



Published in final edited form as:

J Neural Eng. 2018 April 01; 15(2): 026020–. doi:10.1088/1741-2552/aaa365.

ERAASR: An algorithm for removing electrical stimulation artifacts from multielectrode array recordings

Daniel J. O'Shea^{1,2,*} and Krishna V. Shenoy^{1,2,3,4}

¹Neurosciences Program, Stanford University, Stanford, CA, U.S.A

²Departments of Electrical Engineering, Bioengineering, and Neurobiology, Stanford University, Stanford, CA, U.S.A

³Bio-X Program, Stanford Neurosciences Institute, Stanford University, Stanford, CA, U.S.A

⁴Howard Hughes Medical Institute, Stanford University, Stanford, CA, U.S.A

Abstract

Objective—Electrical stimulation is a widely used and effective tool in systems neuroscience, neural prosthetics, and clinical neurostimulation. However, electrical artifacts evoked by stimulation prevent the detection of spiking activity on nearby recording electrodes, which obscures the neural population response evoked by stimulation. We sought to develop a method to clean artifact-corrupted electrode signals recorded on multielectrode arrays in order to recover the underlying neural spiking activity.

Approach—We created ERAASR: an algorithm which performs Estimation and Removal of Array Artifacts via Sequential principal components Regression. This approach leverages the similar structure of artifact transients, but not spiking activity, across simultaneously recorded channels on the array, across pulses within a train, and across trials. The ERAASR algorithm requires no special hardware, imposes no requirements on the shape of the artifact or the multielectrode array geometry, and comprises sequential application of straightforward linear methods with intuitive parameters. The approach should be readily applicable to most datasets where stimulation does not saturate the recording amplifier.

Main results—The effectiveness of the algorithm is demonstrated in macaque dorsal premotor cortex using acute linear multielectrode array recordings and single electrode stimulation. Large electrical artifacts appeared on all channels during stimulation. After application of ERAASR, the cleaned signals were quiescent on channels with no spontaneous spiking activity, whereas spontaneously active channels exhibited evoked spikes which closely resembled spontaneously occurring spiking waveforms.

Significance—We hope that enabling simultaneous electrical stimulation and multielectrode array recording will help elucidate the causal links between neural activity and cognition and facilitate naturalistic sensory prostheses.

*Corresponding Author: djoshea@stanford.edu.

Keywords

Electrical stimulation; multielectrode array; neural populations; neural prosthetics; algorithm

1. Introduction

Electrical stimulation is a method of modulating neural activity that is widely used within neuroscience and neuroengineering, as well as for treatment of chronic neurological pathologies. Within neuroscience, microstimulation is used to probe the functional organization of neural circuits. Whereas electrophysiological recordings provide correlative insights that neural activity in a brain region appears related to a specific behavior, electrical stimulation can demonstrate causal contributions of brain regions to specific cognitive functions (Salzman et al., 1990). In particular, intracortical microstimulation (ICMS) has been used to identify neural networks underlying perception, attention, cognition, and movement (see Tehovnik, 1996; Cohen and Newsome, 2004; Clark et al., 2011; Histed et al., 2013, for a review). ICMS can also drive reliable percepts in humans and animal models, including somatosensory vibration (Romo et al., 2000) and visual phosphenes (Tehovnik and Slocum, 2007), which underscores its utility for delivering artificial sensation in visual (Tehovnik et al., 2009) and motor prosthetic systems (Berg et al., 2013; O'Doherty et al., 2011). In clinical contexts, stimulation, especially delivered to deeper brain structures, has demonstrated efficacy for the treatment of neurological and neuropsychiatric conditions (Wichmann and DeLong, 2006; Holtzheimer and Mayberg, 2011) and recently as an avenue for providing sensory information as will be needed for a range of emerging brain-machine interface systems (Flesher et al., 2016).

As with all perturbations to the nervous system, accurate interpretation of the results relies on a solid understanding of the effects of the perturbation made (e.g., Jazayeri and Afraz, 2017; Otchy et al., 2015). With lesion and pharmacology studies, understanding the size of the affected region, completeness of the intended effect, the timecourse of the effect and recovery, and the influence of compensatory or homeostatic adaptation mechanisms are all critical to drawing correct conclusions from the behavioral impairments observed. Analogously, with ICMS, it is essential to understand the effect that electrical stimulation has on neural activity in order to draw correct inferences from causal experiments. However, whereas the traditional conception of ICMS assumes dense, localized activation near the electrode tip (Stoney et al., 1968; Tehovnik, 1996; Tehovnik et al., 2006), recent evidence suggests that the net effect of ICMS may be significantly more complex and distributed than previously appreciated (e.g. Histed et al., 2009; Logothetis et al., 2010; Graziano et al., 2002; Butovas and Schwarz, 2003).

In the context of neural prosthetics, researchers aim to improve the fidelity of artificial perception with sophisticated spatiotemporal patterning (e.g., Dadarlat et al., 2015) and to optimize stimulation parameters to maximize therapeutic benefit. Obtaining a clear picture of the effects of particular stimulation paradigms will likely prove essential to this goal, enabling closed-loop tuning of stimulation parameters relative to a desired effect on the local neural activity.

The effects on neural firing patterns evoked by ICMS have been difficult to characterize because the electrical artifact induced by stimulation interferes with the electrophysiological recording equipment used to record neuronal responses (Ranck, 1975; Merrill et al., 2005). A variety of approaches have been developed to work around these artifacts, which can be organized into several categories. Online approaches employ special hardware within the signal recording pathway to remove the stimulation artifact online (e.g., Wichmann and Devergnas, 2011; Brown et al., 2008; Müller et al., 2012). While effective, hardware approaches can be expensive or challenging to scale to high channel count recording arrays and typically require very stable artifacts over time for proper removal. Offline approaches employ standard electrophysiology collection hardware to record neural signals and then estimate and subtract the artifact post hoc. The most common approach for artifact estimation is to simply interpolate across the artifact (O'Keefe et al., 2001; Montgomery et al., 2005) and/or to average over repeated stimulation epochs (Hashimoto et al., 2002; Montgomery et al., 2005; Klink et al., 2017) and subtract, exploiting the larger stereotypy of the stimulus artifact relative to variable neural responses. Others have employed curve fitting approaches to estimate the artifact, exploiting the differences in the shape of artifact and action potentials, thereby allowing for small variations in the observed artifact across trials (Wagenaar and Potter, 2002; Erez et al., 2010).

These previous approaches have two major shortcomings addressed in this work. First, these approaches only recover neural signals occurring after the stimulation pulse is completed, with a typical window of several milliseconds during which signal is discarded. However, ICMS is typically delivered in trains of stimulation pulses occurring at frequencies approaching several hundred pulses per second (e.g., Churchland and Shenoy, 2007; Klink et al., 2017). This failure to recover neural activity concomitant with the stimulation pulses renders a large fraction of the peri-stimulation signal unusable for spike detection and waveform discrimination. Secondly, prior work has focused on artifacts detected at a single electrode, not exploiting the similar structure of contaminating artifacts across multiple electrodes using a multielectrode array. A notable exception to both of these shortcomings was developed by Mena and colleagues (2016), where they addressed both of these limitations using a modern statistical framework to isolate electrical artifacts from neural spiking signals in multielectrode array recordings of the retina *in vitro*. However, this approach currently requires the availability of electrical images for each neuron's spike waveform on multiple surrounding electrodes. Unfortunately, for the multielectrode array configurations commonly employed in *in vivo* cortical electrophysiology, the electrical image of most neurons is confined to a single electrode. Additionally, since the set of recorded neurons is biased towards a highly active subset of neurons (Barth and Poulet, 2012), we anticipate that some neurons that are driven to spike by electrical stimulation may be quiet outside of the stimulation period, and thus omitted in the dictionary of available electrical images.

In this paper, we present a fast method for removing electrical stimulation artifacts from multielectrode array recordings. We term this approach ERAASR: Estimation and Removal of Array Artifacts via Sequential principal components Regression. ERAASR does not require any special hardware and can recover the full timecourse of neural signals during and after stimulation, requiring only that the electrical artifact does not saturate the amplifier

on the recording electrodes. Our algorithm exploits the similarity of artifacts across multiple electrodes, multiple pulses within a pulse train, and multiple trials of repeated stimulation. The ERAASR algorithm is structured as a sequence of cascaded, linear operations—PCA and linear regression—across each axis of the data tensor. We validate our method on neural signals collected during an arm-movement task with 24-channel linear multielectrode arrays (“V probes”, Plexon Inc.) in macaque dorsal premotor cortex, demonstrating that spiking waveforms extracted during the stimulation period closely resemble spontaneous spiking waveforms on each electrode.

2. Methods

2.1. Subjects

Animal protocols were approved by the Stanford University Institutional Animal Care and Use Committee. The subject was one adult male macaque monkey (*Macaca mulatta*), monkey P. After initial training, we performed a sterile surgery during which the macaque was implanted with a head restraint and a recording cylinder (NAN Instruments) which was located over left, caudal, dorsal premotor cortex (PMd). The cylinder was placed normal to the skull and secured with methyl methacrylate. A thin layer of methyl was also deposited atop the intact, exposed skull within the chamber. Before stimulation sessions began, a small craniotomy (3 mm diameter) was made under ketamine/xylazine anesthesia, targeting an area in PMd which responded during movements and palpation of the upper arm (17 mm anterior to interaural stereotaxic zero).

2.2. Reaching task

For the purposes of this study, we engaged the monkey in a reaching task in order to drive task-related neural activity in premotor cortex where our electrodes were located. We trained the monkey to use his right hand to grasp and translate a custom 3D printed handle (Shapeways, Inc.) attached to a haptic feedback device (Delta.3, Force Dimension, Inc.). The other arm was comfortably restrained at the monkey's side. The monkey was trained to perform a delayed reaching task by moving the haptic device cursor towards green rectangular targets displayed on the screen. Successful completion of each movement triggered a juice reward.

2.3. Electrophysiology and stimulation configuration

At the start of each experimental session, both a stimulation electrode and a recording probe were secured to two independently controllable, motorized micromanipulators (NAN Instruments). Both probes were lowered simultaneously through blunt, non-penetrating guide tubes into dorsal premotor cortex at 3 $\mu\text{m/s}$ to an approximate depth of 2 mm (Figure 1a). The stimulation electrode was a single tungsten microelectrode with approximately 1 M Ω impedance at 1 kHz (Part #UEWLGC-SECN1E, Frederick Haer Company). The recording probe was a linear electrode array consisting of 24 contact sites located at 100 μm spacing along the length of the shank (V-probe PLX-VP-24-15ED-100-SE-100-25(640)-CT-500, Plexon Inc.). The recording probe penetration site was located at an approximate distance of 0.75 mm to 1.5 mm from the stimulation site. As best possible, the stimulating

electrode was inserted at the same location and to the same depth, whereas we changed the location of the recording electrode for each recording session.

The stimulation was performed using a StimPulse electrical stimulation system used as a combined function generator and isolated current source (Frederick Haer Company). The electrical stimulation current flowed through the brain via the electrode to a “ground screw” located at the posterior pole of the implant whose tip was in contact with the dura below the skull. Microstimulation trains consisted of 20 biphasic pulses delivered at 333 Hz for 57 ms (150 μ s cathodic, 100 μ s pause, 150 μ s anodic). Stimulation amplitude was varied between 5 μ A to 40 μ A. Stimulation delivered at the start of most experimental sessions, while the arm was in a passive, resting position, triggered brief movements of the contralateral upper arm and shoulder at thresholds typically $> 140 \mu$ A. Stimulation onset was triggered via TTL pulse delivered by the task engine (Simulink Real-time Target, The Mathworks; NIDAQ digital to analog card, National Instruments). Stimulation was delivered on 20–40% of trials. Stimulation amplitude was fixed within each block of trials, where we collected at least 8–10 repetitions of each stimulation trial type for each block. Within each session, we typically began with blocks of lower amplitude stimulation before proceeding to higher amplitude stimulation.

