Long-term stability of neural prosthetic control signals from silicon cortical arrays in rhesus macaque motor cortex

2011 J. Neural Eng. 8 045005

(http://iopscience.iop.org/1741-2552/8/4/045005)

J. Neural Eng. **8** (2011) 045005 (11pp) doi:10.1088/1741-2560/8/4/045005

Long-term stability of neural prosthetic control signals from silicon cortical arrays in rhesus macaque motor cortex

Cynthia A Chestek[1], Vikash Gilja[2], Paul Nuyujukian[3,4], Justin D Foster[1], Joline M Fan[3], Matthew T Kaufman[5], Mark M Churchland[1], Zuley Rivera-Alvidrez[1], John P Cunningham[1], Stephen I Ryu[6] and Krishna V Shenoy[1,3,5,7]

[1] Department of Electrical Engineering, Stanford University, Stanford, CA, USA
[2] Department of Computer Science, Stanford University, Stanford, CA, USA
[3] Department of Bioengineering, Stanford University, Stanford, CA, USA
[4] Stanford Medical School, Stanford University, Stanford, CA, USA
[5] Neurosciences Program, Stanford University, Stanford, CA, USA
[6] Department of Neurosurgery, Palo Alto Medical Foundation, Palo Alto, CA, USA

E-mail: shenoy@stanford.edu

Received 29 September 2010
Accepted for publication 13 January 2011
Published 20 July 2011
Online at stacks.iop.org/JNE/8/045005

Abstract
Cortically-controlled prosthetic systems aim to help disabled patients by translating neural signals from the brain into control signals for guiding prosthetic devices. Recent reports have demonstrated reasonably high levels of performance and control of computer cursors and prosthetic limbs, but to achieve true clinical viability, the long-term operation of these systems must be better understood. In particular, the quality and stability of the electrically-recorded neural signals require further characterization. Here, we quantify action potential changes and offline neural decoder performance over 382 days of recording from four intracortical arrays in three animals. Action potential amplitude decreased by 2.4% per month on average over the course of 9.4, 10.4, and 31.7 months in three animals. During most time periods, decoder performance was not well correlated with action potential amplitude ($p > 0.05$ for three of four arrays). In two arrays from one animal, action potential amplitude declined by an average of 37% over the first 2 months after implant. However, when using simple threshold-crossing events rather than well-isolated action potentials, no corresponding performance loss was observed during this time using an offline decoder. One of these arrays was effectively used for online prosthetic experiments over the following year. Substantial short-term variations in waveforms were quantified using a wireless system for contiguous recording in one animal, and compared within and between days for all three animals. Overall, this study suggests that action potential amplitude declines more slowly than previously supposed, and performance can be maintained over the course of multiple years when decoding from threshold-crossing events rather than isolated action potentials. This suggests that neural prosthetic systems may provide high performance over multiple years in human clinical trials.

1. Introduction

Neural prostheses, also termed brain–machine interfaces (BMIs) or brain–computer interfaces (BCIs), are an emerging

[7] Author to whom any correspondence should be addressed.

class of medical technology with the potential to improve the quality of life for severely disabled patients [1]. In particular, intracortical signals from multi-electrode arrays that penetrate 1–2 mm into cortex can provide useful control signals. Performance from these systems can be high, which motivates determining how long it can be maintained. For example, discrete selections of visual targets on a screen can be made at a rate of 6.5 bps [2], which approximately corresponds to typing on a keyboard at approximately 15 words per minute. For continuous control, cortical signals have been used to guide a computer cursor to targets on a screen [3–14]. Recent work has demonstrated increased speed and reduced variability [14–16]. Further motivating clinical translation, cortical signals have also been used to control the arm and hand on an anthropomorphic robotic arm for self-feeding in primates [17]. Offline, information from other degrees of freedom, such as fingers and wrist, has been extracted as well [18]. The clinical need for such control signals is high, since there is a substantial population of paralyzed patients, and state-of-the-art prosthetic limbs currently have more degrees of freedom than one can easily control with conventional approaches [19].

However, to conduct a risk-benefit assessment for human patients, it is important to quantify the stability of cortical BMI performance. The major focus of previous work on this question has been to characterize the tissue response to electrode implantation. On the shortest timescale, after implantation, effects such as edema and hemorrhage are observed [20]. Within weeks, the brain largely recovers from the acute effects [21, 22], but it is possible that the extent of this trauma determines the number of viable neurons surrounding the electrode for long-term recording [23]. This early period is associated with minimal recorded neural activity [24]. On the timescale of 1–4 months, there is also an encapsulation response, in which astrocytes and microglia form a tight insulative sheath around the electrode [22, 25], which can remain in place for years [26]. This has been suggested as the primary mechanism for the degradation of single unit activity over time [27]. One study reported large declines in single unit activity using microwire probes on similar timescales [24]. This effect may reach steady state as early as 12 weeks post implant [22, 25, 26, 28]. However, another study using microwires demonstrated recording capability for up to 7 years [29]. Also, many BMI studies have been completed using the 'Utah' array for over 6 months. The Utah array, which is a monolithic array of 100 silicon microelectrodes, is particularly important to understand because it has been FDA approved for use in human patients [8, 9, 30]. This long-term use raises the question of how histological and electrophysiological measurements are related. More generally, it also raises the question of what fraction of overall signal quality degradation is caused by biological response as opposed to materials and engineering failures.

The discrepancy regarding timescale might be related to the perceived importance of single unit activity (clearly differentiable action potentials arising from individual neurons). Many histological and electrophysiological studies have assumed that single unit activity will be well correlated with BMI performance. Until recently, this was a reasonable assumption since early BMI results emphasized the use of single units [2–4]. In these systems, the activity of individual neurons is differentiated by separating action potentials with distinctive waveforms, a mathematical process termed 'spike sorting'. However, with present algorithms, spike sorting leads to very little improvement in BMI performance [12, 31–33]. In fact, high performance studies recently reported have used only 'threshold-crossing' events, in which the firing rate of one unit per electrode is determined by counting the number of times that the voltage falls below some multiple of the RMS voltage, regardless of whether this activity results from more than one neuron [12, 14–16]. Since these studies used at least one array implanted for more than 1 year, it is important to revisit whether decline in the amplitude of action potentials (Vpp) actually corresponds to performance loss.

Several previous studies have examined these questions. One previous study examined day-to-day stability in recordings from the Utah array in a feline model and saw substantial interday changes in neural waveforms which correlated with measures of inflammation [34]. However, they did not quantify the effects on decoder performance with this animal model. A number of papers have attempted to track single units across a few days [30, 35–37]. Most relevant to the current study, one study examined the ratio of Vpp to the standard deviation in voltage in three macaques with Utah arrays, and found no significant decline [38]. However, if the decline is more subtle than previously supposed, the particular metric used to measure it may not detect the effect over the time periods studied, and it would not be noted subjectively during experiments if it was not also reflected in decoder performance. This study also did not attempt to systematically examine the information content on these arrays over many months, as measured by decoder performance over time.

