Statistical assessment of the stability of neural movement representations

Ian H. Stevenson¹, Anil Cherian², Brian M. London², Nicholas A. Sachs³, Eric Lindberg², Jacob Reimer⁴*, Marc W. Slutzky¹,²,⁵, Nicholas G. Hatsopoulos⁴, Lee E. Miller¹,²,³, and Konrad P. Kording¹,²

1. Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, 345 E. Superior Ave., Chicago, IL 60611, USA
2. Department of Physiology, Feinberg School of Medicine, Northwestern University 303 E. Chicago Ave., Chicago, IL 60611, USA
3. Department of Biomedical Engineering, Northwestern University 2145 Sheridan Rd., Evanston, IL 60208, USA
4. Department of Organismal Biology and Anatomy, University of Chicago 1027 E. 57th St. Chicago, IL 60637, USA
5. Department of Neurology, Feinberg School of Medicine, Northwestern University 303 E. Chicago Ave., Chicago, IL 60611, USA

*Current Address: ¹Department of Neuroscience, Baylor College of Medicine, TX 77030

CORRESPONDING AUTHOR
Konrad P. Kording
Department of Physical Medicine and Rehabilitation Northwestern University Feinberg School of Medicine 345 East Chicago Ave Attn: ONT-931 Chicago, IL USA Phone: 312-238-6966 e-mail: kk@northwestern.edu

ACKNOWLEDGEMENTS
IHS and KPK were supported by the Chicago Community Trust and US National Institute of Health grants 1R01NS063399 and 2P01NS044393. AC, BML, and LEM were supported by NIH R01NS048845. NAS was supported by the NCRR-NIH Clinical and Translational Science Awards Program (CTSA), part of the Roadmap Initiative, "Re-Engineering the Clinical Research Enterprise," UL1RR0254741. EL and MWS were supported by NIH K08NS060223. JR and NGH were supported by NIH-NINDS Grant R01NS45853.
ABSTRACT

In systems neuroscience, neural activity that represents movements or sensory stimuli is often characterized by spatial tuning curves that may change in response to training, attention, altered mechanics, or the passage of time. A vital step in determining whether tuning curves change is accounting for estimation uncertainty due to measurement noise. In this study, we address the issue of tuning curve stability using methods which take uncertainty directly into account. We analyze data recorded from neurons in M1 using chronically implanted multi-electrode arrays in four monkeys performing center-out reaching. Using simulations, we demonstrate that under typical experimental conditions, the effect of neuronal noise on estimated preferred direction can be quite large, and is affected by both the amount of data and the modulation depth of the neurons. In experimental data we find that, after taking uncertainty into account using bootstrapping techniques, the majority of neurons appear to be very stable on a timescale of minutes to hours. Lastly, we introduce adaptive filtering methods to explicitly model dynamic tuning curves. In contrast to several previous findings suggesting that tuning curves may be in constant flux, we conclude that the neural representation of limb movement is, on average, quite stable and that impressions to the contrary may be largely the result of measurement noise.

INTRODUCTION

Many experiments in neuroscience aim at measuring how experimental manipulations affect the tuning properties of neurons. Tuning curves typically characterize how firing rates depend on a single relevant property of a stimulus or movement. For example, visual neurons are often characterized by their dependence on stimulus orientation, auditory neurons by their dependence on pitch, and motor cortical neurons by their dependence on the direction of movement of an animal’s hand. A common method for studying how these coding properties change is to characterize the tuning curve, introduce an experimental manipulation, and characterize the tuning curve again to assess if changes have occurred (Cronin et al. 2010; Li et al. 2001; Rokni et al. 2007; Schummers et al. 2007). Tuning curves in a variety of brain areas have been shown to change during adaptation and learning. In the absence of an experimental manipulation we can ask how stable tuning curves are and whether they drift due to the passage of time. Importantly, determining whether changes in tuning are statistically significant depends on our ability to take into account the effects of noise and limited data.

Since noise in the recorded data (e.g. Poisson noise in spike counts) will affect tuning curve estimates, it is important to characterize any tuning curve parameters with confidence bounds. Without such bounds it is impossible to determine whether an observed change is due to actual changes in tuning or simply measurement noise and uncertainty in the estimation. We will tend to overestimate the average (absolute) magnitude of any existing tuning curve changes, since the measured change consists of the real change plus apparent changes due to measurement noise. Therefore, measurement noise must be considered in order to obtain reasonable estimates of the magnitude of changes in tuning curves.

Here we focus on the role of measurement uncertainty in estimating the tuning curves of neurons in primary motor cortex. In particular, we examine the discharge properties of neurons during a well-learned center-out task, where a monkey is trained to reach from the workspace center to several peripheral target locations (Georgopoulos et al. 1982; Kakei et al. 1999; Kalaska and Hyde 1985; Morrow et al. 2007; Scott and Kalaska 1997). Typically, tuning curves that characterize the discharge of neurons in primary motor cortex (M1) as a function of reaching direction are well fit by a cosine function (Georgopoulos et al. 1982; Georgopoulos et al. 1988). The peak of this cosine indicates the direction of movement for which the neuron's firing rate is maximal, the “preferred direction” (PD).
A central question in the neural control of movement is whether the preferred directions of M1 neurons are stable during normal behavior, from reach to reach. If tuning curves fluctuate from reach to reach the motor system must either be redundant, such that fluctuations do not affect behavior (Pohlmeyer et al. 2007; Rokni et al. 2007), or must rapidly adapt to allow for stable movement. By recording many trials from the same neurons over the course of several days, a number of studies have shown that neuronal tuning appears to be relatively stable (Chestek et al. 2007; Dickey et al. 2009). However, there is also some contrasting evidence to suggest that neurons fluctuate rapidly on timescales on the order of minutes to hours (Rokni et al. 2007). One recent study found that the preferred directions of movement-related neurons in M1 varied substantially over time with a majority of neurons shifting as much as 30 degrees within 15 minutes (Rokni et al. 2007). These results suggested that neurons may be quite unstable and that there must be substantial redundancy in the nervous system and down-stream processing to ensure stable behavior. However, these estimates were obtained with methods that did not explicitly correct for measurement noise.

