THE ROYAL SOCIETY

Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat

Claus-C. Hilgetag[1]*, Gully A. P. C. Burns[2], Marc A. O'Neill[1], Jack W. Scannell[1] and Malcolm P. Young[1]

[1]*Neural Systems Group, Department of Psychology, University of Newcastle upon Tyne, Ridley Building, Newcastle upon Tyne NE1 7RU, UK*

[2]*University of Southern California Brain Project, University of Southern California, Hedco Neuroscience Building, Los Angeles, CA 90007, USA*

The number of different cortical structures in mammalian brains and the number of extrinsic fibres linking these regions are both large. As with any complex system, systematic analysis is required to draw reliable conclusions about the organization of the complex neural networks comprising these numerous elements. One aspect of organization that has long been suspected is that cortical networks are organized into 'streams' or 'systems'. Here we report computational analyses capable of showing whether clusters of strongly interconnected areas are aspects of the global organization of cortical systems in macaque and cat. We used two different approaches to analyse compilations of corticocortical connection data from the macaque and the cat. The first approach, optimal set analysis, employed an explicit definition of a neural 'system' or 'stream', which was based on differential connectivity. We defined a two-component cost function that described the cost of the global cluster arrangement of areas in terms of the areas' connectivity within and between candidate clusters. Optimal cluster arrangements of cortical areas were then selected computationally from the very many possible arrangements, using an evolutionary optimization algorithm. The second approach, non-parametric cluster analysis (NPCA), grouped cortical areas on the basis of their proximity in multidimensional scaling representations. We used non-metric multidimensional scaling to represent the cortical connectivity structures metrically in two and five dimensions. NPCA then analysed these representations to determine the nature of the clusters for a wide range of different cluster shape parameters.

The results from both approaches largely agreed. They showed that macaque and cat cortices are organized into densely intra-connected clusters of areas, and identified the constituent members of the clusters. These clusters reflected functionally specialized sets of cortical areas, suggesting that structure and function are closely linked at this gross, systems level.

Keywords: corticocortical connections; anatomical connectivity; evolutionary optimization; cost functions; non-parametric cluster analyses; multidimensional scaling

1. INTRODUCTION

A complex network of anatomical fibres links areas and nuclei in the brain. This network largely determines each region's inputs and its outputs, and so its structure frames the contribution any region can make to information processing in the whole nervous system, and hence presumably to behaviour. Neuroanatomical experiments have generated a wealth of information about the existence or absence, and sometimes also about finer details, of connections linking particular cortical areas with others. Reliable information on how neural systems are organized, however, cannot be gained simply by inspecting these complex data, in the same way that conclusions drawn from inspection of complex spike raster patterns without proper analysis would be prone to error. Data analysis is required. In previous studies (e.g. Young 1992, 1993; Scannell *et al*. 1995; Young *et al*. 1995) we have used simple inferential statistics, multidimensional scaling, matrix transforms and seriation algorithms to investigate the general topology of neural systems in primates and cats, and we have also used computational methods to implement hierarchical analysis (Hilgetag *et al*. 1996) to study other aspects of neural systems' organization.

An influential idea in cortical neuroscience has been that the cortex is not homogeneous but divided into 'streams', 'systems' or 'clusters' of areas, which are considered to be related functionally and anatomically. These three terms have been used somewhat interchangeably. Although these aspects of organization have long been argued to be apparent in the primate visual system (e.g. Ungerleider & Mishkin 1982; Goodale & Milner 1992),

*Author and address for correspondence: Boston University School of Medicine, Department of Anatomy and Neurobiology, 700 Albany Street W746, Boston, MA 02118, USA (claush@bu.edu).

the existence of streams in cortical systems has been debated repeatedly (e.g. Simmen *et al*. 1994; Young *et al*. 1994, 1995; Goodhill *et al*. 1995).

These debates have often focused on whether there are or are not distinguishable streams within the primate cortical visual system (Merigan & Maunsell 1993). We have previously tried to bring a quantitative approach to this issue, and in doing so, have examined whether the incidence of connections is or is not as predicted under the competing hypotheses (Young *et al*. 1995). If the system is split, for instance, then the elements of the putative dorsal stream should be significantly more connected with their associates than with the elements of the ventral stream, and vice versa. Furthermore, there should also be significantly more confirmed-absent connections between dorsal and ventral areas than within each of these groupings. In both analyses, a simple χ^2-test rejected the null hypothesis decisively, indicating that the system is indeed divided (Young *et al*. 1995).

The requirement to define in what way real data might quantitatively support or deny these hypotheses forced us to specify an explicit definition of a connectionally differentiated cluster (or system or stream). According to our definition, a cluster is a set of structures that are more connected among themselves than they are with any other structures, and more disconnected from other structures than they are among themselves. This definition is sufficiently general that it could apply within or without the primate visual system, for which it was defined. In the specific case of the primate visual system, however, while the χ^2-tests show that the incidence of connections is as predicted by the preclassification of structures into two streams, the analysis does not explore the possibility that better classifications of the structures into other clusters could be found. Hence, the χ^2-analysis cannot itself rule out the possibility that there are more than two streams within the system. We were interested in whether clusters of areas that optimally fit the explicit criteria could be found computationally, both in the primate visual system and in cortical systems in general.

We developed two independent approaches to explore this question. In the first, optimal set analysis (OSA), clusters of areas that optimally fit the explicit criteria were sought computationally by an optimization algorithm that used an explicit cost function derived from our cluster definition. We believe that this approach is novel, and that it may be useful for other data-analytical problems involving complex networks in other disciplines. In the second approach, we combined existing data-analysis methodologies, and used non-parametric cluster analysis (NPCA) to produce automatic classifications of groups of structures by examining relationships in the connection data. Some of the results have been reported previously in abstract form (Hilgetag *et al*. 1998).

2. METHODS

(a) *Data*

We analysed information from four different sets of cortical connectivity data: (i) a global compilation of primate cortical connectivity (Young 1993) with 834 connections between 73 areas; (ii) data for 319 interconnections of 32 areas in the macaque visual system (Felleman & Van Essen 1991) (this set also included explicit information about 382 potential links between areas that have been reported absent (Felleman & Van Essen 1991)); (iii) connectivity data for the somatosensory-motor system in the same species (66 connections between 15 areas, Felleman & Van Essen 1991); as well as (iv) a global collation of cat cortical connectivity (892 interconnections of 55 areas, see figure 1). The first three data sets comprised information about the existence or absence of a particular connection. The cat data also included ordinal information about the relative connection densities. The collation of cat cortical data was developed from the data set described in Scannell *et al*. (1995) and forms part of a larger database of thalamocortical connectivity of the cat (Scannell *et al*. 1999). This set of connectivity data is available for detailed scrutiny at http://www.flash.ncl.ac.uk/ptrs/cor_thal.html.

In the following sections, we describe the two analytical strategies that we used to delineate the organization of connectionally defined clusters in the mammalian cortex.

(b) *Optimal set analysis*

We developed a new optimization method (OSA) akin to one we previously employed in the analysis of hierarchical relationships in the cat and macaque cortex (Hilgetag *et al*. 1996; Hilgetag, O'Neill & Young, this issue). In our approach, we specified a cost function that evaluated the overall cluster arrangement of cortical areas. The cost was then optimized, by rearranging the network of areas, using a modified simulated annealing algorithm (Laarhoven & Aarts 1987). Our method is, generally speaking, a stochastic optimization technique. Stochastic optimization is an efficient approach for circumnavigating the problems associated with evaluating complex, inconsistent and incomplete data sets by means of simple integer cost functions (Hilgetag *et al*. 1997). It can find optimal solutions when the solutions space is too large for exhaustive search and inaccessible to analytical methods. Such conditions presented themselves here. As regards the size of the search space, sets of 73 areas (as in the data for the primate cortex) can be repartitioned into different subsets (that is, clusters) in more than $\lambda = 10^{77}$ ways, according to the formula

$$\lambda = \sum_{m=1}^{n} S_n^{(m)} = \sum_{m=1}^{n} \left[\frac{1}{m!} \sum^{m} (-1)^{m-k} \binom{m}{k} k^n \right]. \quad (1)$$

$S_n^{(m)}$ is a Stirling number of the second kind describing all possible ways of partitioning a set of n elements (here $n = 73$) into m non-empty subsets (Abramowitz & Stegun 1972), m ranges from 1 to the largest possible number of clusters, n; and k is an index variable. There are fewer possibilities for partitioning the smaller visual, somatosensory and cat connectivity data sets with $n_{vis} = 32$, $n_{som} = 16$, and $n_{cat} = 55$ areas, but these data still allow huge numbers of potential candidates, in the order of $\lambda_{vis} > 10^{26}$, $\lambda_{som} > 10^{10}$ and $\lambda_{cat} > 10^{53}$, respectively. Search spaces of these magnitudes present insurmountable problems for exhaustive approaches.

We defined a two-component integer cost that allowed us to evaluate if a particular grouping of the areas represented a 'good' overall cluster arrangement, according to our definition of clusters as sets of areas that are much more connected with each other than with other cortical areas, and much less connected with other areas than among themselves. Such a concept is mirrored by recent definitions given by other authors (e.g. Tononi *et al*. 1998). A particular partitioning of areas into clusters was considered optimal, if it fulfilled the following two conditions:

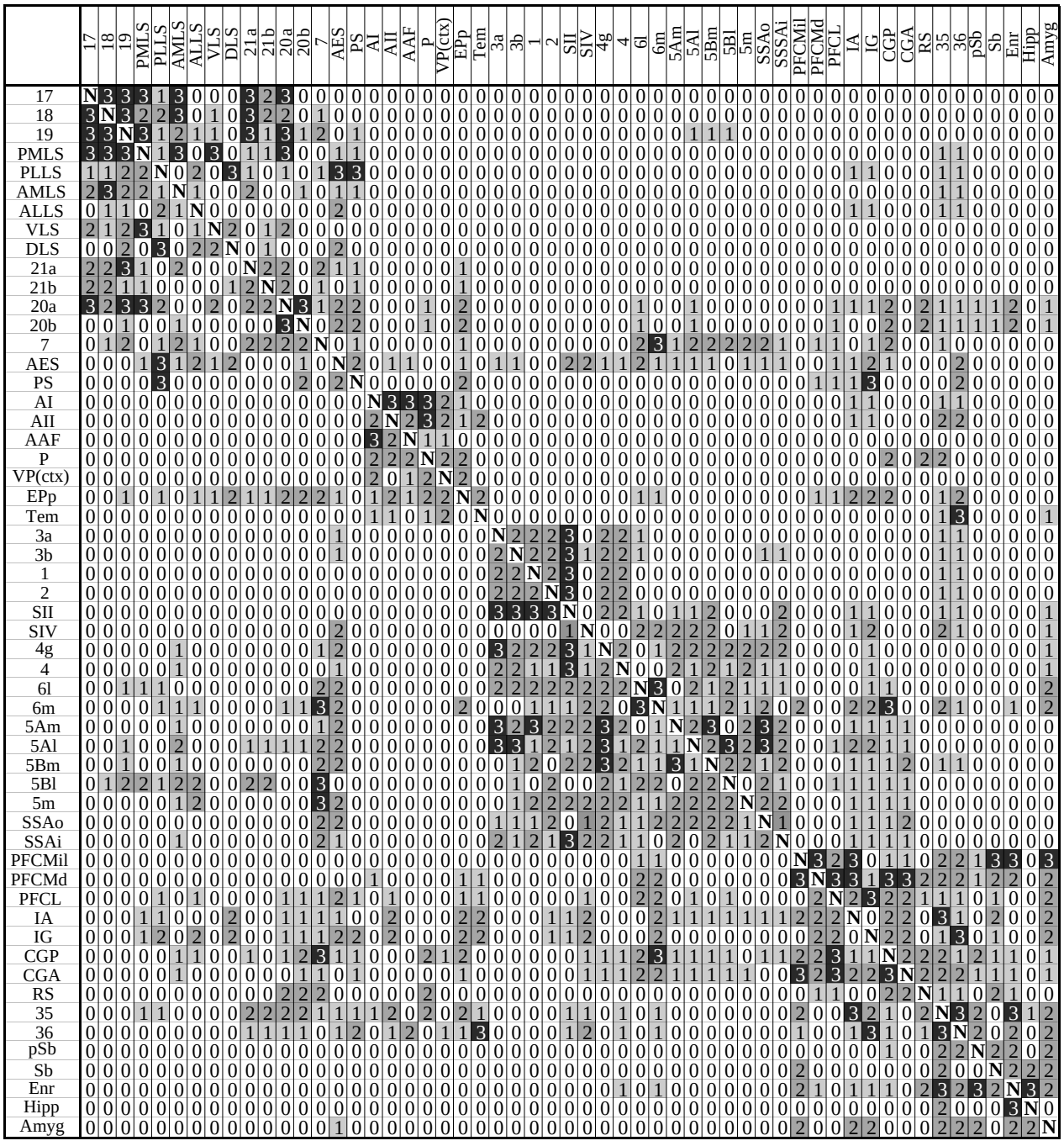


Figure 1. Cortical connectivity of the cat. The entries indicate: 0, absent or unknown connection; 1, weak connection; 2, moderately strong connection; 3, strong connection; N, intrinsic connections of the areas, which are not considered here. These data are part of a larger compilation of cat thalamocortical connectivity (Scannell *et al*. 1999). For abbreviation of area names see Scannell *et al*. (1995).