The recording probe was connected to a 3-headed switching headstage, a component of the Blackrock StimSwitch system (Blackrock Microsystems). The probe connects to one bank of inputs, control lines from the StimSwitch control box connect to the second bank of inputs, and outgoing voltage signals leave via the third bank of outputs. During these experiments, this component was configured to act as a unity gain buffer (head stage) which simply relays the voltage signals to a shielded ribbon cable (Samtec Inc.) to the amplifier (Blackrock Microsystems). Each electrode is differentially amplified relative to a common reference line in the V-probe itself, which is also shorted to surrounding guide tube for better noise rejection.

Broadband voltages were recorded from all 24 electrode channels of the recording probe. We also recorded from the stimulation electrode during initial penetration to verify that we could see nearby neurons on the electrode. We then disconnected the stimulation electrode from the recording system and connected it to the StimPulse stimulator. Broadband signals were filtered at the amplifier (0.3 Hz 1 pole high-pass filter, 7.5 kHz 3 pole low-pass filter), digitized to 16 bit resolution over ± 8.196 mV (resolution = 0.25 μ V), and sampled at 30 kHz.

2.4. Unit selection

During experiments, we recorded any units or multi-units that could be isolated on the multielectrode array, without regard for their responsiveness or modulation by the task. Before analysis, we applied a screening procedure which looked only at non-stimulated trials in order to remove neurons that were very unreliable or very weakly modulated by the task. Briefly, we filtered non-stimulated trials spike trains with a 30 ms Gaussian window, then aligned trials separately from 100 ms pre-target onset to 70 ms post go cue, and then from 300 ms pre-movement to 600 ms post, then averaged within groups of trials with the same reach target. We defined each unit's “signal” to be the range of firing rates over time

and conditions, and “noise” to be the maximum standard error of the mean firing rate. We included units where the ratio of signal to noise exceeded at least 2, a total of 138 units in monkey P. Note that this selection process looked only at non-stimulation trials.

2.5. Artifact amplitude model

The amplitude of the electrical artifact was measured on each recording channel as the peak to peak voltage concomitant with the first stimulus pulse, averaged across trials. We then devised a model relating spatial location to artifact size that allowed us to pinpoint the location of the recording probe in each session relative to the stimulation source. Amplitude measurements were used to fit a model relating artifact amplitude to the distance of each site to the stimulation source. This model jointly optimizes parameters defining the distance-dependent amplitude scaling and the location of the recording probe in each session (relative to the stimulation source). We define x_s as the distance measured laterally from the probe to the closest point on the stimulating electrode; we define y_s as the vertical distance along the penetration path from the most superficial electrode on the probe to the tip of the stimulating electrode. The location of each electrode e in each session s relative to the probe location $(x, y)_s$ is given by

$$(x, y)_{e,s} = (x_s, y_s + e\Delta) \quad (1)$$

where $\Delta = 100 \mu\text{m}$ spacing between sites. This contact site is located at a distance

$$d_{e,s} = \sqrt{x_{e,s}^2 + y_{e,s}^2} \quad (2)$$

from the stimulation source which lies at the origin. We then assume that the amplitude of the artifact on each electrode for a specific stimulation amplitude I_a falls off as the reciprocal of this distance.

$$V_{e,s,a} = \frac{v_0 * I_a}{d_{e,s}} \quad (3)$$

where v_0 has units $\text{V}/(\mu\text{A mm})$. The parameters of this model are this single scaling parameter v_0 as well as a pair of location coordinates for the probe on each session. We then fit this model using constrained non-linear curve fitting (`lsqcurvefit` in Matlab) to minimize the squared error between empirical and modeled artifact amplitude across electrodes, sessions, and amplitudes. Confidence intervals of the parameters were obtained using the Jacobian of the fitted model.

We note that ERAASR makes no assumptions regarding the artifact amplitudes across electrode channels. This model is fit solely to demonstrate basic features of the evoked artifacts in our datasets and to infer the location of the recording probes.

2.6. Synthetic dataset generation

We generated synthetic datasets on which to test the performance of ERAASR when evoked spiking activity was highly reliable relative to the electrical stimulus. We applied ERAASR to real datasets and extracted the estimated artifacts for an individual channel. We then added a known, representative spiking waveform on top of this artifact at specific times during stimulation. These spikes were added to a subset of pulses at offsets drawn from a normal distribution with a fixed mean and specified standard deviation or jitter. The proportion of pulses with synthetic spikes and the jitter were jointly varied systematically. White noise was then added to the resultant traces to match that channel's RMS voltage. This synthetic channel then replaced the original channel, resulting in a synthetic dataset with 23 real channels and 1 synthetic channel for which "ground truth" spike times and the spiking waveform were known. We then applied ERAASR to the synthetic data tensor, followed by high pass filtering at 250 Hz and thresholding at 5x RMS voltage to extract spiking waveforms. The mean recovered waveform was then compared to the true spiking waveform filtered by the same high-pass filter.

2.7. Reduced channel count comparison

To probe the dependence of ERAASR performance on channel count, we applied ERAASR to subsets of channels selected from the original datasets. We first identified "spike-free" channels that had no spontaneous or evoked spiking activity, which were typically very superficial. For a given channel count, we randomly selected groups of channels of that size from the original 24, ensuring that at least one of the chosen channels was spike-free. For each spike-free channel, we computed the RMS voltage of the channel post cleaning and high-pass filtering, against the RMS value computed for that channel using non-stimulation trials passed through the same filter. This provided direct quantification of the effectiveness of ERAASR in removing corrupting artifact. We averaged this normalized RMS value over channels and over 10 random groups of channels for each channel count (except 24, where all channels were included).

3. The ERAASR algorithm

We describe the operation of the ERAASR algorithm as applied to a representative example dataset recorded in macaque PMd. Pseudocode is provided in Algorithm 1.

3.1. Precise temporal alignment and tensor construction

We began with the raw, unfiltered voltage traces for each stimulation trial, collected simultaneously on the 24 channels, sampled at 30 kilosamples/sec. We first corrected the small amount of jitter between triggering the stimulator to the beginning of the stimulation train, typically on the order of a few milliseconds. We began by precisely aligning the stimulation trials to each other in time. This process was performed separately for each group of trials sharing a common stimulation amplitude. We extracted the raw voltage data off of one electrode channel in a 60 ms window surrounding the triggering event, high-pass filter the signal with a fourth-order, 250 Hz corner frequency Butterworth filter, to remove slow drift. We then detected the onset of the artifact due to the first biphasic pulse on each trial using an appropriately low negative threshold, set to avoid spiking activity but reliably

detect artifacts at the lowest stimulation amplitude. We then realigned the voltage data for each trial to this threshold crossing event. An outlier detection procedure was performed to detect rare instances of stimulator failure, including trials where 19 or 21 pulses were delivered instead of 20. Outlier detection was performed by projecting the realigned, unfiltered voltage traces onto the first 10 principal components of that group, and detecting trials where the absolute value of the z-scored scores in any of the first 10 PCs exceeded a value of 5. Outlier trials (typically < 2% of trials) were excluded from subsequent analysis. The traces were then up-sampled 10x to 300 kHz using spline interpolation and the set of trials were precisely aligned in time. We manually verified the results of this procedure to ensure successful artifact removal and accurate alignment of the artifact pulse trains. The aligned traces are then resampled at the original 30 kHz sampling rate.

We constructed a data tensor for each group of trials sharing a common stimulation amplitude and extract a 60 ms window of the trial-aligned broadband voltage data across the trials, yielding a data tensor \mathbf{X}_{raw} with size C (number of channels) \times T (number of timepoints per pulse) \times P (number of pulses) \times R (number of trials). In our datasets $C = 24$ channels, $T = 90$ timepoints per pulse (3 ms at 30 kHz sampling), $P = 20$ pulses, and R was typically on the order of 100–200 trials. We lightly high-pass filtered the raw signals in time with a fourth-order, 10 Hz corner frequency Butterworth filter, which removed slow drifts from the traces and makes subsequent processing more robust. We first attempted to clean each channel by exploiting the similarities of the artifact across simultaneous channels (Figure 2a).

3.2. Removal of shared structure over channels

We then unfolded the data tensor into an $RTP \times C$ matrix \mathbf{M}^c (Figure 2b), and performed principal components analysis. This allows us to re-express each channel's response vector ($RTP \times 1$) as a linear combination of principal components, each $RTP \times 1$. As the artifact waveform is very large relative to the interesting spiking activity and shared across all channels, one would expect the top few principal components to capture preferentially shared variance due to the artifact. Empirically, we determined that the top $K_C = 4$ principal components (PCs) captured much of the artifact shape (Figure 2c). We then reconstructed each channel's response omitting the contribution from the artifact. The simplest method to accomplish this would be to reconstruct the channel's response by using a linear combination of all PCs except the first K_C . However, in some cases, especially when lower currents were used and the magnitude of the artifact relative to the spiking activity was smaller, we observed that some of the top PCs would begin to incorporate a small amount of spiking activity from individual channels. By reconstructing those channels from the remaining PCs, the spiking activity itself would be partially distorted because a fractional portion of those spikes would be subtracted along with the artifact. We addressed this issue by again exploiting the locality of spiking activity, by assuming that spiking activity from any one channel would not be present on other channels except the immediately adjacent channels. However, the artifact is shared on a much larger spatial scale and can be separated from the spiking activity. Therefore, for each channel c , we used the following procedure: reconstruct the top K_C PCs using all channels except c and its immediately adjacent neighbors $c - 1$ and $c + 1$, that is, using a modified version of the loading weights $w_{(k)}^C$ for

the k th principal component in which we set the weights for c ; $c-1$, $c+1$ to 0. More generally, we define parameter λ_C that dictates the number of adjacent channels excluded from the reconstruction. We refer to this vector of modified loading weights as $w_{(k),\overline{c}}^C$.

$$\left(w_{(k),\overline{c}}^C\right)_i = \begin{cases} 0 & \text{if } c - \lambda_C \leq i \leq c + \lambda_C \\ \left(w_{(k)}^C\right)_i & \text{else} \end{cases} \quad (4)$$

Using these loading weights, we reconstruct the matrix of the top K_C artifact-capturing PCs, where each column is given by

$$\left(A_c\right)_{(k)} = M^C w_{(k),\overline{c}}^C \quad k=1, \dots, K_C; c=1, \dots, C \quad (5)$$

We then regress the channel's response vector against these artifact components and subtract this reconstructed artifact from the channel's response:

$$\hat{M}_c^C = M_c^C - \left(A_c^T A_c\right)^{-1} \left(A_c^T\right) M_c^C \quad (6)$$

This cleaning across channels captured (Figure 2d) and removed (Figure 2e) much of the artifact from each pulse. However, when examining the responses on an individual channel across each of the 20 pulses superimposed, we observed consistent structure in the traces (Figure 2f).