In this study, we address these issues using methods which take measurement uncertainty and potential instability directly into account. We analyze data recorded from neurons in M1 using chronically implanted multi-electrode arrays. Using simulations, we demonstrate that under typical experimental conditions, the effect of neuronal noise on estimated preferred direction can be quite large. Using experimental data we find that, after taking uncertainty into account, the average patterns of neuronal activity expressed by tuning curves are quite stable on a timescale of minutes to hours. We then explore ways of explicitly modeling dynamic tuning curves using adaptive filtering methods. These methods reveal that there may be tiny but real fluctuations of the tuning properties of neurons but that such changes are very hard to detect robustly. We conclude that the neural representation of limb movement is actually quite stable, and that impressions to the contrary may be the result largely of measurement noise.

METHODS

Tasks

Four monkey subjects (designated F, K, C, and R) were trained on an eight-target center-out reaching paradigm. Monkeys F, K, and C were seated in a primate chair, with movement constrained to a horizontal plane. The monkey grasped the handle of a two link planar manipulandum that moved within a 20 cm by 20 cm workspace. Feedback about movement was given on a computer screen in front of the monkey. Handle position was displayed as a circular cursor, 1-2 cm in diameter. The experiments with monkey R used a KINARM device (BKIN Technologies, Kingston, ON, Canada) in which the monkey’s arm rested on cushioned troughs secured to links of a two-joint robotic arm (Scott 1999). The shoulder joint was abducted 90 degrees such that shoulder and elbow flexion and extension movements were made in the horizontal plane. A cursor coincident with the handle position of the robotic arm was projected onto a horizontal screen placed above the monkey’s arm. All trials began with the acquisition of a square center target that the monkey was required to hold for 0.3-1.1s (0.6-1.1s for monkey F, 0.5-0.6s for monkey C, 0.3-0.5s for monkey K). After this hold period, subjects reached to one of eight equally spaced, peripheral targets. Subjects had approximately 1.25 s to acquire the outer target and were required to hold this outer target for at least 0.2-0.5s. Each success was rewarded with juice or water.

Surgery

Once the monkey was able to perform the center-out task satisfactorily, we performed a surgery to implant a 100 electrode intra-cortical array (Blackrock Microsystems). We made a craniotomy centered above the arm area of M1. After opening the dura, we identified the area on the crown of the precentral gyrus, just medial to the spur of the arcuate sulcus. In some cases we stimulated within this area using a ball electrode array (monopolar, biphasic, 50 Hz, 100 µsec pulse width, < 6mA) to locate proximal limb movements. All surgery was performed under isoflurane gas anesthesia except during intra-operative stimulation. In order to increase
cortical excitability, approximately 30 minutes prior to stimulation, we began infusing remifentanil (0.4 µg/kg/min, i.v.) while gradually reducing the concentration of isoflurane to 0.25%. We positioned the array within the identified area, taking care to avoid any major surface vessels and inserted it rapidly with a pneumatic inserter (supplied by Blackrock Microsystems). The dura was closed over the array, with a piece of artificial pericardium (Preclude ePTFE membrane, Gore and Associates, Inc.) under the dura to prevent the back of the array from adhering to the dura. Another piece of Preclude was placed over the dura. The original bone flap was thinned and replaced, and the skin was closed over the craniotomy. All leads from the array were routed to a percutaneous connector secured to the monkey’s skull (Nordhausen et al. 1996).

All animal use procedures were approved by the institutional animal care and use committees at Northwestern University (datasets F, K, and C) or the University of Chicago (dataset R).

**Tuning curve estimation and simulations**

Spike trains from each subject were recorded during center-out reaching for 30-40 minutes resulting in at least 300 successful trials for datasets F, R, and K and 290 trials for dataset K. Neural signals were classified as single-units based on action potential shape and minimum inter-spike intervals of 1.6 ms. Spike sorting was performed off-line by manual cluster cutting. Trial-by-trial spike counts from 100 ms prior to movement onset until 300 ms after movement onset are used throughout the analysis.

Following the conventional cosine tuning model, we assume that the firing rate of each neuron depends on the direction of hand movement, $\theta$ as

$$ y = b_0 + b_1 \cos(\theta - \theta^*) $$  \hspace{1cm} (1)

where the parameters $b_0$, $b_1$, and $\theta^*$ denote the baseline firing rate, modulation, and the preferred direction (Georgopoulos et al. 1986). Traditionally, these parameters are estimated by minimizing the squared error between predicted and actual firing rates (Swindale 1998). To simplify the optimization, we can use the sum-difference formula for cosines to rewrite (1) as

$$ y = b_0 + c_1 \cos(\theta) + c_2 \sin(\theta) $$  \hspace{1cm} (2)

giving $\theta^* = \text{atan2}(c_2, c_1)$ and $b_1 = (c_1 - c_2)/(\cos\theta^* - \sin\theta^*)$. In this form we can efficiently estimate the parameters using linear regression. Note that, since minimizing the squared error is equivalent to maximizing the log-likelihood with a Gaussian noise model, this optimization implicitly assumes Gaussian noise. In the low firing rate limit, other noise models, such as Poisson, may be more appropriate (Cronin et al. 2010), and several studies have suggested that exponential-cosine tuning functions may be a better description of neurons in M1 (Amirikian and Georgopoulos 2000; Hatsopoulos et al. 2007). An analysis of these alternative tuning models is presented in the Supplementary Material. However, for the main analysis we assumed cosine tuning with Gaussian noise as a standard model.

