Large-scale spatiotemporal spike patterning consistent with wave propagation in motor cortex

Kazutaka Takahashi1,*, Sanggyun Kim3,*, Todd P. Coleman3, Kevin A. Brown4, Aaron J. Suminski5, Matthew D. Best2 & Nicholas G. Hatsopoulos1,2

Aggregate signals in cortex are known to be spatiotemporally organized as propagating waves across the cortical surface, but it remains unclear whether the same is true for spiking activity in individual neurons. Furthermore, the functional interactions between cortical neurons are well documented but their spatial arrangement on the cortical surface has been largely ignored. Here we use a functional network analysis to demonstrate that a subset of motor cortical neurons in non-human primates spatially coordinate their spiking activity in a manner that closely matches wave propagation measured in the beta oscillatory band of the local field potential. We also demonstrate that sequential spiking of pairs of neuron contains task-relevant information that peaks when the neurons are spatially oriented along the wave axis. We hypothesize that the spatial anisotropy of spike patterning may reflect the underlying organization of motor cortex and may be a general property shared by other cortical areas.

1 Department of Organismal Biology and Anatomy, University of Chicago, 1025 E 57th Street Culver Room 206, Chicago, Illinois 60637, USA. 2 Committee on Computational Neuroscience, University of Chicago, 1025 E 57th Street Culver Room 206, Chicago, Illinois 60637, USA. 3 Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive MC 0412, San Diego, La Jolla, California, 92093, USA. 4 Center for Neuroscience, New York University, 4 Washington Place, New York City, New York 10003, USA. 5 Department Electrical Engineering and Computer Science, Milwaukee School of Engineering, 1025 N Broadway, Milwaukee, Wisconsin 53202, USA. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to N.G.H. (email: nicho@uchicago.edu).

DOI: 10.1038/ncomms8169
Propagating waves of neural activity are ubiquitous and have been documented at different spatial resolutions in a number of different neocortical areas including visual1–6, somatosensory7–9, auditory10,11 and motor cortices12–14 as measured via multielectrode local field potential (LFP) recordings, voltage-sensitive dyes (VSDs) and multiunit activities. Oscillatory LFPs and electroencephalograms in the beta frequency range (15–40 Hz) are ubiquitous in the motor cortex of mammals including monkeys15–18 and humans19,20. In particular, we have previously demonstrated that across the precentral gyrus of the upper-limb area of primary motor cortex (MI), these oscillations are not perfectly synchronized but rather exhibit phase gradients that indicate planar propagating waves along what we define as a beta wave axis, a rostro-caudal axis in monkeys13 and a medio-lateral axis in humans14 at a range of propagating speeds that were consistent across subjects.

However, as both LFPs and VSD measure aggregate potentials from groups of neurons near the recording site, it has never been shown whether action potentials from individual neurons demonstrate spatiotemporal patterning consistent with wave propagation. This is important because it is still debated as to what aggregate signals such as LFPs and VSD signify physiologically, whereas single-unit action potentials are understood to mediate interneuronal communication. Moreover, the functional significance of this wave propagation for motor control is unclear (but see recent computational studies21,22). Here we first show that MI neurons can be classified, based on the spike waveform widths, into two groups of neurons exhibiting distinct spectral properties. We then estimate effective connectivity of networks of spiking neurons based on this classification using a Granger causality analysis applied to point processes, and demonstrate that a class of simultaneously recorded, single-motor cortical neurons with narrow spike waveforms in non-human primates spatially coordinates their spiking activity in a manner that closely matches the orientation of prominent beta wave propagation. We also demonstrate that sequential spiking activity of that class of neuron pairs contains task-relevant, target-direction information whose magnitude varies according to the spatial orientation of the constituent neurons in a manner consistent with the beta wave axis.

Results

Beta waves in the motor cortex. We recorded multiple single-unit and LFP activity from MI using chronically implanted high-density microelectrode arrays while three rhesus monkeys (Rs, Mk and Rj) made planar reaching movements using a two-link robotic exoskeleton (BKIN Technologies, ON, Canada). The monkeys performed a random target-pursuit (RTP) task23 that required them to move a cursor (aligned with the position of their hand) through a sequence of randomly positioned targets. Movement durations from target to target ranged from 300 to 450 ms with mean speeds (±s.d.) of 22.33 ± 11.17 (Rs), 14.12 ± 6.27 (Mk) and 6.11 ± 7.29 cm s⁻¹ (Rj). Planar beta wave activity measured from spatially distributed LFP sites was evident at particular intervals of time throughout the performance of this task (Fig. 1a). We used a method described previously13 to characterize the properties of planar beta waves. We found that the degree of planar wave propagation as measured by a quantity called phase gradient directionality (PGD) was strongest ~100–150 ms after the target onset (Fig. 1b) when beta power was high (Fig. 1c), and when visual target information reached the motor cortex13 followed by movement initiation to the new target (see wrist speed in Fig. 1b). Consistent with our previous findings using a center-out task13, wave propagation directions during the RTP task exhibited either a bimodal distribution (monkey Rs) or unimodal distribution with a small secondary mode (monkeys Mk and Rj), with one mode oriented in the rostral-caudal direction and a secondary mode oriented in the opposite direction (Fig. 1d). We denoted the caudal wave and rostral wave directions defined by the mean direction of the first or only mode of the wave propagation distribution and the opposite direction oriented roughly along the rostro-caudal axis. The distribution of propagation speeds was always unimodal with means and medians ranging from 23.2 to 26.7 and from 10.1 to 13.5 cm s⁻¹, respectively (Fig. 1e).