3.3. Removal of shared structure over pulses

This structure indicated that there remained artifact structure that is unique to individual channels but shared among the pulses in each trial. To address this, we rearrange \hat{M}^C into a new matrix M^P with size $TRC \times P$, a set of responses to each individual pulse, concatenated over trials and channels (Figure 3a). We determined that the top $K_P = 2$ principal components (PCs) captured much of the artifact shape. Consequently, we used the same PC regression procedure to estimate and subtract the artifact on each pulse, omitting the contribution of the pulse under consideration and optionally adjacent pulses within λ_P . For our dataset, $\lambda_P = 0$ was used. If $w_{(k)}^P$ is a vector of P loading weights describing the k th principal component, and for each pulse p , $w_{(k),\overline{p}}^P$ is given by:

$$\left(w_{(k),\overline{p}}^P\right)_i = \begin{cases} 0 & \text{if } p - \lambda_P \leq i \leq p + \lambda_P \\ \left(w_{(k)}^P\right)_i & \text{else} \end{cases} \quad (7)$$

For each pulse, we reconstructed the matrix of the top K_P artifact-capturing PCs, whose columns are given by:

$$(A_p)_{(k)} = M^P w_{(k,p)}^P \quad k=1, \dots, K_p; p=1, \dots, P \quad (8)$$

We then regressed the channel's response vector against these artifact components and subtract this reconstructed artifact from the channel's response:

$$\hat{M}_p^P = M_p^P - (A_p^T A_p)^{-1} (A_p^T) M_p^P \quad (9)$$

This cleaning procedure over pulses captured much of the common artifact structure observed within each train (Figure 3b). It is also possible to perform this cleaning over pulses separately for each channel, following an approach like that described below for cleaning over trials.

We next examined the set of cleaned traces for individual channels and pulses in the train, over the full set of trials. We observed that there remained common artifactual structure in these traces, which suggests that there is artifact structure unique to that channel and pulse number but shared among many trials (Figure 3c). In most cases, this shared structure was similar among nearby trials but exhibited a slow drift in the shape of the artifacts over the experimental session.

3.4. Removal of shared structure across trials

To remove these residual artifacts, we again employed a principal components regression approach, exploiting the shared structure of the artifact over multiple trials. We performed this operation separately for each channel. We rearranged the cleaned \hat{M}^P into a set of C data matrices M_c^R for $c \in \{1, \dots, C\}$, each with size $TP \times R$ (Figure 4a). We determined that the top $K_R = 4$ PCs captured much of the artifact shape over trials. If $w_{(c,k)}^R$ is a vector of R loading weights describing the k th principal component for channel c , we defined:

$$\left(w_{(c,k),r}^R \right)_i = \begin{cases} 0 & \text{if } r - \lambda_R \leq i \leq r + \lambda_R \\ \left(w_{(c,k)}^R \right)_i & \text{else} \end{cases} \quad (10)$$

where λ_R adjacent trials were excluded from the reconstruction. For our dataset, $\lambda_R = 0$ was used.

For each trial, we reconstructed the matrix of the top K_R artifact-capturing PCs, whose columns are given by:

$$\left(A_{(c,r)} \right)_{(k)} = \left(M_c^R \right) \left(w_{(c,k),r}^R \right) \quad k=1, \dots, K_R; r=1, \dots, R; c=1, \dots, C \quad (11)$$

We then regressed the trial's response vector against these artifact components and subtract this reconstructed artifact from the trial's response:

$$\hat{M}_{(c,r)}^R = M_{(c,r)}^R - \left(A_{(c,r)}^T A_{(c,r)} \right)^{-1} \left(A_{(c,r)}^T \right) M_{(c,r)}^R \quad (12)$$

We then rearranged the set of cleaned \hat{M}_c^R back into a $C \times T \times P \times R$ tensor $\mathbf{X}_{\text{cleaned}}$. This cleaning procedure captured much of the residual artifact (Figure 4c). The cleaned signals in $\mathbf{X}_{\text{cleaned}}$ did not display features obviously indicative of residual stimulation artifact, but did on some channels contain readily detectable spiking activity (Figure 4c, bottom panel). We then spliced the cleaned data back into the full trials, adding appropriate offsets to preserve continuity at the first and last timepoints of the inserted segment.

3.5. Post-stimulation transient cleaning

Following stimulation offset, a slower transient was observed in all channels. These transients were highly similar among trials. Consequently, we employed a similar principal components regression procedure to remove these slow transients. We performed the following procedure for each stimulation amplitude separately. We began by aligning the trials to stimulation offset and extracting voltage data in a 30 ms time window beginning after stimulation end, yielding a tensor \mathbf{X}_{post} with size R by T by C . We then rearrange this tensor into a set of C data matrices M_c^{post} for $c \in \{1, \dots, C\}$, each with size $T \times R$. We determined that much of the transients' temporal structure was present in each trial for a given channel and that the top $K_{\text{post}} = 2$ principal components (PCs) of these M_c^{post} matrices captured much of the transient shape over trials. Consequently, we repeated the same principal components regression approach described in eqs. (10) to (12) to clean the post-stimulation transients over trials. We then rearranged the set of cleaned \hat{M}_c^{post} back into a R by T by C tensor $\mathbf{X}_{\text{post-cleaned}}$. The signals in $\mathbf{X}_{\text{post-cleaned}}$ were then reinserted into the full voltage traces for each stimulation trial.

3.6. Spike thresholding and sorting

The cleaned voltage traces, along with the original voltage traces on non-stimulated trials were then high-pass filtered using a fourth-order 250 Hz corner frequency Butterworth filter, as is typically done online before spike detection. We then thresholded each signal at 5x RMS voltage, using a greedy procedure in which the spike threshold crossings were detected in order of size (largest to smallest), and no further threshold crossings were considered within a lockout period extending 0.3 ms prior to and 1.0 ms after the current threshold crossing. We then hand-sorted the spiking waveforms using MKsort (Matthew Kaufman and Ripple, Inc., <https://github.com/ripple-neuro/mksort>).

4. Results

We delivered high-frequency (333 Hz) biphasic electrical stimulation in macaque dorsal premotor cortex through a tungsten microelectrode, while simultaneously recording neural spiking activity using a second 24-channel linear multielectrode array (Figure 1a). When a

stimulus train was delivered through the electrode, a highly stereotyped stimulation artifact appeared on all channels of the recording array (Figure 1b). Amplitudes could exceed ± 7500 μV which is nearly two orders of magnitude larger than spiking waveform amplitudes (~ 100 μV). The artifact traces resembled filtered versions of the original current stimulus; similar transients were evoked for each of the 20 biphasic pulses. These artifacts increased in amplitude with larger stimulation current and smoothly varied over the 24 channels of the recording probes (Figure 5a,b).

4.1. Array location inference from artifact size

We used these amplitude profiles over channels for each session to fit a model relating artifact amplitude and the reciprocal of the distance of each electrode to the stimulation source. The model jointly optimizes a scaling parameter of the amplitude-distance relationship as well as parameters specifying the location of the recording probe on each session (See Methods). We fit this model to data collected in 9 recording and stimulation sessions in macaque prefrontal cortex. Figure 5c shows the recorded artifact amplitude (normalized by stimulation current) under the model against the model prediction, in which artifact size falls off with distance from the stimulation site (squared correlation $R^2 = 0.91$). The fitted probe locations are plotted in Figure 5d; these closely aligned with the approximate locations noted during probe insertion. We note that ERAASR itself does not assume any particular relationship among the artifact sizes among electrode channels; instead, ERAASR uses PC regression to identify and remove shared structure over the array channels. The close fit of the model to the empirical artifact size is consistent with a distance-dependent attenuation of the electrical stimulus, but we expect that ERAASR could be applied to datasets with arbitrary shared artifact structure among channels.

4.2. Artifact removal via ERAASR

We applied ERAASR to remove the artifacts from these traces. The ERAASR algorithm leverages several features of these recording datasets. First, it assumes that the true neural signal is corrupted by an additive stimulation artifact, which depends critically on the assumption that no information is lost due to amplifier saturation during the stimulation period. For the 5 μA to 40 μA range of stimulation amplitudes used, the stimulation artifact produced did not saturate the amplifier's range (± 8196 μV) on any of the recording channels in any experimental session. Therefore, if the artifact shape can be properly estimated, it can be removed via simple subtraction. Second, because we recorded simultaneously on multiple, closely spaced channels, we have multiple simultaneous observations of the electrical stimulation artifact at different points in space, whereas the individual channels do not share the same spiking waveforms, especially if we exclude immediately adjacent channels. Third, the stimulation train contains 20 repeats of identical biphasic pulses, which produced highly similar artifact transients. Finally, we deliver the stimulation train over many trials, yielding similar artifacts on each repetition. Because spiking activity evoked during the stimulation period is unlikely to occur at precisely the same time relative to each pulse and on each trial, we can use these repeated artifact measurements to build an estimate of the artifact shape for subsequent removal.

Figure 6a shows a set of traces for a single trial across the array on one of the recording sessions before high-pass filtering is performed; Figure 6b shows the same data after high-pass filtering before spike detection. The raw, pre-cleaned voltage traces for this same trial are shown in Figure 1b, demonstrating that the vast majority of the artifact initially present has been removed. These high-pass filtered traces were then thresholded to extract spikes and then manually sorted into different units.

To evaluate the artifact removal process, we use prior knowledge about the multielectrode array channels. The most superficial (low numbered) and deepest (high numbered) channels were located above cortex and in white matter respectively, which we inferred from the signals observed on these channels, the density of neural signals on intermediate channels presumed to lie within cortex, and the known thickness of premotor cortex relative to the depth span of the probe. During non-stimulated trials, no detectable spiking waveforms were observed on these superficial and deep channels, therefore the presence of threshold crossings in the cleaned stimulation voltage traces on these channels would be surprising, and suggest that the cleaning procedure had failed to fully remove the stimulation artifact. However, as expected, the presence of evoked spiking activity during the stimulation window is limited to the intermediate channels, where spontaneous spiking activity was also observed (Figure 6). For closer inspection, we superimposed the raw voltage traces with cleaned, artifact-removed, high-pass filtered voltage traces for individual trials. Figure 7a–b shows such a comparison for the most superficial channel, where neither spontaneous nor evoked spiking activity was detectable. Figure 7c–d shows the same comparison for a channel with spontaneous spiking as well as robust evoked spiking activity.

Next, it is essential to ensure that threshold crossings detected during the stimulation period are likely due to real neural spiking activity, as opposed to residual transients due to stimulation artifact that persists through the cleaning procedure. We reasoned that threshold crossings created by residual artifact on a specific channel would not be expected to resemble the spontaneous spiking waveforms collected during non-stimulated trials on a per-trial basis. For each session, we generated a visual comparison of the spiking waveforms detected on non-stimulated trials with those detected within the stimulation window (putatively “evoked” spikes). These evoked spiking waveforms detected during the stimulation period were highly similar to the spiking waveforms detected during non-stimulation trials. Figure 8 demonstrates this similarity for a representative pair of experimental sessions. We note that these waveforms need not be identical, as evoked spiking activity from other neurons located further from the electrode could superimpose to corrupt the waveforms from nearby neurons, especially when this faraway spiking activity is highly synchronized by stimulation.

We can also quantify the amount of residual artifact directly, by again utilizing channels that displayed no spontaneous spiking activity. Figure 9a compares the RMS voltage of these individual channels sampled from the post-cleaning stimulation window against a time window taken before stimulation. This metric captures the amount of baseline noise (largely thermal noise) on each channel, thereby providing an expectation for the variance of a given channel if the entirety of the corrupting artifact were successfully removed. For low amplitudes, the RMS voltage is often slightly lower than the pre-stimulation RMS, which is

possible due to the greedy nature of the cleaning procedure. For larger amplitudes, the stimulation RMS remains acceptably close to pre-stimulation RMS. Figure 9b summarizes the similarity of spiking waveforms on each channel observed both spontaneously and during the stimulation window.

