

A high-performance brain-computer interface

Gopal Santhanam^{1*}, Stephen I. Ryu^{1,2*}, Byron M. Yu¹, Afsheen Afshar^{1,3} & Krishna V. Shenoy^{1,4}

Recent studies have demonstrated that monkeys^{1–4} and humans^{5–9} can use signals from the brain to guide computer cursors. Brain-computer interfaces (BCIs) may one day assist patients suffering from neurological injury or disease, but relatively low system performance remains a major obstacle. In fact, the speed and accuracy with which keys can be selected using BCIs is still far lower than for systems relying on eye movements. This is true whether BCIs use recordings from populations of individual neurons using invasive electrode techniques^{1–5,7,8} or electroencephalogram recordings using less⁶ or non-invasive⁹ techniques. Here we present the design and demonstration, using electrode arrays implanted in monkey dorsal premotor cortex, of a manifold higher performance BCI than previously reported^{9,10}. These results indicate that a fast and accurate key selection system, capable of operating with a range of keyboard sizes, is possible (up to 6.5 bits per second, or ~15 words per minute, with 96 electrodes). The highest information throughput is achieved with unprecedentedly brief neural recordings, even as recording quality degrades over time. These performance results and their implications for system design should substantially increase the clinical viability of BCIs in humans.

Most BCIs translate neural activity into a continuous movement command, which guides a computer cursor to a desired visual target^{1–3,5–9}. If the cursor is used to select targets representing discrete actions, the BCI serves as a communication prosthesis. Examples include typing keys on a keyboard, turning on room lights, and moving a wheelchair in specific directions. Human-operated BCIs are currently capable of communicating only a few letters per minute (~1 bits per second (bps) sustained rate⁹) and monkey-operated systems can only accurately select one target every 1–3 s (~1.6 bps sustained rate¹⁰), despite using invasive electrodes.

An alternative, potentially higher-performance approach is to translate neural activity into a prediction of the intended target and immediately place the cursor directly on that location. This type of control is appropriate for communication prostheses and benefits from not having to estimate unnecessary parameters such as continuous trajectory^{4,11}. We conducted a series of experiments to investigate how quickly and accurately a BCI could operate under direct end-point control.

We used a standard instructed-delay behavioural task¹² to assess neural activity in the arm representation of monkey premotor cortex (PMD), as shown in Fig. 1a and described in Methods. As previously

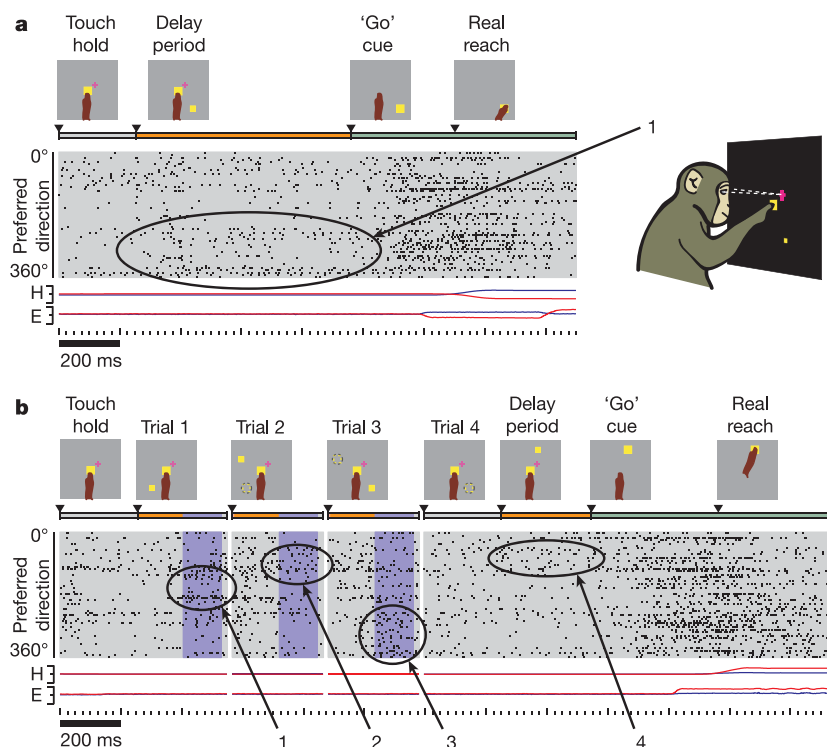


Figure 1 | Instructed-delay (real reach) and BCI (prosthetic cursor) tasks, with accompanying neural data. Large numbered ellipses draw attention to the increase in neural activity related to the peripheral reach target.