For simulations, the firing rate is assumed to be constant over the observation window, drawn from a Poisson distribution with a firing rate given by (1). For the example neurons we fixed the parameters by hand to a physiologically plausible range, and for a more complete analysis we incrementally varied the modulation depth. For comparisons to recorded data, the parameters $b_0$, $b_1$, and $\theta^*$ were estimated from the entire recording session for each recorded neuron. We then simulated "stable" neurons with the same tuning properties, and compared these results, where tuning curves are fixed, with the observed data. This allowed us to test whether results that previously seemed to suggest tuning curve drift might be consistent with noisy, stable neurons. It is important to note that, although tuning curve estimation assumed Gaussian noise, the
simulations used here assumed Poisson noise. This is somewhat non-standard from the point of view of statistical inference, but we have used this approach since the Gaussian noise assumption is so commonly used for estimating motor tuning curves.

Bootstrapping

To estimate the uncertainty associated with each of the parameters in (1), we used bootstrapping (Davison and Hinkley 1997; Efron and Tibshirani 1997). Briefly, we randomly resampled the original data with replacement and estimated the parameters from each of these resampled datasets. Bootstrapping creates a distribution of parameters consistent with the data, from which confidence intervals can be estimated and significance can be assessed. In practice, the results are robust to the choice of number of bootstrap samples and in this case were nearly identical for >500 samples. In the following analyses we used 1000 bootstrap samples.

The variability of the bootstrap samples gives an indication of how uncertain the parameter estimates are. Confidence intervals for each of the parameters can then be assessed using the quantiles of the bootstrapped distributions over the parameters. For example, 95% confidence intervals are estimated by the 2.5 and 97.5 percentiles of the distribution. After sorting the samples, the lowest and highest 2.5% of the samples are discarded. The range of the remaining samples then defines the error margin. For 1000 bootstrap samples these would be the 26th and 975th samples from the ranked list. Importantly, this method does not need to make the assumption that the distribution is Gaussian, symmetric, or even uni-modal.

Since the preferred direction parameter is in circular coordinates, special care is needed to estimate the confidence interval. A number of methods exist for dealing with small, <100, sample sizes (Fisher 1996; Otieno and Anderson-Cook). However, in our case (~1000 bootstrap samples, from well-tuned neurons) it suffices first to center the samples by the median, and then apply the quantile procedure described above.

Estimating confidence intervals allows us to assess statistical significance. Mainly, we want to determine whether the change in preferred direction from one block of trials to the next is statistically significant. In this case we examined the distribution of differences between the bootstrap samples for preferred direction from each block. If the 95% confidence interval for this distribution of differences did not contain 0 then we reject the null-hypothesis (that the difference actually is 0) at the 5% significance level. For a previous application of this method to tuning curves see (Churchland and Shenoy 2007).

In general, the size of confidence intervals is determined by the number of spike observations. For example, the standard error of the mean for a Poisson distributed random variable scales as \(1/\sqrt{N}\) where \(N\) is the number of observations. For cosine tuned neurons, the confidence interval about the preferred direction is the main quantity of interest. For simplicity we can reparameterize the tuning curves (1) in terms of the total number of spikes emitted by a neuron and the depth of modulation relative to baseline (\(M = b_t/b_0\)). Given the assumptions of Poisson noise and cosine tuning, uncertainty about preferred direction is affected only by these two factors.

To assess how these factors influence the measurement uncertainty, we parametrically varied the total expected number of spikes as well as the modulation depth. For each simulated neuron we sampled spikes from a Poisson distribution in 8 reach directions with the firing rate given by a cosine tuning curve. Without loss of generality, we can assume that the preferred direction is 0 deg. We then used the bootstrapping approach described above to generate bootstrap PD distributions for each modulation depth, spike count combination and determined the 95% confidence bounds of the distribution.
In order to compare our results with previously published data, we read out the maximal (max) and minimal (min) firing rates across direction, from the graphs published in several papers (Kalaska et al. 1989; Rokni et al. 2007; Wise et al. 1998). In these cases we estimated modulation depth from the tuning curves as 

\[ \hat{M} = (\text{max} - \text{min}) / (\text{max} + \text{min} / 2) \]

and the total number of spikes was estimated as 

\[ \hat{N} = Tw(\text{max} + \text{min} / 2) \]

where \( T \) is the number of trials and \( w \) is the integration window.

Corrected estimate of the variance of changes in preferred direction

In the presence of measurement noise, the differences that we observe in preferred direction between blocks will be due partially to a true (hidden) drift and partially due to the measurement noise. Therefore, we will tend to overestimate fluctuations in preferred direction. Using \( PD \) to denote the true, but unobserved, preferred direction and \( PD_{\text{est}} \) to denote the estimated preferred direction

\[
\text{var}(PD_{1,\text{est}} - PD_{2,\text{est}}) = \text{var}((PD_1 + \epsilon_1) - (PD_2 + \epsilon_2))
\]

where \( \epsilon_1 \) and \( \epsilon_2 \) denote instantiations of the measurement noise, which is assumed Gaussian. Since the variance of a sum of independent random variables is equal to the sum of the variances, the corrected estimate of the fluctuations in preferred direction is given by

\[
\text{var}(PD_1 - PD_2) = \text{var}(PD_{1,\text{est}} - PD_{2,\text{est}}) - \text{var}(\epsilon_1) - \text{var}(\epsilon_2)
\]

which captures the excess variance (above the measurement noise) that may be attributed to an actual change in preferred direction. We estimated \( \text{var}(\epsilon) \) empirically by using the variance of the bootstrap samples (mean across the population of neurons). In some cases, the corrected value \( \text{var}(\Delta PD) \) was negative. These values were set to 0 to calculate summary statistics, and suggest that there was no detectable PD drift after accounting for measurement noise. The mean and difference statistics reported in the text were all calculated using circular statistics (Berens 2009). However, the equations above use a linear rather than circular coordinate system for correcting the variance. In the results presented here the distribution of \( \Delta PD \) is narrow enough that (4) and (5) provide an accurate approximation to the fully circular variance correction.