4.3. Effects of spike reliability and channel count

Fundamentally, ERAASR estimates electrical artifacts by exploiting the stereotypy of the artifact relative to variable spiking activity over channels, pulses, and trials. Over channels, this assumption is satisfied provided that the electrodes are spatially distributed enough to sample different populations of neurons; the regression step is performed so as to exclude channels sharing common spiking activity. Over pulses and trials, however, ERAASR relies on the assumption that evoked spiking activity is not perfectly reliable, such that spikes may be present on a subset of pulses with non-zero jitter in the pulse to spike latency. In order to better assess the distortion introduced for stably evoked activity that violates these assumptions, we performed a set of simulations. Here, we applied ERAASR to real datasets and extracted the estimated artifact on a single channel with no spiking activity. We then superimposed a known representative spiking waveform on this artifact at certain times, thereby generating synthetic datasets in which we varied the fraction of pulses with evoked spikes (the reliability) and the standard deviation of the spike timing relative to the previous pulse (the jitter). As expected, when spikes were inserted on nearly every pulse with a very narrow latency distribution, ERAASR recovered partially distorted versions of the true waveform (Figure 10a,b) and failed to recover an increasing proportion of the spike times (Figure 10c).

We also explored the dependency of channel count on ERAASR cleaning performance. Beginning with a real dataset, we applied ERAASR on subsets of channels with varying size. We then evaluated ERAASR performance by comparing the RMS of cleaned channels with no spiking activity, where the cleaned signal should be quiescent, against the RMS on those channels during non-stimulated trials. Figure 10d demonstrates that for these datasets, cleaning performance is poor and highly variable for low channel counts, but begins to saturate above 15 channels. However, additional channels would likely be more helpful in datasets where the channels exhibit more heterogeneity. Additionally, more channels may be beneficial when variability in the artifact over pulses and trials is increased, reducing the utility of these subsequent cleaning steps.

5. Discussion

We discuss the main features of our method for estimating and removing electrical stimulation artifacts in comparison with existing approaches, as well as limitations of our approach and possible improvements. We also highlight a set of interesting neuroscientific arenas where the ability to observe electrically perturbed neural activity might be particularly illuminating, underscoring the utility of artifact removal methods.

Here, we have applied the ERAASR algorithm to artifacts recorded on a linear multielectrode array (1D). However, we reiterate that ERAASR makes no assumptions regarding the electrode array geometry. The PCA regression step allows for artifacts on each

electrode to be described as linear combinations of individual, shared components. Consequently, we anticipate that ERAASR should be readily applicable to 2D and 3D arrays with arbitrary geometry, provided that the artifact remains within the amplifier linear regime on the channels of interest so that no information is lost.

We also note that variability in electrode impedance should also not degrade ERAASR performance. The algorithm does not assume any relation between the artifact amplitudes over channels. Changes in the artifact size due to drifting impedance over experimental sessions could simply be dealt with by applying ERAASR independently on each dataset. If the artifact size/shape is sufficiently unstable within a recording session, the cleaning step over trials will play a larger role in artifact removal. However, extensions of the algorithm leveraging a PC regression step in which the components and channel loading weights are allowed to vary smoothly in time across a session could be developed. Differences in the shape of the artifact could also result due to variability in impedance over channels. This transformation should be linear in nature and result in different frequency bands having electrode-specific gains and phase delays, which should be decomposable into individual components across channels. Therefore, ERAASR should deal gracefully (and automatically) with differences in impedance across electrodes, although our recordings did not provide sufficient variability to verify this. The artifact shapes in our dataset were largely similar across channels, and the remaining channel-specific structure was cleaned subsequently by exploiting shared structure across pulses and trials.

5.1. Comparison to other methods

ERAASR exploits similarities in the electrical artifact observed across multiple channels, pulses within a stimulus train, and trials with repeated delivery of the same stimulus. Of these, the shared structure of simultaneously recorded artifact across multiple channels was most useful. Most previously described attempts exploit only the shared structure across trials, typically using an averaging or moving average approach (Hashimoto et al., 2002; Montgomery et al., 2005; Klink et al., 2017) to estimate artifacts on an individual channel. However, these approaches are necessarily quite sensitive to variability in artifact shape or amplitude across trials. While this variability may be small in relative terms, the much larger amplitude of artifacts relative to neural spiking waveforms can create meaningfully large errors in the residual signal, requiring that several milliseconds of the signal in the vicinity of the stimulation pulse be discarded. Another common approach is to exploit differences in the shape of stimulation artifacts relative to spiking waveforms, by using curve-fitting algorithms to capture artifacts in the voltage signal but exclude spiking activity (Wagenaar and Potter, 2002; Erez et al., 2010). Similarly, these fits reliably capture artifact shape at a fixed delay from the stimulation pulse, but struggle to capture earlier portions of the transient (Erez et al., 2010), requiring peri-stimulus signal to be discarded. Motivated by the popularity of high-frequency pulse trains in ICMS experiments, our method addresses the problem of observing neural activity through the entire stimulation pulse without discarding signal.

Our algorithm combines a series of simple, intuitive, linear operations along the channels, pulses, and trials axes of the data tensor. The core operation is principal components

regression, in which PCA is used to identify a template describing shared structure across a given axis of the tensor, followed by a regression step in which the artifact on a given channel (or pulse, or trial) is reconstructed by excluding that channel (pulse, trial) and its neighbors. This approach mitigates the risk of removing spiking signals during the cleaning process, exploiting the local nature of spiking responses across channels and the inherent biological variability in spiking responses on each pulse and each trial. Furthermore, each of these stages is intuitive and the results easily understood. A small set of key parameters, including the number of principal components and the set of neighbors excluded from the reconstruction at each stage, can be adjusted to trade off between more aggressive cleaning and more veridical preservation of spiking waveform shape.

Our method operationally separates the process of artifact estimation from spike detection and extraction, which is performed as usual using typical high-pass filtering and thresholding on all voltage data after the cleaning process is complete. In contrast, a promising alternative method for multi-channel artifact removal is presented by Mena and colleagues (2016) for multielectrode array recordings of the retina. This approach employs a statistical model describing both the artifact and the spike generation process, jointly estimating artifact and spikes present in the voltage signals. This method exploits shared structure of the artifact across a local group of channels around the stimulation source, but also requires that the electrical image of each neuron (the shape of spiking waveforms over several nearby electrodes) be known as an input to the cleaning algorithm. While this approach is highly effective and appropriate in the context of retinal recordings, where individual neurons are recorded on many densely spaced electrodes, the electrical images of neurons in primate cortex using typical multielectrode arrays are often limited to one or two electrodes.

5.2. Limitations

First, our approach relies critically on the assumption that stimulation artifact linearly superimposes with neural spiking signals. This requirement is difficult to avoid; when the amplifier saturates, information about the spiking activity that would superimpose with the artifact is lost. Therefore, ERAASR's utility is limited to stimulation configurations and amplitudes which occupy the linear, non-saturated regime of the amplifier, or at least to timepoints within the stimulation period where the artifact is not saturating. Practically, this sets a minimum distance between the electrode array and the stimulation source for a given stimulation amplitude. This distance could be designed such that the return path of the electrical current steers the electric field so as to minimize the recorded artifact amplitude (Rattay and Resatz, 2004). This problem can also be effectively managed in hardware, employing special circuitry to estimate and partially cancel artifacts online (Wichmann and Devergnas, 2011; Brown et al., 2008; Müller et al., 2012). A hybrid approach employing multi-channel artifact removal circuitry to prevent saturation using a predictable transformation of the recorded signal, supplemented with *post hoc* artifact cleaning procedure like the one described here could be effective across a much larger range of stimulation amplitudes.

Second, our algorithm sequentially removes shared structure across channels, pulses, and trials. We accomplish this by reshaping the voltage data tensor into a matrix (or set of matrices) so that familiar methods like PCA can be employed. This problem of finding shared structure naturally lends itself to tensor decomposition (Kolda and Bader, 2009), which could identify shared structure jointly over each of these axes and potentially improve the artifact estimation. Tensors naturally arise in neuroscientific data collection contexts, and decomposition methods are becoming increasingly useful for identifying population structure (Seely et al., 2016; Elsayed and Cunningham, 2017; Williams et al., 2017).

Third, our algorithm operates by greedily removing any shared structure as artifact, which can inadvertently remove or distort spiking waveform signals as well. In our dataset, we did not observe this while removing shared structure across channels, as we explicitly excluded adjacent channels from the reconstruction process. However, removing shared structure across pulses and trials assumes that spiking responses evoked by a stimulation pulse are variable in time. Moreover, with increasing stimulation amplitude, the temporal precision of evoked spikes may increase (Ranck, 1975; Butovas and Schwarz, 2003), which would lead to waveform distortion due to overzealous reconstruction and removal of spiking signals as artifact. We experimented with using matched filters for robust spike detection while relaxing the artifact estimation technique (data not shown), though we did not explore whether this would be applicable at higher currents and more synchronous evoked spiking. In these circumstances, we expect that a joint estimation of artifact and spiking as proposed by Mena et al. (2016) could be adapted to more effectively recover highly regular evoked spiking by exploiting other differences in the structure of spikes from artifact.

Fourth, the computational complexity of our algorithm could likely be improved. For a dataset with 24 channels, 60 ms time windows at 30 kHz, and 150 trials, the full alignment and cleaning procedure can be performed in approximately 75 seconds on a laptop. The computational time complexity of the ERAASR algorithm as implemented is $\mathcal{O}(C^3 + R^3 + P^3 + C(RTP)^2 + P(RTC)^2 + R(TPC)^2)$ where C is channel count, R is trial count, P is pulse count, and T is timepoints per pulse. The algorithm comprises three rounds of PCA and linear regression on the rearranged artifact tensor. Faster, approximate versions of these algorithms exist and could be used to improve the computation time. To support online neural signal estimation, many of these steps could be precomputed and artifact subtraction applied to new observations one pulse at a time.

5.3. Motivation for direct observation of electrically perturbed neural activity

The initial conception of ICMS suggests that stimulation activates most neurons within a sphere surrounding the electrode tip (Stoney et al., 1968; Tehovnik, 1996; Tehovnik et al., 2006). This idea derives originally from the findings of Stoney et al. (1968), who cleverly side-stepped the issue of stimulation artifacts by using collisions of antidromic and orthodromic spikes in a dual-stimulation paradigm as an indirect measure of local neuronal activation. A large body of subsequent work (e.g. Ranck, 1975; Asanuma et al., 1976; Tehovnik, 1996; Rattay, 1999; Tehovnik and Slocum, 2007; Marcus et al., 1979; Nowak and Bullier, 1996; Nowak and Bullier, 1998b; Nowak and Bullier, 1998a; Kimmel and Moore, 2007) has demonstrated that an individual current pulse primarily evokes spikes at axons-in

particular, the excitable axon initial segment—and that the likelihood of evoking a spike in a neuron depends on distance, pulse duration, and current according to a relatively simple power law. The spatial localization of ICMS is consistent multiple behavioral studies as well (e.g. Murasugi et al., 1993; Tehovnik et al., 2004; Tehovnik et al., 2005; Bartlett et al., 2005; Tehovnik and Sommer, 1997). These studies derive an estimate of the effective activation volume and dependency on stimulation parameters by combining behavioral readouts with known features of cortical spatial organization. For example, Murasugi et al. (1993) found that low current ICMS delivered to monkey area MT could bias perceived dot motion direction, but that ICMS currents above 20 μA less effectively biased perception. The authors conclude that below this threshold current, electrical activation is confined to a single “directional” column in MT, about 0.2 mm in diameter (Albright et al., 1984).

However, recent evidence challenges this traditional view of localized, dense activation. Using two-photon calcium imaging, Histed et al. (2009) demonstrated that low-intensity ICMS activated a sparse collection of neurons distributed in a relatively large volume extending several hundred μm in mice, and as far as 4 mm away in cats with 10 μA stimulation, perhaps resulting from lateral axons extending multiple millimeters within layer 2/3 of visual cortex. By advancing the electrode 30 μm , the authors found that an entirely non-overlapping sparse set of cells was activated, suggesting that ICMS at low currents (10–25 μA) activates axonal processes within $\approx 15 \mu\text{m}$ of the electrode tip. The directly activated population is therefore potentially sparse, widely distributed, and stochastic due to a precise dependence on axonal proximity to the stimulating electrode.