(i) CO1 ('attraction'): there exist as few as possible connections between all different clusters;
(ii) CO2 ('repulsion'): there are as few as possible known absent connections within all clusters.

Condition CO1 on its own would be perfectly fulfilled by an arrangement that collects all individual areas into one universally inclusive cluster, which would allow no connections at all between different clusters. This demonstrates that trying to satisfy condition CO1 tends to reduce the overall number of clusters, for which reason we termed it the 'attraction' component. On the other hand, condition CO2 alone can be perfectly satisfied by an arrangement in which every individual area forms a separate cluster. Therefore, this condition tends towards the subpartitioning of clusters, and we termed it the 'repulsion' component. Minimizing the attraction and repulsion conditions together can lead to a number of clusters that lies between just one and the number of individual areas, depending on the distribution of connections among the areas and the weights that were assigned to the two cost components (see § 3).

The evolutionary algorithm we used for minimizing the cost function is illustrated in figure 2, and is now described.

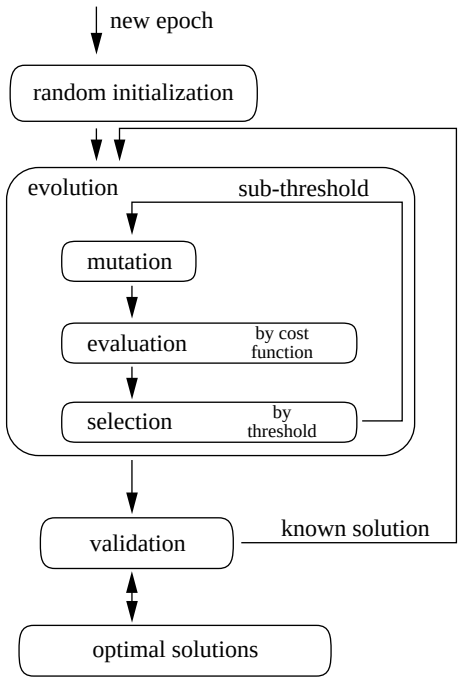


Figure 2. Evolutionary optimization algorithm. The diagram outlines the logical flow in the algorithm used for optimal set analysis. The two-component cost function used for the evaluation of candidate area arrangements is based on an explicit cluster definition (equation (3)). See text for details of the single steps in the procedure.

(i) *Initialization*

The computation of optimal arrangements proceeded in epochs that began with an arbitrary assignment of all areas to randomly created clusters. Typically, 50 epochs, each with different initial random configurations, were computed through, and up to 50 different optimal solutions could be collected per epoch. Generally, however, the number of optimal arrangements found per epoch was smaller than this (see § 3).

(ii) *Evolution*

The processor automatically determined the composition and the number of clusters in the optimal solutions by means of random, evolutionary modifications. Cluster arrangements were modified by two different kinds of mutations: an area could be assigned to a different cluster, which already existed or was newly created by this mutation. Alternatively, areas in two different clusters could be swapped. These simple modifications ensured that descendant structures varied little from their parents, so that the search space was examined at high resolution. The search could be restricted to optimal arrangements with a given number of clusters, if required.

The value of a candidate structure produced during an epoch was evaluated in integer units according to the two-component cost function set out above. For the data sets that contained only information about the existence or absence of connections, each single deviation from a perfect cluster arrangement was counted as one unit. The cat data set was evaluated in two different ways. First, the data were just considered as information about existing or absent connections ('binary classification'). These data were then analysed in the same way as the data on all other cortical systems. Second, information about the three different strengths of existing connections was incorporated ('quaternary classification'). In this approach, departures from a perfect arrangement under the attraction part of the cost function were counted in the respective units of connection densities: 1 (for weak), 2 (for intermediate), or 3 (for dense). Where this was possible on the basis of the available data (that is, for the primate visual data), we distinguished between connections reported absent and those hitherto unreported, excluding the latter ones from the analysis. For all other data sets, unknown and explicit absent connections were treated as equal and were thus counted in the repulsion component.

Additionally, the attraction and repulsion components of the cost function were multiplied with weighting factors, w_{attr} and w_{rep}, so that the total cost, C_{total}, resulted as

$$C_{\mathrm{total}} = w_{\mathrm{attr}} \times C_{\mathrm{attr}} + w_{\mathrm{rep}} \times C_{\mathrm{rep}}, \tag{2}$$

and

$$C_{\mathrm{attr}} = \sum m_{ab}, \forall a, b, A, B\colon a \in A, b \in B, A \cap B = 0, m_{ab} = \{1|2|3\}, \tag{3a}$$

$$C_{\mathrm{rep}} = |\{m_{ab}\}|, \forall a, b, A\colon a \in A, b \in A, m_{ab} = \{0\}, \tag{3b}$$

$$m_{ab} \in M, a \neq b.$$

Hence, C_{attr} was computed as the sum of all positive connections strengths, m_{ab}, for all the connections (a, b) that ran between all non-overlapping clusters A and B; whereas C_{rep} was computed as the (cardinal) number of the all absent connections within all clusters. The connection strengths m_{ab} are elements of the connectivity matrix M.

We computed series of different optimal solutions by varying either of the weights of attraction or repulsion, w_{attr} and w_{rep}, in the combined cost function. Weights for the repulsion component were increased in steps of one. For the cat connectivity data, the attraction weights were increased at steps of one-third, to obtain a sufficiently fine transition between differently weighted cluster solutions, despite the influence of the individual interaction strengths. The attraction weights in all other data sets were increased in steps of one.

The weights in the series ranged from $w_{\mathrm{attr}} = 1$ to maximally $w_{\mathrm{attr}} = 40$ (while w_{rep} being kept as 1), and from $w_{\mathrm{rep}} = 2$ to maximally $w_{\mathrm{rep}} = 30$ (with w_{attr} kept constant as 1). Higher w_{attr} led to fewer, bigger clusters (in the limit case to just one), whereas higher w_{rep} led to more and smaller clusters.

(iii) *Selection*

If a descendant structure possessed a cost that was not higher than 125% of the parental cost, it was considered for further optimization, provided that in turn a (secondary) descendant structure could be found within the same cost limit. This upper threshold value was established experimentally. We found that for values above 135% for the selection threshold, the algorithm did not descend to lower cost arrangements efficiently. A selection threshold of higher than 100% allowed the algorithm to withdraw from local minima. If the requirements for the descendant cost could not be met, the algorithm tracked back and introduced a different mutation.

(iv) *Validation*

A computed structure was accepted as new, and kept in the set of solutions, if (i) it possessed a cost that was lower than, or not more than 1% higher than, the cost of any solution encountered up to that moment; and (ii) was different from all other stored solutions. We programmed the processor to collect optimal solutions within a 1% band of the lowest cost found, because relatively few optimal solutions were detected, and their

costs were not so low that it seemed acceptable to ignore structures of only slightly higher cost.

Often, more than one optimal arrangement was found by this procedure. To summarize all optimal structures, we used a cluster-count representation as described in § 2(d). The optimization algorithm was programmed in ANSI C/C++ and ran on a DEC ALPHA 3000–600 UNIX workstation with 256 Mb of RAM. Central processing unit time for completing 50 optimization epochs for one data set ranged from several hours to a few days, depending on the size and density of the connectivity data set being analysed.

To restate the goal of our approach, we applied evolutionary optimization techniques to investigate the organization of networks. This, however, does not mean that we assume that these biological networks necessarily possess any clustered aspect to their organization, or that existing clusters should be optimally arranged according to the criteria we use. However, if cortical systems actually do possess a clustered organization, then the object of our approach is that the clusters evolved by our processor should optimally reflect the cluster structure inherent in the real brain.

(c) *Multidimensional scaling and non-parametric cluster analyses*

In an alternative approach, we employed statistical cluster procedures (NPCA) that detect spatial relationships among distributions of points in metric coordinate systems. For this analysis, the ordinal connectivity data which represent pairwise anatomical relationships between areas had to be converted into metric distances that preserve as closely as possible the connectional relationships between areas. We used non-metric multidimensional scaling (NMDS) (e.g. Kruskal 1964*a*,*b*) for this transformation. The combined approach of NMDS and NPCA is described in detail in the companion paper by Burns & Young (this issue), and we refer interested readers to this description.

Following Burns & Young (this issue), we generated two- and five-dimensional (2D and 5D) NMDS configurations for all given connectivity matrices, using the FIT = 1, 2 and 0.5 cost functions under the tied and untied approaches in the SAS/STAT MDS procedure (SAS Technical Report P-229 SAS/STAT Software; SAS Institute, Inc. 1990). We used the SAS MODECLUS function for NPCA with significance testing. Ten separate MODECLUS jobs (using the standard algorithm, method = 1) were run on each configuration according to four different kernel paradigms; see Burns & Young (this issue) for details of the paradigms and parameter settings. The results from these analyses were summarized in cluster-count matrices, as described in the following section.

(d) *Representation and validation of results*

To compare the results from the different analysis approaches, we defined a common format for the representation of the obtained results. Several individual results were grouped together in symmetrical 'cluster-count' matrices, K, to demonstrate consistent features across the different runs within the same analysis approach. The cluster-count matrices showed the relative frequency by which any two areas were found in shared clusters in different runs, and thus indicated the association of the respective areas. The order of areas within each cluster-count matrix was re-sorted in such a way that areas were followed by other areas with which they had the highest association possible. This procedure tended to group together the main clusters and so allowed easier interpretation of the overall cluster structure. The association values in the matrices were shaded according to a greyscale distribution from white (for 0% association) to black (100%).

Additionally, we transformed some of the matrices K into cluster trees to show the overlap and consistency of the cluster arrangements for different significance thresholds. In this representation, areas were given a lighter shading, starting from the significance threshold at which they completely disconnected from any other cluster. To investigate the linkage of the OSA clusters, we also computed matrices that indicated the relative frequency of existing connections between clusters and non-existing connections within clusters for all pairs of areas across all optimal solutions. From these matrices we produced diagrams to show the areas that most frequently contributed to the attraction or repulsion cost. All routines for the representation of cluster-count matrices or trees were written in Visual-Basic for Excel (v. 7).

For the OSA approach, cluster-count matrices summarized the multiple optimal configurations obtained for specific settings of the attraction and repulsion cost weights. The settings most straightforward to interpret were the 'balanced' case (attraction and repulsion weights equal, and equal to one) and the 'high-repulsion' condition, where all non-existing connections have been moved outside of the clusters due to sufficiently high-repulsion weight settings. We also computed 'averaged' cluster-count matrices which summarized the optimal configurations obtained for all weight settings between 'high repulsion' and 'high attraction' (the setting which led to all areas being assembled in one big cluster).

In the case of the NMDS–NPCA analyses described in § 2(c), cluster counts were computed for method-specific subtypes of the analyses. In particular, for the dimensionality of the NMDS representations (2D or 5D), the NPCA kernel paradigms (fixed-radius or nearest-neighbour) and, in the case of the cat connectivity data, for the different classifications of connection strengths (binary or quaternary classification). We also computed matrices summarizing all different conditions. Within the averaged counts, the cluster counts of each non-parametric cluster tree were weighted equally so each paradigm or parameter setting contributed by the same amount to the overall cluster-count score. The association values accounted for the significance of clusters. If a cortical area had been in a cluster that was subsequently dissolved under the SAS JOIN option, the area would not contribute further to the cluster count (that is, it would not be counted as being in the same cluster as any other structures, including itself). This explains how some representations resulting from the NPCA indicate associations between an area and itself of less than 100%.

We compared the results of the different approaches with each other, and against the background of randomly created data. The random data were produced by randomly reshuffling entries in the connectivity matrices, thus preserving the distribution of existing and absent connections as well as the strengths of connections, and by subsequent OSA of the reshuffled matrices (treating a population of 20 random matrices, under the 'balanced' weights setting). OSA and NPCA results were compared by correlation analysis. The respective cluster-count matrices were rearranged so as to correspond to each other in the same order of areas (the 'balanced' OSA results were considered the standard for a particular data set), and R- and R^2-values for all $n \times n$ corresponding entries in the matrices K_{OSA} and K_{NPCA} were computed, using Excel's CORREL function.

Table 1. '*Small-world*' *coefficients for cortical connectivity networks of the cat and the macaque monkey*

(The indices were computed according to the definitions given in (Watts & Strogatz 1998). Values for $p(L)$ and $p(C)$ are the probabilities that the coefficients for the real connectivity data stem from the respective populations of random values. The probabilities indicate that the coefficients for the real data are highly significant.)