To investigate spatiotemporal patterns of spiking activity (see Supplementary Figs. 1 and 2b,d,f for many examples of Peristimulus time histograms (PSTHs) of neurons used in the analysis), we first partitioned the population of recorded neurons into two potentially distinct, functional classes (Fig. 2a) based on the observation that their spike waveform widths exhibited a bimodal distribution. For each data set separately as well as for pooled data, a two-Gaussian mixture model (that is, bimodal) was found to be the most parsimonious fit to the distributions based on the Bayesian Information Criterion (BIC; Fig. 2b). Narrow-spiking neurons tended to exhibit higher mean spike rates as compared with broad spiking neurons for all three monkeys (however, only Mk’s data sets showed a statistically significant difference in means, unpaired two-sample t-test with unequal variances): 4.89 ± 4.10 spikes s⁻¹ versus 2.96 ± 1.94 (P = 0.0662) for Rs; 11.94 ± 11.48 spikes s⁻¹ versus 4.22 ± 4.48 (P = 0.0419) for Mk; and 11.68 ± 9.00 spikes s⁻¹ versus 5.59 ± 4.45 (P = 0.0097) for Rj. A few examples of PSTHs of neurons for monkey Mk from both narrow and wide classes used are shown in Fig. 2c (more in Supplementary Fig. 1 and for other monkeys in Supplementary Fig. 2). By summing spikes over all neurons within a class, the population spike rates of the two classes for monkey Mk were different particularly around the time that LFP beta power increased (Fig. 2d, and for other monkeys in Supplementary Fig. 2). Temporally, the spectrogram of the narrow-spiking population showed sustained power over the beta range, and the power increased when the LFP beta power increased (Fig. 2e). In contrast, the wide population exhibited a systematic power decrease in the beta and higher frequency bands around the onset of LFP beta power increase (Fig. 2f). Moreover, the spectrum of the population activity associated with narrow spike waveforms exhibited a spectral peak that was very close to the LFP beta frequency peak, whereas the population spectrum associated with wide spike waveforms exhibited no prominent beta peak (right side of Fig. 2f).

A number of studies have classified cortical neurons based on their extracellular spike widths, and some have suggested that narrow-spiking neurons represent local inhibitory neurons23–28. However, the range of spike widths that defined our narrow class is larger than that of many of these studies. Given that local inhibitory neurons comprise only ~20% of all neurons in neocortex29,30, the large proportion of spiking neurons that we classified as narrow suggests that this class includes excitatory pyramidal neurons as well as local inhibitory neurons that, as a whole, have a tendency to synchronously oscillate in the beta frequency range.

Directed effective connectivity. We developed a generative model to characterize the spiking responses of a given neuron based on the past spiking of all other simultaneously recorded neurons in the network. This model is a generalization of Granger causality to point processes, and can be used to infer excitatory- and inhibitory-directed effective connections among multiple neurons, unlike popular pairwise network estimation procedures31. Specifically, we estimated the conditional intensity
function (CIF) of each neuron (a receiving neuron), \( \lambda(t|H(t)) \), given the spiking history, \( H(t) \), of all other neurons (sending neurons) in the network using a generalized linear model (GLM). The spike history from each sending neuron was composed of five 3 ms time windows for a total of 15 ms. The history length was selected based on the time it would take for the beta waves to propagate across the array at the mean estimated beta wave propagation speeds (Fig. 1e). The underlying assumptions of the model are twofold: (1) the present spike count of a receiving neuron conditioned on the past spike counts of the sending neurons is Poisson distributed, and (2) the log spike count of the receiving neuron is a linear function of the spike counts of the sending neurons. To check the sufficiency of the models used in the analysis, we performed a model-order selection analysis using the BIC, and found that the majority of the neurons showed an optimal spike history duration of \(<15\, \text{ms} \,(\approx 96\% ; \text{ref. 32})\). To assess model goodness-of-fit (GOF) for each neuron, we used the time-rescaling theorem\(^{33}\) and Kolmogorov–Smirnov (KS) plots to define the GOF area (GOFA) and GOFA ratio (GOFR; Methods and Supplementary Fig. 3a–c). More than 85% of neurons exhibited GOFA to be \(<0.05\) (Supplementary Fig. 3d) and \(>68\%\) of neurons showed GOFR to be \(<0.2\) (Supplementary Fig. 3e). Thus the majority of the neurons used in the analysis were reasonably well characterized as Poisson GLM.