Many mapping studies have focused on direct activation in response to a single isolated current pulse, whereas multiple pulses at high frequency are much more common in scientific and prosthetic experiments. When longer pulse trains are used, particularly at high currents, electrical stimulation might readily engage a large collection of neural circuits while driving temporally complex neural responses (Strick, 2002; Graziano et al., 2002). ICMS induces transsynaptic spiking in connected neurons (Stoney et al., 1968; Butovas and Schwarz, 2003), and as larger currents activate more neurons, the effect of transsynaptic activation may be enhanced due to the summation of many from this larger directly activated population. Cortically evoked spiking may also depend on subthreshold activity in neurons, which is continuously modulated by ongoing task or stimulus-evoked responses (Kara et al., 2002). Short-term synaptic depression commonly seen in cortex would also modulate the effectiveness of transsynaptically evoked spiking (Thomson and Deuchars, 1994; Stratford et al., 1996; Deisz and Prince, 1989; Varela et al., 1997).

Larger stimulation currents can recruit inhibitory interneurons and result in inhibition of cortical responses lasting hundreds of milliseconds after stimulation (Berman et al., 1991; Watanabe et al., 1966; Li et al., 1960; Chung and Ferster, 1998; Creutzfeldt et al., 1966; Kara et al., 2002; Masse and Cook, 2010; Seidemann et al., 2002; Butovas et al., 2006; Butovas and Schwarz, 2003). Additionally, in an experiment combining cortical stimulation with fMRI and extracellular recording, Logothetis et al. (2010) conclude that high-frequency ICMS disrupts normal information flow along cortico-cortical pathways, by recruiting strong inhibition that prevents the spread of activation beyond the regions whose direct afferents were stimulated. Consequently, a population of cells perturbed electrically may

undergo a complex, time-varying pattern of modulation that reflects circuit micro-architecture, local circuit dynamics, short-term synaptic plasticity, recruitment of inhibition, and an interaction with ongoing task-related activity.

With this evident complexity, directly recording the effect of ICMS on neural activity will likely clarify the circuit mechanisms underlying intriguing aspects of the interaction between microstimulation and behavior. An accurate picture of this interaction could address many longstanding questions, such as how electrically evoked eye movements interact with visual guided eye movements (Sparks and Mays, 1983), how precise timing of electrical pulses modulates saccadic effects (Kimmel and Moore, 2007), how perceptual decisions are biased by stimulation of LIP (Hanks et al., 2006), how motor preparation recovers subsequent to electrical disruption (Churchland and Shenoy, 2007; Shenoy et al., 2011), and how motor cortical stimulation can evoke complex movements converging on a specific endpoint (Graziano et al., 2011) and effectively nullify the contribution of goal-directed behavior in the motor cortical output (Griffin et al., 2011). A detailed comparison with the effect of optogenetic stimulation could clarify the differences (Gerits and Vanduffel, 2013; Diester et al., 2011; Ohayon et al., 2013) and similarities (Dai et al., 2013) observed between the two modalities' effects on behavior. Lastly, a detailed model of microstimulation's interaction with cortical dynamics could facilitate delivery of more reliable or realistic artificial sensory perception in visual, auditory, and motor prostheses (Otto et al., 2005; Tehovnik et al., 2009; O'Doherty et al., 2011; Bensmaia and Miller, 2014; Dadarlat et al., 2015).

6. Conclusion

We developed ERAASR, an algorithm for estimating and removing artifacts caused by electrical stimulation on multielectrode array experiments. ERAASR assumes that artifact is shared over many channels and that evoked transients are highly repeatable across pulses and trials, whereas spiking activity is highly local and temporally jittered. Shared structure in the voltage signals is identified and removed sequentially across channels, across pulses in a stimulus train, and across trials, using straightforward linear methods. We believe our technique will be useful to neuroscientists in drawing precise causal links between perturbations and the effect of stimulation on neural activity and behavior and aid in the design and implementation of enhanced neural prosthetic systems capable of restoring lost sensation and facilitating precise control.

7. Code availability

Full Matlab code for ERAASR is available at <https://github.com/djoshea/eraasr> along with a link to a representative example dataset.

Acknowledgments

We thank Tirin Moore, Rob Franklin, and Paul Venable for helpful discussions. We thank Peiley Lau for devising the ERAASR acronym. This article was supported by the following grants: Christopher and Dana Reeve Paralysis Foundation, Burroughs Wellcome Fund Career Awards in the Biomedical Sciences., Defense Advanced Research Projects Agency Reorganization and Plasticity to Accelerate Injury Recovery N66001-10-C-2010, US National Institutes of Health Institute of Neurological Disorders and Stroke Transformative Research Award R01NS076460, US National Institutes of Health Director's Pioneer Award 8DP1HD075623-04, US National Institutes of Health Director's Transformative Research Award (TR01) from the NIMH 5R01MH09964703, Defense Advanced

Research Projects Agency NeuroFAST award from BTO W911NF-14-2-0013, Howard Hughes Medical Institute, NSF GRFP, and NSF IGERT.

References

- Albright TD, Desimone R, Gross CG. Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of neurophysiology*. 1984; 51(1):16–31. [PubMed: 6693933]
- Asanuma H, Arnold A, Zarzecki P. Further study on the excitation of pyramidal tract cells by intracortical microstimulation. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale*. 1976; 26(5):443–461. [PubMed: 1010001]
- Barth, Alison L., Poulet, James Fa. Experimental evidence for sparse firing in the neocortex. *Trends in neurosciences*. 2012; 35(6):345–355. DOI: 10.1016/j.tins.2012.03.008 [PubMed: 22579264]
- Bartlett, John R., DeYoe, Edgar A., Doty, Robert W., Lee, Barry B., Lewine, Jeffrey D., Negrão, Nubio, Overman, William H, Jr. Psychophysics of electrical stimulation of striate cortex in macaques. *Journal of neurophysiology*. 2005; 94(5):3430–3442. DOI: 10.1152/jn.00406.2005 [PubMed: 16079195]
- Bensmaia, Sliman J., Miller, Lee E. Restoring sensorimotor function through intracortical interfaces: progress and looming challenges. *Nature reviews Neuroscience*. 2014; 15(5):313–325. DOI: 10.1038/nrn3724 [PubMed: 24739786]
- Berg JA, Dammann JF 3rd, Tenore FV, Tabot GA, Boback JL, Manfredi LR, Peterson ML, Katyal KD, Johannes MS, Makhlin A, Wilcox R, Franklin RK, Vogelstein RJ, Hatsopoulos NG, Bensmaia SJ. Behavioral demonstration of a somatosensory neuroprosthesis. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*. 2013; 21(3):500–507. DOI: 10.1109/TNSRE.2013.2244616
- Berman NJ, Douglas RJ, Martin KA, Whitteridge D. Mechanisms of inhibition in cat visual cortex. *The Journal of physiology*. 1991; 440:697–722. DOI: 10.1113/jphysiol.1991.sp018731 [PubMed: 1804983]
- Brown EA, Ross JD, Blum RA, Nam Y, Wheeler BC, DeWeerth SP. Stimulus-Artifact Elimination in a Multi-Electrode System. *IEEE transactions on biomedical circuits and systems*. 2008; 2(1):10–21. DOI: 10.1109/TBCAS.2008.918285 [PubMed: 23852629]
- Butovas, Sergejus, Hormuzdi, Sheriar G., Monyer, Hannah, Schwarz, Cornelius. Effects of Electrically Coupled Inhibitory Networks on Local Neuronal Responses to Intracortical Microstimulation. *Journal of neurophysiology*. 2006; 96(3):1227–1236. DOI: 10.1152/jn.01170.2005 [PubMed: 16837655]
- Butovas, Sergejus, Schwarz, Cornelius. Spatiotemporal effects of microstimulation in rat neocortex: a parametric study using multielectrode recordings. *Journal of neurophysiology*. 2003; 90(5):3024–3039. DOI: 10.1152/jn.00245.2003 [PubMed: 12878710]
- Chung S, Ferster D. Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. *Neuron*. 1998; 20(6):1177–1189. [PubMed: 9655505]
- Churchland, Mark M., Shenoy, Krishna V. Delay of movement caused by disruption of cortical preparatory activity. *Journal of neurophysiology*. 2007; 97(1):348–359. DOI: 10.1152/jn.00808.2006 [PubMed: 17005608]
- Clark, Kelsey L., Armstrong, Katherine M., Moore, Tirin. Probing neural circuitry and function with electrical microstimulation. *Proceedings of the Royal Society of London B: Biological Sciences*. 2011; 278(1709):1121–1130. DOI: 10.1098/rspb.2010.2211
- Cohen, Marlene R., Newsome, William T. What electrical microstimulation has revealed about the neural basis of cognition. *Current opinion in neurobiology*. 2004; 14(2):169–177. DOI: 10.1016/j.conb.2004.03.016 [PubMed: 15082321]
- Creutzfeldt, Otto D., Watanabe, Satoru, Lux, Hans D. Relations between EEG phenomena and potentials of single cortical cells. I. Evoked responses after thalamic and epicortical stimulation. *Electroencephalography and clinical neurophysiology*. 1966; 20(1):1–18. DOI: 10.1016/0013-4694(66)90136-2 [PubMed: 4161317]
- Dadarlat, Maria C., O'Doherty, Joseph E., Sabes, Philip N. A learning-based approach to artificial sensory feedback leads to optimal integration. *Nature neuroscience*. 2015; 18(1):138–144. DOI: 10.1038/nn.3883 [PubMed: 25420067]