	L_{real}	μL_{ran}	δL_{ran}	$p(L)$	C_{real}	μC_{ran}	δC_{ran}	$p(C)$
macaque: visual	1.69	1.65	0.005	8×10^{-15}	0.594	0.321	0.009	0
macaque: somatosensory	1.77	1.72	0.044	0.244	0.569	0.312	0.04	8×10^{-11}
macaque: whole cortex	2.18	1.95	0.005	0	0.49	0.159	0.005	0
cat: whole cortex	1.79	1.67	0.002	0	0.602	0.302	0.006	0

The R^2-values represented the percentage of variability in the balanced OSA results that accounted for the variability of the result sets under investigation.

3. RESULTS

Before entering the detailed analyses described in the previous sections, we sought preliminary but independent assurance that cortical connectivity implies a clustered cortical organization. This assurance can be derived from computing network indices that have recently been proposed by Watts & Strogatz (1998) to investigate sparsely connected networks. Table 1 presents local clustering coefficients, C, and characteristic path lengths, L, for the actual anatomical connectivity data and for populations ($n=20$) of the same, but randomly reshuffled data sets, respectively.

The coefficient L, which stands for the characteristic path length of the actual and randomly rearranged connectivity data, indicates the average shortest path, measured as the number of intermediate areas, between any two areas. We return to this feature of cortical connectivity in § 4.

The clustering index C describes the number of connections, out of all those possible, that have been reported to exist among the neighbours of any node in the network, averaged over all nodes. This measure informs on the 'cliquishness' of all local neighbourhoods in a network (Watts & Strogatz 1998). For completely randomly connected networks, C is close to the number of actual connections in the network, normalized by the number of connections that could be realized between all network nodes. Comparing the entries in columns 6–8 of table 1, which present the indices of the real connectivity data and the mean, μ, and standard deviation, δ, for the random data demonstrates that the local clustering of the real data is significantly greater than that of randomly reshuffled data. These results underline the fact that the actual connectivity data clearly possess dense local clusters, with an average local connectivity nearly twice as high as the average global connectivity among the areas. The nature and membership of these local clusters can be determined from the results of the two varieties of cluster analyses we developed, and we now turn to a detailed description of these results.

(a) *Macaque cortex—OSA*

Optimal set analysis of the whole macaque connectivity data set yielded clusters shown in figure 3*a*. This diagram represents 50 different optimal cluster arrangements that resulted from optimizing the cost in equation (3) with identical weights for attraction and repulsion ($w_{attr}=w_{rep}=1$). We refer to this setting as the 'balanced' condition. The optimal arrangements for this condition consisted of 22 to 25 clusters and carried a combined cost of 628–632. The cost component for connections between clusters ranged between 545 and 568, while the component for non-existing connections within the clusters varied between 60 and 78. Many clusters in the optimal solutions contained only very few areas, or consisted of individual areas. Five larger clusters, however, also appeared consistently (with a relative frequency of $\geqslant 60\%$) throughout the optimal arrangements. These clusters were formed by the areas: (i) AITv, Al3, ER, TSl, TS2, TGd, A35, TGv, the amygdala; (ii) LIP, VIP, DP, PO, FST, MSTd, MSTl, FEF; (iii) PIP, V3A, MT(V5), Vl, V2, V3, V4, VP; (iv) A3, A5, A6, Rl, S2, TPt, A7b; and (v) TF, A9, TH, A23, A24, TS3, STPa, A7a, STPp, A46. At higher relative attraction weights, these clusters, as well as individual areas, joined up to form fewer, larger groups. A small cluster that was consistently separated from the other ones and persisted after the remaining clusters joined up was formed by areas PAL, PAAR, PROA, REIT, KA and PAAC.

Figure 3*b*,*c* shows the areas that were the main contributors to the attraction and repulsion costs under the balanced condition. Figure 3*b* presents the 20 leading originators or recipients of connections running between optimal clusters. Such areas could be considered important nodes of communication in the network of area clusters. The averaged frequency of inter-cluster connections for all 50 optimal configurations was normalized by the absolute number of connections an area sends or receives. The diagram therefore indicates the percentage of connections, of all those an area possesses, that link the area to other areas in different clusters. Figure 3*c* shows the 20 areas contributing most to the cost of non-existing connections within clusters. Again the frequency of connections is averaged over all 50 optimal solutions and normalized by the total number of non-existing connections the areas possess. The diagram demonstrates that the relative frequency of the area's non-connections within its cluster is small compared to the frequency of connections it makes with other clusters. The latter component would therefore be the more influential factor in the optimizations that determined the overall configuration of clusters.

All areas were grouped together in a single cluster at an attraction weight of more than 40. On the other end of the spectrum, all non-existing and unknown connections were located outside of the clusters for all optimal

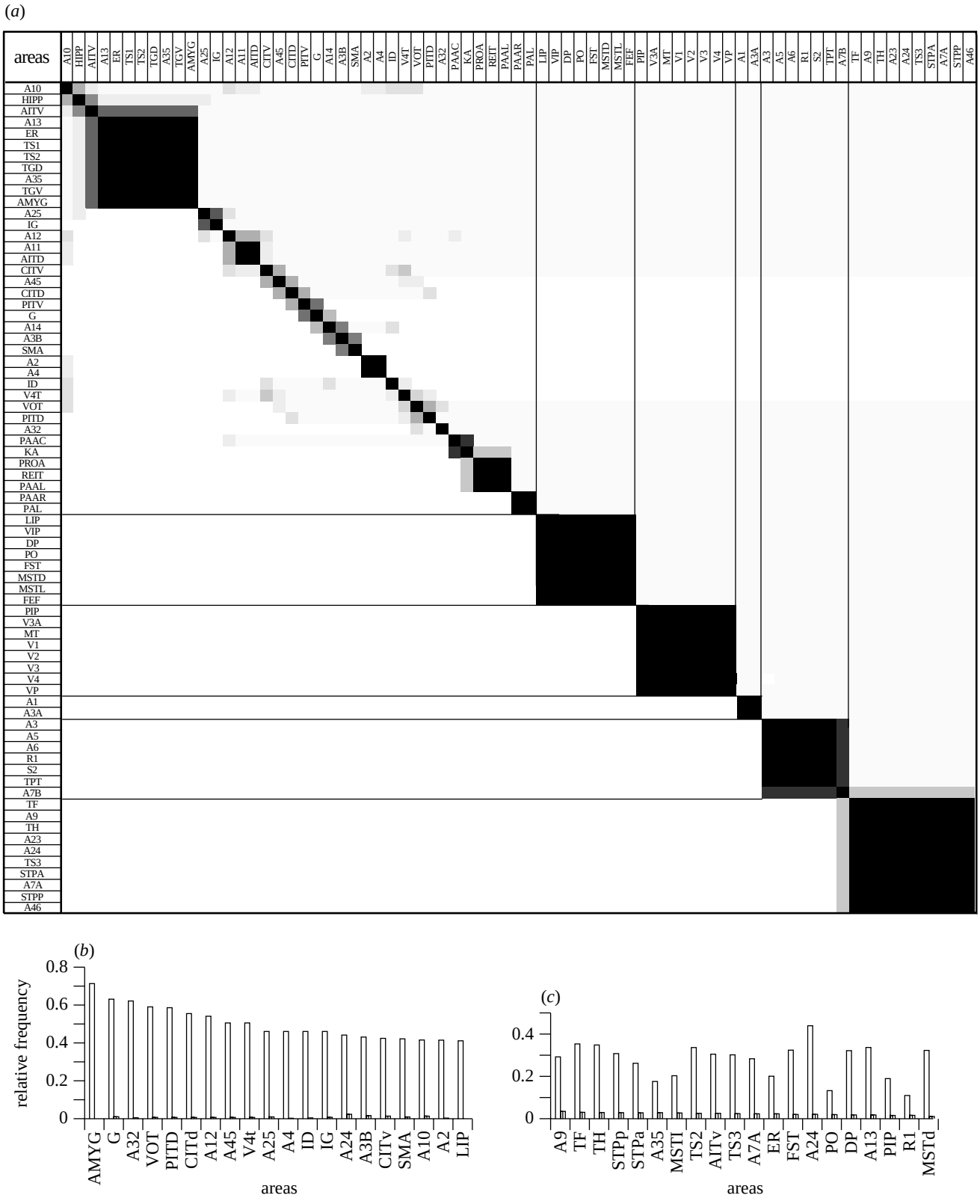


Figure 3. (*a*) Global optimal sets of areas for the primate cortex, derived with balanced attraction and repulsion weighting of the cluster-cost components. Auxiliary lines are plotted to facilitate the recognition of area assignments to clusters. (*b*) Top 20 areas that were most strongly contributing to inter-cluster connectivity in all optimal configurations summarized in figure 3*a*. The frequencies of originating or received connections were normalized by the number of optimal solutions (50) and by the total number of connections the respective area possesses. Frequencies of inter-cluster connections are given by lighter bars, frequencies of absent connections within the clusters by darker ones. (*c*) Top 20 areas ranked according to their contribution to repulsion by non-existing connections within the clusters in the optimal arrangements in figure 3*a*. The connection frequencies are normalized by the number of optimal solutions and all non-existing connections an area is involved in. Since the frequencies of absent connections were comparatively small, it was mainly the structure of existing connections that determined the optimal organization of the overall arrangement.

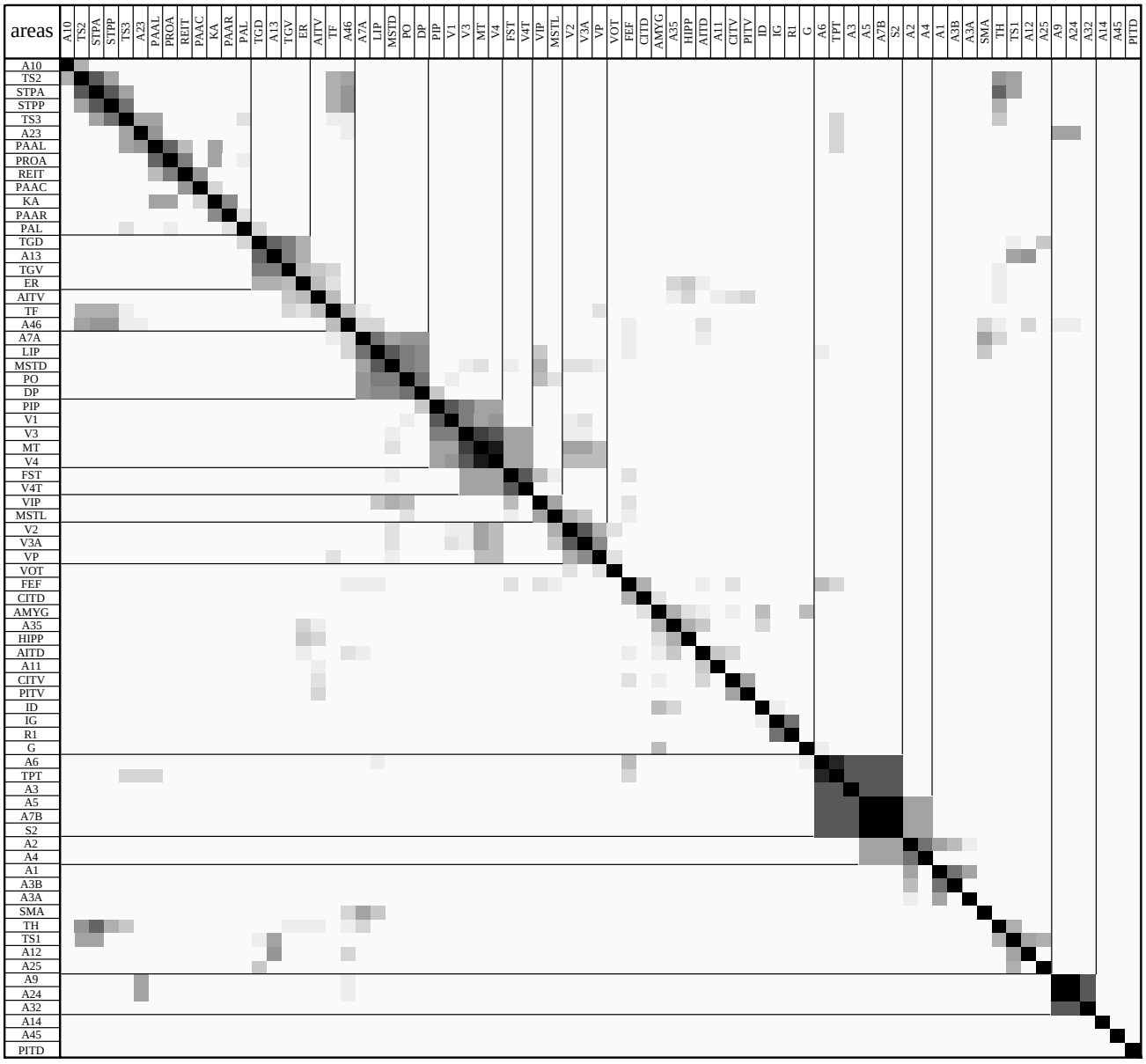


Figure 4. Summary of 50 optimal 'building-block' configurations for the macaque cortex for high-repulsion weight setting ($\geqslant 6$).

arrangements at repulsion weights equal to, or higher than, six. Such 'high-repulsion' settings produced 50 optimal solutions with 37 to 41 clusters, which were interconnected with 692 to 696 connections. Figure 4 shows a summary of these optimal configurations. As there were no established absent connections left within the clusters, these represent the most compact and irreducible groupings of areas possible, and we have previously termed such groups 'building blocks' (Hilgetag *et al.* 1998). Building blocks that consistently appeared in these solutions were: (i) TF, STPa, A7a, STPP, A46; (ii) V4, V3A, Vl, V2, V3; (iii) LIP, VIP, MT, FST, MSTd, FEF; (iv) Al3, ER, TSl, TGd, TGv; (v) A5, Rl, A7b, S2, TPt; (vi) A9, A24, A32; and (vii) ID, IG, AMYG.