As in the standard Granger causality protocol, an effective connection from the sending neuron to the receiving one was inferred by comparing the log-likelihoods of the data under models with and without a particular sending neuron’s spiking history. If the performance of the model decreased significantly (\( P<0.05, \chi^2\)-test) when a sending neuron was removed, an effective connection was inferred, and the sign of the sum of the five history terms determined whether the connection was excitatory or inhibitory\(^{34}\). We estimated the complete network of significant directed connections between recorded neurons with narrow spike waveforms and mean spike rates of at least 1 Hz (to ensure numerical stability) in different time windows. The temporal evolution of the networks in 150 ms windows incremented by 50 ms steps is illustrated in Fig. 3a for monkey Rs with 38 narrow neurons, Fig. 4a for monkey Mk with 21 narrow neurons and Supplementary Fig. 4a for monkey Rj with 33 narrow neurons. The largest number of connections was observed between 100 and 350 ms after the appearance of a new visual target at a time slightly after LFP beta power and the PGD reached their peak values. The temporal evolution of network topology estimates, based on the number of excitatory, inhibitory and all connections, was highly reliable as determined by calculating networks on non-overlapping subsets of data within each data set (Fig. 5 and Supplementary Fig. 5). Furthermore, in order to check whether our connectivity results depended on the precise spike timing among neurons or were a consequence of rate modulation (neurons with very high spike rates are more likely to have effective connections), we randomly shuffled the trial order for each neuron and within each of eight target directions, and recomputed network interactions. We found that...
the number of connections were significantly lower as compared with the unshuffled networks (Supplementary Fig. 6 for monkeys Rs and Mk), suggesting that effective connections reflect single-trial coordination across neurons, rather than overall rate modulations or tuning properties.

To place the estimated networks in a spatial context, we computed the orientation of each connection on the cortical surface. For each cortical orientation, we summed up each connection for that orientation. This normalization was performed to account for anisotropies in the spatial arrangement of the recorded neurons. Between –50 and 100 ms relative to visual target onset, few connections were observed and very little spatial structure was evident. The number of connections increased in subsequent epochs, and peaked between 50 and 250 ms after the onset of a new visual target. The clear bimodality in the circular distribution of the excitatory connection network emerged between 150 and 250 ms after the visual target onset (Fig. 3c for monkey Rs, Fig. 4c for monkey Mk and Supplementary Fig. 4c for monkey Rj, and Supplementary Fig. 7).

Furthermore, this bimodality remained even when we used one spike-width threshold value, the weighted average of the means of the two Gaussians fit on all spike widths (Supplementary Fig. 8).

By comparing the network topology during early and late time windows (Fig. 6a, left and right respectively), we observed the emergence of a larger number of significant connections and a strong anisotropy of directed excitatory connections. The circular distribution of excitatory connection directions became bimodal and closely oriented to the beta wave axis during the later time window in all three monkeys. By fitting the beta wave directional distributions with a mixture of two von Mises distributions, we determined the mode and the circular s.d. (CSD). We found that the modes of excitatory connections fell within 1.1 times the CSD of the beta wave modes. In particular, the beta wave primary/secondary modes measured with respect to the anatomical medial direction (± CSD) versus excitatory connection modes were 123/279 ± 34/52 degrees versus 130/290 degrees (Rs), beta wave modes of 120/258 ± 18/60 degrees versus 140/50 degrees (Mk) and beta wave modes of 297/104 ± 72/51 degrees versus 270/90 degrees (Rj; Fig. 6b, left). The circular distributions of inhibitory connection directions were also bimodal but oriented almost orthogonal to the beta wave axis (Fig. 6b, right). In contrast to the...
narrow-spiking neurons, there were few, if any, statistically significant connections among wide-spiking neurons in each time window (Fig. 6c).

For excitatory connections between narrow-spiking neurons, we took the mean of the five time delays weighted by their model coefficient values, and computed the average time delay between each excitatory connection. Dividing the distance of each excitatory cell pair projected along the beta wave axis by the average time delay provided an estimate of the 'conduction' velocity of each excitatory connection along the beta wave axis. Estimated velocities ranged from 10 to 30 cm s\(^{-1}\) across all six time windows and four data sets, in rough agreement with our estimated wave propagation speeds.