- Dai, Ji, Brooks, Daniel I., Sheinberg, David L. Optogenetic and Electrical Microstimulation Systematically Bias Visuospatial Choice in Primates. *Current biology: CB*. 2013; 24(1):63–69. DOI: 10.1016/j.cub.2013.11.011 [PubMed: 24332543]
- Deisz RA, Prince DA. Frequency-dependent depression of inhibition in guinea-pig neocortex in vitro by GABAB receptor feedback on GABA release. *The Journal of physiology*. 1989; 412:513–541. [PubMed: 2557431]
- Diester, Ilka, Kaufman, Matthew T., Mogri, Murtaza, Pashaie, Ramin, Goo, Werapong, Yizhar, Ofer, Ramakrishnan, Charu, Deisseroth, Karl, Shenoy, Krishna V. An optogenetic toolbox designed for primates. *Nature neuroscience*. 2011; 14(3):387–397. DOI: 10.1038/nn.2749 [PubMed: 21278729]
- Elsayed, Gamaleldin F., Cunningham, John P. Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nature neuroscience*. 2017; 20(9):1310–1318. DOI: 10.1038/nn.4617 [PubMed: 28783140]
- Erez, Yaara, Tischler, Hadass, Moran, Anan, Bar-Gad, Izhar. Generalized framework for stimulus artifact removal. *Journal of neuroscience methods*. 2010; 191(1):45–59. DOI: 10.1016/j.jneumeth.2010.06.005 [PubMed: 20542059]
- Flesher, Sharlene N., Collinger, Jennifer L., Foldes, Stephen T., Weiss, Jeffrey M., Downey, John E., Tyler-Kabara, Elizabeth C., Bensmaia, Sliman J., Schwartz, Andrew B., Boninger, Michael L., Gaunt, Robert A. Intracortical microstimulation of human somatosensory cortex. *Science translational medicine*. 2016; 8(361):361ra141. doi: 10.1126/scitranslmed.aaf8083
- Gerits, Annelies, Vanduffel, Wim. Optogenetics in primates: a shining future? In: *Trends in genetics: TIG*. 2013; 29(7):403–411. DOI: 10.1016/j.tig.2013.03.004 [PubMed: 23623742]
- Graziano, Michael S A., Aflalo, Tyson N S., Cooke, Dylan F., Griffin, Darcy M., Hudson, Heather M., Belhaj-saïf, Abderraouf, Cheney, Paul D., Sponberg, Simon, Spence, Andrew J., Mullens, Chris H., Full, Robert J., Graziano, Michael S A., Aflalo, Tyson N S., Cooke, Dylan F. Arm Movements Evoked by Electrical Stimulation in the Motor Cortex of Monkeys Arm Movements Evoked by Electrical Stimulation in the Motor Cortex of Monkeys. 2011 Aug.:4209–4223. 2005. DOI: 10.1152/jn.01303.2004
- Graziano, Msa, Taylor, Csr, Moore, Tirin. Complex Movements Evoked by Microstimulation of Precentral Cortex. *Neuron*. 2002; 34:841–851. [PubMed: 12062029]
- Griffin DM, Hudson HM, Belhaj-Saif A, Cheney PD. Hijacking Cortical Motor Output with Repetitive Microstimulation. *Journal of Neuroscience*. 2011; 31(37):13088–13096. DOI: 10.1523/JNEUROSCI.6322-10.2011 [PubMed: 21917792]
- Hanks, Timothy D., Ditterich, Jochen, Shadlen, Michael N. Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature neuroscience*. 2006; 9(5):682–689. DOI: 10.1038/nn1683 [PubMed: 16604069]
- Hashimoto, Takao, Elder, Christopher M., Vitek, Jerrold L. A template subtraction method for stimulus artifact removal in high-frequency deep brain stimulation. *Journal of neuroscience methods*. 2002; 113(2):181–186. DOI: 10.1016/S0165-0270(01)00491-5 [PubMed: 11772439]
- Histed, Mark H., Bonin, Vincent, Reid, R Clay. Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. *Neuron*. 2009; 63(4):508–522. DOI: 10.1016/j.neuron.2009.07.016 [PubMed: 19709632]
- Histed, Mark H., Ni, Amy M., Maunsell, John H R. Insights into cortical mechanisms of behavior from microstimulation experiments. *Progress in neurobiology*. 2013; 103:115–130. DOI: 10.1016/j.pneurobio.2012.01.006 [PubMed: 22307059]
- Holtzheimer, Paul E., Mayberg, Helen S. Deep Brain Stimulation for Psychiatric Disorders. *Annual review of neuroscience*. 2011; 34(1):289–307. DOI: 10.1146/annurev-neuro-061010-113638
- Jazayeri, Mehrdad, Afraz, Arash. Navigating the Neural Space in Search of the Neural Code. *Neuron*. 2017; 93(5):1003–1014. DOI: 10.1016/j.neuron.2017.02.019 [PubMed: 28279349]
- Kara, Prakash, Pezaris, John S., Yurgenson, Sergey, Reid, R Clay. The spatial receptive field of thalamic inputs to single cortical simple cells revealed by the interaction of visual and electrical stimulation. *Proceedings of the National Academy of Sciences of the United States of America*. 2002; 99(25):16261–16266. DOI: 10.1073/pnas.242625499 [PubMed: 12461179]

- Kimmel, Daniel L., Moore, Tirin. Temporal patterning of saccadic eye movement signals. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2007; 27(29):7619–7630. DOI: 10.1523/JNEUROSCI.0386-07.2007 [PubMed: 17634356]
- Klink, P Christiaan, Dagnino, Bruno, Gariel-Mathis, Marie-Alice, Roelfsema, Pieter R. Distinct Feedforward and Feedback Effects of Microstimulation in Visual Cortex Reveal Neural Mechanisms of Texture Segregation. *Neuron*. 2017; 95(1):209–220.e3. DOI: 10.1016/j.neuron.2017.05.033 [PubMed: 28625487]
- Kolda T, Bader B. Tensor Decompositions and Applications. *SIAM Review*. 2009; 51(3):455–500. DOI: 10.1137/07070111X
- Li CL, Ortiz-Galvin A, Chou SN, Howard SY. Cortical intracellular potentials in response to stimulation to lateral geniculate body. *Journal of neurophysiology*. 1960; 23:592–601. [PubMed: 13761815]
- Logothetis, Nikos K., Augath, Mark, Murayama, Yusuke, Rauch, Alexander, Sultan, Fahad, Goense, Jozien, Oeltermann, Axel, Merkle, Hellmut. The effects of electrical microstimulation on cortical signal propagation. *Nature neuroscience*. 2010; 13(10):1283–1291. DOI: 10.1038/nn.2631 [PubMed: 20818384]
- Marcus S, Zarzecki P, Asanuma H. An estimate of effective spread of stimulating current. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale*. 1979
- Masse, Nicolas Y., Cook, Erik P. Behavioral time course of microstimulation in cortical area MT. *Journal of neurophysiology*. 2010; 103(1):334–345. DOI: 10.1152/jn.91022.2008 [PubMed: 19864437]
- Mena, Gonzalo E., Grosberg, Lauren E., Hottowy, Pawel, Litke, Alan, Cunningham, John, Chichilnisky, EJ., Paninski, Liam. Electrical Stimulus Artifact Cancellation and Neural Spike Detection on Large Multi-Electrode Arrays.” en. 2016
- Merrill, Daniel R., Bikson, Marom, Jefferys, John G R. Electrical stimulation of excitable tissue: design of efficacious and safe protocols. *Journal of neuroscience methods*. 2005; 141(2):171–198. DOI: 10.1016/j.jneumeth.2004.10.020 [PubMed: 15661300]
- Montgomery, Erwin B, Gale John T, Huang He. Methods for isolating extracellular action potentials and removing stimulus artifacts from microelectrode recordings of neurons requiring minimal operator intervention. *Journal of neuroscience methods*. 2005; 144(1):107–125. DOI: 10.1016/j.jneumeth.2004.10.017 [PubMed: 15848245]
- Müller, Jan, Bakkum, Douglas J., Hierlemann, Andreas. Sub-millisecond closed-loop feedback stimulation between arbitrary sets of individual neurons. *Frontiers in neural circuits*. 2012; 6:121.doi: 10.3389/fncir.2012.00121 [PubMed: 23335887]
- Murasugi CM, Salzman CD, Newsome WT. Microstimulation in visual area MT: effects of varying pulse amplitude and frequency. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 1993; 13(4):1719–1729. [PubMed: 8463847]
- Nowak LG, Bullier J. Spread of stimulating current in the cortical grey matter of rat visual cortex studied on a new in vitro slice preparation. *Journal of neuroscience methods*. 1996; 67(2):237–248. [PubMed: 8872891]
- Nowak LG, Bullier J. Axons, but not cell bodies, are activated by electrical stimulation in cortical gray matter. I. Evidence from chronaxie measurements. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale*. 1998a; 118(4):477–488. [PubMed: 9504843]
- Nowak LG, Bullier J. Axons, but not cell bodies, are activated by electrical stimulation in cortical gray matter. II. Evidence from selective inactivation of cell bodies and axon initial segments. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale*. 1998b; 118(4):489–500. [PubMed: 9504844]
- O'Doherty, Joseph E., Lebedev, Mikhail A., Ifft, Peter J., Zhuang, Katie Z., Shokur, Solaiman, Bleuler, Hannes, Nicolelis, Miguel A L. Active tactile exploration using a brain-machine-brain interface. *Nature*. 2011; 479(7372):228–231. DOI: 10.1038/nature10489 [PubMed: 21976021]
- O'Doherty, Joseph E., Lebedev, Mikhail a, Ifft, Peter J., Zhuang, Katie Z., Shokur, Solaiman, Bleuler, Hannes, Nicolelis, Miguel a L. Active tactile exploration using a brain-machine-brain interface. *Nature*, pp. 2011; :1–5. DOI: 10.1038/nature10489

- Ohayon, Shay, Grimaldi, Piercesare, Schweers, Nicole, Tsao, Doris Y. Saccade modulation by optical and electrical stimulation in the macaque frontal eye field. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2013; 33(42):16684–16697. DOI: 10.1523/JNEUROSCI.2675-13.2013 [PubMed: 24133271]
- O'Keefe DT, Lyons GM, Donnelly AE, Byrne CA. Stimulus artifact removal using a software-based two-stage peak detection algorithm. *Journal of neuroscience methods*. 2001; 109(2):137–145. DOI: 10.1016/S0165-0270(01)00407-1 [PubMed: 11513948]
- Otchy, Timothy M., Wolff, Steffen B E., Rhee, Juliana Y., Pehlevan, Cengiz, Kawai, Risa, Kempf, Alexandre, Gobes, Sharon M H., Ölveczky, Bence P. Acute off-target effects of neural circuit manipulations. *Nature*. 2015; 528(7582):358–363. DOI: 10.1038/nature16442 [PubMed: 26649821]
- Otto, Kevin J., Rousche, Patrick J., Kipke, Daryl R. Microstimulation in auditory cortex provides a substrate for detailed behaviors. *Hearing research*. 2005; 210(1–2):112–117. DOI: 10.1016/j.heares.2005.08.004 [PubMed: 16209915]
- Ranck JB Jr. Which elements are excited in electrical stimulation of mammalian central nervous system: a review. *Brain research*. 1975; 98(3):417–440. [PubMed: 1102064]
- Rattay F. The basic mechanism for the electrical stimulation of the nervous system. *Neuroscience*. 1999; 89(2):335–346. [PubMed: 10077317]
- Rattay, Frank, Resatz, Susanne. Effective electrode configuration for selective stimulation with inner eye prostheses. *IEEE transactions on bio-medical engineering*. 2004; 51(9):1659–1664. DOI: 10.1109/TBME.2004.828044 [PubMed: 15376514]
- Romo R, Hernández A, Zainos A, Brody CD, Lemus L. Sensing without touching: psychophysical performance based on cortical microstimulation. *Neuron*. 2000; 26(1):273–278. [PubMed: 10798410]
- Salzman CD, Britten Kenneth H, Newsome William T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*. 1990; 346:174–177. [PubMed: 2366872]
- Seely, Jeffrey S., Kaufman, Matthew T., Ryu, Stephen I., Shenoy, Krishna V., Cunningham, John P., Churchland, Mark M. Tensor Analysis Reveals Distinct Population Structure that Parallels the Different Computational Roles of Areas M1 and V1. *PLoS computational biology*. 2016; 12(11):e1005164.doi: 10.1371/journal.pcbi.1005164 [PubMed: 27814353]
- Seidemann, Eyal, Arieli, Amos, Grinvald, Amiram, Slovin, Hamutal. Dynamics of Depolarization and Hyperpolarization in the Frontal Cortex and Saccade Goal. *Science*. 2002; 295(5556):862–865. DOI: 10.1126/science.1066641 [PubMed: 11823644]
- Shenoy, Krishna V., Kaufman, Matthew T., Sahani, Maneesh, Churchland, Mark M. A dynamical systems view of motor preparation Implications for neural prosthetic system design. *Progress in brain research*. 2011; 192:33–58. DOI: 10.1016/B978-0-444-53355-5.00003-8 [PubMed: 21763517]
- Sparks DL, Mays LE. Spatial localization of saccade targets. I. Compensation for stimulation-induced perturbations in eye position. *Journal of neurophysiology*. 1983; 49(1):45–63. [PubMed: 6827303]
- Stoney SD Jr, Thompson WD, Asanuma H. Excitation of pyramidal tract cells by intracortical microstimulation: effective extent of stimulating current. *Journal of neurophysiology*. 1968; 31(5):659–669. [PubMed: 5711137]
- Stratford KJ, Tarczy-Hornoch K, Martin KA, Bannister NJ, Jack JJ. Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. *Nature*. 1996; 382(6588):258–261. DOI: 10.1038/382258a0 [PubMed: 8717041]
- Strick, Peter L. Stimulating research on motor cortex. *Nature neuroscience*. 2002; 5(8):4–5. [PubMed: 11753409]
- Tehovnik EJ. Electrical stimulation of neural tissue to evoke behavioral responses. *Journal of Neuroscience Methods*. 1996; 65(1):1–17.
- Tehovnik EJ, Sommer MA. Electrically evoked saccades from the dorsomedial frontal cortex and frontal eye fields: a parametric evaluation reveals differences between areas. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale*. 1997; 117(3):369–378. [PubMed: 9438704]