a, Standard instructed-delay reach trial. Data from selected neural units are shown (grey shaded region); each row corresponds to one unit and black tick marks indicate spike times. Units are ordered by angular tuning direction (preferred direction) during the delay period. For hand (H) and eye (E) traces, blue and red lines show the horizontal and vertical coordinates, respectively. The full range of scale for these data is ± 15 cm from the centre touch cue. **b**, Chain of three prosthetic cursor trials followed by a standard instructed-delay reach trial. T_{skip} is denoted by the orange parts of the time line. Neural activity was integrated (T_{int}) during the purple shaded interval and used to predict the reach target location. After a short processing time ($T_{\text{dec+rend}} \approx 40$ ms), a prosthetic cursor was briefly rendered and a new target was displayed. The dotted circles represent the reach target and prosthetic cursor from the previous trial, both of which were rapidly extinguished before the start of the trial indicated. Trials shown here are from experiment H20041106.1 with monkey H.

¹Department of Electrical Engineering, Stanford University, 330 Serra Mall, 319 Paul G. Allen Center for Integrated Systems Annex, Stanford, California 94305-4075, USA.

²Department of Neurosurgery, Stanford University School of Medicine, 300 Pasteur Drive, Edwards Building, R-297, Stanford, California 94305-5327, USA. ³Medical Scientist Training Program, Stanford University School of Medicine, 251 Campus Drive, MSOB 309, Stanford, California 94305-5404, USA. ⁴Neurosciences Program, Stanford University School of Medicine, Stanford, California 94305, USA.

*These authors contributed equally to this work.

reported, neural activity during the delay period (time from target appearance until 'go' cue) reflects the end-point of the upcoming reach¹³. This was also the case if the visual target was presented only briefly, thereby confirming that sustained neural activity is motor-related rather than purely sensory (see Supplementary Information). The reach end-point can be decoded from delay-period activity using maximum-likelihood techniques¹⁴.

Before conducting BCI experiments, we analysed the neural activity from instructed-delay control experiments to set parameters essential for high-performance BCI operation. We subdivided the delay period into two epochs: a time to skip (disregard) after target onset while waiting for reach end-point information to become reliable and therefore readily decodable (T_{skip}), and a time to integrate the neural data that will be used to predict the desired target selection (T_{int}).

The first epoch, T_{skip} , includes the time for visual information about the target to arrive in PMd (50–70 ms), the time for the subject to select among targets if more than one are present, and the time for neural activity reflecting the desired target to be generated. Despite being of considerable scientific interest^{15,16}, neural activity during these early periods is discarded in the present BCI design. Some activity during this period may already be predictive of the desired target, but it is not yet clear how best to decode this information. T_{skip} was chosen to be 150 ms based on control experiments including a multi-target task where the monkey was trained to reach for one of many simultaneously presented targets (see Supplementary Information).

The second epoch, T_{int} , directly follows the first and provides the

neural data used to predict the desired BCI cursor position. Given the Poisson-like noise in the spike timing of cortical neurons¹⁷ (confirmed with our data), a longer T_{int} will average away more noise and result in more accurate predictions of reach end-point. However, a longer T_{int} will also reduce the total number of cursor positionings that can be made per second. Herein lies the fundamental speed–accuracy trade-off that we must optimize in order to increase BCI performance.

To determine the best T_{int} to be used in BCI experiments, we analysed the effect of this parameter on two performance metrics. The first is single-trial accuracy, which is the percentage of targets correctly predicted. We found that accuracy rises and largely saturates around 85–90% as T_{int} increases to 200–250 ms. Figure 2a illustrates this effect as a function of total trial length, which is defined to be the sum of T_{skip} (150 ms), T_{int} (variable) and a small system overhead time associated with decoding and rendering the prosthetic cursor on the screen ($T_{\text{dec+rend}} \approx 40$ ms). Should a minimum level of single-trial accuracy be required for a particular application, a corresponding minimum T_{int} can be chosen.