Because areas were joined in one inclusive cluster only for very large attraction weights (>40), representations that summarized over all the weight settings were strongly skewed towards the configurations resulting from high atttraction weights. For this reason we refrained from presenting such a global summary for this data set.

(b) *Macaque cortex—NMDS–NPCA*

We analysed the data set examined by OSA in the § 3(a) with the non-parametric clustering approach, as described in § 2. We then computed cluster-count summaries across the various parameter combinations for this approach, and compared these summaries against the cluster-count plot of configurations obtained for the balanced OSA approach. The correlation coefficients for comparing with the different NMDS–NPCA summaries ranged between $R = 0.41$ and $R = 0.47$. The agreement between the independent analysis approaches was, therefore, only slightly worse than for different conditions within the OSA approach itself, and much better than for random data, see § 3(h).

(c) *Macaque visual system—OSA*

In the case of the primate visual system, we obtained six different optimal solutions, at a total cost of 135, with the attraction cost component ranging between 77 and 90 and the repulsion cost between 42 and 58 (for atttraction

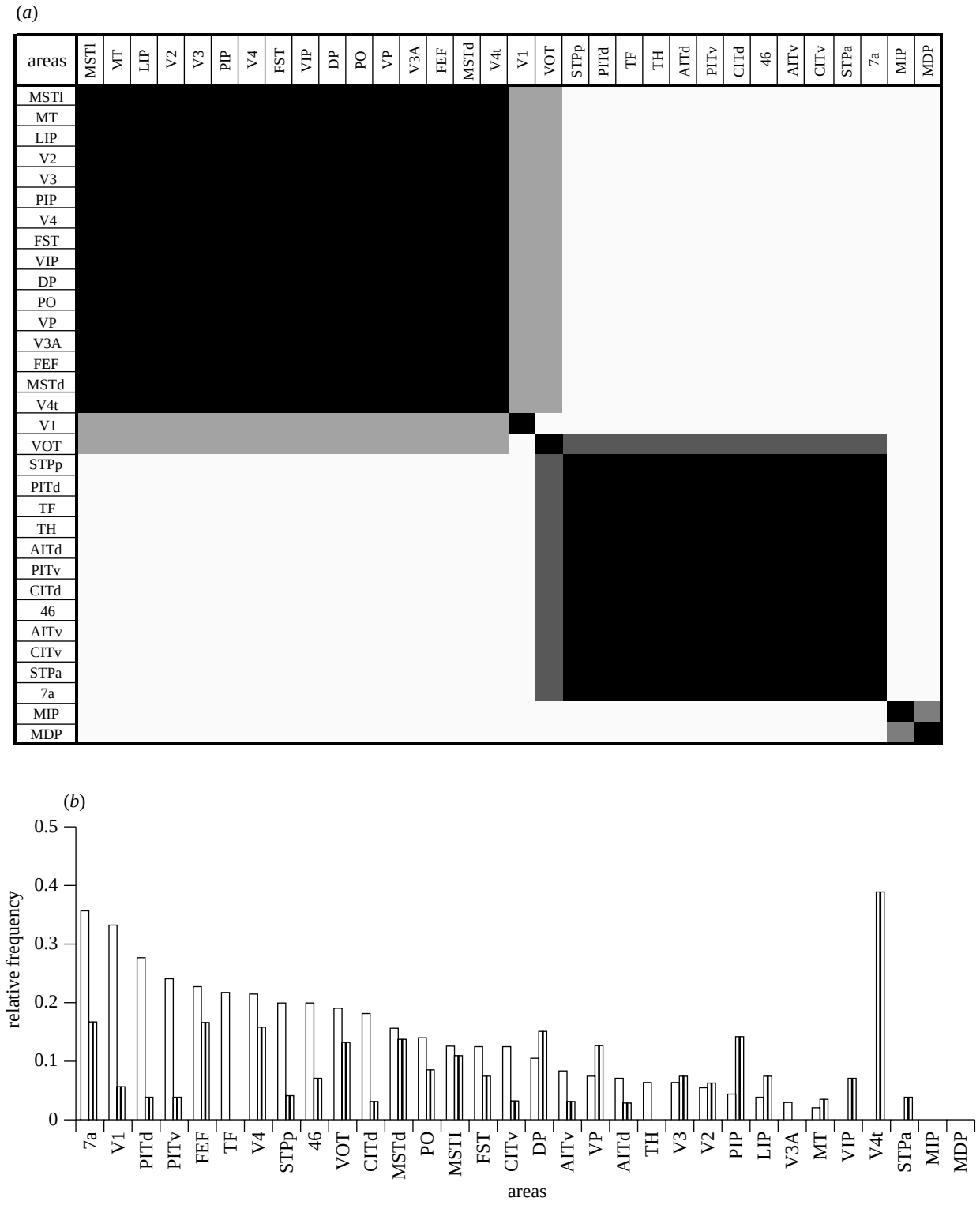


Figure 5. (*a*) Optimal sets of primate visual areas for balanced attraction and repulsion weighting. This summary is based on six optimal area arrangements. (*b*) Ranking of visual areas according to their contribution to inter-cluster connections in the arrangements shown in figure 5*a*. The relative frequencies are normalized in the same way as those in figure 3*b*. The figure also indicates the normalized frequencies of any non-existing linkages that the areas were involved in. The latter ones are given in darker bars, whereas frequencies of inter-cluster connections are shown in lighter shading.

and repulsion weights both set to one). The arrangements in these solutions, depicted in figure 5*a*, show a clear separation of the visual areas into two main clusters. The first group comprised areas MSTl, MT, LIP, V2, V3, VP, V3A, PIP, V4, V4t, FST, VIP, DP, PO, FEF and MSTd, and the second STPp, PITd, TF, TH, AITd, PITv, CITd, 46, AITv, CITv, STPa and 7a. Primary cortical area V1 associated with the first group in one-third of the solutions and formed a separate cluster in the others. The only area with an alternating preference for the two groups was area VOT, which appeared in the first cluster in one-third of the solutions and in the second cluster in

the remaining solutions. Areas MIP and MDP formed another separate cluster, which was split in 50% of the solutions. Little is known, however, about the connectivity of these areas. Only two efferents are listed for each of the areas by Felleman & Van Essen (1991). We consequently considered separate clusters formed by MIP and MDP to be artefactual.

Figure 5*b* shows the contribution of the individual areas to connections between the clusters, and non-connections within the clusters, in all optimal configurations. These relative frequencies are again normalized by the total number of connections and non-connections an area possesses. Low relative frequencies for the inter-cluster connections, of less than 0.5, indicate that the areas are relatively closely integrated in their respective clusters, as they form most of their existing connections within the clusters.

For increasingly higher attraction weights, first V1 and VOT joined the first of the clusters given above, while at attraction weights of 17 and above, a single cluster of all 32 areas emerged as the single optimal solution. This solution carried a repulsion cost of 382, that is, identical to the number of established non-existing connections.

The number of clusters in the optimal arrangements increased for higher repulsion weights. At weights equal or larger than seven for this part of the cost function, the repulsion cost was zero, indicating that all non-existing connections were located outside the clusters. A further increase of the weight for this cost component would, therefore, not lead to different optimal arrangements, and the resulting high-repulsion weight clusters represent building blocks of the visual system's connectional organization. Figure 6*a* summarizes the six optimal configurations that resulted from a repulsion weight of seven. These solutions consisted of six to eight clusters and possessed attraction costs of 192 or 193. The main groups of areas, as outlined by figure 6*a* are: (i) STPp, STPa, TF, TH, 46, 7a; (ii) LIP, DP, V4, V4t, FST; (iii) VOT, PITd, PITv, CITd, AITv, CITv, AITd; and (iv) MT, V2, PIP, V3, VIP, PO, VP, V3A, FEF, MSTd; while area V1 formed a separate cluster. Area MSTl tended to associate with the first of the listed groups. The last cluster in the list, (iv), existed in this composition in all of the optimal solutions, whereas some areas in the other groups showed alternating memberships. These building-block groups are the most compact clusters that fulfil our definition of optimal cluster arrangements. Figure 6*b* shows the ranking of areas according to the relative frequency with which they sent or received connections between the building blocks.

Figure 7*a* summarizes the whole spectrum of possible solutions, obtained for weights between repulsion equal to 7 and attraction equal to 17 (while keeping the opposite cost component weighted at one). The figure shows the associations between 155 optimal cluster arrangements. Despite some variability in this summary, a number of distinct groups can be recognized. The hierarchical tree diagram in figure 7*b* presents the data in another format and shows directly which areas consistently formed clusters, given a particular association threshold.

(d) *Macaque visual system—NPCA*

The general picture of the visual system's organization was confirmed by the results from the NPCAs. Figure 8 shows the pooled data from all 2D clustering computations for the primate visual system using a wide range of parameters for describing the fixed-radius kernels of potential clusters. The figure clearly reproduces the general dichotomy of the primate visual system. It also suggests further subdivisions of the two main clusters, most clearly for the 'earlier' areas V1, V2, V3, MT, V3A, V4t, VP and PIP, and the 'later' ones VIP, MSTl, PO, LIP, MSTd, FST, DP and FEF. The correlative agreement of this summary with the balanced OSA solutions was $R = 0.63$, the cluster-count summary of all NMDS–NPCA configurations agreed with the main OSA results by $R = 0.55$–0.63.

(e) *Macaque somatosensory-motor system—OSA*

We analysed a compilation of somatosensory-motor area connectivity published by Felleman & Van Essen (1991). When attraction and repulsion were balanced, there was a unique optimal arrangement of four clusters and one isolated area, area 36 (see figure 9*a*). The clusters were: (i) Id, 35; (ii) SMA, 4, 5, 6; (iii) 1, 2, 3a, 3a, SII; and (iv) Ig, 7b, Ri. This arrangement carried a weight of 34 (30 connections between the clusters, four non-existing connections within the clusters). We found that only one cluster resulted from attraction weights of 40 or above. At higher repulsion weights, of six and above, all non-existing connections were located outside the clusters, while 40 connections linked the building-block clusters shown in figure 9*b*. We also computed NMDS–NPCA representations for this data set, which agreed well with the results outlined here. The different representations are compared statistically in § 3(h).

(f) *Cat cortex—OSA*

The cat cortical connectivity data, which were graded in three categories of density strengths, offered the chance to explore the effect of such finer categorizations on the structure of optimal connectivity clusters. We first report the results for the more finely graded classification, then for a classification of the data simply into binary categories of existing or apparently absent connections.

For the graded data, 108 different optimal solutions existed for the case of balanced attraction and repulsion weights. These configurations consisted of five to nine clusters and possessed a combined cost of 700 to 707 (acknowledging the different density strengths), with 348 to 394 connections running between the clusters, and 172 to 233 tested non-existing or untested connections remaining within the clusters. The main clusters, which appeared consistently across all the solutions were (i) 5Al, 5m, 5Am, SII, SSAi, SIV, SSAo, 4g, 6l, 5Bl, 6m, 5Bm, 1, 2, 4, 3a, 3b, 7, AES; (ii) PFCL, pSb, 35, 36, Amyg, 2b, Sb, Enr, RS, IA, PFCMd, CGA, IG, CGP, PFCMil; (iii) P, AAF, AI, VP(cortex), AII; (iv) PLLS, 17, 18, 19, AMLS, 2a, 21a, 21b, VLS, PMLS (see figure 10*a*). Figure 10*b*,*c* displays the areas contributing most to connectivity between the optimal clusters in all solutions, and to the repulsion of non-existing connections within the clusters.