In this report, we analyzed neural data from 382 days in three monkeys and four intracortical arrays over various periods during 3.6, 9.4, 10.4 and 31.7 months of implantation. These data were characterized in terms of the peak-to-peak voltage of action potentials as well as offline decoder performance. Voltage decline over long timescales was slow but significant, with several noticeable 'events' in which there were rapid changes. Over time, decoder performance was rarely correlated with Vpp, when simple threshold crossings were used. On shorter timescales, large but reversible changes in Vpp were quantified. This study empirically demonstrates not only that average Vpp decays more slowly than previously supposed, but that decoder performance may decline even more slowly. In particular, it declines in a less direct fashion when using threshold crossings (i.e. multi-unit activity) instead of relying on well-isolated action potentials originating from single neurons.

2. Methods

2.1. Behavioral tasks

All protocols were approved by the Stanford University Institutional Animal Care and Use Committee. We trained three rhesus macaques to perform 2D reaches to visual

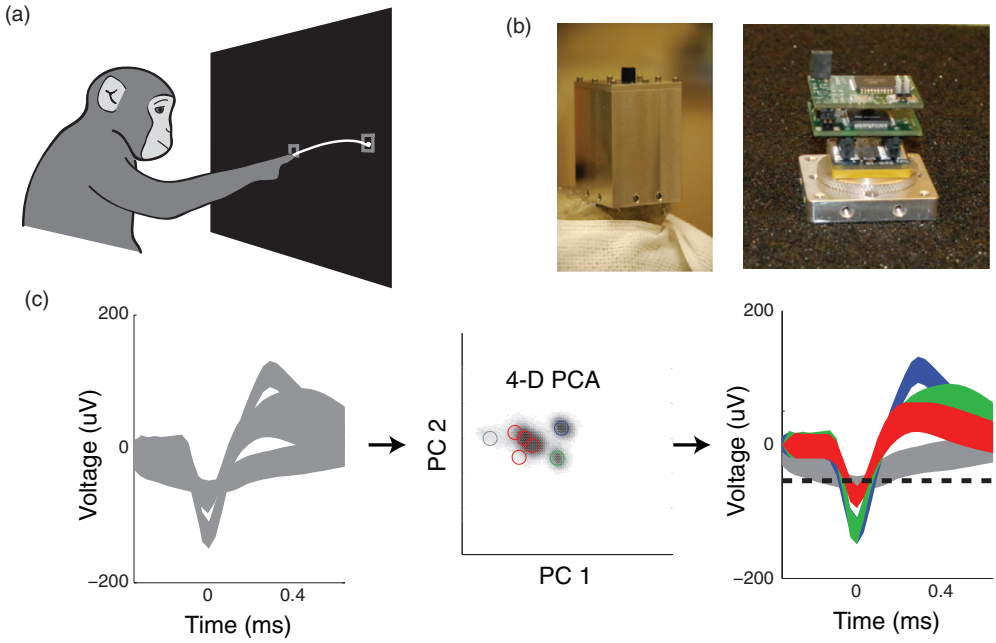

Figure 1. (a) Illustration of the experimental setup. Animals made 2D reaches to radial targets. (b) Wireless recording device used for monkey L [39]. (c) Neural data were processed through various filters, and through a principal component based automatic spike sorter [32]. Data were first filtered, peak-aligned and noise whitened in the leftmost panel. Then, waveforms were projected into 4D principle component space, where units were clustered using an expectation–maximization based mixture of Gaussian models, producing the classification shown on the right.

targets, as shown in figure 1(a). Two monkeys, I and L, were prospectively selected for this study after having completed previous studies [39–41]. A third animal, J, was retrospectively selected for this study on account of having an observed Vpp decline during the course of a previous study [42] as well as substantial longitudinal data from an ongoing study [14]. For these three animals, we were interested in quantitatively comparing prosthetic performance over time. Thus, we analyzed data from stereotypical tasks; during each time such a task was run on a subset of days for more than 30 days.

These data consist of 382 daily datasets from three animals and four arrays across 3.6, 9.4, 10.4 and 31.7 months of recording, as shown in figure 2. To our knowledge, this represents the largest cortical dataset analyzed for array stability. Waveform voltage data could be obtained from all 382 datasets. A subset of this data, 184 datasets, included long periods of identical tasks. These periods are shown in the dashed boxes in figure 2. During these periods, decoder performance could be quantitatively compared over time. Two animals had multiple periods where they performed a consistent task over more than 30 days. Point-to-point reaches were performed in complete darkness except for the illuminated target. Hand position was optically tracked with a reflective bead attached to the distal joint of the index or middle finger and measured at 60 samples s^{-1} (nominal sub-millimeter resolution) using a Polaris system (NDI, Waterloo, Ontario, Canada). For the non-BMI datasets, the task was sequenced using Tempo software (Reflective Computing, St Louis, MO). For the pre-BMI training trials, which consisted of real center-out reaches, the task was sequenced using custom Matlab and Simulink code executed on a real-time xPC target. All monkeys were highly trained for many months prior to implantation, such that there was likely little to no learning occurring during the present study. For all animals the task was a simple center-out reach task [43]. However, animals varied in the number of targets (4,7,8) and radius of the targets (80–120 cm), in addition to other small changes. However, performance measures for a specific animal were internally consistent, allowing quantitative comparisons over time, but not between animals. The number of trials used was identical for each animal, and included at least 25 trials per condition. All animals received a liquid reward for successful trials.

2.2. *Electrophysiology recordings*

One or two commercially available silicon 'Utah' arrays (Blackrock Technologies Inc., Salt Lake City, UT, USA) containing 100 electrodes each were implanted into PMd/M1 of these animals using standard neurosurgical techniques [44]. Neural data were recorded using a Cerebus system (Blackrock Microsystems, Salt Lake City, UT) while the monkey was participating in a neuroscience experiment for 1–5 h during the day. For monkey L only, wireless data were recorded using HermesC [39, 45], pictured in figure 1(b). These data were recorded on 39 particular days while the animal was in the home cage from 20 channels at a time. Wireless data were obtained 91% of the time except during the in-rig experiments, animal transfers, and rare equipment failures. While this system has the capability of recording threshold-crossing data from all 20 channels simultaneously, only the single broadband channel was used in this study. The device was programmed

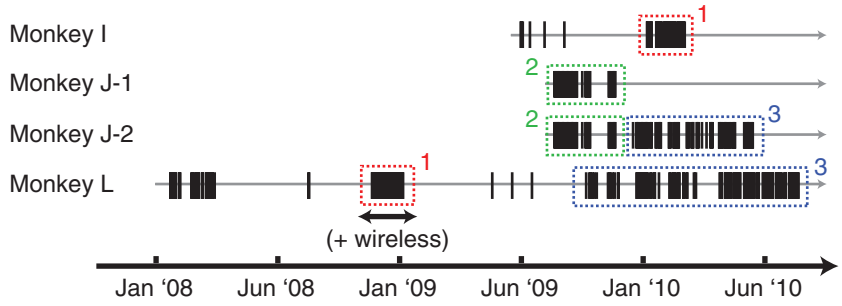

Figure 2. Chart showing when datasets were collected. The tick marks represent 1 day of recording. The gray line represents the time each animal was implanted. The dashed boxes indicate times where a consistent center-out reach task was performed over more than 30 days such that offline BMI performance could be analyzed. The color and number denote different but self-consistent tasks. Period with wireless data is denoted by black arrow.

to switch between each of 20 channels every 23 s and record data at 15.7 ksps such that waveform shape could be recorded 24 h a day at a duty cycle of approximately 4%. Therefore, the wireless data were obtained at half of the sampling rate as the wired recordings and were only recorded for 23 s every 8 min.