Adaptive filtering

Although the methods described above for assessing and describing drift are fairly robust, one potential issue with these approaches is that they assume that tuning curves are stable within each block and that each trial is independent. Ideally, we should use a method that explicitly assumes that tuning curve parameters are drifting over time, from trial to trial. Adaptive filtering is one approach that makes this assumption explicit. Several recent studies have used point-process adaptive filtering to describe time-varying tuning of place cells in hippocampus (Brown et al. 2001; Eden et al. 2004). Here, for consistency with the initial analysis, we analyze trial-by-trial spike counts assuming cosine tuning with Gaussian noise and use least-mean-squares (LMS) or steepest-decent adaptive filtering (Haykin 1996). Briefly, the parameters of the tuning model \( \beta \) are updated every trial following

\[
\hat{\beta}_k = \hat{\beta}_{k-1} - \epsilon \frac{\partial l_k(\beta)}{\partial \beta} \bigg|_{\beta = \hat{\beta}_{k-1}}
\]

(5)

where \( l_k(\beta) \) denotes the instantaneous, negative log-likelihood (squared error) and \( \epsilon \) denotes a learning rate which must be optimized separately. Assuming Gaussian noise and writing (2) in matrix form such that
where $X_k = [1 \cos(\theta_k) \sin(\theta_k)]$ and $\beta = [b_0, c_1, c_2]^T$ we have the updates

$$\hat{\beta}_k = \hat{\beta}_{k-1} + \epsilon(y_k - X_k \hat{\beta}_{k-1}) X_k^T.$$  \hspace{1cm} (6)

The aim of this analysis is to update the parameters at each time-step to minimize prediction errors. The learning rate for each neuron and parameter is chosen to maximize the total likelihood of the time-varying parameter estimates. The final time-varying parameter estimates then allow us to extract the per-trial drift rate and fluctuations for each parameter. For the analyses presented here we focus solely on the dynamics of the preferred direction. In this case the tuning curve is parameterized following (1), and the baseline and modulation are held constant, while the preferred direction is updated as

$$\hat{\theta}_k^* = \hat{\theta}_{k-1}^* + \epsilon(y_k - (b_0 + b_1 \cos(\theta_k - \hat{\theta}_{k-1}^*))) \sin(\theta_k - \hat{\theta}_{k-1}^*).$$  \hspace{1cm} (7)

The LMS algorithm works well for processes where the parameters follow a Gaussian random walk. To examine the performance of adaptive filtering we simulated cosine-tuned Poisson neurons with fixed baseline firing rate and modulation, but where the preferred direction followed a random walk

$$\theta_k^* = \theta_{k-1}^* + \eta \quad \eta \sim N(\mu, \sigma).$$  \hspace{1cm} (8)

In this case the preferred direction changes by a random amount $\eta$ each trial, drawn from a Gaussian distribution with mean $\mu$ and standard deviation $\sigma$.

Time-varying tuning can lead to surprising results. For instance, a neuron that appears very broadly tuned using traditional block-estimation techniques may, in reality, be a sharply tuned but drifting neuron (Brown et al. 2001). Finally, it is important to note that when the parameters do not have linear dynamics, but instead change suddenly, recursive methods may be more appropriate to estimate these changes accurately (see Eden et al. 2004). Here we used the more basic LMS algorithm as a first step in analyzing trial-by-trial fluctuations in tuning parameters.

---

RESULTS

Measuring uncertainty in estimates of preferred direction

In order to illustrate how measurement noise affects estimates of neuronal preferred direction, we simulated the discharge of a typical cosine-tuned Poisson neuron (Fig 1A) during repeated reaches to 8 targets. We assume the simulated neuron has a movement related discharge with a baseline firing rate of 20 Hz and a modulation of 5 Hz. Without loss of generality we assume that the neuron has an actual preferred direction (PD) of 180 degrees. Due to Poisson noise in the simulation, the number of spikes generated during different reaches to the same target differed considerably (Fig 1B). This type of variability is typical of experimental recordings. To characterize measurement uncertainty in this situation we used the well-established technique of bootstrapping. Essentially, bootstrapping is a procedure that produces alternative fits that could have been obtained if one could repeat a given experiment many times. In the case of the simulated neuron in Fig 1, we initially simulated 40 reaches. With bootstrapping, we resampled from these data with replacement, each time computing a tuning curve and preferred direction (Fig 1C). Ultimately, this process can be used to
obtain a distribution of possible preferred directions that is be compatible with the original spike data and that can be used to estimate how reliable each tuning parameter is (Fig 1D). In this example, the 95% confidence interval spanned approximately 40 degrees. The parameters we used were well within the range of typical M1 experiments (Georgopoulos et al. 1982; Kalaska et al. 1989; Morrow et al. 2007; Wise et al. 1998); however, even after 40 trials there was considerable uncertainty in the preferred direction, and this uncertainty directly determines whether or not a change in tuning can be detected.

To determine whether the tuning curve of a neuron has changed we can use this same bootstrapping approach on two successive blocks of data. For illustration, we simulated a hypothetical neuron with a tuning curve that changed following an experimental manipulation (Fig 2A). We again assumed that this neuron has a baseline firing rate of 20 Hz and a modulation of 5 Hz. In this case we assumed that the preferred direction changes from 90 deg in the first condition to 135 deg in the second condition. After simulating 40 reaches for each condition (Fig 2B), we calculated distributions over potential PDs for each of the two conditions using bootstrapping (Fig 2C). The two resulting distributions of PDs overlapped only slightly, and the distribution of differences between the preferred directions was significantly different from 0. For this simulation, bootstrapping allowed us to conclude that the PD changed as a result of the experimental manipulation. However, the width of the resulting histograms (~40 degrees) indicates that changes in PD for a given neuron need to be rather large in order to be visible with this amount of data.