Spatial anisotropy and target information. In order to investigate any behavioural relevance to the spatial anisotropy of excitatory connections among narrow neurons, we next investigated the information content of the sequential spiking activities between neuron pairs with excitatory connections. Sequential spiking activity at time \(t\) was defined as the product of the receiving neuron's activity at time \(t\) and the integrated activity of the sending neuron's activity in the 15-ms time interval preceding time \(t\). We found systematic differences in sequential spiking activity as a function of target direction. To assess this quantitatively, we calculated the mutual information between target direction and the sequential spiking of narrow neuron pairs that possessed excitatory effective connections as a function of time (Fig. 7a). For each of 18 cortical orientation bins, the information was summed across all cell pairs with excitatory connections and divided by the total number of possible connections in that orientation. It should be noted that we added information across cell pairs in order to capture relative information trends in time and across different cortical orientations. Absolute information values cannot be inferred because addition assumes a lack of redundancy across cell pairs, an assumption that may not be warranted. Spike sequencing provided peak target-direction information \(\sim 150–250\) ms for monkey Rs, 100–150 ms for monkey Mk and 150–300 ms for monkey Rj. The circular distribution of mutual information exhibited spatial anisotropy
whose peaks were oriented along the beta wave axis, 130 degree for monkey Rs, 140 degree for monkey Mk and 90 degree for monkey Rj with respect to the medial direction, and they were all within the beta wave primary mode direction ± CSD (Fig. 7b). Cross-correlation coefficient peaks (within ± 1 orientation bin) between the mutual information and beta wave distributions were 0.520, 0.353 and 0.252 for monkey Rs, Mk and Rj, respectively. Spike sequencing of narrow-spiking neurons that were not effectively connected showed an order of magnitude of smaller amount of total mutual information compared with those pairs with excitatory effective connections, and the spatial pattern of information did not resemble the distribution of beta wave directions (Fig. 7c). To verify that this spatial pattern of information was specific to the sequential firing of neuron pairs, we computed target-direction information from single-unit spiking activity (Supplementary Fig. 9a,b). The circular distribution of orientations of neuronal pairs weighted by the mean single-neuron mutual information values averaged over the two neurons with an excitatory connection between them revealed a weaker alignment to the beta wave axis (Supplementary Fig. 9c) than that of mutual information of sequential spiking of neurons with excitatory connections. Cross-correlation peaks (within ± 1 orientation bin) between mean single-neuron information and beta wave distributions were 0.165, 0.135 and 0.102 for monkey Rs, Mk and Rj, respectively (Supplementary Fig. 9d,e).

Discussion
Propagating wave activity is ubiquitous throughout the cortex but has never been characterized at the single-neuron level. Although our results do not directly demonstrate that single units propagate their spiking activity in a wave-like manner across the motor
cortical sheet, we show for the first time spatiotemporal spike patterning that closely matches propagating wave activity as measured by LFPs in terms of both its spatial anisotropy and its transmission velocity. Although the horizontal anisotropy of directed effective connections among the narrow-spiking neurons has never before been experimentally documented (however, effective anisotropy was used in a modelling study22), anatomical data on the horizontal connectivity of motor cortex may shed light on these observations. By staining degenerating axons and terminals after localized lesions within MI, Gatter and Powell documented a preponderance of horizontal axons in layers 2/3 and 5 in non-human primates that were more spatially extensive along the rostro–caudal dimension as compared with the medio–lateral dimension35. In addition, Huntley and Jones36 used retrograde labelling injections in primate motor cortex and found weak strips of labelled cell bodies arranged in the rostro–caudal dimension. However, these data do not provide evidence for stronger (anatomical or functional) connections in the rostro–caudal dimension.

Physiological data also suggest that the functional connections we observed may reflect properties of horizontal connectivity within motor cortex. Electrical stimulation and current-source density analysis in rat motor cortical slices found long-distance excitatory, monosynaptic interactions mediated by horizontal connections in layers 2/3 and 5 with a conduction velocity of ~10 cm s\(^{-1}\), which is on the same order of magnitude as the propagation velocities we documented with the LFP wave and the excitatory connections in our effective connectivity analysis37. Based on spatially smoothed multunit responses recorded from the auditory cortex, propagating activity was documented with comparable propagation speeds of ~20 cm s\(^{-1}\) using a similar gradient field method11.

Although the beta wave propagation axes and propagation speeds were consistent across all the animals used in the study, there are two issues to be resolved in future studies. First, the orientation of the primary beta wave direction was not consistent across the three monkeys. It is not obvious why this would be the case, but it may be owing to the exact placement of the arrays in the medio/lateral dimension along the motor strip. Second, our monkeys exhibited a fairly wide range of hand movement speeds while wave propagation speeds were fairly consistent. We would like to investigate in more detail whether there is a relationship between the kinematics, kinetics and muscle activity of the limb, and the statistical properties of the beta waves.

Our observation that population activity among narrow-spiking neurons exhibited a pronounced beta spectral peak suggests that beta oscillations in the motor cortex may be mediated by a network of functionally distinct cells38. Simulation studies have suggested that beta oscillations emerge in network interactions of local inhibitory neurons22, which in turn entrain excitatory pyramidal neurons to fire rhythmically39. These local inhibitory and excitatory pyramidal cells may represent the population of narrow-spiking neurons that tend to oscillate at the beta frequency range. In contrast, wide-spiking neurons as a population do not resonate at any particular frequency, and, moreover, exhibit a reduction in oscillatory power in the beta and higher frequency bands around the time at which LFP beta power increases.