For a repulsion weight of 17, all non-existing connections in the optimal arrangements were located between the clusters. The seven optimal configurations which obtained in this case possessed 657 to 675 connections

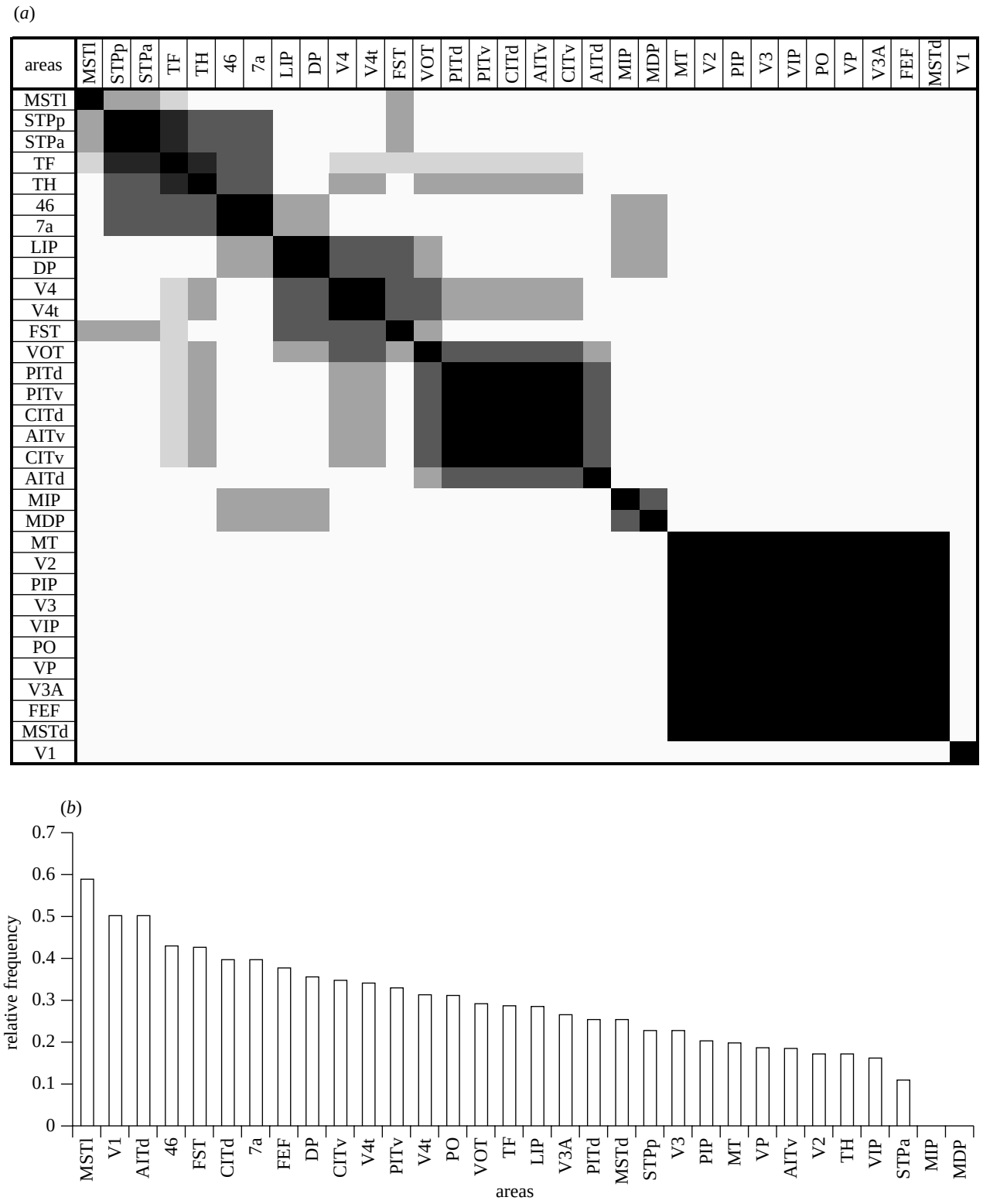


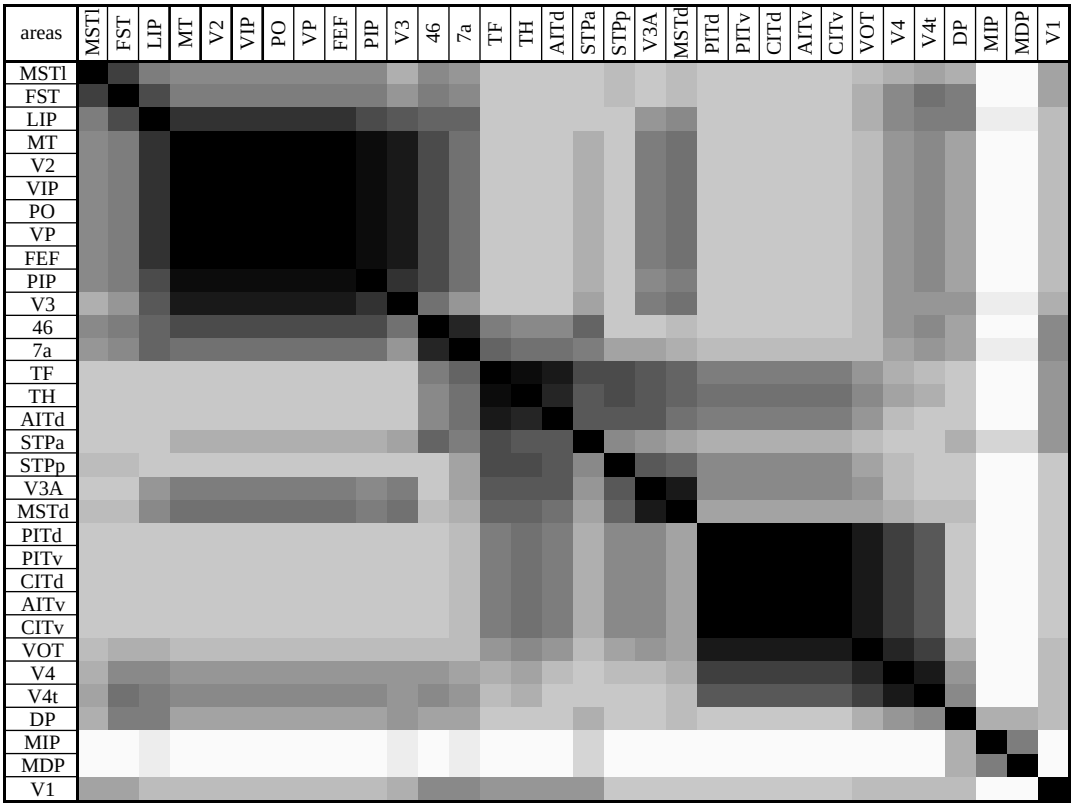
Figure 6. (*a*) Summary of six optimal building-block arrangements (repulsion weight ⩾seven) for the primate visual system. (*b*) Ranking of areas contributing to inter-cluster connections for the configurations summarized in figure 6*a*. The relative frequencies were obtained by normalization analogous to the process for figure 3*b* and figure 5*b*.

between the clusters, which resulted in a total weighted cost of 995 to 104.

Attraction weights for these data were stepped up in 0.3 intervals, to account for the different degrees of density classification. At an attraction cost of 9.6, all areas formed one inclusive cluster.

Figure 11 shows a cluster-tree summary of 2069 optimal solutions for the cat cortex computed over all different settings for the attraction and repulsion weights. Despite the changes in cluster sizes and composition, which led to the many weakly associated sets in the plot, a number of stable clusters can be easily recognized.

(*a*)



(*b*)

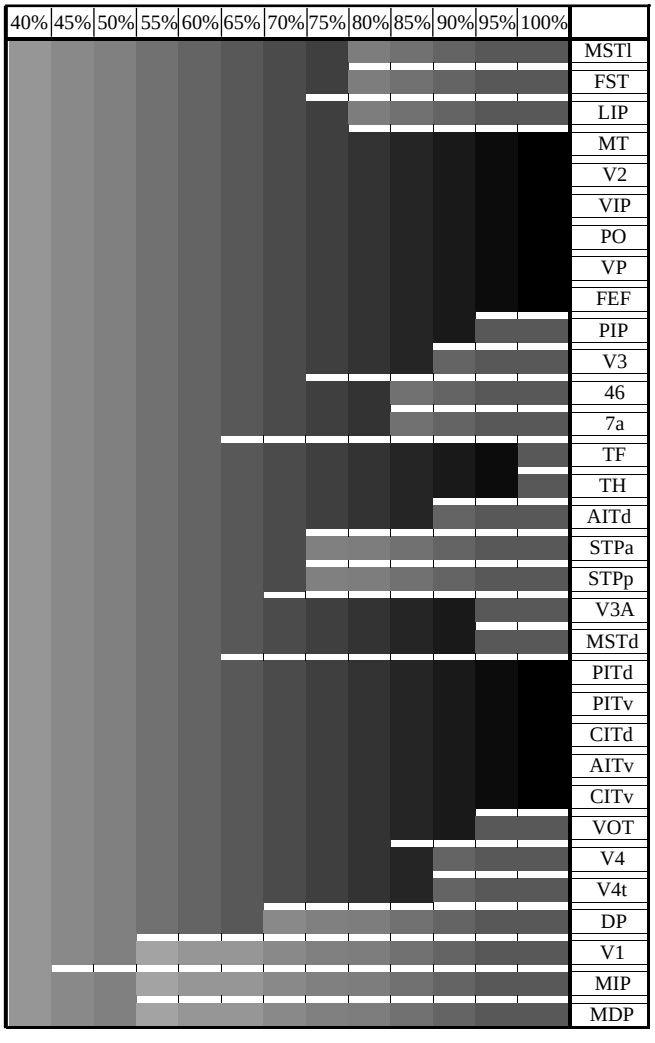


Figure 7. (*a*) Summary of all 155 optimal configurations for the primate visual system obtained for series of different attraction and repulsion weightings (from repulsion weight of seven to an attraction weight of 17). (*b*) Cluster-tree diagram for the optimal arrangements presented in figure 7*a*. The tree illustrates at which relative frequency, given at the top, clusters separate. Lighter shading indicates singular areas completely separated from other clusters.

