

ONLINE METHODS

Human participants, electrophysiology and behavioral tasks. An investigational device exemption (IDE) for these studies was obtained from the US Food and Drug Administration and all studies were performed with approval from Institutional Review Boards: Spaulding Rehabilitation Hospital Institutional Review Board, New England Institutional Review Board, Rhode Island Hospital Institutional Review Board, Partners Human Research Committee. The recording device, preamplifiers, data acquisition systems and computer are part of the BrainGate Neural Interface System (Cyberkinetics Neurotechnology Systems, Inc). CAUTION: Investigational device. Limited by Federal Law to Investigational Use. The sensor is a 10×10 array of silicon microelectrodes that protrude 1 mm (hS1) or 1.5 mm (hS3) from a 4.2×4.2 mm platform (**Supplementary Fig. 1**). For signal acquisition, 96 electrodes are available, with minimum inter-electrode distance of 400 μm . Participant 1 (hS1) was a 24-year-old male with tetraplegia (C4 ASIA A). Participant 3 (hS3) in the pilot clinical trial is a 56-year-old female who sustained a pontine stroke 9 years before trial enrollment, resulting in loss of speech and locked-in syndrome, which later resolved to incomplete tetraplegia. After obtaining informed consent and carrying out the medical and surgical screening procedures, the array was implanted in the dominant M1 hand/arm area, identified anatomically as the ‘knob’ region^{31,33} of the precentral gyrus in pre-operative magnetic resonance imaging.

Each participant used M1 spiking activity to control a cursor displayed on a computer screen^{31,32}. The participant was instructed to imagine moving a circular cursor displayed on the screen to one of four peripheral targets, positioned at 0, 90, 180 and 270°. In each session, 20 trials were collected for each of

the four pseudo-randomly presented radial targets. The 3 s period after target appearance was included in the analysis. Two datasets per participant were used, corresponding to two research sessions conducted on two different days. Data sets were collected on 2004.09.16 and 2004.09.20 (hS1) and 2006.01.23 and 2006.01.24 (hS3). We used spike-sorting utilities in Offline Sorter (Plexon) to identify and sort neuronal units in all of the human and nonhuman primate recordings. We did not distinguish whether a single unit sorted from the same electrode on different days corresponded to the same neuron or not.

Nonhuman primate subjects, electrophysiology and behavioral tasks. Two datasets were recorded in two experimental sessions from each of four rhesus monkeys (*Macaca mulatta*). All recordings were obtained via single or dual cortically implanted 10×10 microelectrode arrays (electrode length, 1.0 mm), similar to the array described above. M1 neurons from monkeys mLA and mCL were recorded while they performed point-to-point planar movements. The monkeys used a manipulandum to move a position feedback cursor that was presented on the monitor. Targets were randomly placed (that is, uniform in two-dimensional space), one at a time, on the workspace. After the successful acquisition of a random number (3–9) of targets, the monkeys received a juice reward. Only segments of data recorded during the reaching phases of the tasks, from two experimental sessions, were included in the analyses (datasets: mLA, 2004.03.25 and 2004.03.26; mCL, 2004.03.25 and 2004.03.29). M1 and PMv neurons were simultaneously recorded from monkey mCO while this monkey performed reaching and grasping movements toward objects moving in the workspace (C.E. Vargas-Irwin, P. Yadollahpour, G. Shakhnarovich, M.J. Black and J.P. Donoghue, *Soc. Neurosci. Abstr.* 673.18, 2008). A motion-capture system was used to record arm-hand configurations and related behavioral epochs. Only data segments corresponding to 1-s segments during the reach-grasp phase before the final object grasp, from two sessions, were included in the analyses (datasets: 2008.03.19 and 2007.12.12). M1 and 5d neurons were simultaneously recorded via dual arrays from monkey mAB. The monkey performed visually guided pursuit tracking of a circular cursor projected on a horizontal screen while wearing an external device for kinematic measurements (Kinarm, BKIN Technologies). This cursor followed randomly generated trajectories of varying speeds over the planar, horizontal workspace (B.A. Philip and J.P. Donoghue, *Soc. Neurosci. Abstr.* 672.22, 2008). Only data segments recorded during tracking, from two sessions, were included in the analyses (datasets 2008.05.08 and 2008.05.09). All procedures were in accordance with Brown University Institutional Animal Care and Use Committee approved protocols and the Guide for the Care and Use of Laboratory Animals.