It is important to understand how measurement uncertainty varies as a function of these simulation parameters. Given cosine tuning and Poisson noise, there are two relevant parameters that determine the precision of PD estimates: the total number of spikes and the modulation depth relative to baseline (see Methods). We used the bootstrapping methods described above to calculate the average confidence interval size as a function of these two parameters (Fig 3). In this case a modulation depth of 1 corresponded to a cosine between 0 and the maximal firing rate, and a modulation depth of 0 corresponded to an un-tuned cell. In general, more strongly modulated neurons yield more precise PD estimates, and as more data are observed the estimates of all tuning curve parameters become more precise.

Based on these simulated results, the precision of PD estimates in typical studies is likely to vary substantially. Here we have shown total spike numbers and modulation depths for exemplar neurons in several published studies (Kalaska et al. 1989; Rokni et al. 2007; Wise et al. 1998), as well as our own data (Fig 3B). Some of these studies allowed PDs to be estimated with a precision of approximately 10 degrees while others were probably precise only to 50 or 60 degrees. For our data the average modulation depth was 0.49±0.02 with 6.23±0.39 spikes/trial on average (400 ms integration window). Mapping these values onto the simulation results suggested an average confidence interval size of 33.5 deg for blocks of 40 trials, 20.6 deg for blocks of 120 trials, and 16.9 deg for blocks of 240 trials.

Here we have assumed that the only source of variability in addition to directional tuning is Poisson noise. Because there are certainly other noise sources, the values reported in Figure 3 should be seen as a lower bound on the uncertainty. However, this figure can be used to estimate measurement uncertainty both for data that are already published, and also provides a type of power analysis for future studies. Given a
baseline firing rate and modulation depth, this figure provides an approximate number of trials necessary to achieve a desired precision in preferred direction estimates. For the purposes of this study, it is important to note that uncertainty in estimates of preferred direction can be of the same order of magnitude as the size of changes reported in previous papers -- changes that have been attributed to either experimental manipulations or instability.

Assessing tuning curve stability, taking uncertainty into account

While examples and power analysis illustrate how bootstrapping can be used to assess measurement uncertainty in the presence of simulated Poisson noise, bootstrapping is also a powerful technique for the analysis of actual data. Several recent studies have suggested that the preferred directions of neurons in primary motor cortex may be drifting rapidly over the course of tens of minutes (Carmena et al. 2005; Rokni et al. 2007). However, in describing changes in preferred directions, these studies did not explicitly account for measurement uncertainty. As noted above, estimated changes in preferred direction can easily be inflated, since they include effects due to both any actual changes and measurement noise. Here we used bootstrapping to assess the stability of primary motor cortical neurons during well-learned, center-out reaching.

To address these issues experimentally, we analyzed how the properties of recorded neurons in primary motor cortex from four animal subjects evolved over the course of a recording session. We recorded four datasets C (79 neurons), F (36 neurons), K (92 neurons), R (78 neurons), each consisting of at least 290 successful trials with sessions lasting 30-40 minutes. After collecting spike counts from 100 ms prior to movement onset to 300 ms following movement onset, 70.5% of the recorded neurons (201/285 total -- 38/79, 34/36, 72/92, and 57/78 in datasets C, F, K, and R, respectively) showed significant cosine tuning across all trials ($\alpha = 0.05$, Hotelling's $T^2$-test, see Supplementary Note). In the following sections, we limited our analysis to these significantly tuned neurons.

Following previous approaches (Rokni et al. 2007) we divided our dataset of center-out movements into blocks of trials. For each block, we estimated the neuron’s preferred direction, baseline firing rate, and modulation (see Methods), and determined the 95% confidence interval for the preferred direction estimates using bootstrapping. To compare these empirical results to the previous power analysis we first varied the number of trials included in each block (block size) and noted the average confidence interval size. As suggested by the power analysis, the size of the confidence interval decreases as the block size increases (Fig 3). In this case, where modulation is measured from the data, confidence interval size dropped approximately as 1/sqrt(number of trials). In close approximation to the power analysis estimates we found that in real data the average confidence interval size is 43.7 deg for blocks of 40 trials, 24.4 deg for blocks of 120 trials, and 16.3 deg for blocks of 240 trials. Simulated Poisson neurons that were stable and matched for tuning showed a very similar decrease in confidence interval size as a function of the number of trials (Fig 4, red curves).

---

Figure 4 about here

---

Keeping in mind that measurement uncertainty (confidence interval size) is related to the block-size, we then assessed stability in preferred direction and examined changes in tuning between blocks. Using bootstrapping analysis and examining successive blocks of 120 trials we found that many of neurons appear to be quite stable (Fig 5A). However, there were neurons in each of the four datasets that appeared to drift and have relatively large changes in preferred direction, modulation, and baseline firing between successive blocks (Fig 5B). To determine whether these changes are real effects or simply artifacts of high measurement uncertainty, it is important to consider whether the observed changes are statistically significant. Greater drift
in preferred direction is correlated with smaller baseline firing rates (Fig 5C) as well as smaller modulation (Fig 5D). In general, the neurons with larger apparent changes were the ones that also had large confidence intervals and high uncertainty about preferred direction (Fig 5E).