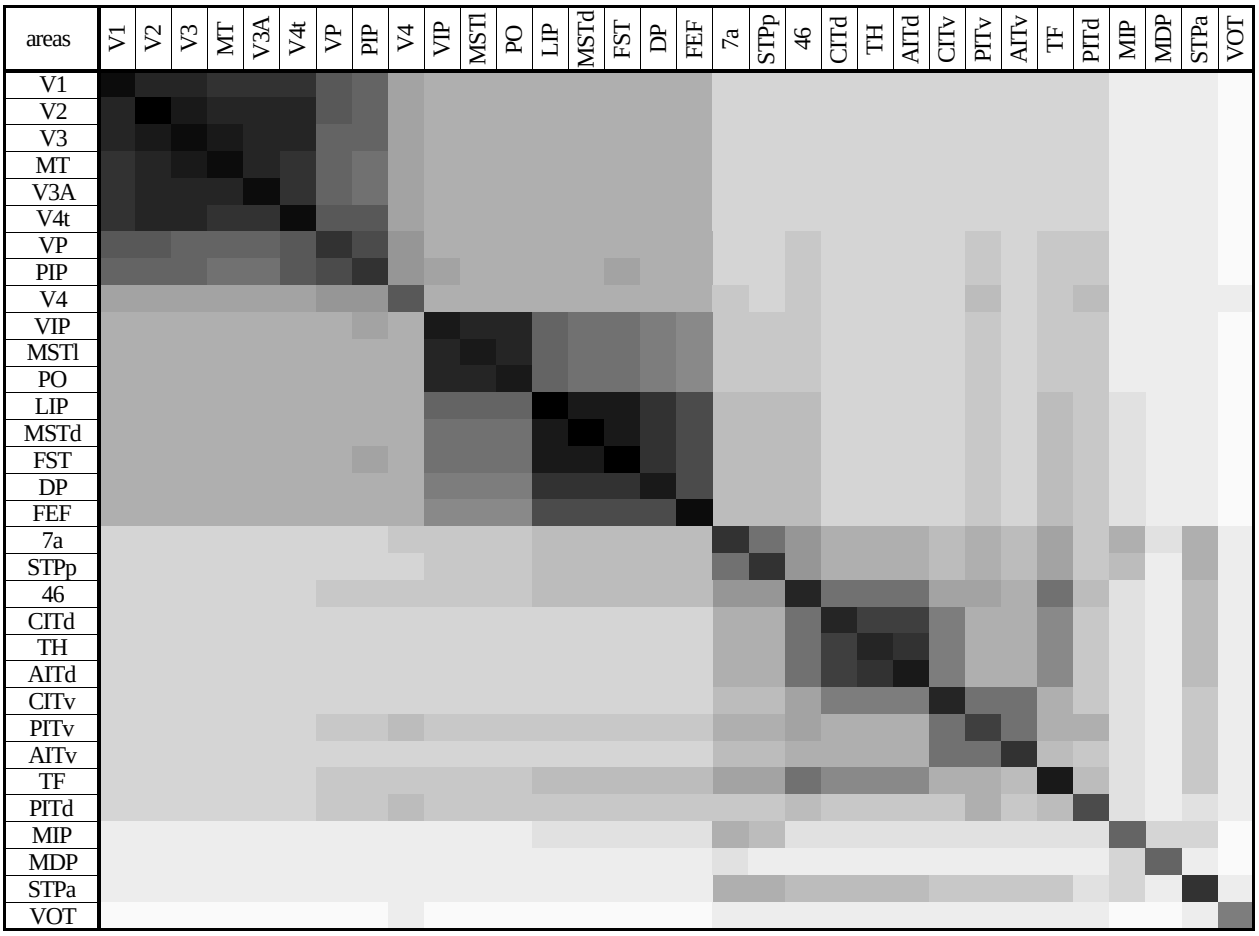


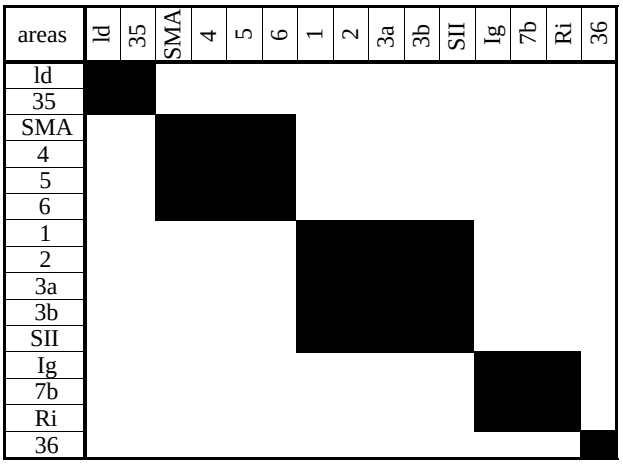
Figure 8. Cluster-count summary of all clusters identified by fixed-radii NPCA in 2D-NMDS representations of primate visual connectivity.

Particularly striking were the very stable clusters of somatosensory-motor areas.

The binary classification of these connectivity data, and their subsequent OSA, led to very similar cluster structures as the categorization of existing connections in three classes. The main subdivisions detailed above for the balanced OSA condition were also recognizable for balanced OSA arrangements of the binary data, even though the first two groups of somatosensory-motor and 'limbic' areas were more strongly subdivided in the binary data. The correlation between the cluster-count summaries for the quaternary and binary data (655 optimal arrangements) under balanced OSA was $R = 0.61$. The cluster arrangements of the binary data were reduced to building-block structures at a repulsion weight of ten, and 248 optimal arrangements obtained under this condition. All areas came together in one cluster for an attraction weight of 21.

There are two principal ways of interpreting the graded strengths of existing connections for this set of cat connectivity data. In the 'optimistic' interpretation, which we followed in the above analysis of the quaternary data, the numerical values assigned to the different strengths directly reflect the absolute anatomical strengths or densities of the connections. On this assumption, it is possible to evaluate the different connections as if, for instance, any intermediate-strength connection (with a value of two) is indeed about twice as strong as two weak connections (with a strength value of one). In this view, the connection strength classes approximate interval data. In a more conservative interpretation, on the other hand, the strength values are solely numerical labels for ranks of different connection densities, which are considered ordinal or nominal. In this case, a direct numerical comparison of different connection categories would not be possible. We considered this possibility by modifying our OSA algorithm in such a way that different connection categories would be considered independently from each other during the optimizations. This was achieved by giving all connections of a particular category such high weights that the arrangement of areas possessing this connection could not be affected, even if all connections of lower ranks together exerted a contrary influence. In this way, the overall arrangement was first shaped by the optimization of the highest ranks, then by next highest and so on, while the different ranks did not directly interfere with each other. The weights of the absent and unknown connections were determined as means of the weights for the existing connections, weighted by the frequency of connections in a particular strength category. Applying such a modified OSA to the finely graded cat connectivity data yielded a unique cluster arrangement, which resembled the main subdivisions obtained for the balanced OSA of the quaternary data. Differences between the

(*a*)



(*b*)

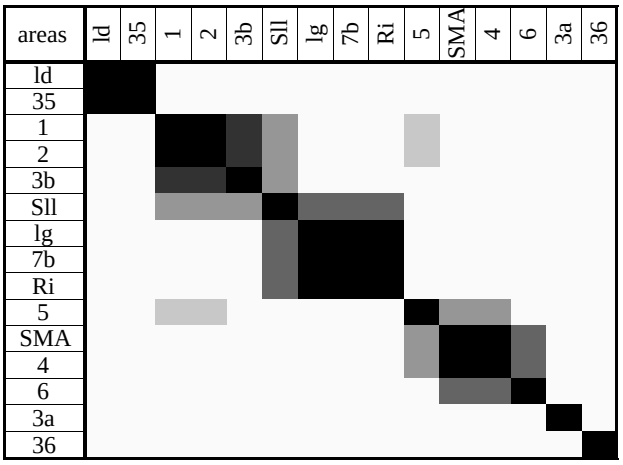


Figure 9. (*a*) Optimal sets of primate somatosensory-motor areas, obtained for balanced attraction and repulsion weighting. (*b*) Optimal building block sets of somatosensory-motor areas in the macaque (repulsion weight of six). The diagram summarizes five different optimal arrangements.

arrangements existed in secondary somatosensory-motor and visual area clusters becoming apparent in the latter approach. The agreement between cluster-count plots of the two approaches was $R = 0.56$.

(g) *Cat cortex—NPCA*

The main features of the organization of the cat connectivity data were again confirmed by independent analysis of the same data by the NMDS–NPCA approach. For this method, as for the OSA analyses, the data were classified in two different ways: as existing or absent connections, or with three different grades of strength for the existing data. Figure 12 shows a cluster-count summary for all solutions derived with this approach, using 2D and 5D NMDS representations as well as various clustering paradigms and parameters for the binary and quaternary data. Despite the averaging effect of the different settings, the general picture resembles that resulting from the OSA analyses, and summaries of the two different methodological approaches show a correlation of $R = 0.73$ (compared to balanced OSA of quaternary data).

(h) *Validation of cluster configurations*

We tested the significance of the optimal cluster configurations by comparing the lowest costs that we obtained from analysis of the actual data and for analyses of 20 randomly reshuffled data sets, respectively. All the compared sets were analysed under the balanced OSA condition. A comparison of these minimal costs for all data sets is given in table 2. The minimal cost from the random data sets had means that were much, and statistically significantly, greater (between six standard deviations for the somatosensory set and 42 standard deviations for the cat data set) than the minimal cost of actual connectivity clusters. This confirms the conclusion from our initial small-world analysis, that cortical connectivity is organized in local clusters of densely interconnected areas.

Correlations between clusters obtained from the real data and from randomly reshuffled data further showed that the actual clusters had no similarity with any of the random cluster solutions: all correlation coefficients were very close to zero (average correlation: 0.01). There was, however, a good correlation between the configurations obtained for real data under the different OSA conditions of balanced weights and high repulsion, with the correlation coefficients ranging between 0.49 (cat quaternary) and 0.71 (macaque somatosensory areas). Correlations between configurations from balanced conditions and summaries of solutions pooled across all attraction and repulsion weight settings were also in this region.

4. DISCUSSION

(a) *Small-world characteristics of cortical networks*

Our results demonstrated that cortical networks were organized in local clusters of densely interconnected areas. This aspect of organization was initially indicated by the local cluster indices we computed for all data sets (see table 1, columns 6–8). These indices demonstrated that the average degree of local connectivity was almost twice as high as the general level of interconnectedness between cortical areas. We also computed coefficients for the characteristic path lengths in all data sets. These indices describe the average shortest paths between any two areas in the cortical network. The values in table 1 (columns 2–5) showed that the characteristic path lengths of the actual connectivity data were somewhat larger, but still in the same order of size, as the short characteristic path lengths of randomly redistributed data. This result, together with the significant local clustering of the networks, suggests that cortical connectivity possesses attributes of 'small-world' networks, as defined by (Watts & Strogatz 1998), that is, $C_{\mathrm{real}} \gg C_{\mathrm{ran}}$ and $L_{\mathrm{real}} \approx L_{\mathrm{ran}}$. The organization of small-world networks allows efficient processing and propagation of signals, as such systems combine very short average path lengths with a tight local integration of processing nodes. The original paper by Watts & Strogatz (1998) investigated a first example of a completely mapped neural network, the connectivity of the nematode *Caenorhabditis elegans*, and found it to possess small-world features. Our study shows that mammalian cortical connectivity possessed even shorter characteristic path lengths than the *C. elegans* network ($L_{C.\ elegans} = 2.65$), combined with a higher

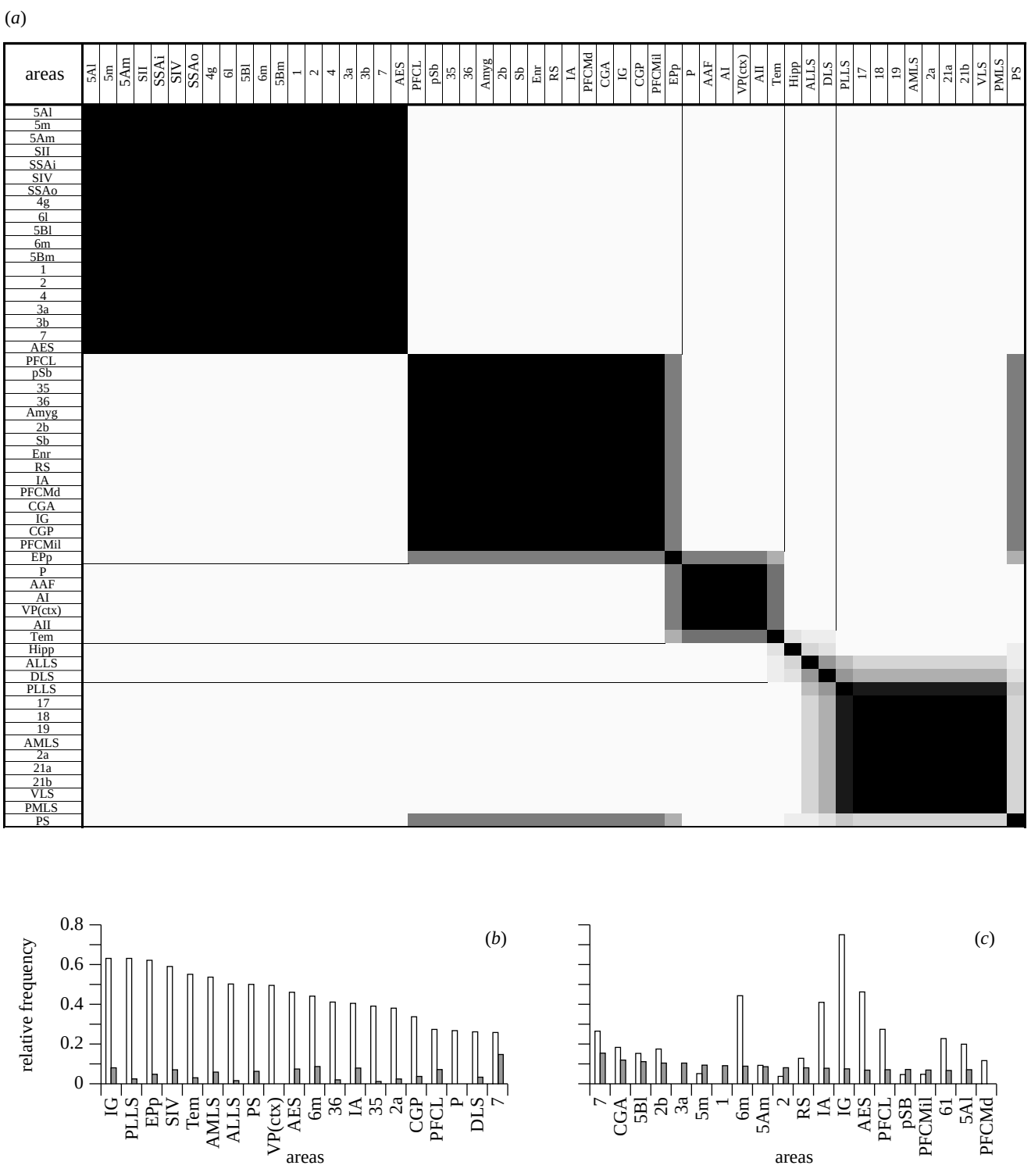


Figure 10. (*a*) Summary of 108 optimal area sets for the cat cortex, using balanced attraction and repulsion weights. Connections entering this analysis were specified in three strength categories. (*b*) Ranking of the 20 areas that were contributing most to inter-cluster connectivity in all optimal configurations summarized in figure 10*a*. All frequencies are normalized as for figure 3*b*. Frequencies of inter-cluster connections are shown in lighter shading, those of absent connections within the clusters in darker shading. (*c*) Top 20 areas ranked for relative frequency of non-existing connections within the optimal sets displayed in figure 10*a*.

absolute local clustering ($C_{C.\ elegans} = 0.28$). These attributes may be partly due to a higher degree of connectedness in the cortical networks. Indeed, the relative degree of local clustering ($C_{\mathrm{real}}/C_{\mathrm{ran}}$) is higher in *C. elegans* (5.6) than it is for the cortical data (between 1.8 and 3). The companion paper by Stephan *et al.* (this issue) confirms that the efficient organization of cortical connectivity has functional implications, as the spread of task-independent excitation in the primate cortex also displays small-world characteristics.