Vpp from spike sorted neurons was required for many analyses presented in this paper. Due to the large size of the dataset, an automated spike sorting algorithm was required. Neural units were isolated offline using noise whitened principal components and a mixture of Gaussian models for clustering [32, 46] as illustrated in figure 1(c). This spike sorting algorithm was augmented with a bipolar shape heuristic to eliminate any non-neural artifacts that might appear as sorted units. A large proportion of the data were manually verified to ensure that no artifacts were included. When well-isolated single units were required for an analysis, unit quality was rated manually for a small number of datasets. Units were classified as putative single units, single units with some contamination by other units, and multi-units. The number of putative single units in early datasets was 34, 27, and 28/38 for animals I, L, and J/J, respectively (monkey J had two arrays).

For measuring offline decoder performance, as described below, data were not spike sorted. Instead, a threshold was set at $-4.5\times$ the RMS voltage on a particular electrode, and this multi-unit's activity was decoded similarly to a single unit. Several prior studies have reported little or no gain in performance using sorted spikes rather than threshold crossings [12, 32], as discussed above. This was verified using 39 datasets that were manually spike sorted, resulting in only a 4% performance improvement. Using threshold crossings also simplified standardization across animals and time periods, as data for monkeys L and J were collected as part of ongoing online BMI experiments, in which $-4.5\times$ RMS threshold crossings were used.

2.3. Neural data analysis

For estimating prosthetic decoder performance offline, two different decoders were used: a discrete and a continuous decoder. For the discrete decoder, a naive Bayes classifier was used, similar to one we have used previously [2]. The average firing rate of the threshold crossings on a particular electrode on a particular day was modeled as a Poisson distribution with a characteristic mean firing rate λ for each angle, measured with 500 ms of neural data after the target appears. For each block of single trials in each direction, all other reaches were used to train a model which was then tested on those trials. Decoders were trained and tested on data within the same day. To decode the target angle, the likelihood of each angle was calculated as

$$P(\Theta = \theta | y_t) = \sum_{n=1}^{N} \frac{\lambda_{n,\theta}^{y_t} \, e^{-\lambda_{n,\theta}}}{y_t!} \tag{1}$$

where θ is the target angle, N is the number of neurons, and y_t is the number of spikes that occurred during the integration window. The angle with the highest probability was then selected. Using this type of decoder, performance can be evaluated with percent correct.

A continuous decoder was also constructed offline using a linear filter [4, 7, 8, 35]. The firing rate and position were averaged over 100 ms bins during each trial. The linear model used the firing rates of the units at ten sequential 100 ms time lags. Position was modeled as a function of firing rate using a linear Wiener filter. In equations (2) and (3), the firing rate matrix X has ten columns for each unit, and horizonal and vertical position is found in matrix Y:

$$Y = XB \tag{2}$$

$$\begin{bmatrix} \mathrm{pos}(1) \\ \dots \\ \mathrm{pos}(t) \end{bmatrix} = \begin{bmatrix} \mathrm{neur}_1\mathrm{lag}_1(1) & \mathrm{neur}_1\mathrm{lag}_2(1) & \dots & \mathrm{neur}_N\mathrm{lag}_{10}(1) \\ \dots & & & \\ \mathrm{neur}_1\mathrm{lag}_1(t) & \dots & \dots & \mathrm{neur}_N\mathrm{lag}_{10}(t) \end{bmatrix} \times \begin{bmatrix} B_{\mathrm{neur}_1\mathrm{lag}_1} \\ \dots \\ B_{\mathrm{neur}_N\mathrm{lag}_{10}} \end{bmatrix}. \tag{3}$$

Both X and Y have rows corresponding to the total number of 100 ms bins in the experiment. The resulting linear decoder, matrix B, is computed through linear regression. The model was tested using cross validation with 20 folds. That is, for each 5% of the data, neural decoders were generated using the other 95% of the data, and tested on that 5%. The final

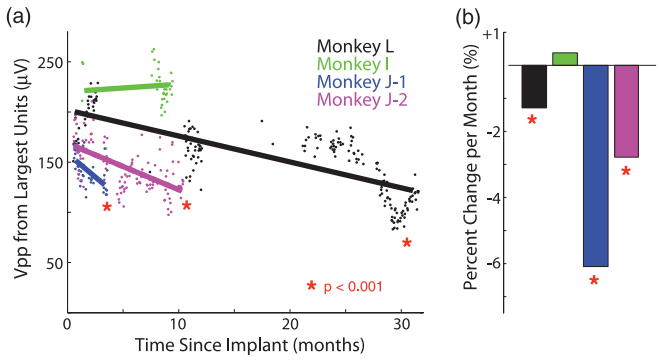

Figure 3. (a) Voltage from largest unit on each channel averaged across channels for four arrays in three animals, where color denotes a specific array. (b) Average change per month in each array.

performance value was averaged across all such 5% pieces. Continuous performance was quantified using a correlation coefficient with actual hand position, as well as the average distance to the target during the decoded reach.

3. Results

3.1. Long timescales

The simplest way to characterize electrode stability across many months is to examine the peak-to-peak amplitudes (Vpp) of the largest action potentials on each channel. Figure 3 shows this Vpp averaged across electrodes (with high noise electrodes removed) for four arrays in three animals. Three out of four arrays showed a statistically significant negative trend during this time ($p < 0.001$). The size of the effect was +0.3% (I, not significant) , −6.1% (J-1), −2.8% (J-2), and −0.6% (L) per month. The arrays with fewer months of recording show larger effects, which is consistent with a trend that slows down over time.

These data are characterized by long periods with a small negative change and shorter periods with large changes, which may be relevant to determining the mechanism of these changes. For example, in monkey J, both arrays showed decline during the first 2 months after implantation. This can be seen more easily by viewing only the channels with observed isolated action potentials on day 1, shown in figure 4(a), rather than all channels as in figure 3. The overall Vpp declined by 47.0% and 28.2% ($p < 0.001$). Subjectively, during experiments well-isolated action potentials were observed to decline substantially in number. However, in one of these arrays that continued to be used for online BMI experiments over the following year, it is clear that this trend did not continue at the same rate. For the 8.5 months following, voltage declined by only 2.0% per month. Another period of decline was observed at the end of 27 months (2.2 years) of recording from monkey L. During the initial 27 months, this array's Vpp declined by 17.2% (0.6% per month). However, during the following 2.5 months, as shown in figure 4(b), it declined by 35% (and then recovered by 15%). This event was concurrent with an infection at the implant site, but there is no obvious mechanism to explain the changes in voltage. In all animals, Vpp declined slower than what would be expected if the glial scar response sealed off the electrode in the first 12 weeks.