The second performance metric is information transfer rate (ITRC, in bps). This quantity measures the rate at which information is conveyed from the subject, through the BCI, to the environment^{10,18}. It is the information per trial, which is closely related to single-trial accuracy, divided by the total trial length. As shown in Fig. 2a, the optimal ITRC occurs at short trial lengths, despite relatively low single-trial accuracy at these trial lengths. The highest ITRC is 7.7 bps at a total trial time of 260 ms, which corresponds to a T_{int} of 70 ms ($T_{\text{skip}} = 150$ ms, $T_{\text{dec+rend}} = 40$ ms). As further confirmation that neural responses are reflecting motor intention even at a rapid pace, we repeated the above analysis with the multi-target task. Despite this more-challenging task, requiring selection among many visual targets, the ITRC was diminished by a modest 30% (see Supplementary Fig. S4).

The performance curves in Fig. 2a are extrapolations using experimental data from individual trials that had long delay periods and long times between trials (Fig. 1a). To measure directly the ITRC performance when actually presenting trials at high speeds, we conducted a series of BCI experiments using a real-time system capable of rapidly decoding neural information. BCI experiments began with the collection of delay-period activity preceding reaches to different target locations (Fig. 1a) and fitting statistical models to the activity (model training). Then, during BCI prosthetic cursor trials (Fig. 1b), the intended target was decoded and a circular cursor was rendered on the screen at the predicted location. If the prediction was correct, the next target was displayed immediately. In this manner, a sequence of high-speed prosthetic cursor trials could be generated. Figure 1b illustrates three successful prosthetic cursor trials followed by a standard real reach trial.

Using this paradigm, we varied the number of locations at which a target could appear on any given trial. This allowed task difficulty to be varied, which contributes to the ITRC metric. Performance values were calculated by averaging data from several hundred trials per

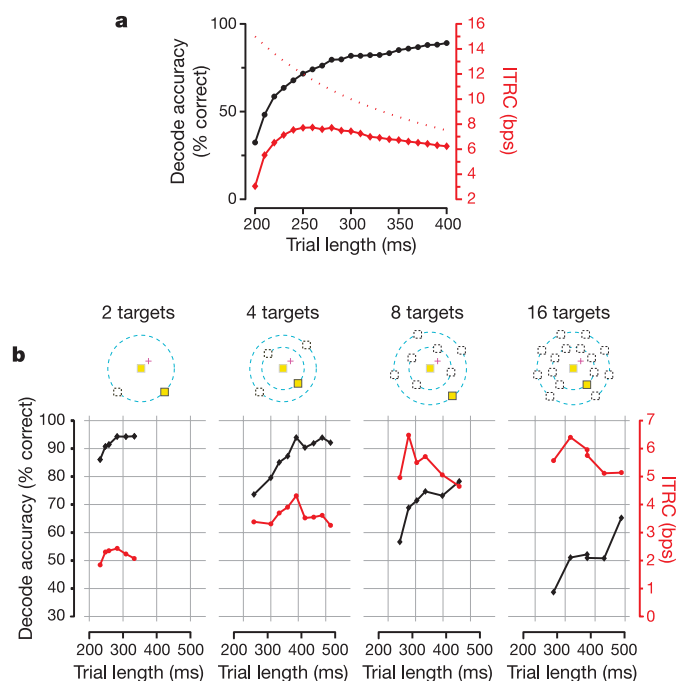


Figure 2 | Single-trial accuracy and ITRC with monkey H. **a**, Performance curves investigating the dependence on T_{int} were calculated from control experiment H20041118 (8-target configuration). The trial length was $T_{\text{skip}} + T_{\text{int}} + T_{\text{dec+rend}}$ with $T_{\text{skip}} = 150$ ms and $T_{\text{dec+rend}} \approx 40$ ms. T_{int} was varied and performance was computed. Performance metrics were very consistent day after day and between monkeys (data not shown). The theoretical maximum ITRC in bps, assuming 100% accuracy regardless of T_{int} , is plotted as the dotted red curve. **b**, Performance measured during BCI experiments. Performance is plotted for each target configuration and across varying total trial lengths. Each data symbol represents performance calculated from one experiment (many hundreds of trials). Across target configurations, single-trial accuracy decreases and ITRC increases as more targets locations are used.