In all datasets we found that the average change in preferred direction from one block to the next was small. For a block size of 40 trials the mean change was -0.8±1.1 deg, and for blocks of 120 trials was 2.3±1.5 deg for all neurons. Despite the fact that the average PD change is close to 0, there was substantial variability in the changes, which, at first glance, appears to be an indication that preferred directions fluctuate (Fig 6). The standard deviation of PD changes was 39.0±2.2 deg for blocks of 40 trials and 22.9±3.2 deg for blocks of 120 trials. However, based on the bootstrapping analysis only a small number of these changes were significant (5.8% for 40 trials and 5.5% for 120 trials, bootstrapping test, \( \alpha = 0.05 \), not correcting for multiple comparisons). In fact, simulated Poisson neurons that were stable and matched for tuning had a distribution of changes from one block to the next that was almost identical to the distribution of observed PD changes. These results suggest that the majority of neurons are stable on timescales on the order of minutes or, if they are unstable, drift at a rate that is undetectable given the amount of measurement noise (<43.7 deg over 40 trials and <16.3 deg over 120 trials).

For the population, we can attempt to estimate the true variability in PD changes using a corrected measure of standard deviation that attributes the observed fluctuations to a combination of real changes in PD and measurement noise (see Methods). Using this method we found that the corrected standard deviation for blocks of 40 trials was 0.4±1.1 deg and for blocks of 120 trials was 1.8±2.5 deg, compatible with absolutely stable preferred directions. Our findings are thus consistent with the results from several past studies that incorporated enough data to reduce measurement errors (Chestek et al. 2007; Ganguly and Carmena 2009) and provide statistical insights into previous findings of apparent instabilities.

Explicitly modeling dynamic tuning curves with adaptive filtering

In the methods described above, tuning curves are assumed to be static within each block with changes occurring only between blocks. So far our results have suggested that any changes in preferred direction that do exist are below the level of measurement uncertainty (16.3 deg over 120 trials on average). These block-by-block methods gain statistical power by incorporating more trials, but, at the same time, by including so many trials, they average over any potential fluctuations that might exist. Ideally, we would like to use a method that explicitly models dynamic tuning curves that change from trial to trial. Adaptive filtering is one approach for modeling exactly this type of relationship, and several recent studies have used adaptive filtering to model dynamic place fields in hippocampal data (Brown et al. 2001; Eden et al. 2004). Here we use a basic version of adaptive filtering techniques to assess dynamic tuning curves in primary motor cortex.

Rather than look for differences in tuning between blocks of trials, adaptive filtering methods update the tuning curve parameters after each observation. At each time-step, the parameters are adjusted to reduce the size of prediction errors. Here we used the least-mean-squares algorithm (LMS) based on steepest-descent adaptive filtering. Larger errors result in larger parameter changes, and the size of the parameter updates is determined by a learning rate that is optimized separately. By adjusting the tuning parameters dynamically,
Adaptive filtering allows potential changes in tuning to be tracked from trial to trial. Here we assumed that only the preferred direction changes from trial-to-trial while the baseline firing rate and modulation are constant (see Methods for details).

Adaptive filtering is able to track changes in the preferred directions of simulated cosine tuned neurons. The analysis is built on the assumption that the process of spike generation is noisy and that neurons undergo both drift (fixed changes) and diffusion or fluctuation (random changes). In simulation, cosine tuned neurons with a constant drift in preferred direction were accurately tracked for a range of drift rates (Fig 7A). For a drift rate in the range ±2 deg/trial and no fluctuations, the estimated drift rate had an RMS error of only 0.07 deg/trial (Fig 7B). Fluctuations in preferred direction could also be tracked, albeit less accurately (Fig 7C). For fluctuations with a standard deviation between 0 and 20 deg/trial the estimated fluctuation had an RMS error of 2.5 deg/trial (Fig 7D). It is important to note that, while adaptive filtering excels at tracking large changes in the parameters, estimates of small fluctuations are substantially less accurate and sensitive to the learning rate. With an RMS error of 2.5 deg/trial it will be difficult to detect the 0.16 deg/trial variation (1.8±2.5 deg over 120 trials) suggested by the corrected standard deviation analysis above. However, these simulation results suggest that if neurons in primary motor cortex are truly unstable with large, random fluctuations in preferred direction, adaptive filtering methods may reveal this instability.

In actual data, adaptive filtering revealed some degree of fluctuation in the preferred directions of a subset of primary motor cortical neurons. Of the 201 significantly tuned neurons from our four datasets, 177 neurons were best fit by a stable rather than dynamic tuning curve. The remaining 24 neurons (~12%) showed some degree of fluctuation, with a median fluctuation of 2.1±0.3 deg/trial (Fig 8A). Assuming that preferred directions follow a Gaussian random walk, results from Rokni et al. as well as the uncorrected block-by-block analysis (Fig 6) suggest fluctuations on the order of 2.1 to 2.2 deg/trial on average (22.9±3.2 over 120 trials for our analysis and 29±3 deg over 160 trials (Rokni et al. 2007)). At first glance, this subset of neurons thus appears to be fluctuating consistent with previous results. However, even in this minority of neurons, it is possible that the observed fluctuations are false positives.

The fact that these estimates are close to the RMS error observed when estimating fluctuations in simulated data suggests that some care is needed in interpreting these values. The fluctuations found by adaptive filtering did improve spike prediction. The average log likelihood ratio for the dynamic tuning model was 8.66±1.6 bits/trial relative to a homogenous Poisson process, while the average log likelihood ratio for the static tuning model was 8.56±1.6 bits/trial relative to a homogenous Poisson process (both on training data). However, with any finite amount of data, adaptive filtering methods can find small false-positive fluctuations even for simulated, stable neurons.