To determine how the voltage changes might affect decoder performance, offline performance was evaluated in two ways. First, a discrete prediction of which target the animal was reaching to provided a percent correct value. Second, an offline implementation of a continuous linear decoder mapped neural firing rates to hand position. The quality of the continuous decode was measured by correlation coefficient between predicted and actual hand position. However, correlation coefficient can remain stable in the face of large absolute changes in decoder performance. Therefore, a second metric, mean distance to target, was also calculated during the decoded reach as in [47, 48].

For three out of four animals, decoder performance was not significantly correlated with Vpp. Figure 5 shows all three performance metrics below the voltage amplitude traces in monkeys I and J. Performance variation was not significantly correlated with voltage variation in these arrays ($p > 0.05$ except one $p = 0.02$, which is not significant when corrected for multiple comparisons). In the remaining data from the second array in monkey J, it continued to be uncorrelated with voltage over months 4–11, during which

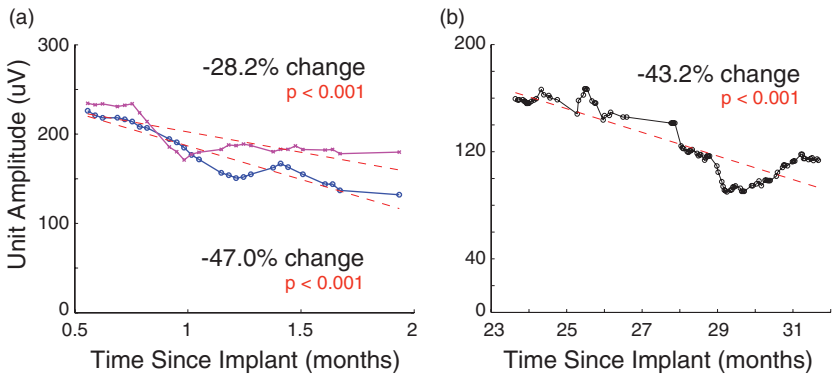

Figure 4. (a) Voltage from largest unit averaged across the 28 and 38 channels on two arrays from monkey J that had observed single unit activity in initial dataset during first 2 months of recording; (b) similar data from monkey L after 2 years of recording. All data are smoothed across a 10 day averaging window to reveal trends.

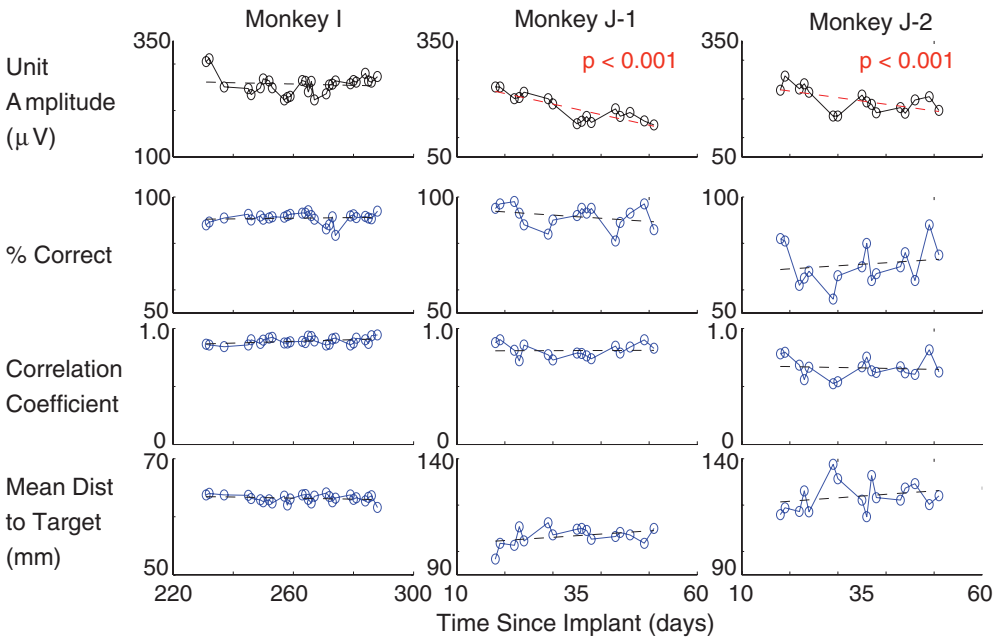

Figure 5. Top row shows average voltage amplitude across electrodes with observed single unit activity on day 1 on three arrays in two animals during periods where the same task was performed for more than 20 days. The number of single unit channels was 34, 28, and 38 electrodes for I, J-A, and J-B, respectively. Lower rows show offline decoder performance. For discrete target decoders, performance is measured by percent correct. For a continuous linear decode, performance is measured with correlation coefficient to actual hand position and mean distance to target over the course of the trial. The red lines denote significant trends. No smoothing filters were used.

a different task with more targets was run and decoded offline (23 datasets). Performance by itself during these periods showed no significant negative trends, despite the negative trends in voltage during the first 2 months for both arrays in monkey J. These data show that action potential amplitude cannot be used as a proxy for measuring BMI performance.

Similar data for monkey L are shown in figure 6. This array is older than the others, having been implanted for over 2.5 years. There were two periods with identical tasks across periods longer than 30 days, which are graphed on either side of the ellipses near day 350 in figure 6, such that absolute performance cannot be compared across this boundary, but trends can be examined within each period. During the first period, variation in performance was correlated with variation in voltage for all three measures ($\rho = 0.45$, $p < 0.01$). However, there were no significant trends in performance during this time. During the second period, performance using all three metrics was highly correlated with voltage ($\rho = 0.67$, $p < 0.001$) and decreased along with the voltage ($p < 0.001$). During the final 5 days of recording, performance had nearly recovered, with an average discrete percent correct of 93% predicting which of eight targets the animal was reaching toward and an average correlation coefficient with hand position of 0.85. This is despite the fact that Vpp had only recovered from approximately half of the loss.

3.2. Short timescales

In addition to long-term trends, short-term variation in neural waveforms was observed over the course of hours and between days. Figure 7 shows data from the wireless in-cage recordings. The example electrode in figure 7(A) had single unit action potentials that were sometimes visible and sometimes undifferentiable from the multi-unit hash. These short-term changes occur despite the fairly stable average unit voltage shown in figures 3–6. Voltage data for 19 channels of wireless recording are shown in figure 7(B). For each electrode, the 95% envelope of voltage waveform snippets was found and normalized to the average voltage on that electrode across all days so that percentage deviations could be examined. Peak-to-peak voltages ranged from 66% to 236% of these mean values. The speed of changes is apparent from the abrupt changes between adjacent 1 h bins.