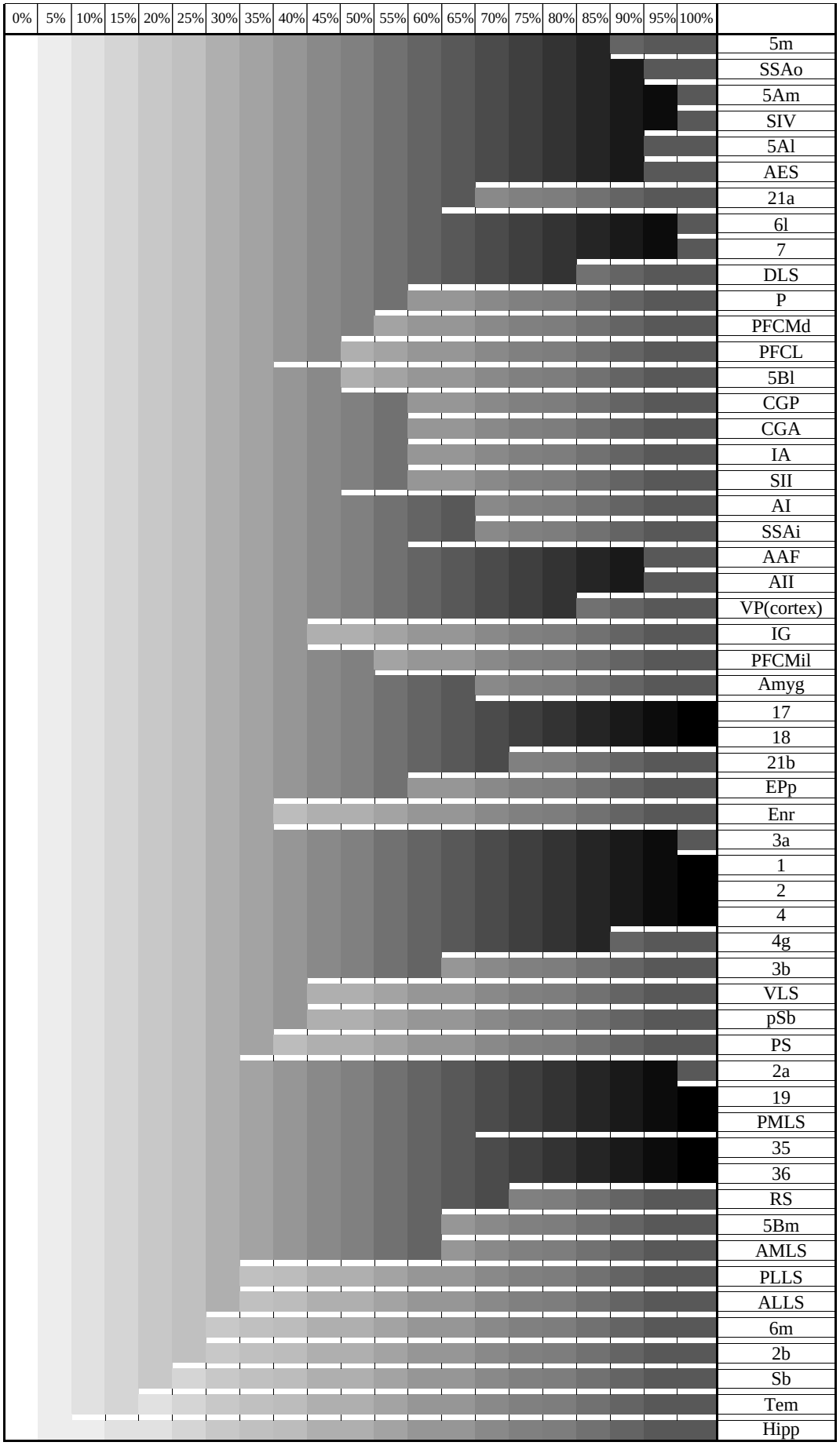


Figure 11. Cluster-tree diagram summarizing 2069 optimal arrangements that obtained over the range of all different attraction–repulsion weight settings used in the OSA of the cat cortical connectivity set (quaternary classification). Settings ranged from a repulsion weight of 17 to an attraction weight of 9.6.

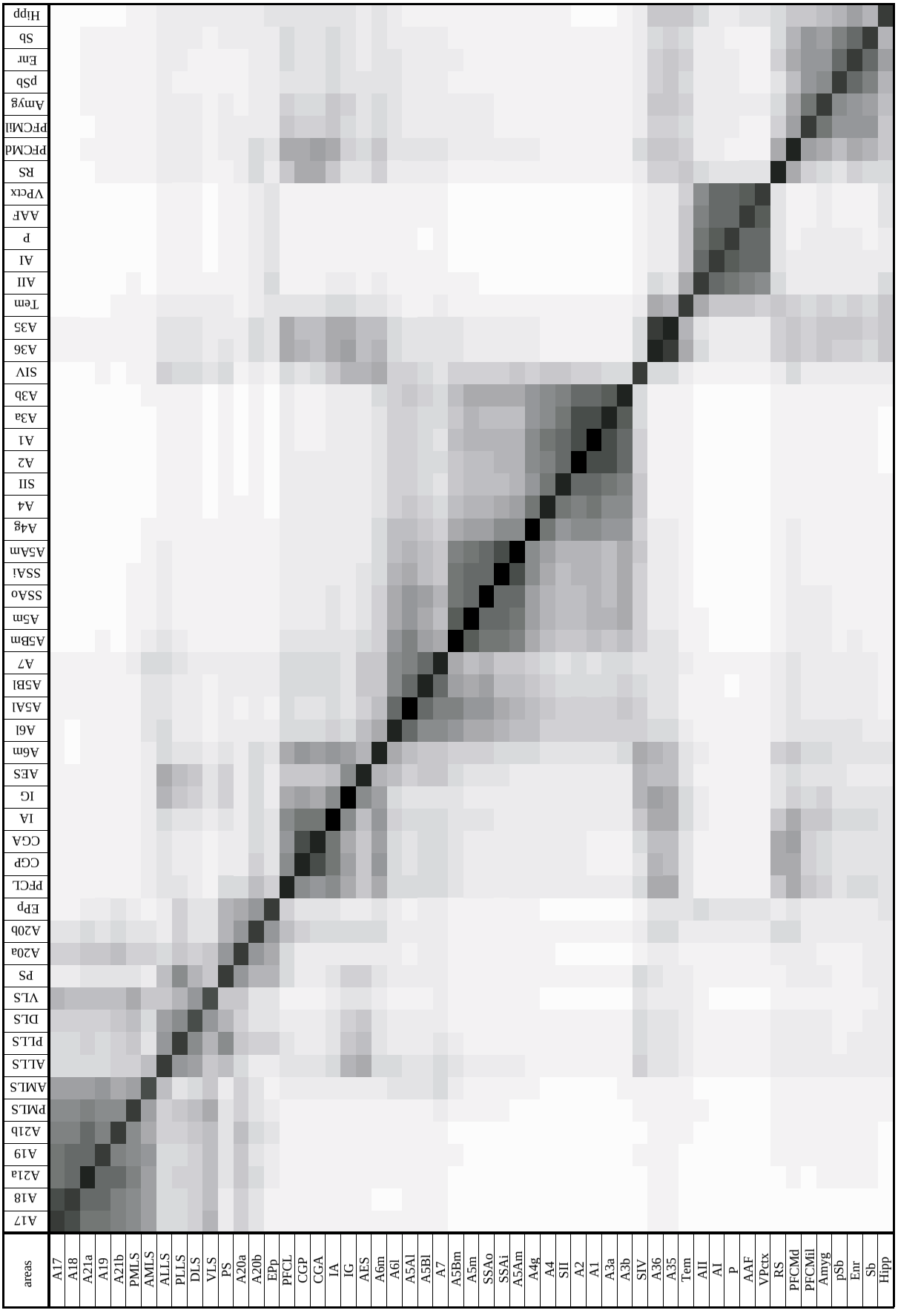


Figure 12. Summary of complete set of solutions obtained for the cat cortical connectivity data by application of all combinations and settings of NMDS and NPCA. The diagram also averages across the two differently classified cat data sets (binary and quaternary categorization of connection strengths).

Table 2. *Minimal costs obtained in the balanced OSA condition for actual and 20 randomly reshuffled data set*

(The means, μ and standard deviations, δ, of the solutions for randomly reshuffled data show that the actual connectivity data carry a cost which is far, and significantly, smaller. The clusters detected are therefore not spurious.)

	$\text{cost}_{\text{real}}$	$\mu\text{cost}_{\text{ran}}$	$\delta\text{cost}_{\text{ran}}$
macaque: visual	135	244	4
macaque: somatosensory	34	53	3
macaque: whole cortex	628	830	6
cat: whole cortex	700	1245	13

(b) *Optimal set analysis*

We developed a new technique, optimal set analysis, for investigating the cluster structure of complex networks. This method has a number of advantages over other methods that have been applied to the analysis of cortical connectivity (e.g. Young *et al*. 1995; Jouve *et al*. 1998; Tononi *et al*. 1998). OSA is founded on an explicit definition for the overall cluster structure of a network (conditions CO1 and CO2). Apart from the concept expressed in this definition, OSA makes no further assumptions about the structure of the data and does not require any restructuring or transformation of experimental values. OSA is conceptually simple, and in this sense, more straightforward to interpret than other methods. In the course of its derivation of optimal cluster arrangements, OSA also automatically determines the optimal number of clusters. The method can analyse incomplete data sets and can be applied (with current computer technology) to the analysis of very large networks (largest tested >2000 nodes). The technique takes into account that there might be several, or very many, equally good solutions for the optimization problem. The method is not restricted to the analysis of neuronal connectivity, and in principle any problem that can be mathematically expressed as a graph (that is, as an interconnected system) can be analysed. OSA can accommodate metric, ordinal and categorical information as weights for the 'edges' of a graph, unlike many other types of graph analysis presently. We exemplified this by means of the two different assumptions we used for OSA analysis of the cat connectivity data, and our hierarchical analysis, which used a similar data-analytic approach (Hilgetag, O'Neill & Young, this issue).

To make an even wider range of neural data accessible to analysis, various other cost functions could be integrated with the optimization algorithm used by OSA. The method described by Tononi *et al*. (1998) provides a more general framework than the NPCA used here. Their method works on object coordinates and not object relations or distances. It can identify nonlinear relationships, but it requires the testing of large subsets of the data. OSA provides a systematic approach to this latter problem, and a combination of both methods (i.e. the integration of Tononi *et al*.'s clustering indices as OSA cost functions) could allow a general and flexible approach to functional data.

(c) *Organization of cortical systems*

Our analyses broadly confirmed the subdivisions of cortical systems in the macaque and the cat inferred from earlier approaches (Young 1993; Scannell *et al*. 1995). There were, however, a number of differences and methodologically derived idiosyncracies, which we will discuss in relation to the primate visual system, where they emerged particularly clearly.

Our analyses confirmed the dichotomy of the primate visual system. The separation of the areas is exemplified by the optimal set arrangement obtained for balanced attraction and repulsion weighting. Here, one cluster contained the occipito-parietal and parietal areas V2, V3, VP, V3A, MT, V4t, V4, PIP, LIP, VIP, DP, PO, MSTl, MSTd, FST and FEF, and a second cluster the inferior-temporal and prefrontal areas PITv, PITd, CITv, CITd, AITv, AITd, STPa, STPp, 7a, TF, TH, VOT and 46. The unique position of primary visual cortex (V1) is expressed by its separation from either of the large clusters. This grouping, which obtained for the balanced case of our cluster definition, shows considerable agreement, but also some notable differences, compared with the widely assumed structure of the dorsal and ventral visual streams. Most conspicuously, the prototypical 'ventral' area, V4, appeared in a group with otherwise 'dorsal' areas, and the second 'ventral' cluster contained an area, 7a, which most systems neuroscientists would associate with the 'dorsal' stream.