The temporal dynamics of planar wave activity, spatially oriented effective connections and motor behaviour indicate the following chain of events. The degree of planar wave activity

---

**Figure 6 | Spatiotemporal patterns of network connectivity of narrow and wide classes of neurons in two time windows.** Data from monkeys Rs (top), Mk (middle) and Rj (bottom). (a) Networks of significant directed connections on the array at different time windows. Red and blue arrows represent excitatory and inhibitory connections, respectively. The black dots represent the relative positions of the electrodes on the array where the neurons were detected. If >1 neuron was detected on a given electrode, a given dot contains all units on that electrode. Neurons with firing rates ≥1 spike s\(^{-1}\) and narrow spike widths (≤ ~0.2667 ms) in MI were analysed. The green scale bar on the right-bottom equals 400 μm on the cortical surface. R, rostral; C, caudal; M, medial and L, lateral denote the anatomical orientation. (b) Circular distributions of directed excitatory (left) and inhibitory (right) connections weighted by their strength and normalized by the total number of possible connections. The circular distribution of excitatory connections for time window [−50, 100] ms in purple overlies that for time window [150, 300] ms in red. The circular distributions of inhibitory connections for the same windows are in teal and blue, respectively. The dashed black lines represent the beta wave axis. (c) Same as in a, but using neurons with firing rates ≥1 spike s\(^{-1}\) and wide spike widths (≥0.4000 ms).
begins increasing at target appearance and reaches a maximum at 100–150 ms later. Thus, this enhancement of planar wave activities in the beta oscillation range appears to be triggered by arrival of information about the visual target in MI. At roughly 150 ms after target onset, the rostro-caudal anisotropy of excitatory effective connections emerges (Figs 3 and 4, and Supplementary Fig. 4), and sequential firing among these connected neurons begins carrying information about target direction. Finally, movement speed begins to increase 200 ms after target appearance resulting in movement to the target. It is our view that spatiotemporal dynamics of the LFP provides a mesoscopic perspective of patterning that is generated by spatiotemporal arrangements among groups of single units. Therefore, planar wave activity reflects the underlying spatiotemporal distribution of spiking activity among the units. A recent modelling study22 showed spatial anisotropy among inhibitory cells, while our study showed spatial anisotropy in excitatory connections. However, a recent study using an immature cerebral cortex preparation showed that cortical waves in layer III were blocked by glutamatergic receptor antagonists, but not by GABAergic receptor antagonists49. Furthermore, a modelling study showed that long-range coherent oscillations with some delay were partially induced by interactions of excitatory cells41.

These results may have implications for the development of cortically controlled brain machine interfaces. Current decoding algorithms for brain machine interfaces have ignored the spatial organization of cortical neurons and their functional interactions. In principle, by augmenting the GLM framework, we used here by including external movement covariates, we could invert these spike prediction models to decode movement parameters. Furthermore, a recent study has speculated that spatiotemporally patterned electrical stimulation in motor cortex (similar to the recorded patterns we have observed in our study) could replace deep brain stimulation to treat symptoms of Parkinson’s disease42.

Our information analysis of sequential firing may have important behavioural implications. First, we have shown that functionally connected, narrow-spiking neurons carry more task-relevant information in their sequential firing than do unconnected, narrow-spiking neurons. This implies that there is a subnetwork of neurons within the motor cortex that may be playing a special role in representing and transmitting information to guide behaviour. Second, this behavioural information is most pronounced in the sequential firing of neurons that are spatially aligned along the rostro-caudal axis, implying that the motor cortex possesses information-carrying, spatial structure beyond somatotopy. Future work will investigate whether downstream targets such as the muscles are sensitive to and may take advantage of this spatial structure when driving motor behaviour.

Methods

Behavioural tasks. Three male macaque monkeys (Macaca mulatta) were operantly trained to move a cursor appearing above the monkey’s hand location to targets projected onto a horizontal, reflective surface in front of the monkey. The monkey’s arm rested on arm troughs secured to links of a two-joint exoskeletal
A silicon-based electrode array composed of 100 electrodes (1.0 mm electrode length and 400 µm interelectrode distance) was implanted in the arm area of MI of each monkey. During a recording session, signals from up to 96 electrodes were amplified (gain, 5,000), band-pass filtered between 0.3 and 7.5 kHz, and recorded digitally (14-bit) at 30 kHz per channel using a Cerebus acquisition system (Blackrock Microsystems, Inc., UT). Only waveforms (1.6 ms in duration; 48 sample time points per waveform) that crossed a threshold were stored and spike sorted using Offline Sorter (Plexon, Inc., TX). Signal-to-noise ratios were defined as the difference in mean peak-to-peak voltage divided by twice the mean s.d. The mean s.d. was computed by measuring the s.d. of the spike waveform at each of the 48 sample time points of the waveform and then averaging. All isolated single units used in this study possessed signal-to-noise ratios of 3:1 or higher. The data for each neuron were converted to a binary time series at a 1 ms time resolution. A total of four data sets (one data set for animal Rs, two data sets from animal Rj), and one data set for animal Rk,) were analysed, where a data set is defined as all simultaneously recorded neural data collected in one recording session. Each data set contained between 59 and 115 simultaneously recorded units from MI. A total of 302 single-unit samples were recorded from MI over all four data sets. We use the term ‘samples’ because the recordings were made over all four data sets. We use the term ‘samples’ because the recordings were made.