Point process history models. The distribution of the point process sample paths of a given i^{th} neuron is completely specified^{16,18,34} by the conditional intensity function (instantaneous conditional spiking rate)

$$\lambda_i(t | \mathcal{H}(t), z(t)) = \lim_{\Delta \rightarrow 0} \frac{\Pr(N_i(t + \Delta) - N_i(t) = 1 | \mathcal{H}(t), z(t))}{\Delta} \quad \text{equation (3)}$$

where $\Pr(\cdot | \cdot)$ is a conditional probability, $N_i(t)$ denotes the sample path (that is, a right-continuous function that jumps 1 each time a spike occurs), $\mathcal{H}(t)$ denotes the conditioning intrinsic and ensemble spiking histories up to, but not including, time t , and $z(t)$ denotes other relevant extrinsic covariates, such as stimuli and behavioral variables. We focused on intensity function models conditioned on spiking histories (see Supplementary Fig. 3 for analyses involving conditional intensity function models that also included extrinsic covariates such as hand position and velocity). The sample path distribution for the discrete time point process belongs to the exponential family with canonical parameter $\log(\lambda(t | \mathcal{H}_t)\Delta)$, which we modeled as

$$\log(\lambda_i(t | \mathcal{H}_t)\Delta) = \mu_i + \sum_{k=1}^{10} K_{1,i,k} \cdot x_i + \sum_{j \neq i} \sum_{k=1}^4 K_{2,i,j,k} \cdot x_j \quad \text{equation (4)}$$

where t indexes discrete time, $\Delta = 1$ ms, μ_i relates to a background level of spiking activity, and x_i denotes the spiking history (spike train) in the time interval $(t - 100 \text{ ms}, t)$ for the i^{th} neuron, with $x_{i,t} \in \{0, 1\}$ for $i = 1, 2, \dots, n$ recorded neurons. $K_{1,i,k}$ and $K_{2,i,j,k}$ consisted of temporal basis functions of the raised cosine type⁹ with coefficients to be estimated. Ten and four basis functions were used for the intrinsic and ensemble history filters, respectively. Thus, $K_{1,i}$ and $K_{2,i,j}$ in equation (2) consisted of nonparametric temporal filters for the intrinsic and ensemble spiking histories, respectively. Consistent with known neurophysiology and measured autocorrelation functions of the recorded spike trains, we enforced an absolute refractory period of 2 ms in the intrinsic history component. A history model for a particular neuron did not include the spiking history of other neurons recorded by the same electrode. That was done to avoid potential negative correlation artifacts, especially at zero and short time lags, commonly introduced by current spike thresholding and sorting algorithms. This rule was also adopted in the computation of the distribution of pair-wise correlation coefficients. Model parameters were estimated via gradient-ascent maximization of the penalized log-likelihood functions¹⁸. A regularization term in the form of a ridge regression penalty was added for the model parameters related to the

ensemble history effects. After estimating a conditional intensity function model, the probability of a given neuron spiking at any given 1-ms time bin, conditioned on past spiking histories, was obtained as¹⁶

$$\Pr(x_{i,t} = 1 | \mathcal{H}_t) = \lambda_i(t | \mathcal{H}_t)\Delta + o(\Delta) \approx \lambda_i(t | \mathcal{H}_t)\Delta \quad \text{equation (5)}$$

The term $o(\Delta)$ relates to the probability of observing more than one spike in a 1-ms time interval.

Instantaneous collective states: pair-wise maximum-entropy point process models. The total interdependence in multivariate stochastic processes can be decomposed into two main components³⁵: a time ‘causal’ component (that is, the statistical dependence of current states on past events) and an instantaneous component (for example, instantaneous dependencies among neurons). We considered instantaneity at two time resolutions: 1- and 10-ms time bins. When using the 10-ms resolution, the rare cases of time bins with more than one spike were represented as 1-spike events. We estimated statistical interdependencies in these instantaneous collective states by first fitting zero time-lag pair-wise maximum entropy distribution models^{7,22}. Estimation of probability distributions for high-dimensional systems without further constraints is typically an intractable problem. On the other hand, second-order statistics in the form of pair-wise correlations are still feasible to compute and maximum entropy distributions constrained on pair-wise correlations can then be estimated. This maximum entropy distribution model constrained on mean rates and pair-wise zero time-lag correlations^{7,23} is given by

$$\Pr(x_{1,t}, x_{2,t}, \dots, x_{n,t}) = \frac{1}{Z_{(\alpha, \beta)}} \exp \left\{ \sum_i \alpha_i x_{i,t} + \frac{1}{2} \sum_{i \neq j} \beta_{i,j} x_{i,t} x_{j,t} \right\} \quad \text{equation (6)}$$

where $x_{i,t} \in \{-1, 1\}$, corresponding to no spike and spike, respectively, $Z_{(\alpha, \beta)}$ is a normalization constant, $\{\alpha_i\}$ reflects constraints imposed by the empirical mean spiking rates, and $\{\beta_{ij}\}$ reflects constraints imposed by the zero time-lag pair-wise correlations, with $\beta_{ij} = \beta_{ji}$. The conditional spiking probability of a given neuron under this pair-wise maximum entropy distribution model is given by

$$\Pr(x_{i,t} = 1 | x_{\setminus i,t}) = \frac{1}{2} \frac{\exp \left\{ \alpha_i + \sum_{j \neq i} \beta_{i,j} x_{j,t} \right\}}{\exp \left\{ \alpha_i + \sum_{j \neq i} \beta_{i,j} x_{j,t} \right\} + \exp \left\{ -\alpha_i - \sum_{j \neq i} \beta_{i,j} x_{j,t} \right\}} \quad \text{equation (7)}$$