Why are V4 and area 7a apparently misassigned in these optimal clusters? On the one hand, this may illustrate the limitations of the approach. The mismatch between well-founded neurophysiological ideas about the properties of these areas and the clusters with which they associated here could suggest that our analysis of a single type of data, albeit with an explicit and intuitive definition of cost, is insufficient. It may be that a satisfactory classification of primate visual areas will require a wider variety of information about connectivity, and perhaps more complex criteria. On the other hand, specific aspects of data structure may themselves give rise to the apparent misassignment. For the χ^2-test in Young *et al*. (1995), the ventral stream was defined as V4, VOT, PITd, PITv, CITd, CITv, AITd and AITv, and the dorsal as MT (V5), MSTd, MSTl, FST, PO, LIP, VIP, DP and 7a. All other areas were classified as 'early' or 'late'. We re-examined the connectivity between the Young *et al*. (1995) groupings and found that they represented the optimal (and unique) arrangement for these areas according to the OSA cost functions. This arrangement had a cost of 30, while the same arrangement with V4 and 7a swapped carried a cost of 51. V4 and 7a swapped into the 'incongruous' assignments derived by OSA only when the connections of all 32 areas were included. In this case, the cost went from 135 for our apparently paradoxical arrangement to 141, when V4 was placed in the ventral and 7a in the dorsal stream. Hence, adding the 'early' and the 'late' areas of Young *et al*. (1995) drives V4 and 7a to the anomalous groupings. Our explanation is that the earlier visual areas tend to cluster with the dorsal stream. These areas share many interconnections with V4, and so draw V4 towards the dorsal cluster when all areas are included. Correspondingly, all the 'late' areas except FEF tend to cluster with the ventral structures, and share many connections with 7a, drawing this area towards the ventral grouping. Hence, the apparent misassignments of these areas reflect a feature of data

structure: the dorsal–ventral dichotomy is not orthogonal to the early–late organization of areas. The dorsal stream tends to be 'lower' and the ventral stream 'higher', and the positions of V4 and 7a in the latter respect affect their clustering assignments. V4 is lower than 7a, and tends to cluster with its early associates, while the opposite is true for 7a. This aspect of data structure is evident in the balanced solution, but is less apparent when the pooled solutions from all attraction–repulsion weight settings are considered. In this more general case, area V4 associates more appropriately with ventral stream areas (figure 7*a*,*b*). This result is due to V4's separation from dorsal stream areas for higher repulsion weights, indicating that V4 possesses many connections with the dorsal group of areas, but that is also separated from them by many non-existing connections. Consequently, the analyses confirm that V4 is integrated with both ventral and (particularly early) dorsal areas.

A division of visual cortex has recently also been suggested for the cat by Lomber *et al*. (1996). These researchers used cooling to reversibly deactivate cortex along the middle suprasylvian sulcus and the posterior suprasylvian gyrus, and to compare the functions of these cortices. They found a dissociation of functions, which they linked to differences in the underlying visual pathways. Our analysis of global connectivity in the cat cortex, however, found no evidence for different connectivity-based streams (see figures 10–12), and cat visual areas formed a rather homogeneous cluster. It remains to be established whether and how these functional differences can be related to the underlying anatomical organization (Young *et al*., this issue).

(d) *Inter-cluster connection analysis*

We performed an analysis of the connections running between the optimal sets of areas. The results of this analysis suggested the areas most often involved in making connections to other clusters. It can thus help to identify potentially important points of contact between the clusters, or areas that are only loosely integrated in clusters. The results were again of particular interest for the primate visual cortex, where they may help to point out regions of reconvergence or cross-talk between the two visual streams.

Looking at the ranking for connections between optimal primate visual clusters (obtained for balanced OSA, figure 5*b*), the strongest contributor to inter-cluster connectivity was area 7a. This area is also identified by Jouve *et al*. (1998) as an important mediator between the dorsal and ventral streams. However, V4t, which was also identified as a possible mediator between the streams by Jouve *et al*. (1998), was among the most unlikely candidates possessing no inter-stream connections. The ventral and dorsal subdivisions of posterior inferior temporal cortex, the frontal eye fields, TF, STP, VOT and area 46 were more communicative between the optimal clusters. It would seem therefore that there are multiple opportunities for cross-talk and reconvergence of the visual streams (see Young 1992; Young *et al*. 1995). However, the present analysis contained no information on strengths of connections nor on the laminar destinations of projections, both of which factors could be important determinants of functional impact.

Results on interconnections between clusters from the primate visual cortex building blocks (i.e. higher repulsion analyses) most prominently involved area 46 and areas of the STS. Inferior temporal areas, however, also appeared unusually communicative between clusters, and it would therefore seem that this region represents a widely connected set of nodes for cortical traffic, rather than the end-point of an analysis sequence (see Young *et al*. 1995).

(e) *Predictive aspects*

Many analyses of connectivity suggest where undiscovered connections, sometimes of particular types (e.g. Hilgetag *et al*. 1996), should be found (e.g. Young 1993). However, on the present basis, the number of predictions of connections that are likely to be present is significantly lower than is predicted by a recent missing data estimation procedure (Jouve *et al*. 1998). In that treatment, interpolation generated large numbers of new connections. The frontal eye fields (FEF), for example, were predicted to be connected to every visual cortical area except V1, and V4t to exhibit about 40 new connections when it is examined in more detail (Jouve *et al*. 1998). Indeed, the configuration derived by Jouve and colleagues from factor analysis of their interpolated matrix is almost identical ($R^2 > 0.92$ by Procrustes rotation) to a configuration derived by Young (1992) from the very unlikely limit-case assumption that all unreported connections exist. Few researchers would credit that such profuse but unobserved connectivity could be awaiting discovery. We have wondered, therefore, whether the interpolation algorithm employed by Jouve and colleagues can be realistic. Their procedure used a parameter derived from large, well-defined and well-studied cortical areas to interpolate connections for less well-studied, often poorly defined cortical areas, which in general are small. But what may hold for V1, V2, V4 and other large, well-studied areas may not hold for small strips of cortex such as V4t, or for poorly defined regions. Similarly, the interpolation procedure did not appear to take account of the probability that false negatives are less likely than false positives if one considers anatomical methodology, in which a good deal of the brain is often sectioned, and one can see transported label, even in surprising places, if it is present. Further experimental neuroanatomy, motivated by expectations from the analysis of existing experiments, would clarify these uncertainties. The predictions of undiscovered connectivity made by Jouve *et al*. (1998) are particularly worthy of experimental study because they have led to the supposition of a theoretically interesting 'relay stream' that they believe lies between the dorsal and ventral streams of processing. The principal members of this relay stream, however, are V4t and FEF, which are also the principal recipients of interpolated new connections. It is possible, therefore, that the 'relay stream' is a property of interpolated data structure and not of the real visual system, and it is the case that analyses of real data see no sign of it.

The present results also allow predictions for further neuroanatomical experiments. For example, members of a cluster are more often connected to one another than would be expected for cortical connectivity in general. Hence, if two areas whose interconnections have not yet

been studied are consistently found together in a cluster, there is a high likelihood that the areas will prove to be connected. Specific predictions of this type can be generated by comparing cluster cohabitees, visible in the cluster diagrams, with as yet unreported connections, which have been detailed elsewhere (e.g. Felleman & Van Essen 1991; Hilgetag *et al*. 1996).

5. CONCLUSIONS

Our results show that clusters of differentially connected areas can be detected reliably in mammalian cortical systems by independent statistical and optimization approaches. The cluster arrangements were clearest for the primate visual system and the cat cortical system. In the primate visual system, the results confirmed a subdivision into three main groups of primary, 'ventral' and 'dorsal' areas. Interestingly, such an organization was also borne out by cluster analyses of the spread of task-independent activity in the same system (Stephan *et al.*, this issue). The larger variability in the results for other systems might be due to the characteristics of the respective data, which did not, for example, distinguish between known absent and so-far unstudied connections. This re-emphasizes the importance of explicitly reporting absent connections from anatomical experiments, but it is clear that new experimental data will be required to clarify many of these issues (Young *et al*. 1995). The density of the global macaque connectivity data was considerably lower than the density of all other data sets. We suspect that existing experimental information and collations of corticocortical connections in this system are markedly incomplete. We have begun efforts to create a more formalized, detailed, extensive and up-to-date collation system (Burns & Young, this issue; Stephan *et al*., this issue), but the likelihood that significant connections in the monkey remain to be discovered further reinforces the importance of further experimental studies.

Supported by the Wellcome Trust and the University of Newcastle upon Tyne. We thank Klaas Stephan and Rolf Kötter for helpful discussions about the OSA and statistical techniques.

REFERENCES

Abramowitz, M. & Stegun, I. A. (eds) 1972 *Handbook of mathematical functions*. New York: Dover.

Felleman, D. J. & Van Essen, D. C. 1991 Distributed hierarchical processing in the primate cerebral cortex. *Cerebr. Cortex* **1**, 1–47.

Goodale, M. A. & Milner, A. D. 1992 Separate visual pathways for perception and action. *Trends Neurosci*. **15**, 20–25.

Goodhill, G. J., Simmen, M. W. & Willshaw, D. J. 1995 An evaluation of the use of multidimensional scaling for understanding brain connectivity. *Phil. Trans. R. Soc. Lond.* B **348**, 265–280.

Hilgetag, C.-C., O'Neill, M. A. & Young, M. P. 1996 Indeterminate organization of the visual system. *Science* **271**, 776–777.

Hilgetag, C.-C., O'Neill, M. A. & Young, M. P. 1997 Optimization analysis of complex neuroanatomical data. In *Computational neuroscience, trends in research, 1997* (ed. J. M. Bower), pp. 925–930. New York: Plenum Press.

Hilgetag, C.-C., Burns, G. A. P. C., O'Neill, M. A. & Young, M. P. 1998 Cluster structure of cortical systems in mammalian brains. In *Computational neuroscience: trends in research, 1998* (ed. J. M. Bower), pp. 41–46. New York: Plenum Press.

Jouve, B., Rosenstiehl, P. & Imbert, M. 1998 A mathematical approach to the connectivity between the cortical visual areas of the macaque monkey. *Cerebr. Cortex* **8**, 28–39.

Kruskal, J. B. 1964*a* Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **29**, 1–27.

Kruskal, J. B. 1964*b* Nonmetric multidimensional scaling: a numerical method. *Psychometrika* **29**, 115–129.

Laarhoven, P. J. M. V. & Aarts, E. H. L. 1987 *Simulated annealing: theory and applications*. Dordrecht, The Netherlands: Kluwer.

Lomber, S. G., Payne, B. R., Cornwell, P. & Long, K. D. 1996 Perceptual and cognitive visual functions of parietal and temporal cortices in the cat. *Cerebr. Cortex* **6**, 673–695.

Merigan, W. H. & Maunsell, J. H. 1993 How parallel are the primate visual pathways? *A. Rev. Neurosci*. **16**, 369–402.

SAS Institute, Inc. 1990 *SAS/STAT user's guide*, v. 6, 4th edn. SAS Institute, Inc.

Scannell, J. W., Blakemore, C. & Young, M. P. 1995 Analysis of connectivity in the cat cerebral cortex. *J. Neurosci.* **15**, 1463–1483.

Scannell, J. W., Burns, G. A. P. C., Hilgetag, C.-C., O'Neill, M. A. & Young, M. P. 1999 The connectional organization of the cortico-thalamic system of the cat. *Cerebr. Cortex*. **9**, 277–299.

Simmen, M. W., Goodhill, G. J. & Willshaw, D. J. 1994 Scaling and brain connectivity. *Nature* **369**, 448–449.

Tononi, G., McIntosh, A. R., Russell, D. P. & Edelman, G. M. 1998 Functional clustering: identifying strongly interactive brain regions in neuroimaging data. *NeuroImage* **7**, 133–149.

Ungerleider, L. G. & Mishkin, M. 1982 Two cortical visual systems. In *Analysis of visual behaviour* (ed. D. G. Ingle, M. A. Goodale & R. J. Q. Mansfield), pp. 549–586. Cambridge, MA: MIT Press.

Watts, D. J. & Strogatz, S. H. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442.

Young, M. P. 1992 Objective analysis of the topological organization of the primate cortical visual system. *Nature* **358**, 152–155.

Young, M. P. 1993 The organization of neural systems in the primate cerebral cortex. *Proc. R. Soc. Lond.* B **252**, 13–18.

Young, M. P., Scannell, J. W., Burns, G. A. P. C. & Blakemore, C. 1994 Scaling and brain connectivity—reply. *Nature* **369**, 449–450.

Young, M. P., Scannell, J. W., O'Neill, M. A., Hilgetag, C.-C., Burns, G. & Blakemore, C. 1995 Non-metric multidimensional scaling in the analysis of neuroanatomical connection data and the organization of the primate cortical visual system. *Phil. Trans. R. Soc. Lond.* B **348**, 281–308.