During experiments in which the animal sat in a primate chair with a fixed head position, smaller changes in neural waveforms were also observed. Figure 8 shows normalized voltage data from hand verified single units in one dataset per animal. Over the course of the experiment, the Vpps progressively deviate from their initial value, as shown by the standard deviation traces. The final standard deviation from the initial values was 9.0%, 6.4%, and 14.8% over 1, 2, and 5 h, respectively.

For all experimental days and animals, the data were spike sorted in an automated fashion, and the Vpp of the largest sorted unit on each channel was calculated. Changes in these waveform amplitudes were measured within single experimental sessions, as well as across many days. A histogram of changes that occurred (per hour) over 1–5 h while the animals were seated in the rig is shown in figure 9(a). With 21 310 unit-day pairs, this distribution shows an average absolute change of 3.2% per hour, though 6% of the waveforms changed by over 10%. Across days, changes were more

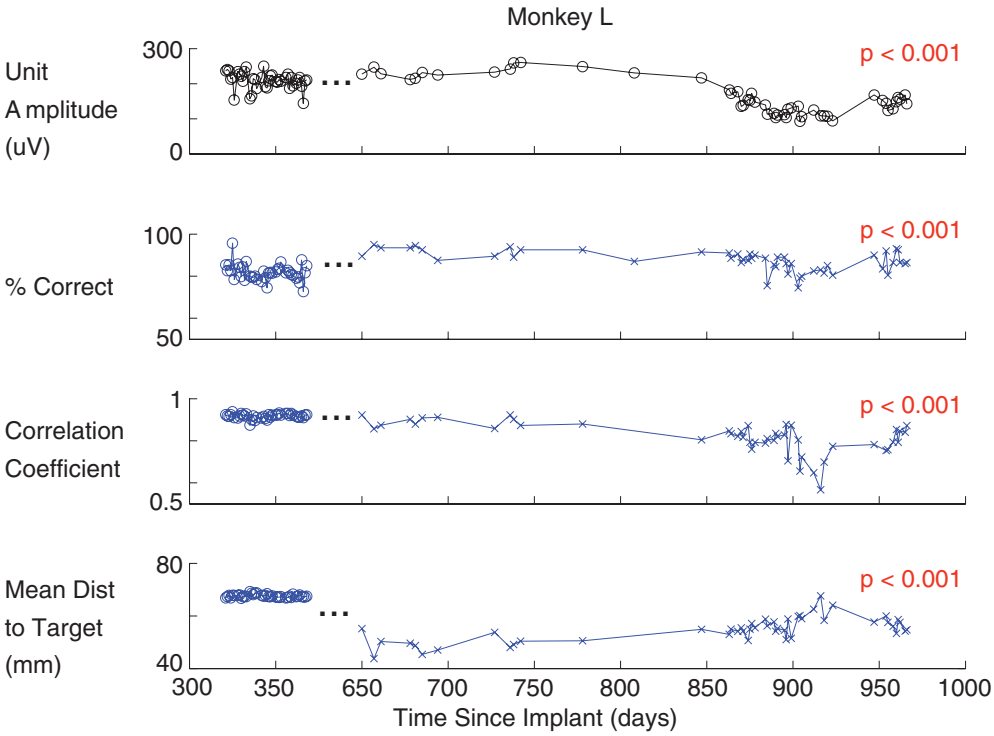

Figure 6. Top row shows voltage amplitude from largest unit on each electrode averaged across electrodes for monkey L. Ellipses denote discontinuity on x axis near month 12. The performance data to the left and right of these ellipses come from different tasks, denoted by symbols, with different average performance, but are self-consistent on either side. Data are more sparse at the beginning of the second period because various training paradigms were attempted, and only a subset of the days included an identical 8 cm center-out task with more than 200 trials, though this became standard later on. The same performance metrics as figure 5 are shown below. On the first task, no significant trends are present. On the second task, voltage declines and recovers significantly ($p < 0.001$), and is significantly correlated with changes in performance using all three metrics ($p < 0.001$). No smoothing filters were used.

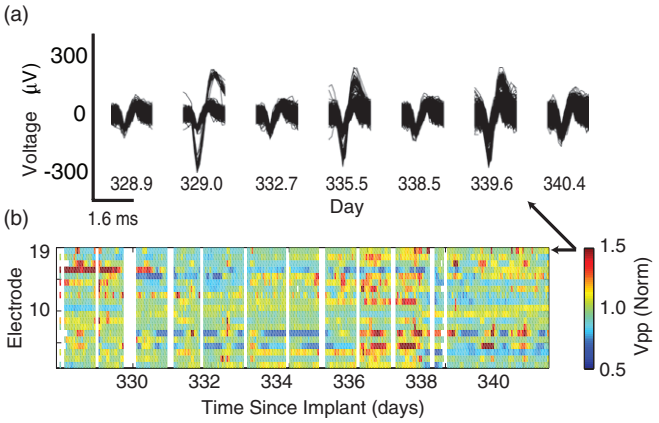

Figure 7. (a) Example waveforms across days from wireless dataset. Unit is regularly visible and occasionally disappears. (b) Average voltage in 1 h bins calculated on 19 of 20 active channels across 13 days of wireless recording starting 328 days after implantation. Voltage is normalized to mean voltage on that channel. White space denotes time without wireless recording, usually during daily experimental session. The arrows denote the channel which was plotted in (a).

substantial with an average absolute change from the mean voltage of 23.7%, as shown in figure 9(b). Again, there were outlying changes that could be very large with 14.6% of the waveforms changing by over 50% between two given days. These results are consistent with previous studies in which large changes were observed between days [34, 38]. The wireless data above provide the fast timecourse of these changes.

4. Discussion

Action potential amplitude (Vpp) declines reported in this study were slow, with an average decline of 2.4% per month. Older implants had slower rates of decline, consistent with a slowing trend. These data are consistent with a prior study [38], which described no trend presumably because these changes are rare, slow, and subtle enough to require many months of data . Also, no significant trend is visible using all sorted units rather than only the largest units on each electrode. For monkey J, these data may also be consistent with tissue response studies in which there is a period of rapid decline in the amplitudes of single units during the first 12 weeks [22, 24, 28]. To our knowledge this is the first study to correlate decoder performance with array voltage over multi-month timescales, though preliminary work has appeared previously [49, 50]. Vpp was not well correlated with performance when using threshold crossings, which suggests that the presence and amplitude of single units cannot be used as a measure of BMI performance. While the performance discussed here was assessed with offline decodes, monkeys J and L were simultaneously participating in online BMI experiments.