### Beta wave characterization

We used the wave characterization method as defined in Rubino et al. Briefly, we first applied a band-pass filter bidirectionally whose center frequency for each data set was defined to be the most prominent peak frequency over the beta oscillation range (10–40 Hz) and the width was ± 3 Hz, then applied the Hilbert transform to a filtered voltage signal recorded by an electrode located at (x,y) on the array (V(x,y,t)) to obtain the instantaneous amplitude A(x,y,t) and phase ϕ(x,y,t) follows:  

$$V(x, y, t) + iHB[V(x, y, t)] = A(x, y, t)e^{iϕ(x,y,t)},$$  

(1)

where H is the Hilbert transform operator and A(x,y,t) is the instantaneous amplitude of the voltage signal, and $f^2 = 1$. The wave velocity $\mathbf{v}(t) = (dx/dt, dy/dt)$ was computed by taking the derivative of a constant of 

$$\phi(x,y,t) = \text{const.}$$  

(2)

with respect to time: 

$$\frac{d\phi}{dt} = \nabla \cdot \mathbf{v} + \frac{\partial \phi}{\partial t} = 0.$$  

(3)

Since the direction of the velocity, $-\nabla\phi$, is perpendicular to the constant phase contour, the speed is 

$$\frac{d\phi}{dt} = \frac{\nabla\phi}{\|\nabla\phi\|}.$$  

(4)

Note that velocity is well defined only when the phase gradient is non-zero and when the signal exhibits a well-defined propagation direction. In order to measure how well the phase gradients align across the array, we used the quantity called the PGD(t):

$$\text{PGD}(t) = \frac{\|\nabla\phi\|}{\|V\|}.$$  

(5)

The bar here denotes the spatial average. If the phase gradients at all spatial locations align at a given time, t, then PGD(t) = 1. If the phase gradients are randomly distributed, then PGD will be close to 0. Thus, we used a threshold value of 0.2, which is based on the tail of the PGD distribution when channel assignments are randomly shuffled. As long as PGD(t) was above the threshold, then we used the estimated values of mean wave direction and speed across the array, which were computed as: 

$$\text{wave direction}(t) = -\nabla\phi,$$ 

$$\text{speed}(t) = \frac{|\nabla\phi|}{\|\nabla\phi\|}.$$  

(6)

### Unit spiking data selection

We created trial-based data sets using a time interval from −100 to 350 ms relative to visual target onset. We had 2,547, 2,746, 1,999 and 1,526 trials from each of the four binned waveforms. We selected only neurons whose average firing rates during the interval of interest, −100 to 350 ms relative to visual target onset, were higher than 1 spike s$^{-1}$. Then, we fitted a mixture of Gaussians to the distribution of waveforms widths where the width was defined as the duration from trough-to-peak of the mean waveform of each sorted unit. We used mixture Gaussian models from 1 to 4 Gaussians for each data set and the pooled data. Then, we computed the BIC for each data set and the pooled data set for each model to find a most parsimonious Gaussian mixture model. Every data set and the pooled data set had a minimum BIC at 2 (Fig. 2b). Thus, we decided to fit each data set with a mixture of two Gaussians. For each data set, we computed the average of the two means fell between 0.3667 and 0.4000 ms so that neurons with waveform widths of 0.4000 ms or less were classified as narrow-spiking units. Similarly, the largest average of the two means fell between 0.3667 and 0.4000 ms so that neurons with waveform widths of 0.4000 ms or more were classified as wide-spiking units. Thus, we had two thresholds. We decided to take this conservative threshold scheme to clearly separate the two populations of neurons. The numbers of neurons that satisfied the narrow condition were 38 neurons (4.89 ± 4.10 spikes s$^{-1}$, mean ± s.d.) among a total of 115 neurons for the first data set (monkey Rs), 14/21 neurons (11.94 ± 11.48 and 9.71 ± 8.59 spikes s$^{-1}$, respectively) among a total 59/65 neurons for the second and third data sets (monkey M), and 33 neurons (11.68 ± 6.00 spikes s$^{-1}$) among a total 63 neurons for the fourth data set (monkey Rj). The numbers of wide-spiking neurons were 29 neurons (2.96 ± 1.94 spikes s$^{-1}$) among a total of 115 neurons for the first data set (monkey Rs), 13/10 neurons (4.22 ± 4.48 and 4.88 ± 5.69 spikes s$^{-1}$, respectively) among a total of 59/65 neurons for the second and third data sets (monkey M), and 6 neurons (3.59 ± 4.45 spikes s$^{-1}$) among a total 63 neurons for the fourth data set (monkey Rj). Neurons that did not fall into either the narrow or wide class were not further analysed. Peri-stimulus time histograms binned for two example neurons used in the analyses are presented based on the spike-width class for each data set in Fig. 2c, and Supplementary Figs 1 and 2. We also broke each data set into three subsets of data consisting of an equal number of trials for the first two monkeys, and into two subsets of data for the third monkey to check for temporal consistency. We analysed each trial using 150 ms time window.

### Power spectra and spectrograms for spiking of population of neurons

We binned each neuron’s spike train into 1 ms bins and summed the binned spike trains over all neurons within the narrow or wide spike waveform population. Then, we used a multivariate spectrum estimation method (time-bandwidth product and number of tapers were 3 and 5, and moving window of 250 ms in 2 ms increments) to compute average power spectra over [−100, 300] ms relative to the visual cue onset and average spectrograms over the time windows specified above using Chronux toolbox.