Table 1 | BCI experiments with highest ITRC for monkeys H and G

Monkey	Number of targets	Accuracy (%)	Trials per second	bps
H	2	94.3	3.5	2.4
H	4	94.5	2.8	4.7
H	8	68.9	3.5	6.5
H	16	51.1	2.9	6.4
G	2	84.2	3.6	1.3
G	4	93.0	2.5	3.8
G	8	76.8	2.5	5.3
G	16	26.4	2.2	3.1

Each row lists the experiment with highest performance (ITRC) for a given target layout. Other experiments yielded higher single-trial accuracy or involved faster cursor rates, but did not achieve the highest ITRC for the corresponding target layout (not shown).

condition. Table 1 lists the highest ITRC results during BCI experiments with 2, 4, 8 or 16 targets. The best overall performance was achieved with the 8-target task (6.5 and 5.3 bps, monkeys H and G). This performance corresponds to typing ~ 15 words per minute with a basic alphanumeric keyboard.

Although a sustained performance rate of 6.5 bps is manifold greater than reported previously, it is lower than the extrapolated result (7.7 bps). Furthermore, the ITRC peak was expected at a total trial length of 260 ms, but our BCI experiments yielded 5 bps with this timing (monkey H). These discrepancies are due to the limitations inherent when using control experiments to extrapolate performance for speeds at which the subject must quickly recognize new targets and rapidly change neural activity (that is, change reach plans). The differences between extrapolated and directly measured performance were present despite specific model training methods that allowed for a fair comparison (see Supplementary Information).

Having confirmed that large BCI performance gains are possible with a direct end-point control strategy, we investigated two additional performance aspects. First, we varied T_{int} in BCI experiments with monkey H to verify experimentally the trends seen in Fig. 2a. Figure 2b also demonstrates an increase in single-trial accuracy with increasing trial length (black curves) as well as a peak in each ITRC curve (red curves). These results reveal how two or four target tasks restrict ITRC by virtue of the lower number of maximum bits per trial (1 and 2, respectively). Furthermore, given the numbers of neural units available in these experiments, it appears that ITRC is approaching a saturation point beyond which adding more target locations may not produce an appreciable increase in performance (doubling targets from 8 to 16 does not increase ITRC; although the latter layout requires distance tuning, which is known to be weaker than direction tuning¹³). Additional target locations should improve ITRC when more neurons are available.

Second, a common concern for BCIs such as ours is that as the electrode implant ages the number of recordable neurons declines, leading to a drop in overall performance¹⁹. To investigate the impact of neuronal loss, we performed analyses of single-trial accuracy and ITRC using data from control experiments. As expected, single-trial accuracy falls as neuron ensemble size decreases. However, it is possible to compensate partially for this performance loss by increasing T_{int} ; BCI speed may be compromised as a result, but single-trial accuracy can be preserved (Supplementary Fig. S5). Figure 3 plots ITRC as a function of the number of neural units and T_{int} . For small

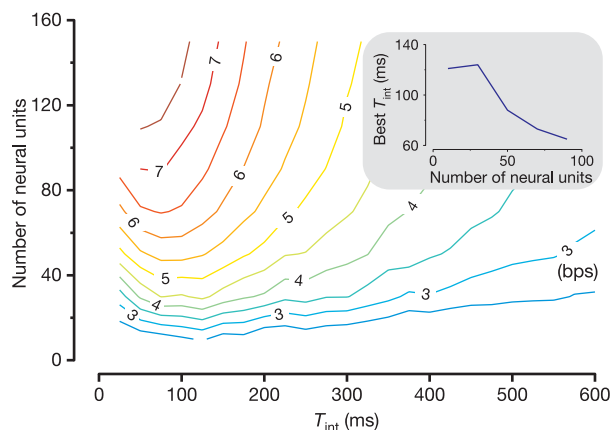


Figure 3 | ITRC as a function of number of neural units and T_{int} . All data are from experiment H20041118, which used an 8-target configuration and contained over 1,300 trials. T_{skip} was fixed at 150 ms. The main panel shows contours of ITRC (bps) as a function of the number of neural units available and T_{int} . The inset shows the value of T_{int} that achieves the maximum ITRC for each neural ensemble size that we tested. Similar results were obtained for data set G20040508 from monkey G.