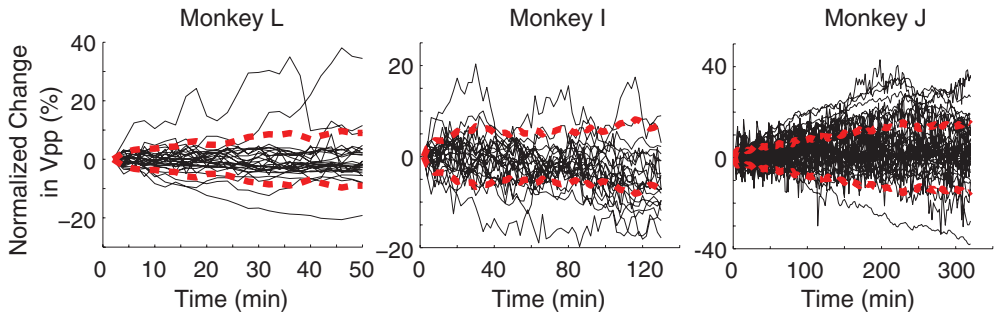

Figure 8. Changes in single unit amplitude during normal experiments while seated with a fixed head position. Voltage normalized to the size of the single unit at the beginning of the experiment to show percentage change in three animals. The red lines denote the standard deviation from zero.

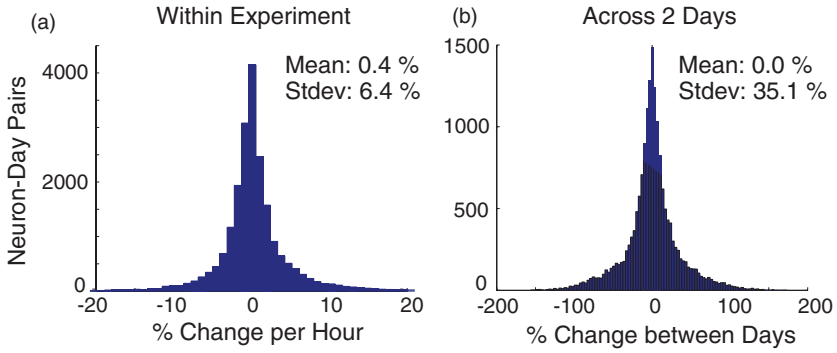

Figure 9. Distribution of percent changes in waveform voltage amplitude from largest unit on an electrode: (a) per hour during 1–5 h experiments, (b) between 2 experimental days on the same electrode.

During this time, no noticeable decline in online decoder performance was noted subjectively, except during the period shown in figure 6, where offline performance declined as well. While offline decodes may not be a good predictor of absolute online performance [47], they may still be useful for estimating the information content over time. Also, the arrays presented here were representative of the arrays observed in our laboratory.

In this study, the performance was analyzed while retraining the decoder every day. This improves performance because of the short-term waveform changes shown in figures 7–9, which lead to a changing subset of neurons comprising the multi-unit. It has been suggested that this may be an obstacle for clinical systems. During online monkey experiments, this requires approximately 15 min of training data and is fully automated in software. No manual supervision of spike sorting is required when using threshold crossings. High performance has also been demonstrated with fewer training trials [11]. Other signal sources such as ECoG have demonstrated stable performance using the same model parameters [51, 52] for up to 8 months [53], but this may reflect a tradeoff between performance and stability. It might be possible that after some number of years, ECoG signals would outperform intracortical array signals on the same task, but that remains unconfirmed. However, with micro-ECoG implants, it might be possible to increase the amount of cortical coverage beyond what is possible using intracortical arrays for applications such as finger decoding [54–56].

To improve intracortical arrays so that they can last for decades, it is important to characterize the type and speed of signal changes. On short timescales, reversible changes in action potential amplitude can occur very quickly. Motion is likely responsible for some of these changes [57]. However, their speed would suggest movements of many tens of μm in a homogenous medium [58]. This does not include nonlinear effects, such as touching a neuron. If there are substantial movements of the array, this underscores the importance of the glial sheath, which enables an electrode to move with less damage to surrounding neural tissue [22]. Changes in electrode or tissue impedance could also result in waveform changes [25]. No matter what the cause, these changes make it very difficult to maintain single units over time.

Fortunately, spike sorting does not substantially improve decoder performance [12]. This could be because only a subset of channels have multiple units, and only a subset of those have opposed directional tuning [32]. On average, combining units may tend to decrease the Poisson noise. Also, human spike sorters have been shown to ignore units containing useful information on account of poor isolation [31]. In this study, threshold-crossing events were detected whenever voltage went below $-4.5\times$ the RMS voltage on that particular channel. In addition to eliminating the need for supervised spike sorting, this approach enabled online BMI experiments to be conducted with older arrays than previous work [14].

Neural data from all arrays in this study were recorded for more than 12 weeks, at which time the glial sheath has probably

already formed. The slow long-term decline in action potential voltage across many months may reflect consolidation of this process, neuron migration away from the site, or accumulating adverse events such as head accelerations or infections. Importantly, Vpp declines on the timescales of months and years may have as much or more to do with materials and engineering failures, and future research is needed to assess this contribution. In this study, large changes in voltage were observed without concurrent changes in performance when using threshold crossings. A number of effects, for example, gliosis or loss of electrode insulation, could cause action potential amplitude to decrease. However, if there was no cell migration away from the electrode, one would still expect the same population of neurons to be recorded at a lower voltage. Therefore, one would expect performance using threshold crossings to remain constant, as described in section 3. While single units might no longer be clearly differentiable, the remaining activity on these channels would still be primarily neural, and probably above the inherent noise of the electrode and amplifiers. If the voltage declined into this noise floor or an alternative process became dominant, the performance would eventually correlate with voltage, similar to the effect observed in monkey L at 27 months post implant. Of course, while this description is consistent with data presented here, only a small number of arrays were examined. Future work examining multi-year old arrays may be warranted. This study also predicts that one could simulate older arrays by using threshold crossings even while single units are present. Wireless systems may be helpful in running array lifetime studies concurrently with conventional experiments [40]. Additionally, fully implantable devices may become crucial for long-term device stability [45, 59, 60]. In this and previous studies, array stability at multiple years is likely confounded by the ability to maintain a transcutaneous link.

This study suggests that the average action potential amplitude on electrode arrays may decay more slowly than previously supposed and that decoder performance can decay at an even slower rate. This is enabled by the use of simple threshold-crossing events rather than well-isolated action potentials from single neurons. This further enables and motivates the translation of cortical BMI work into quality of life improvements for disabled patients.

Acknowledgments

We thank M Risch and S Kang for expert surgical assistance and veterinary care, D Haven for technical consultation, and S Eisensee for administrative support. This work was supported by National Science Foundation graduate research fellowships (CAC, MTK, JMF, VG), NDSEG Fellowship (VG), Burroughs Wellcome Fund Career Awards in the Biomedical Sciences (MMC, KVS), the Christopher Reeve Paralysis Foundation (SIR, KVS), a Stanford University Graduate Fellowship (JPC, ZAR, CAC, JMF, JDF), a CONACYT and DARE fellowship (ZAR), a Stanford NIH Medical Scientist Training Program grant and Soros Fellowship (PN) and the following awards to KVS: Stanford CIS, Sloan Foundation, NIH CRCNS R01-NS054283, McKnight Endowment Fund for Neuroscience, NIH Director's Pioneer Award 1DP1OD006409, DARPA Revolutionizing Prosthetics program contract N66001-06-C-8005, and DARPA REPAIR contract N66001-10-C-2010.