### Granger causality measure for point process models

We utilized a model-based methodology for analysing the causal interactions of multiple neural spike trains. Discrete, all-or-nothing nature of a sequence of action potentials together with their stochastic structure suggests that neural spike trains may be regarded as point processes. Let Ni(t) denote the sample path that counts the number of spikes of neuron i in the time interval (0,t] for t ∈ (0,T] for i = 1,…,M recorded neurons. A point process model of a spike train for neuron i can be completely characterized by its CIF, λi(t|H(t)), defined as:

$$\lambda_i(t|H(t)) = \lim_{\Delta \to 0} \frac{\Pr[N_i(t + \Delta) - N_i(t) = 1|H(t)]}{\Delta},$$  

(7)

where H(t) denotes the history of the spiking activity of all neurons up to time t. The CIF represents the instantaneous firing probability and serves as a fundamental component for constructing the likelihoods and probability distributions for point process data. It is a history-dependent function, and reduces to a Poisson process if it is independent of the history. To simplify the notation, we denote λi(t|H(t)) as λi(t). We used a point process GLM for building models whose history terms contain past spiking activity of other neurons. To test whether neuron j causes neuron i in the ‘Granger’ sense, we developed two classes of point process GLMs for neuron i (one given by P(Ni(0:T)|J) that has the past of all neurons as the covariates for the CIF, and the other given by P(Ni(0:T)|J) that has the past of all except for neuron j. The point process likelihood is given by

$$P(Ni(0:T) = exp \left\{ \int_0^T \log \lambda_i(t|dN(t) - \int_0^t \lambda_j(s)ds \right\}.$$  

(8)
In equation (9), we used a point process GLM to model the relationship between the spiking activity and its covariates (the spiking history)\textsuperscript{[17]}. Here, $\beta_i$ relates to a background level of the activity of neuron $i$, $t_{on}$ is a 5-dimensional (5D) vector of parameters that relate the past spiking of neuron $m$ to the current spiking of neuron $i$, and $\mathbf{h}_n$ is a 5D vector each of whose elements counts the spikes in every 3 ms time window of 15 ms spiking history of neuron $n$. These parameters were chosen to obtain a relatively small number of parameters while maintaining the necessary temporal resolution. The $\langle \cdot \rangle$ represents the dot product between vectors. Note that for $P(N(0:T)|\lambda)$, $\lambda(t)$ includes the past spiking of all the neurons. The other process likelihood, $Q(N(0:T)|\theta)$, is given by the same equation (8), but with $\lambda(t)$ replaced with $\lambda_i(t)$:

$$
\log \lambda_i(t) = \beta_i + \sum_{m=1}^{M} \langle f(m) \rangle \mathbf{h}_n.
$$

(9)

In equation (9), we used a point process GLM to model the relationship between the spiking activity and its covariates (the spiking history)\textsuperscript{[17]}. Here, $\beta_i$ relates to a background level of the activity of neuron $i$, $t_{on}$ is a 5-dimensional (5D) vector of parameters that relate the past spiking of neuron $m$ to the current spiking of neuron $i$, and $\mathbf{h}_n$ is a 5D vector each of whose elements counts the spikes in every 3 ms time window of 15 ms spiking history of neuron $n$. These parameters were chosen to obtain a relatively small number of parameters while maintaining the necessary temporal resolution. The $\langle \cdot \rangle$ represents the dot product between vectors. Note that for $P(N(0:T)|\lambda)$, $\lambda(t)$ includes the past spiking of all the neurons. The other process likelihood, $Q(N(0:T)|\theta)$, is given by the same equation (8), but with $\lambda(t)$ replaced with $\lambda_i(t)$:

$$
\log \lambda_i(t) = \beta_i + \sum_{m=1}^{M} \langle f(m) \rangle \mathbf{h}_n.
$$

(10)

which excludes the effect of the past spiking of neuron $j$. Model parameters

$$
\theta_j = \{ \beta_j, f_1, \ldots, f_M \}
$$

for $P(N(0:T)|\theta)$ and

$$
\theta_i = \{ \beta_i, f_1, \ldots, f_M, \mathbf{h}_n \}
$$

and were calculated by

$$
\hat{\theta}_j = \arg \max_{\theta_j} P(N(0:T); \theta_j),
$$

(11)

$$
\hat{\theta}_i = \arg \max_{\theta_i} P(N(0:T); \theta_i).
$$

(12)

If the spiking history of neuron $j$ helps predict the spiking activity of neuron $i$, the log-likelihood ratio should be significantly greater than zero, and then we say that neuron $j$, ‘Granger-causes’ $i$\textsuperscript{[19]}. The equality of the likelihood ratio holds when neuron $j$ has no causal influence on $i$. Excitatory and inhibitory influences from neuron $j$ to $i$ are distinguished by the sign of $\sum_{j} f_j(t)$ in equation (9), which represents the summed influence of the past spiking of $j$ on $i$.