ensembles (for example, 20 neurons), the ITRC peaks at $T_{\text{int}} \approx 120$ ms but does not decline sharply as T_{int} is further increased; accuracy (and bits per trial) is increasing so as to offset the longer trial times. For larger ensembles, the information content at small T_{int} is relatively high such that further lengthening T_{int} has a dramatic effect on ITRC. Finally, plotting the T_{int} value that maximizes ITRC (Fig. 3 inset) for each ensemble size illustrates that the maximum ITRC is achieved with small T_{int} (60–130 ms), over a broad range of ensemble sizes. Thus, high-performance BCIs may require far shorter trials than previously explored.

Using a direct end-point control strategy, we report here a greater than fourfold (6.5 versus 1.6 bps) increase in BCI performance compared to recent studies. Performance is calculated in a conservative fashion, as the entire trial time ($T_{\text{skip}} + T_{\text{int}} + T_{\text{dec+rend}}$) was used; had just T_{int} been used, as is often done, the maximum ITRC would have been 28.4 bps, but this does not reflect an achievable selection rate. As described previously, our system differs from continuous BCI approaches in several ways, which might account for the performance gain. Additionally, continuous BCIs attempt to move the cursor well enough—although at the expense of speed (1–3 s per selection)—to avoid making errors for a given target selection. Conversely, the direct end-point control reported here need not correct errors within a given selection because these errors can be rectified with rapid follow-on selections.

Our performance results far exceed electroencephalogram (EEG)-based non-invasive system performance, and helps motivate the use of invasive, electrode-based systems in clinical BCIs. Although at its fastest, this direct end-point control BCI demonstrates selection speeds (~ 3.5 trials per second) on a par with saccadic eye movements, the ITRC for saccades is much higher due to their exceptional precision and accuracy. Whereas eye or even speech control may be effective in specific settings, BCIs attempting to restore lost motor function must rely on the natural neural signals if they are to avoid commandeering and interfering with another motor modality.

METHODS

See Supplementary Information for additional details on the methods of this study.

We trained two rhesus monkeys (*Macaca mulatta*) (G and H) to perform instructed-delay centre-out reaches. Animal protocols were approved by the Stanford University Institutional Animal Care and Use Committee. Hand and eye position were tracked optically (Polaris, Northern Digital; Iscan). Stimuli were back-projected onto a frontoparallel screen 30 cm from the monkey. Real reach trials (Fig. 1a) began when the monkey touched a central yellow square and fixated on a magenta cross. After a touch hold time (200–400 ms), a visual reach target appeared on the screen. After a randomized (200–1,000 ms) delay period, a 'go' cue (fixation and central touch cues were extinguished and reach target was slightly enlarged) indicated that a reach should be made to the target. Fixation was enforced throughout the delay period to control for eye-position-modulated activity in PMd^{20,21}. This fixation requirement is applicable in a clinical setting if targets are near-foveal, or imagined as in a virtual keyboard set-up. The hand was also not allowed to move until the go cue was presented, providing a proxy for the cortical function of a paralysed subject^{1–3}. Subsequent to a brief reaction time, the reach was executed, the target was held (~ 200 ms), and a juice reward was delivered along with an auditory tone. An inter-trial interval (~ 250 ms) was inserted before starting the next trial. We presented various target configurations (2, 4, 8 or 16 targets) on the screen, including layouts with 2, 4, or 8 directions, and 1 or 2 distances (6–12 cm radially outward).

Neural recordings. Neural activity was simultaneously recorded from a 96-channel electrode array (Cyberkinetics Neurotechnology Systems) implanted in arm representation of PMd, contralateral to the reaching arm (left, monkey G; right, monkey H). We previously reported our set-up for automatic spike sorting²² for monkey G. A similar, but more sophisticated, system was used for monkey H. The use of an automatic spike sorting system ensured a very fast and repeatable method for classifying neural units each day. We recorded 20–30 single neurons and 60–100 multi-neuron units in a typical session.