- Tehovnik EJ, Tolias AS, Sultan F, Slocum WM, Logothetis NK. Direct and indirect activation of cortical neurons by electrical microstimulation. *Journal of neurophysiology*. 2006; 96(2):512–521. DOI: 10.1152/jn.00126.2006 [PubMed: 16835359]
- Tehovnik, Edward J., Slocum, Warren M. Phosphene induction by microstimulation of macaque V1. *Brain research reviews*. 2007; 53(2):337–343. DOI: 10.1016/j.brainresrev.2006.11.001 [PubMed: 17173976]
- Tehovnik, Edward J., Slocum, Warren M., Schiller, Peter H. Microstimulation of V1 delays the execution of visually guided saccades. *The European journal of neuroscience*. 2004; 20(1):264–272. DOI: 10.1111/j.1460-9568.2004.03480.x [PubMed: 15245498]
- Tehovnik, Edward J., Slocum, Warren M., Schiller, Peter H. Delaying visually guided saccades by microstimulation of macaque V1: spatial properties of delay fields. *The European journal of neuroscience*. 2005; 22(10):2635–2643. DOI: 10.1111/j.1460-9568.2005.04454.x [PubMed: 16307605]
- Tehovnik, Edward J., Slocum, Warren M., Smirnakis, Stelios M., Tolias, Andreas S. Microstimulation of visual cortex to restore vision. *Progress in brain research*. 2009; 175:347–375. DOI: 10.1016/S0079-6123(09)17524-6 [PubMed: 19660667]
- Thomson AM, Deuchars J. Temporal and spatial properties of local circuits in neocortex. *Trends in neurosciences*. 1994; 17(3):119–126. [PubMed: 7515528]
- Varela JA, Sen K, Gibson J, Fost J, Abbott LF, Nelson SB. A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 1997; 17(20):7926–7940. [PubMed: 9315911]
- Wagenaar, Daniel A., Potter, Steve M. Real-time multi-channel stimulus artifact suppression by local curve fitting. *Journal of neuroscience methods*. 2002; 120(2):113–120. DOI: 10.1016/S0165-0270(02)00149-8 [PubMed: 12385761]
- Watanabe S, Konishi M, Creutzfeldt OD. Postsynaptic potentials in the cat's visual cortex following electrical stimulation of afferent pathways. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale*. 1966; 1(3):272–283. [PubMed: 5920554]
- Wichmann, Thomas, DeLong, Mahlon R. Deep brain stimulation for neurologic and neuropsychiatric disorders. *Neuron*. 2006; 52(1):197–204. DOI: 10.1016/j.neuron.2006.09.022 [PubMed: 17015236]
- Wichmann, Thomas, Devergnas, Annaelle. A novel device to suppress electrical stimulus artifacts in electrophysiological experiments. *Journal of neuroscience methods*. 2011; 201(1):1–8. DOI: 10.1016/j.jneumeth.2011.06.026 [PubMed: 21745499]
- Williams A, Kim H, Wang F, Vyas S, Shenoy KV, Schnitzer M, Kolda T, Ganguli S. Dimension reduction of multi-trial neural data by tensor decomposition *Frontiers in Neuroscience (T-28)*. 2017

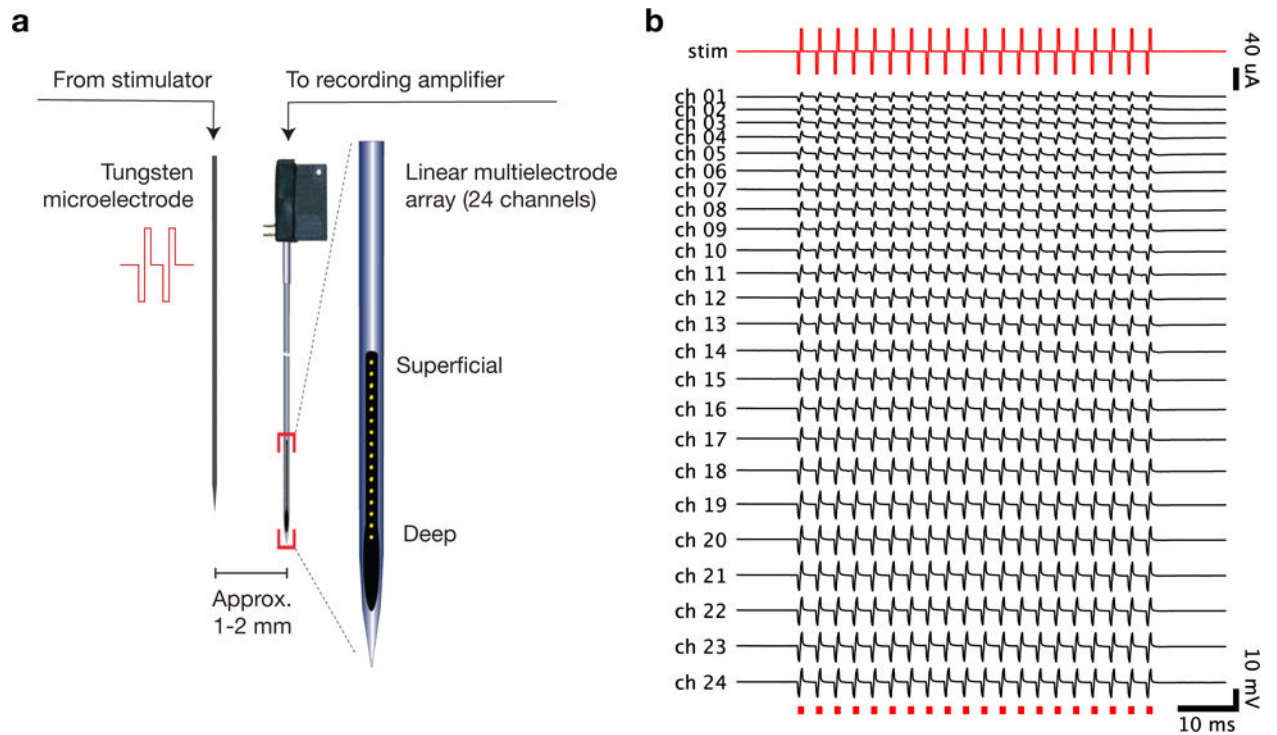


Figure 1.

Our technique facilitates electrical recordings of local neural activity during and after nearby electrical stimulation, provided that the electrical artifact does not saturate the amplifier. (a) A schematic of the recording setup. A tungsten stimulating microelectrode and linear multielectrode array are inserted in parallel into dorsal premotor cortex. The linear array records the electrical activity of nearby neurons spanning the layers of cortex, while electrical current is introduced by the tungsten microelectrode. (b) An example of real electrical recordings on the linear multielectrode array (ch. 1–24) during delivery of a single 40 μ A pulse train (red, stim) through the stimulating electrode. The recording channels span 2.3 mm and are labeled from 1 (most superficial, typically above cortex) through 24 (deep, typically in white matter).

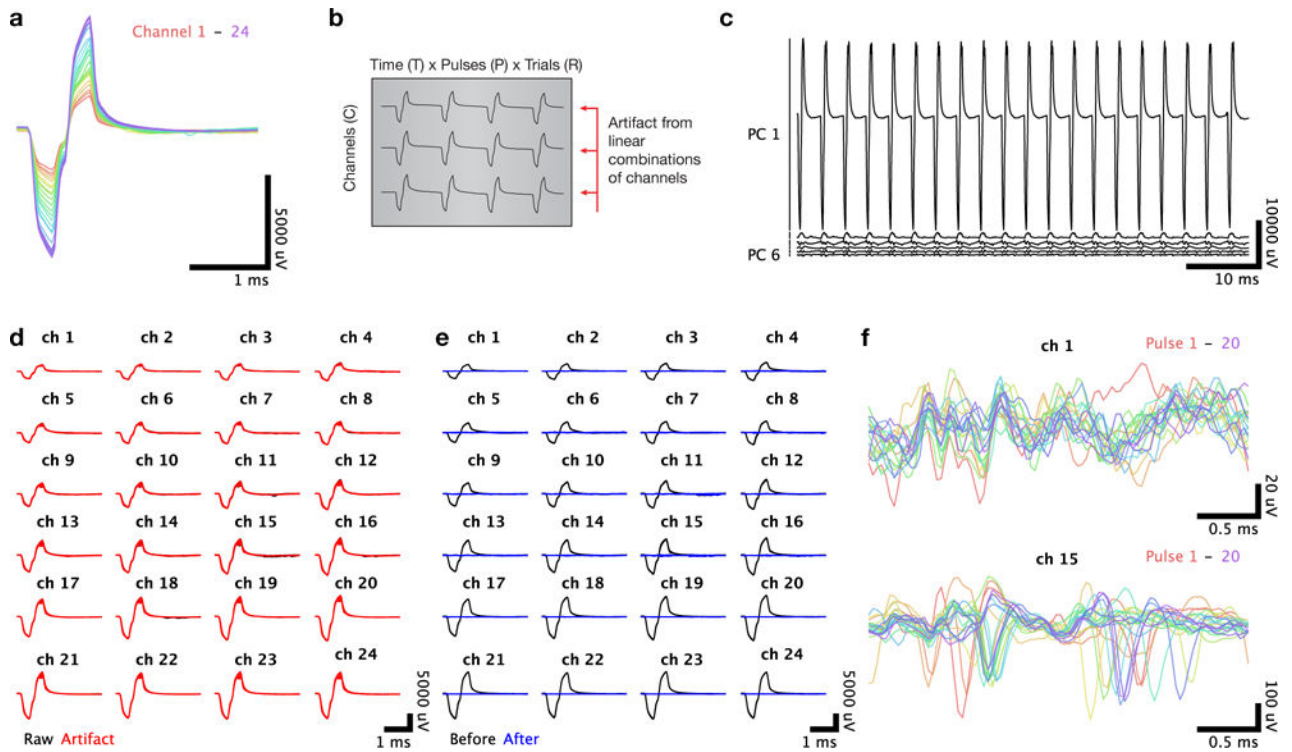


Figure 2.

The first stage of artifact removal removes shared structure recorded simultaneously across linear electrode array channels. (a) The electrical artifact evoked by a single pulse of a stimulus train is highly similar across the 24 channels of the array. (b) To identify the common artifact structure present across channels, we rearrange the artifact data into a matrix with dimensions time \times pulses \times trials ($T \times P \times R$) by channels (C), shown transposed. (c) The principal components of this matrix are linear combinations of the electrode channels. The projections of the recordings along these components clearly display artifact structure shared across channels. (d) Voltage recordings for the first pulse of the stimulus train, plotted separately for each channel with the responses on each trial superimposed. Raw voltage recordings are shown in black, with the artifact inferred via principal components regression shown in red. (e) Same presentation as in (d) but with the residual voltage recordings shown in blue, after the inferred artifact has been subtracted. (f) Cleaned voltage recordings for a single trial with the responses to each individual pulse superimposed. The top panel shows a superficial channel which had no detectable spontaneous neural activity, and the bottom panel shows a lower channel where a well-isolated single unit was spiking spontaneously.