**Significance test.** The Granger causality measure, $D = D_1 - D_2 = 2\Delta g$, provides an indication of the relative strength of causal interaction but gives little insight into whether or not it is statistically significant. We use a statistic based on the log-likelihood ratio test to address this issue. We denote the deviance obtained using the model parameter $\hat{\theta}_i$ as $D_0$, and the deviance obtained using $\theta_i$ as $D_2$. The deviance difference between two models is equivalent to two times log-likelihood ratio given by $D_0 = D_2 - D_1 = 2\Delta g$ (ref. 50). If both models describe the data well, then the deviance difference may be asymptotically described by $\Delta D \sim \chi^2_d$ where $d$ is the dimensionality of two models, and $\Delta g = 5$ in our study. If the value of $\Delta D$ is in the critical region, that is, greater than the upper tail $100(1 - a/2)$ of the $\chi^2_d$ where $a$ determines false-positive rates, then the causal influence is determined as statistically significant. We used a multiple-hypothesis testing error measure called false discovery rate to address the multiple comparison problem\textsuperscript{[23]}. 

**Assessment of GOAF against the data for the point process models.** The validation of a designed statistical model for data is required before making an inference based on the model. The time-rescaling theorem was used to transform point process neural spike train data to a continuous measure appropriate for GOAF assessment\textsuperscript{[33]}. Once a CIF is estimated, rescaled times can be computed using the estimated CIF. These rescaled times will be uniformly distributed random variables on the interval [0,1] if the estimated CIF is a good approximation to the true conditional intensity of the point process. To evaluate whether the rescaled times followed the uniform distribution, we ordered these rescaled times from the smallest to the largest, and then plotted the quantiles of the cumulative distribution function of the uniform distribution on [0,1] against the ordered rescaled times. This form of graphical representation is termed a KS plot. If the model is consistent with the data, then the points should lie on a 45-degree line. Approximate 95% confidence bounds for the degree of agreement between the model and the data may be constructed using the distribution of the KS statistic. We defined the GOAF bounded by the 45-degree line and the cumulative distribution in the KS plot. We also defined the GOAF computed as the ratio of GOAF to an area out of the ~95% confidence bounds for the degree of agreement (see Supplementary Fig. 3a for schematics of the method). These measures were computed for each neuron in each 150 ms time window.

**Causal inference with shuffled trials.** We estimated causality networks after shuffling the trial order of each neuron for the same target direction using the data after the onset of visual cue. Through this analysis, we tested how much the precise spiking times among multiple neurons contributed to significant functional interactions between them. If significant functional interactions were simply caused by an increase in firing rate after the onset of visual cue, shuffling the trials would not affect the network estimates. On the other hand, if they were caused rather by the increase of relative spike timings among multiple neurons, shuffling the trials would break the relative spike timings among neurons and thus lead to a sparser network estimate. For this analysis, we used the data set in the window 5 (after the onset of cue) of monkeys Mk and Rs. For each target direction, we permuted the order of all the trials for each neuron and then estimated the causality network using the shuffled data set. We repeated this procedure 100 times. The mean number of connections was 2.05 with an s.d. of 1.79 compared with 56, 65 and 66 connections for monkey Mk, and the mean number of connections was 12.40 with an s.d. of 1.50 compared with 100, 105 and 128 connections for monkey Rs from the original unshuffled network. This analysis suggests that the increase of significant interactions between neurons after the onset of visual cue is driven by the organized spiking sequencing among neurons rather than an increase of spiking rate.

**Mutual information analysis for single-unit spiking activities and spike sequencing.** We evaluated the information content of single-unit spiking activities and spike sequencing using the mutual information between the target direction and single-cell spiking, or spike sequencing between two neurons for each cortical orientation, respectively. Mutual information between two variables represents the reduction in uncertainty of one variable due to the knowledge of the other. It is defined as

$$
I(X, Y) = H(X) - H(X|Y),
$$

(13)

for random variables $X$ and $Y$ where $H(X)$ is the Shannon entropy of $X$ (ref. 52). In this paper, $Y$ is a random variable representing the target directions, and takes on one of eight directions in the eight octants in 2D plane during the RTP task. For single-neuron analysis, $X$ represents the binary sequence of unit spiking activity. For spike sequence analysis, $X$ represents the spike sequencing from neuron $i$ to $j$ defined by

$$
\tau(X) = \sum_{t=1}^{T} n_i(t) - \mu_i(t),
$$

(14)

which is calculated as the sum of products in spike counts between neuron $i$ at varying time lags $\tau$ (from $\tau_1$ to $\tau_2$) and neuron $j$. For the spike sequencing, we used $\tau_1 = 1$ ms and $\tau_2 = 15$ ms. We chose a 15 ms summation window to match with the duration of history terms used for the GLM described above. Information content of the spike sequencing was calculated using the empirical distributions of $X(t)$ at fixed $i$ and $j$, and $\tau$ over all trials for different target directions. Using the position information of neurons on the multielectrode array, we calculated the cortical orientation of the information content of spike sequencing. To account for biases in estimating information from limited data, the target direction labels were randomly shuffled over all trials so that a trial originally associated with target direction $\xi$ could be randomly assigned one of the eight target directions. Mutual information between target direction and spike sequencing was then calculated. The average information from 10 shuffles was subtracted from the estimated information value.

All the numerical analyses were performed using Matlab (Mathworks, MA). Matlab codes for Granger analysis are available at the website (http://www.neurostat.mit.edu/gcpp).