Models and decoding. For the instructed-delay and BCI tasks, we first collected data across several hundred trials and fit (trained) models based on neural activity collected starting T_{skip} after the target presentation time and extending for a duration T_{int} . For each target location, the distribution of spike counts

across trials was modelled as either a Poisson or multivariate Gaussian. We then used these models to predict the reach target on separate trials, given only the neural data from each trial.

For the instructed-delay control experiments (Figs 2a and 3; see also Supplementary Figs S2–S5), we either used two separate blocks of trials, one for training and one for prediction, or used leave-one-out cross-validation with all the trials in a data set. For BCI experiments, training trials were initially collected to fit the models and during subsequent trials, the target was predicted with the model (Gaussian model for monkey G and Poisson model for monkey H). This prediction was rendered on the screen as a circle around the predicted target location and was visible to the monkey. If the prediction was correct, an auditory tone was delivered and a subsequent target was presented with very little delay. If the prediction was incorrect, the trial was either considered a failure and aborted, or the monkey was allowed to make a real reach to the target. Real reach trials were interspersed to ensure that the monkey remained engaged in the task.

Experiments with monkey G showed consistently lower BCI performance than those with monkey H. This can be attributed to several factors, including statistical models (Gaussian versus Poisson), model training methods, and neural signal quality. When these parameters were equalized in control experiments, performance was comparable for both monkeys.

ITRC calculation. The ITRC of the system was computed using the Blahut–Arimoto algorithm²³. This iterative algorithm produces the theoretical upper-bound on information transmission per use of a communication channel, given only the noise characteristics of the channel. In our experiments, the channel is the entire prosthetic system, with the presented target corresponding to the input and the decoded target corresponding to the output. We divide the information per trial by the average time per trial to obtain the ITRC.

Neuron dropping analysis. For a single day's experiment, we selected all neural units from the array that were responsive to target location within a desired T_{int} using an analysis of variance (ANOVA; $P < 0.05$). For each neural ensemble size of interest, our total set of neural units was subdivided by drawing 100 randomized subsets (without replacement). Performance was computed for each subset and these data were averaged within a given ensemble size. This provided a single ITRC for each T_{int} and ensemble size. We generated the contour plot (Fig. 3) by using linear interpolation across this two-dimensional surface. Furthermore, for each ensemble size tested, a cubic spline interpolation was used to estimate the particular T_{int} that maximized the ITRC (Fig. 3 inset).

Received 1 March; accepted 16 May 2006.