We examined whether the observed fluctuations may have been false positives by again simulating stable cosine tuned neurons of varying modulation depth. In this case, after applying adaptive filtering, ~4-5% of the stable neurons are mistakenly identified as fluctuating. Using these false positives we constructed a null distribution and compared this distribution with the subset of observed fluctuating neurons (Fig 8B, grey). The fluctuations observed in 12% of the recorded neurons are consistent with the results from stable neurons, suggesting that they may indeed be false positives. The null-results suggest that, for typical physiologically realistic tuning curve parameters, fluctuations have to be rather large before they are detectable by adaptive filtering (~5 deg/trial).
DISCUSSION

We have presented results from two approaches aimed at detecting fluctuations in the preferred directions of cosine tuned neurons in primary motor cortex. Using bootstrapping on simulations of stable neurons we have quantified how measurement uncertainty or confidence interval size is affected by modulation depth and the amount of data available. In real data, bootstrapping allows us to estimate changes between blocks of trials and test for significance in a way that directly captures the measurement uncertainty. Finally, we used adaptive filtering techniques to model trial-by-trial changes in preferred direction explicitly. After comparing the results with those from simulated, stable neurons we find no evidence for large fluctuations in preferred direction using either bootstrapping or adaptive filtering. Small fluctuations in preferred direction may exist, but detecting these changes is difficult in the presence of spiking noise.

When examining the properties of neural discharge, the experimental manipulations as well as the statistical approaches used for analysis will influence the interpretation of the results. In this report, we have demonstrated that the uncertainty related to estimates of preferred direction is directly related to the total number of spikes as well as the neuron's modulation depth over different reach directions. The greater the modulation depth and the more data available (number of spikes), the more certain one can be about the tuning of an individual neuron. For typical experiments with limited amounts of data, the uncertainty about preferred direction can be quite large in comparison to the typical effects of a manipulation -- both on the order of a few tens of degrees. Measurement noise is thus of central importance when estimating how tuning curves change over time.

Using bootstrapping we can determine confidence intervals for each of the tuning curve parameters. For the observed neurons the confidence intervals for preferred direction were rather large: 16.3 deg over 120 trials on average. This value fundamentally limits how well we can detect changes in preferred direction. For instance, the standard deviation of changes in preferred direction between blocks of 120 trials was 22.9±3.2 deg -- the same order of magnitude as the confidence interval. Comparing these changes to those from stable, simulated neurons with matched tuning curves and using a corrected measure we find that the standard deviation of changes is likely closer to 1.8±2.5 deg over 120 trials. Adaptive filtering estimates PD changes in a small subset of neurons on the order of 2.1 deg/trial, but after comparing these results to stable simulated neurons we find that these changes are again consistent with stable preferred directions.

A previous study using large blocks of data reported similar stability of tuning parameters (Chestek et al. 2007). However, this study used free reaching rather than a manipulandum and analyzed neurons primarily from PMd, leading the authors to speculate that the differences between their results and those of Rokni et al. could be due to experimental design or the specific population of neurons. The experiments presented here are far closer in design to those by Rokni et al., and yet we find no significant PD changes after corrections for measurement noise. Several studies have examined the stability of neural activity in the context of brain-machine interfaces and observe that firing properties may be highly variable (Carmena et al. 2005) or relatively stable (Ganguly and Carmena 2009) depending on training. However, interpreting these results can be difficult, since it is often unclear how a given decoding scheme relates to the tuning properties of individual neurons. Variability in the decoded behavior tends to be much higher than during normal reaching, and brain-control itself may introduce certain constraints on movement coding.

As a practical matter in data analysis, anything that is not part of the model but affects neural signals is
considered noise. We considered only Poisson noise, which is compatible with the Fano factor typically observed in cortical recordings (Stevens and Zador 1998; Zacksenhouse et al. 2007). However, there are many other factors beyond the monkey's hand direction that potentially influence firing rates (Johnson et al. 2001), including other kinematic variables such as reach speed (Chestek et al. 2007) and limb posture (Caminiti et al. 1990; Scott and Kalaska 1997), as well as, added loads (Kalaska et al. 1989), cortical waves (Rubino et al. 2006), neuromodulator concentrations (Ahern et al. 2002), oxygen concentration (Jiang and Haddad 1994) and circadian rhythms (Barnes et al. 1977). Modeling these other sources of variability, whether they are observed (Paninski et al. 2003; Saleh et al. 2010; Truccolo et al. 2005; Wu and Hatsopoulos 2006) or unobserved (Kulkarni and Paninski 2007; Stevenson et al. 2010), would likely improve the estimates of PD stability. These uncontrolled sources of variability could inflate estimates of both real fluctuations and measurement uncertainty, and the results here thus provide only an upper bound for the instability of PDs of neurons in primary motor cortex. If other sources of measurement noise could be accounted for, the estimated changes in preferred direction may very well be even smaller.

While our results suggest that preferred directions are substantially more stable during normal reaching than some previous reports, there is also convincing evidence that the relationship between a given neuron's activity and hand direction does change over time. Tuning to hand direction changes on very short timescales due to changing kinematics and dynamics (Churchland and Shenoy 2007; Hatsopoulos et al. 2004; Sergio et al. 2005) as well as over longer timescales during sensorimotor learning (Jarosiewicz et al. 2008; Li et al. 2001; Paz and Vaadia 2004). While several studies have shown that tuning curves are sensitive to the measurement epoch and specific task constraints (Hamel-Pâquet et al. 2006; Sergio et al. 2005), here we focused on a fixed, specific portion of the reaches (100 ms prior to, through 300 ms after movement onset). This type of analysis ignores the short-timescale kinematics and dynamics of reaching, and is aimed to test whether tuning curves are stable on longer timescales during a well-learned task. Additionally, we have focused primarily on the stability and uncertainty in estimates of preferred direction. Both the modulation and baseline firing rate of cosine tuned neurons may show a higher degree of instability (Chestek et al. 2007).