where $x_{\setminus i,t}$ denotes the observed neuronal ensemble state not including the i^{th} neuron. Parameters of these pair-wise maximum entropy models were estimated via maximum pseudolikelihood³⁶. The consistency

of this estimator has been previously demonstrated³⁷⁻³⁹, as has its relationship to contrastive divergence methods³⁹. We used Gibbs sampling²⁶ to sample from the estimated maximum entropy models and compute the distributions of multiple-neuron spike coincidences in 1- and 10-ms time bins. For practical reasons, and in contrast with the ensemble history point process models, the data used to fit the maximum entropy distribution models also included neurons isolated in the same recording channel. This could have, if anything, helped to improve the performance of these maximum entropy models in comparison with the ensemble history models.

ROC analysis. ROC curves are a standard tool for the analysis of prediction performance¹⁹. After estimating a conditional intensity function model on training data, the probability of a given neuron spiking at any given 1-ms time bin, conditioned on either past spiking histories or instantaneous states, was computed on test data according to equations (5) and (7), respectively. A tenfold-crossvalidation scheme was used. From this probability, true- and false-positive prediction rates were computed, resulting in the ROC curves. We used the AUC to derive a predictive power measure. Informally, the relationship between AUC and predictive power can be expressed as follows. Consider the case where all of the 1-ms time bin samples are separated into two populations, one consisting of samples with a spike and the other consisting of samples with no spikes. Next, consider randomly drawing two samples, one from each population. The AUC gives the probability that our conditional intensity function model will assign a higher probability (that is, a higher instantaneous spiking rate) to the sample from the spike population than to the sample from the no-spike population²⁰. The AUC therefore provides an assessment of the discriminatory or predictive power for predictive variables under a given model. It also relates to the Wilcoxon-Mann-Whitney U statistics in the case of independent samples. Given the temporal dependencies in our data, we computed the AUC directly from the ROC curve and used random permutation tests to establish the statistical significance of estimated values. Typical confidence intervals were extremely narrow overall. The ROC curve of a chance level predictor asymptotes the diagonal line, resulting in a AUC of 0.5. In our datasets, the AUC corresponding to chance prediction could be slightly larger than 0.5. We defined a predictive power, with respect to this chance level, as $2 \times (AUC - AUC^*)$, where AUC^* is the AUC corresponding to a chance level predictor for a particular neuron and model, estimated via random permutation methods. The scaling by a factor of 2 was introduced so that the predictive power ranged from 0 (no predictive power) to 1 (perfect prediction).

Information rates. We computed an information rate that estimates how much reduction in the uncertainty about whether or not a neuron will spike in any given time bin, conditioned on knowing only the mean spiking rate, is achieved by also knowing spiking histories and their estimated effects under a given history model. Given large enough number of samples and under an ergodicity assumption, this information rate can be approximated as²³

$$\frac{1}{T} \sum_{t=1}^T x_{i,t} \log(\lambda_i(t | \mathcal{H}_t) \Delta) - \lambda_i(t | \mathcal{H}_t) \Delta - \left(\bar{\lambda}_i \Delta \log(\bar{\lambda}_i \Delta) - \bar{\lambda}_i \Delta \right) \quad \text{equation (8)}$$

where $x_{i,t} \in \{0,1\}$, $\bar{\lambda}_i$ is the mean spiking rate and $\lambda_{i,t} = \lambda_i(t | \mathcal{H}_t)$. The summation term corresponds to the average log-likelihood under the history model, with averages computed over T samples (1-ms time bins) and the second term corresponds to the average log-likelihood under a homogeneous Poisson process with the specified mean rate. Computed to base 2 and normalized by Δ , the above quantity corresponds to an information rate in bits per second.

33. Yousry, T.A. *et al.* Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain* **120**, 141–157 (1997).
34. Daley, D.J. & Vere-Jones, D. *An Introduction to Theory of Point Processes* (Springer, New York, 2003).
35. Geweke, J. Measurement of linear dependence and feedback between multiple time series. *J. Am. Stat. Assoc.* **77**, 304–314 (1982).
36. Besag, J. Statistical analysis of non-lattice data. *Statistician* **24**, 179–195 (1975).
37. Geman, S. & Graffigne, C. Markov random field image models and their application to computer vision. *Proc. Int. Congr. Math.* 1496–1517 (1986).
38. Gidas, B. Consistency of maximum likelihood and pseudo-likelihood estimators for Gibbsian distributions. in *Stochastic Differential Systems, Stochastic Control Theory and Applications*. (eds Fleming, W. & Lions, P.-L.) (Springer, New York, 1988).
39. Hyvärinen, A. Consistency of pseudolikelihood estimation of fully visible Boltzmann machines. *Neural Comput.* **18**, 2283–2292 (2006).