- Serruya, M. D., Hatsopoulos, N. G., Paninski, L., Fellows, M. R. & Donoghue, J. Instant neural control of a movement signal. *Nature* **416**, 141–142 (2002).
- Taylor, D. M., Tillery, S. I. H. & Schwartz, A. B. Direct cortical control of 3D neuroprosthetic devices. *Science* **296**, 1829–1832 (2002).
- Carmena, J. M. et al. Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biol.* **1**, E42 (2003).
- Musallam, S., Corneil, B. D., Greger, B., Scherberger, H. & Andersen, R. A. Cognitive control signals for neural prosthetics. *Science* **305**, 258–262 (2004).
- Kennedy, P. R., Bakay, R. A. E., Moore, M. M., Adams, K. & Goldwithe, J. Direct control of a computer from the human central nervous system. *IEEE Trans. Rehabil. Eng.* **8**, 198–202 (2000).
- Leuthardt, E. C., Schalk, G., Wolpaw, J. R., Ojemann, J. G. & Moran, D. W. A brain-computer interface using electrocorticographic signals in humans. *J. Neural Eng.* **1**, 63–71 (2004).
- Hochberg, L. R., Mukand, J., Polykoff, G., Friebs, G. & Donoghue, J. *BrainGate Neuromotor Prosthesis: Nature and Use of Neural Control Signals*. Program No. 520.17, 2005 Abstract Viewer/Itinerary Planner (Society for Neuroscience, Washington DC, 2005).
- Patil, P. G., Carmena, J. M., Nicolelis, M. A. L. & Turner, D. A. Ensemble recordings of human subcortical neurons as a source of motor control signals for a brain-machine interface. *Neurosurgery* **55**, 27–38 (2004).
- Wolpaw, J. & McFarland, D. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proc. Natl Acad. Sci. USA* **101**, 17849–17854 (2004).
- Taylor, D. M., Tillery, S. I. H. & Schwartz, A. B. Information conveyed through brain-control: cursor vs. robot. *IEEE Trans. Neural Syst. Rehabil. Eng.* **11**, 195–199 (2003).
- Shenoy, K. V. et al. Neural prosthetic control signals from plan activity. *Neuroreport* **14**, 591–596 (2003).
- Cisek, P. & Kalaska, J. F. Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature* **431**, 993–996 (2004).
- Messier, J. & Kalaska, J. F. Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. *J. Neurophysiol.* **84**, 152–165 (2000).
- Yu, B. M., Ryu, S. I., Santhanam, G., Churchland, M. M. & Shenoy, K. V. Improving Neural Prosthetic System Performance by Combining Plan and Peri-movement Activity 4516–4519 (Proc. 26th Annu. Int. Conf. IEEE EMBS, San Francisco, California, 2004).
- Churchland, M. M., Yu, B. M., Ryu, S. I., Santhanam, G. & Shenoy, K. V. Neural variability in premotor cortex provides a signature of motor preparation. *J. Neurosci.* **26**, 3697–3712 (2006).
- Yu, B. M. et al. Extracting dynamical structure embedded in neural activity. In *Advances in Neural Information Processing Systems 18* (eds Weiss, Y., Schölkopf, B. & Platt, J.) (MIT Press, Cambridge, Massachusetts, 2006).
- Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, Massachusetts, 2001).
- Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 and 623–656 (1948).
- Polikov, V. S., Tresco, P. A. & Reichert, W. M. Response of brain tissue to chronically implanted neural electrodes. *J. Neurosci. Methods* **148**, 1–18 (2005).
- Cisek, P. & Kalaska, J. F. Modest gaze-related discharge modulation in monkey dorsal premotor cortex during a reaching task performed with free fixation. *J. Neurophysiol.* **88**, 1064–1072 (2002).
- Batista, A. et al. *Heterogeneous Coordinate Frames for Reaching in Macaque PMd*. Program No. 363.12, 2005 Abstract Viewer/Itinerary Planner (Society for Neuroscience, Washington DC, 2005).
- Santhanam, G., Sahani, M., Ryu, S. I. & Shenoy, K. V. An Extensible Infrastructure for Fully Automated Spike Sorting During Online Experiments 4380–4384 (Proc. 26th Annu. Int. Conf. IEEE EMBS, San Francisco, California, 2004).
- Cover, T. & Thomas, J. *Elements of Information Theory* (John Wiley, New York, 1990).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Eisensee for administrative assistance, M. Howard for surgical assistance and veterinary care, and N. Hatsopoulos for surgical assistance (monkey G implant). We also thank M. Churchland and M. Sahani for scientific discussions; E. Knudsen and T. Moore for comments on the manuscript; and A. Batista for collecting the blink-memory data. This study was supported by NDSEG Fellowships (G.S. and B.M.Y.), NSF Graduate Research Fellowships (G.S. and B.M.Y.), the Christopher Reeve Foundation (S.I.R. and K.V.S.), Bio-X Fellowship (A.A.), the NIH Medical Scientist Training Program (A.A.), and the following awards to K.V.S.: a Burroughs Wellcome Fund Career Award in the Biomedical Sciences, the Stanford Center for Integrated Systems, the NSF Center for Neuromorphic Systems Engineering at Caltech, ONR Adaptive Neural Systems, the Sloan Foundation, and the Whitaker Foundation.

Author Contributions G.S. and S.I.R. contributed equally to this work. S.I.R. was responsible for the initial experimental concept and surgical implantation; G.S. and S.I.R. were responsible for experimental design, animal training, data collection and preliminary analysis; G.S. was responsible for infrastructure development, in-depth analysis and writing of the paper. B.M.Y. and A.A. participated in animal training, data collection and analysis. K.V.S. was involved in all aspects of the study.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.V.S. (shenoy@stanford.edu).