The stability of neuronal properties is of central importance to many computational theories. If presynaptic neurons change rapidly then the motor system must either be redundant to the extent that fluctuations do not affect behavior (Rokni et al. 2007) or postsynaptic neurons must adapt to allow for stable movement and decision making. Many behavioral models (e.g. Cheng and Sabes 2007; Wei and Körding 2009) have suggested that learning is an ongoing process where errors are constantly being corrected even during apparently stable behavior. This might suggest that the fluctuations in tuning, if they exist, may actually be functional. Rather than being an artifact of redundancy in the cortical representation of movement, fluctuations may be a reflection of ongoing attempts to correct small reaching errors. While measurement noise makes it difficult to distinguish between these hypotheses, methods are being developed to isolate the effects of reach errors (Chase et al. 2010; Scheidt et al. 2000) and better understand redundancy in the motor system (Jarosiewicz et al. 2008). Statistical techniques that allow modeling of non-stationary data (Kim et al. 2006; Wu and Hatsopoulos 2008) as well as experimental techniques that allow for longer-term recordings (Dickey et al. 2009; Tolias et al. 2007) should both serve to reveal the more detailed structure of tuning curve dynamics during learning as well as stable reaching.
FIGURE LEGENDS

Fig 1: Measurement noise for a simulated neuron with known tuning. (A) Idealized cosine tuning function of a neuron with typical movement-related discharge (20 Hz) and a modulation (5 Hz). (B) 40 trials simulated during movement in 8 directions with Poisson spike noise. (C) Three example bootstrap samples from the initial set of observations in (B). Note that there is substantial variability in the estimated preferred direction (arrows). (D) Using many bootstrap samples we build a distribution of preferred directions that allows estimation of confidence intervals. In this case the 95% confidence interval spans approximately 40 degrees.

Fig 2: Measurement noise affects the ability to detect PD changes. (A) Cosine tuning curves for a pair of simulated neurons having preferred directions of 90 (black) and 135 (gray) degrees. (B) 40 trials simulated from each of these tuning curves during movement in 8 directions with Poisson spike noise. (C) Bootstrap distributions for the preferred direction based on the data in (B). Note that there is overlap between these distributions, and the differences between the two distributions allow us to assess whether there is a significant tuning difference between the two conditions.

Fig 3: Uncertainty in preferred direction as a function of the number of spikes and modulation depth of a simulated neuron. Uncertainty about the preferred direction decreases both with increasing number of observations and increasing modulation. (A) Contour plots illustrating the width of the 95% confidence intervals. Levels denote one-sided 95% confidence intervals in degrees. For instance, for zero observations or a modulation of zero the confidence intervals span ±180 deg. (B) Color plot of same values with representative examples from three published studies -- square: (Rokni et al. 2007), filled circle: (Kalaska et al. 1989), triangle: (Wise et al. 1998). The average confidence intervals for the data used here are shown as open circles for blocks of 40, 120, and 240 trials (from left to right).

Fig 4: Uncertainty in preferred direction with increasing numbers of trials for four sets of data collected from four different monkeys. Average size of the 95% confidence interval for populations of actual (black) and simulated neurons (red). Simulated data were generated by estimating stable, non-drifting tuning curves from the recorded neurons and simulating spiking with Poisson noise. The average number of spikes per neuron is shown to relate these results directly to Fig 3. Dashed lines denote ±SEM across neurons.

Fig 5: Changes in tuning across blocks of 120 trials. Tuning curves for the first (black) and second (red) blocks of 120 trials for the five most stable neurons (A) and the five least stable (B) from each of the four datasets. Error bars denote mean firing rate ±SEM, solid lines denote cosine fits. The size of the PD change is negatively correlated with the baseline firing rate (C) as well as modulation (D). Higher uncertainty in the estimate of preferred direction is correlated with larger apparent drift (E). These correlations suggest that estimated changes in preferred direction could be due to measurement noise.

Fig 6: Estimated changes in preferred direction. (A) Histograms of PD changes for each of the datasets separately and for all neurons for a block size of 40 trials. (B) PD changes for a block size of 120 trials. Black blocks denote changes that were significant at the 95% level by bootstrapping. Note that for blocks of 40 trials the variability of changes between blocks is substantially higher than for blocks of 120 trials, suggesting that measurement noise plays a large role in determining the magnitude of PD changes. Red curves denote the PD changes observed when simulating stable Poisson neurons that were matched to have the same tuning as the observed data.

Fig 7: Adaptive filtering for detecting drift and fluctuation. (A) True and estimated preferred direction for three simulated neurons with PDs drifting 0.9, 0.45, and 0.07 deg/trial. Dashed lines denote the true, underlying preferred direction. (B) Estimated PD drift as a function of the true drift for 512 simulation runs.
using data from 400 trials. (C) True and estimated preferred direction for three simulated neurons with no mean drift, fluctuating 9, 4.5, and 0.7 deg/trial. (D) Estimated PD fluctuation as a function of the true fluctuation for 512 simulation runs. For data from 400 trials the degree of fluctuation is estimated fairly well, but much less accurately than the mean drift.

Fig 8: Adaptive filtering for actual data. For this analysis we combined neurons from all four datasets (C, F, K, and R) and use only the significantly tuned neurons (n = 201). Only 12% (n = 24) of these neurons showed some degree of fluctuation. (A) Trial-by-trial PD changes estimated by adaptive filtering for the subset of neurons that appeared to be fluctuating. Note that 177 of the 201 neurons were better fit by a stable rather than a dynamic tuning curve. (B) Estimated fluctuation as a function of the modulation for the subset of fluctuating neurons (black) and simulated stable neurons (grey). The fluctuations revealed by adaptive filtering are consistent with false positives.


Scott SH, and Kalaska JF. Reaching movements with similar hand paths but different arm orientations. I.


Dataset C
n = 38

Dataset F
n = 34

Dataset K
n = 72

Dataset R
n = 57
Trials

PD [deg]

Trials

PD [deg]

Trials

PD [deg]
A

Observed Fluctuations

$n = 24$ (of 201)

B

Fluctuation [deg/trial]

Observed
False Positives (Simulated Stable)

Modulation [Hz]