis of principles of brain evolution. Behavioral and Brain Sciences 29: 1–12.
52. Bullmore E, Sporns O (2012) The economy of brain network organization. Nature Reviews Neuroscience 13: 336–349.
53. Mitchison G (1991) Neuronal branching patterns and the economy of cortical wiring. Proceedings of the Royal Society of London Series B: Biological Sciences 245: 151.

54. Cherniak C (1992) Local optimization of neuron arbors. Biological cybernetics 66: 503–510.

55. Cherniak C (1994) Component placement optimization in the brain. The Journal of neuroscience 14: 2418.

56. Ramón y Cajal S (1995) Histology of the nervous system of man and vertebrates. New York: Oxford University Press.

57. Cherniak C (1995) Neural component placement. Trends in neurosciences 18: 522–527.

58. Chklovskii D, Schikorski T, Stevens C (2002) Wiring optimization in cortical circuits. Neuron 34: 341–347.

59. Klyachko V, Stevens C (2003) Connectivity optimization and the positioning of cortical areas. Proceedings of the National Academy of Sciences of the United States of America 100: 7937.

60. Buzsaki G, Geisler C, Henze D, Wang X (2004) Interneuron diversity series: circuit complexity and axon wiring economy of cortical interneurons. TRENDS in Neurosciences 27: 186–193.

61. Chklovskii D, Koulakov A (2004) Maps in the brain: What can we learn from them? Annu Rev Neurosci 27: 369–392.

62. Chklovskii D (2004) Exact solution for the optimal neuronal layout problem. Neural computation 16: 2067–2078.

63. Kaiser M, Hilgetag C (2006) Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. PLoS Comput Biol 2: e95.

64. Cherniak C, Mokhtarzada Z, Rodriguez-Esteban R, Changizi K (2004) Global optimization of cerebral cortex layout. Proceedings of the National Academy of Sciences of the United States of America 101: 1081.

65. Chen B, Hall D, Chklovskii D (2006) Wiring optimization can relate neuronal structure and function. Proceedings of the National Academy of Sciences of the United States of America 103: 4723.

66. Ahn Y, Jeong H, Kim B (2006) Wiring cost in the organization of a biological neuronal network. Physica A: Statistical Mechanics and its Applications 367: 531–537.

67. Kaiser M (2007) Brain architecture: a design for natural computation. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365: 3033.

68. Budd J, Kovács K, Ferecskó A, Buzás P, Eysel U, et al. (2010) Neocortical axon arbors trade-off material and conduction delay conservation. PLoS Comput Biol 6: e1000711.

69. Pérez-Escudero A, De Polavieja G (2007) Optimally wired subnetwork determines neuroanatomy of caenorhabditis elegans. Proceedings of the National Academy of Sciences 104: 17180.

70. Koulakov A, Chklovskii D (2001) A wire length minimization approach. Neuron 29: 519–527.

71. Karbowski J (2001) Optimal wiring principle and plateaus in the degree of separation for cortical neurons. Physical review letters 86: 3674–3677.

72. Hilgetag C, Kaiser M (2008) Organization and function of complex cortical networks. In: Lectures in Supercomputational Neuroscience: Dynamics in Complex Brain. pp. 107–133.

73. Ames A (2000) Cns energy metabolism as related to function. Brain Research Reviews 34: 42–68.

74. Von Der Malsburg C (1994) The correlation theory of brain function. In: Models of neural networks II: Temporal Aspects of Coding and Information Processing in Biological Systems. Springer Verlag. pp. 95–119.

75. Destexhe A, Sejnowski T (2001) Thalamocortical assemblies. Oxford: Oxford University Press.

76. Salinas E, Sejnowski T (2001) Correlated neuronal activity and the flow of neural information. Nature reviews Neuroscience 2: 539.

77. Masuda N, Aihara K (2004) Global and local synchrony of coupled neurons in small-world networks. Biological cybernetics 90: 302–309.

78. Faisal A, Selen L, Wolpert D (2008) Noise in the nervous system. Nature Reviews Neuroscience 9: 292–303.

79. Bassett D, Bullmore E, Meyer-Lindenberg A, Apud J,Weinberger D, et al. (2009) Cognitive fitness of cost-efficient brain functional networks. Proceedings of the National Academy of Sciences 106: 11747–11752.

80. van den Heuvel M, Stam C, Kahn R, Pol H (2009) Efficiency of functional brain networks and intellectual performance. The Journal of Neuroscience 29: 7619–7624.

81. Li Y, Liu Y, Li J, Qin W, Li K, et al. (2009) Brain anatomical network and intelligence. PLoS computational biology 5: e1000395.

82. Deary I, Penke L, Johnson W (2010) The neuroscience of human intelligence differences. Nature Reviews Neuroscience 11: 201–211.