References

[1] Gilja V, Chestek C A, Diester I, Henderson J M, Deisseroth K and Shenoy K V 2011 Challenges and opportunities for next-generation intra-cortically based neural prostheses *IEEE Transactions on Biomedical Engineering.* at press doi:10.1109/TBME.2011.2107553

[2] Santhanam G, Ryu S I, Yu B M, Afshar A and Shenoy K V 2006 A high-performance brain–computer interface *Nature* **442** 195–8

[3] Taylor D M, Helms Tillery S I and Schwartz A B 2002 Direct cortical control of 3D neuroprosthetic devices *Science* **296** 1829–32

[4] Serruya M D, Hatsopoulos N G, Paninski L, Fellows M R and Donoghue J P 2002 Instant neural control of a movement signal *Nature* **416** 141–2

[5] Carmena J M, Lebedev M A, Crist R E, O'Doherty J E, Santucci D M, Dimitrov D F, Patil P G, Henriquez C S and Nicolelis M A L 2003 Learning to control a brain–machine interface for reaching and grasping by primates *PLoS Biol.* **1** 193–208

[6] Lebedev M A, Carmena J M, O'Doherty J E, Zacksenhouse M, Henriquez C S, Principe J C and Nicolelis M A L 2005 Cortical ensemble adaptation to represent velocity of an artificial actuator controlled by a brain–machine interface *J. Neurosci.* **25** 4681–93

[7] Carmena J M, Lebedev M A, Henriquez C S and Nicolelis M A 2005 Stable ensemble performance with single-neuron variability during reaching movements in primates *J. Neurosci.* **25** 10712–6

[8] Hochberg L R, Serruya M D, Friehs G M, Mukand J A, Saleh M, Caplan A H, Branner A, Chen D, Penn R D and Donoghue J P 2006 Neuronal ensemble control of prosthetic devices by a human with tetraplegia *Nature* **442** 164–71

[9] Kim S, Simeral J D, Hochberg L R, Donoghue J P and Black M J 2008 Neural control of computer cursor velocity by decoding motor cortical spiking activity in humans with tetraplegia *J. Neural Eng.* **5** 455–76

[10] Mulliken G H, Musallam S and Andersen R A 2008 Decoding trajectories from posterior parietal cortex ensembles *J. Neurosci.* **28** 12913–26

[11] Chase S M, Schwartz A B and Kass R E 2009 Bias, optimal linear estimation, and the differences between open-loop simulation and closed loop performance of spiking based brain–computer interface algorithms *Neural Netw.* **22** 1203–13

[12] Fraser G W, Chase S M, Whitford A and Schwartz A B 2009 Control of a brain–computer interface without spike sorting *J. Neural Eng.* **6** 1–8

[13] Ganguly K and Carmena J M 2009 Emergence of a stable cortical map for neuroprosthetics *PLoS Biol.* **7** 1–13

[14] Gilja V, Nuyujukian P, Chestek C A, Cunningham J P, Yu B M, Ryu S I and Shenoy K V 2010 High-performance continuous neural cursor control enabled by a feedback control perspective *Conference Abstract: Computational and Systems Neuroscience (COSYNE) (Salt Lake City, UT, USA)*

[15] Gilja V, Nuyujukian P, Chestek C A, Cunningham J P, Yu B M, Ryu S I and Shenoy K V 2010 A high-performance continuous cortically controlled prosthesis enabled by feedback control design *Soc. for Neurosci. Abstr.* 20.6

[16] Nuyujukian P, Gilja V, Chestek C A, Cunningham J P, Fan J M, Yu B M, Ryu S I and Shenoy K V 2010 Generalization and robustness of a continuous cortically-controlled prosthesis enabled by feedback control design *Soc. for Neurosci. Abstr.* 20.7

[17] Velliste M, Perel S, Spalding M C, Whitford A S and Schwartz A B 2008 Cortical control of a prosthetic arm for self-feeding *Nature* **453** 1098–101

[18] Vargas Irwin C E, Shakhnarovich G, Yadollahpour P, Mislow J M K, Black M J and Donoghue J P 2010 Decoding complete reach and grasp actions from local primary motor cortex populations *J. Neurosci.* **30** 9659–69

[19] Kuiken T A, Li G, Lock B A, Lipschutz R D, Miller L A, Stubblefield K A and Englehart K B 2009 Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms *J. Am. Med. Assoc.* **301** 619–28

[20] Schmidt S, Horch K and Normann R 1993 Biocompatibility of silicon-based electrode arrays implanted in feline cortical tissue *J. Biomed. Mater. Res.* **27** 1393–9

[21] Stensaas S S and Stensaas L J 1978 Histopathological evaluation of materials implanted in the cerebral cortex *Acta Neuropathol.* **41** 145–55

[22] Turner J N, Shain W, Szarowski D H, Andersen M, Martins S, Isaacson M and Craighead H 1999 Cerebral astrocyte response to micromachined silicon implants *Exp. Neurol.* **155** 33–49

[23] Edell D J, Toi V V, McNeal V M and Clark L D 1992 Factors influencing the biocompatibility of insertable silicon microshafts in cerebral cortex *IEEE Trans. Biomed. Eng.* **39** 635–43

[24] Williams J C, Rennaker R L and Kipke D R 1999 Long-term neural recording characteristics of wire microelectrode arrays implanted in cerebral cortex *Brain Res. Protocols* **4** 303–13

[25] Winslow B D and Tresco P A 2010 Quantitative analysis of the tissue response to chronically implanted microwire electrodes in rat cortex *Biomaterials* **31** 1558–67

[26] Griffith R W and Humphrey D R 2006 Long-term gliosis around chronically implanted platinum electrodes in the rhesus macaque motor cortex *Neurosci. Lett.* **406** 81–6

[27] Polikov V S, Tresco P A and Reichert W M 2005 Response of brain tissue to chronically implanted neural electrodes *J. Neurosci. Methods* **148** 1–18

[28] Szarowski D H, Andersen M D, Spence A J, Isaacson M, Craighead H G, Turner J N and Shain W 2003 Brain responses to micro-machined silicon devices *Brain Res.* **983** 23–35

[29] Kruger J, Caruana F, Volta R D and Rizzolatti G 2010 Seven years of recording from monkey cortex with a chronically implanted multiple microelectrode *Frontiers in Neuroengineering* **3** 1–9

[30] Truccolo W, Friehs G M, Donoghue J P and Hochberg L R 2008 Primary motor cortex tuning to intended movement kinematics in humans with tetraplegia *J. Neurosci.* **28** 1163–78