83. Dehaene S (2002) The cognitive neuroscience of consciousness. MIT Press.

84. He Y, Dagher A, Chen Z, Charil A, Zijdenbos A, et al. (2009) Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. Brain 132: 3366–3379.

85. Lynall M, Bassett D, Kerwin R, McKenna P, Kitzbichler M, et al. (2010) Functional connectivity and brain networks in schizophrenia. The Journal of Neuroscience 30: 9477–9487.

86. van den Heuvel M, Hulshoff Pol H (2010) Exploring the brain network: a review on resting-state fmri functional connectivity. European Neuropsychopharmacology 20: 519–534.

87. Sporns O (2011) The human connectome: a complex network. Annals of the New York Academy of Sciences 1224: 109–125.

88. Bassett D, Greenfield D, Meyer-Lindenberg A,Weinberger D, Moore S, et al. (2010) Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. PLoS computational biology 6: e1000748.

89. Achard S, Bullmore E (2007) Efficiency and cost of economical brain functional networks. PLoS computational biology 3: e17.

90. White J, Southgate E, Thomson J, Brenner S (1986) The structure of the nervous system of the nematode caenorhabditis elegans. Philosophical Transactions of the Royal Society of London B, Biological Sciences 314: 1.

91. Oshio K, Iwasaki Y, Morita S, Osana Y, Gomi S, et al. Database of synaptic connectivity of c. elegans for computation. Technical report, Technical Report 3, CCeP, Keio Future, Keio University,Japan, 2003. Available: http://ims.dse.ibaraki.ac.jp/research/databaseen.html.

92. Kötter R (2004) Online retrieval, processing, and visualization of primate connectivity data from the cocomac database. Neuroinformatics 2: 127–144.

93. Paxinos G, Huang X, Toga A (2008) The Rhesus Monkey Brain in Stereotaxic Coordinates. Academic Press.

94. Kaiser M, Martin R, Andras P, Young M (2007) Simulation of robustness against lesions of cortical networks. European Journal of Neuroscience 25: 3185–3192.

95. Vaishnavi S, Vlassenko A, Rundle M, Snyder A, Mintun M, et al. (2010) Regional aerobic glycolysis in the human brain. Proceedings of the National Academy of Sciences 107: 17757–17762.

96. Lo C,Wang P, Chou K,Wang J, He Y, et al. (2010) Diffusion tensor tractography reveals abnormal topological organization in structural cortical networks in alzheimer's disease. The Journal of Neuroscience 30: 16876–16885.

97. Chalfie M, Sulston J, White J, Southgate E, Thomson J, et al. (1985) The neural circuit for touch sensitivity in caenorhabditis elegans. The Journal of neuroscience 5: 956–964.

98. Schaefer A, Hadwiger G, Nonet M (2000) rpm-1, a conserved neuronal gene that regulates targeting and synaptogenesis in C. elegans. Neuron 26: 345–356.

99. Bono M, Villu Maricq A (2005) Neuronal substrates of complex behaviors in c. elegans. Annu Rev Neurosci 28: 451–501.

100. Gray J, Hill J, Bargmann C (2005) A circuit for navigation in caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America 102: 3184–3191.

101. Chronis N, Zimmer M, Bargmann C (2007) Microfluidics for in vivo imaging of neuronal and behavioral activity in caenorhabditis elegans. Nature methods 4: 727–731.

102. Wadsworth W, Bhatt H, Hedgecock E (1996) Neuroglia and pioneer neurons express unc-6 to provide global and local netrin cues for guiding migrations in c. elegans. Neuron 16: 35–46.

103. Kaiser M, Hilgetag C (2004) Spatial growth of real-world networks. Physical Review E 69: 036103.

104. Felleman D, Van Essen D (1991) Distributed hierarchical processing in the primate cerebral cortex. Cerebral cortex 1: 1.

105. Markov N, Ercsey-Ravasz M, Gomes A, Lamy C, Magrou L, et al. (2012) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cerebral Cortex. In press.

106. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes. 3rd edition. Cambridge University Press Cambridge.

107. Newman M (2004) Fast algorithm for detecting community structure in networks. Physical Review E 69: 066133.

108. Leicht E, Newman M (2008) Community structure in directed networks. Physical Review Letters 100: 118703.

109. Good B, de Montjoye Y, Clauset A (2010) Performance of modularity maximization in practical contexts. Physical Review E 81: 046106.

110. Bassett D, Porter M, Wymbs N, Grafton S, Carlson J, et al. (2012) Robust detection of dynamic community structure in networks. Chaos: In press. arXiv preprint arXiv:12064358.

111. Clauset A, Shalizi C, Newman M (2009) Power-law distributions in empirical data. SIAM review 51: 661–703.