[31] Wood F, Fellows M, Donoghue J and Black M J 2004 Automatic spike sorting for neural decoding *Proc. 26th Annu. Int. IEEE EMBS (San Francisco, CA, USA)* pp 4009–12

[32] Santhanam G, Sahani M, Ryu S I and Shenoy K V 2004 An extensible infrastructure for fully automated spike sorting during online experiments *Proc. 26th Annu. Int. Conf. IEEE EMBS (San Francisco, CA, USA)* pp 4380–4

[33] Stark E and Abeles M 2007 Predicting movement from multiunit activity *J. Neurosci.* **27** 8387–94

[34] Parker R A, Davis T S, House P A, Normann R A and Greger B 2011 The functional consequences of chronic, physiologically effective intracortical microstimulation *Progress in Brain Research* vol 194 at press

[35] Chestek C A, Batista A P, Santhanam G, Yu B M, Afshar A, Cunningham J P, Gilja V, Ryu S I, Churchland M M and Shenoy K V 2007 Single-neuron stability during repeated reaching in macaque premotor cortex *J. Neurosci.* **27** 10742–50

[36] Dickey A S, Suminski A, Amit Y and Hatsopoulos H G 2009 Single-unit stability using chronically implanted multielectrode arrays *J. Neurophysiol.* **102** 1331–9

[37] Serruya M, Hatsopoulos N G, Fellows M R, Paninski L and Donoghue J P 2003 Robustness of neuroprosthetic decoding algorithms *Biol. Cybern.* **88** 219–28

[38] Suner S, Fellows M R, Vargas Irwin C, Nakata G K and Donoghue J P 2005 Reliability of signals from a chronically implanted silicon-based electrode array in non-human primate primary motor cortex *IEEE Trans. Neural Sys. Rehabil. Eng.* **13** 524–41

[39] Chestek C A, Gilja V, Nuyujukian P, Kier R, Solzbacher F, Ryu S I, Harrison R R and Shenoy K V 2009 HermesC: low-power wireless neural recording system for freely moving primates *IEEE Trans. Neural. Sys. Rehabil. Eng.* **17** 330–8

[40] Gilja V, Chestek C, Nuyujukian P, Foster J and Shenoy K V 2010 Autonomous head-mounted electrophysiology systems for freely behaving primates *Curr. Opin. Neurobiol.* **20** 676–86

[41] Rivera-Alvidrez Z, Kalmar R, Ryu S I and Shenoy K V 2010 Low-dimensional neural features predict muscle EMG signals *Proc. 32nd Annu. Int. Conf. IEEE EMBS (Buenos Aires, Argentina)* pp 6027–33

[42] Churchland M M, Cunningham J P, Kaufman M T, Ryu S I and Shenoy K V 2010 Cortical preparatory activity: representation of movement or first cog in a dynamical machine? *Neuron* **68** 387–400

[43] Georgopoulos A P, Kalaska J F, Caminiti R and Massey J T 1982 On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex *J. Neurosci.* **2** 1527–37

[44] Hatsopoulos N, Joshi J and O'Leary J G 2004 Decoding continuous and discrete motor behaviors using motor and premotor cortical ensembles *J. Neurophysiol.* **92** 1165–74

[45] Harrison R R, Kier R J, Chestek C A, Gilja V, Nuyujukian P, Ryu S, Greger B, Solzbacher F and Shenoy K V 2009 Wireless neural recording with single low-power integrated circuit *IEEE Trans. Neural Sys. Rehabil. Eng.* **17** 322–9

[46] Sahani M 1999 Latent variable models for neural data analysis *PhD Thesis, Computational and Neural Systems, California Institute of Technology*

[47] Cunningham J P, Nuyujukian P, Gilja V, Chestek C A, Ryu S I and Shenoy K V 2011 A closed-loop human simulator for investigating the role of feedback-control in brain–machine interfaces *J. Neurophysiol.* at press

[48] Cunningham J P, Nuyujukian P, Gilja V, Chestek C A, Ryu S I and Shenoy K V 2010 A closed-loop human simulator for understanding the role of feedback-control and its relevance for brain–machine interfaces *Soc. for Neurosci. Abstr.* 20.5

[49] Chestek C A, Cunningham J P, Gilja V, Nuyujukian P, Ryu S I and Shenoy K V 2009 Neural prosthetic systems: current problems and future directions *Proc. 31st Annu. Int. Conf. IEEE EMBS (Minneapolis, MN, USA)* pp 3369–75

[50] Chestek C A, Gilja V, Nuyujukian P, Foster J D, Kaufman M T, Ryu S I and Shenoy K V 2010 Impact of spike waveform instability on performance measured across seven weeks using a wireless system *Soc. for Neurosci. Abstr* 85.1

[51] Blakely T, Miller K J, Zanos S P, Rao R and Ojemann J G 2009 Robust long-term control of an electrocorticographic brain–computer interface with fixed parameters *Neurosurg. Focus* **27** 1–5

[52] Schalk G, Miller K J, Anderson N R, Wilson J A, Smyth M D, Ojemann J G, Moran D W, Wolpaw J R and Leuthardt E C 2008 Two-dimensional movement control using electrocorticograph signals in humans *J. Neural Eng.* **5** 75–84

[53] Chao Z C, Nagasaka Y and Fujii N 2010 Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys *Front. Neuroeng.* **3** 1–10

[54] Kubanek J, Miller K J, Ojemann J G, Wolpaw J R and Schalk G 2009 Decoding flexion of individual fingers using electrocorticographic signals in humans *J. Neural Eng.* **6** 1–14

[55] Miller K J, Zanos S, Fetz E E, Nijs M and Ojemann J G 2009 Decoupling the cortical power spectrum reveals real-time representation of individual finger movements in humans *J. Neurosci.* **29** 3132–7

[56] Acharya S, Fifer M S, Benz H L, Crone N E and Thakor N V 2010 Electrocorticographic amplitude predicts finger positions during slow grasping motions of the hand *J. Neural Eng.* **7** 1–13

[57] Santhanam G, Linderman M D, Gilja V, Afshar A, Ryu S I, Meng T H and Shenoy K V 2007 Hermesb: a continuous neural recording system for freely behaving primates *IEEE Trans. Biomed. Eng.* **54** 2037–50

[58] Moffitt M A and McIntyre C C 2005 Model-based analysis of cortical recording with silicon microelectrodes *Clin. Neurophysiol.* **116** 2240–50

[59] Kim S, Bhandari R, Klein M, Negi S, Rieth L, Tathireddy P, Toepper M, Oppermann H and Solzbacher F 2009 Integrated wireless neural interface based on the Utah electrode array *Biomed. Microdevices* **11** 453–66

[60] Borton D A, Song Y K, Patterson W R, Bull C W, Park S, Laiwalla F, Donoghue J P and Nurmikko A V 2009 Wireless, high-bandwidth recordings from non-human primate motor cortex using a scalable 16-ch implantable microsystem *Proc. 31st Annu. Int. Conf. IEEE EMBS (Minneapolis, MN, USA)* pp 5531–4