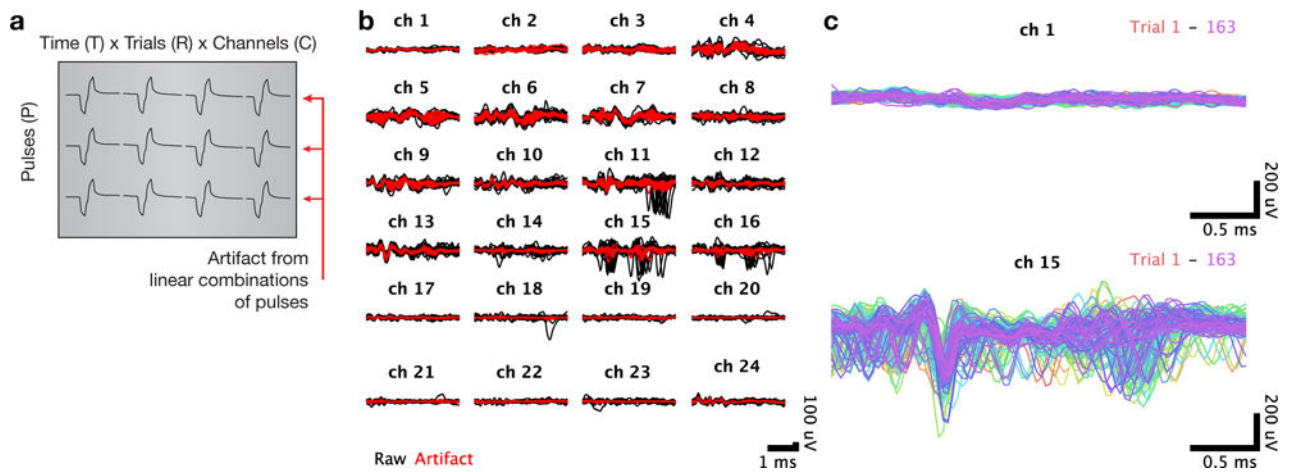


Figure 3.

The second stage of artifact removal removes common artifact structure across the successive pulses of the stimulus train. (a) To isolate structure across pulses, we rearrange the artifact data into individual matrices for each channel, each with dimensions time (T) \times trials (R) \times channels (C) by pulses (P). (b) Voltage recordings for a single trial, plotted separately for each channel with the responses to each pulse superimposed. Raw voltage recordings are shown in black, with the artifact inferred via principal components regression shown in red. (c) Cleaned voltage recordings for a single pulse with the responses to all trials superimposed. The top panel shows a superficial channel which had no detectable spontaneous neural activity, and the bottom panel shows a lower channel where a well-isolated single unit was spiking spontaneously.

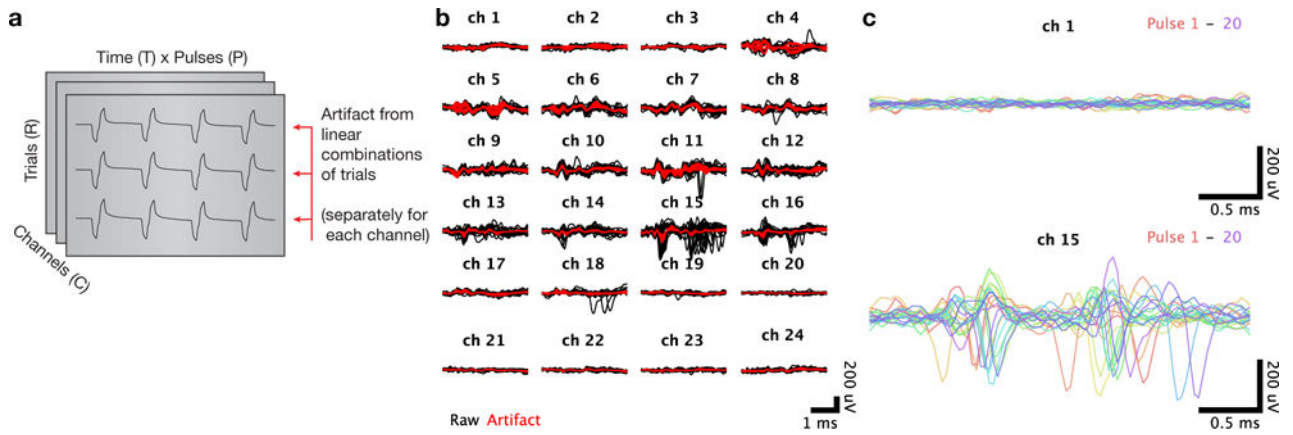


Figure 4.

The third stage of artifact removal removes common artifact structure across trials in which identical electrical stimulus trains were delivered. (a) To isolate structure across trials, we rearrange the artifact data into individual matrices for each channel, each with dimensions trials (R) by time \times pulses ($T \times P$). (b) Voltage recordings for a single pulse, plotted separately for each channel with the responses to each trial superimposed. Raw voltage recordings are shown in black, with the artifact inferred via principal components regression shown in red. (c) Cleaned voltage recordings for a single trial with the responses to all pulses superimposed. The top panel shows a superficial channel which had no detectable spontaneous neural activity, and the bottom panel shows a lower channel where a well-isolated single unit was spiking spontaneously.

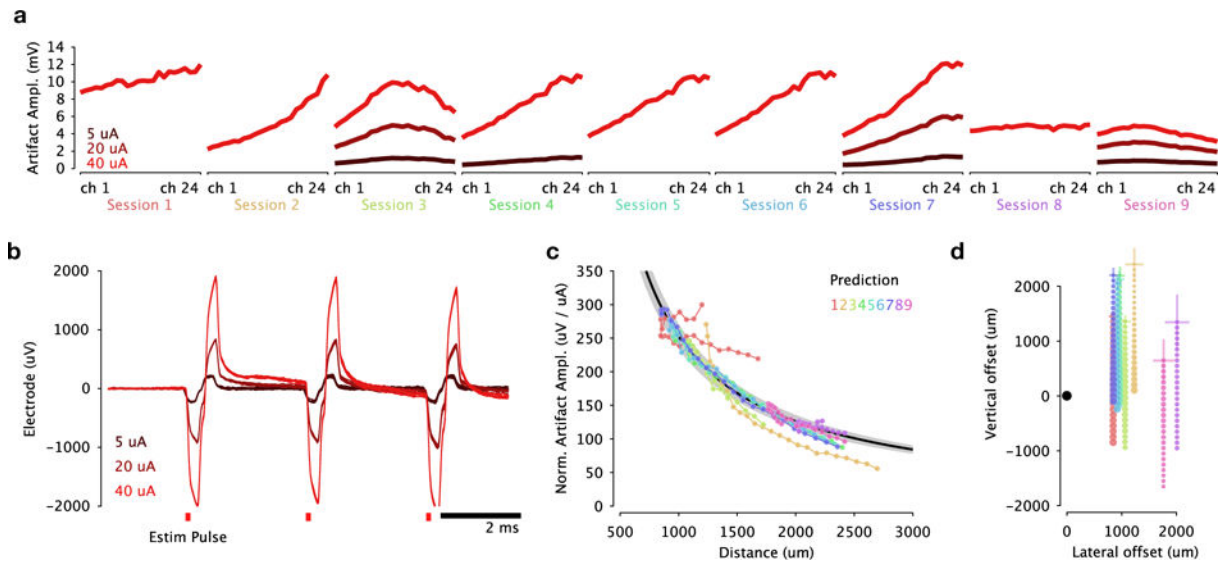


Figure 5.

Electrical artifacts recorded on the electrode array resemble scaled, filtered versions of the original stimulus pulse train. The artifact amplitude varies smoothly with stimulation current and electrode distance, enabling *post hoc* inference of the locations of the recording probes in each session. (a) Electrical artifact amplitude (peak to peak) across multielectrode array channels. Each plot corresponds to a recording session in which the multielectrode array was inserted in a new location. (b) Electrical artifacts recorded on several trials on a superficial electrode with increasing electrical stimulation current. (c) A simple model was fit to artifact amplitudes across recording sessions, which assumes that the amplitude scales linearly with stimulation current and falls off with the reciprocal of Euclidean distance from the stimulating electrode tip. These distances are themselves inferred from the data as parameters describing the position of the electrode array in each session. The fitted model predicted amplitudes (black curve flanked by gray 95% CI) align well with the empirical amplitudes observed on each session (groups of connected dots, each color corresponds to a session). (d) The fitted locations of the recording probe on each session relative to the stimulating electrode tip. Marker size is scaled according to artifact amplitude. Cross-hairs indicate 95% CI for probe location both laterally and vertically.

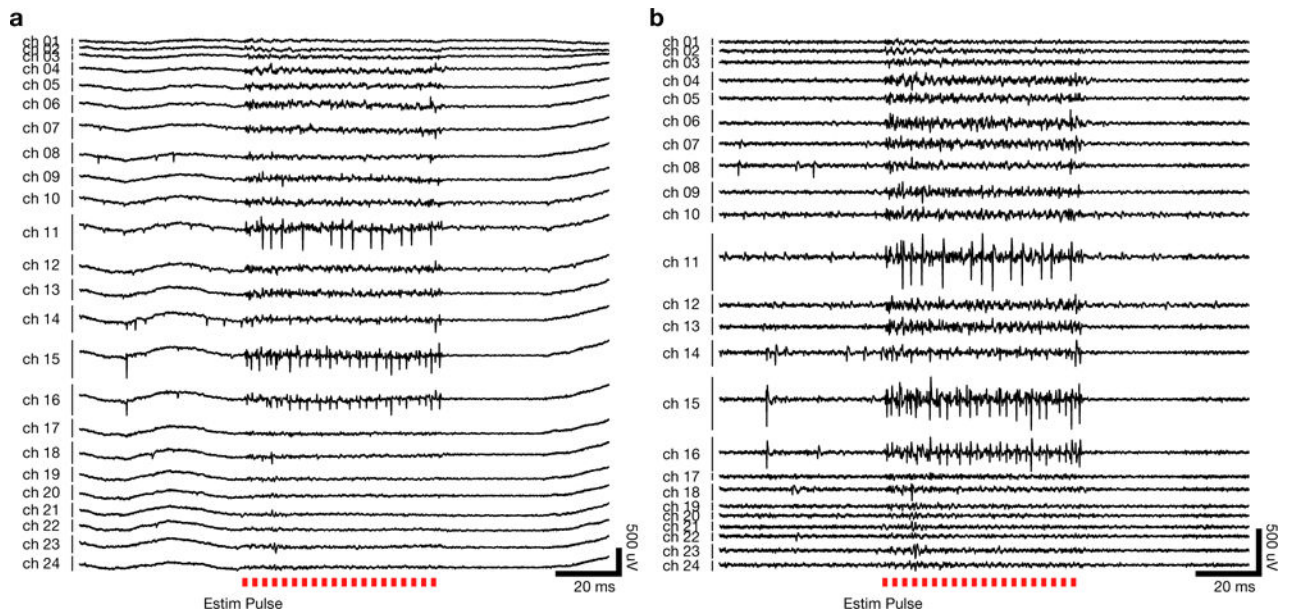


Figure 6.

The cleaned single trial voltage recordings exhibit very little residual stimulation artifact. (a) Electrical recordings on the linear multielectrode array for the same trial shown in 1(b). Spiking activity during the stimulus is readily apparent on multiple channels, including several where spontaneous neural activity was observed. (b) The same recording example as in (a), following high-pass filtering to isolate spiking activity.

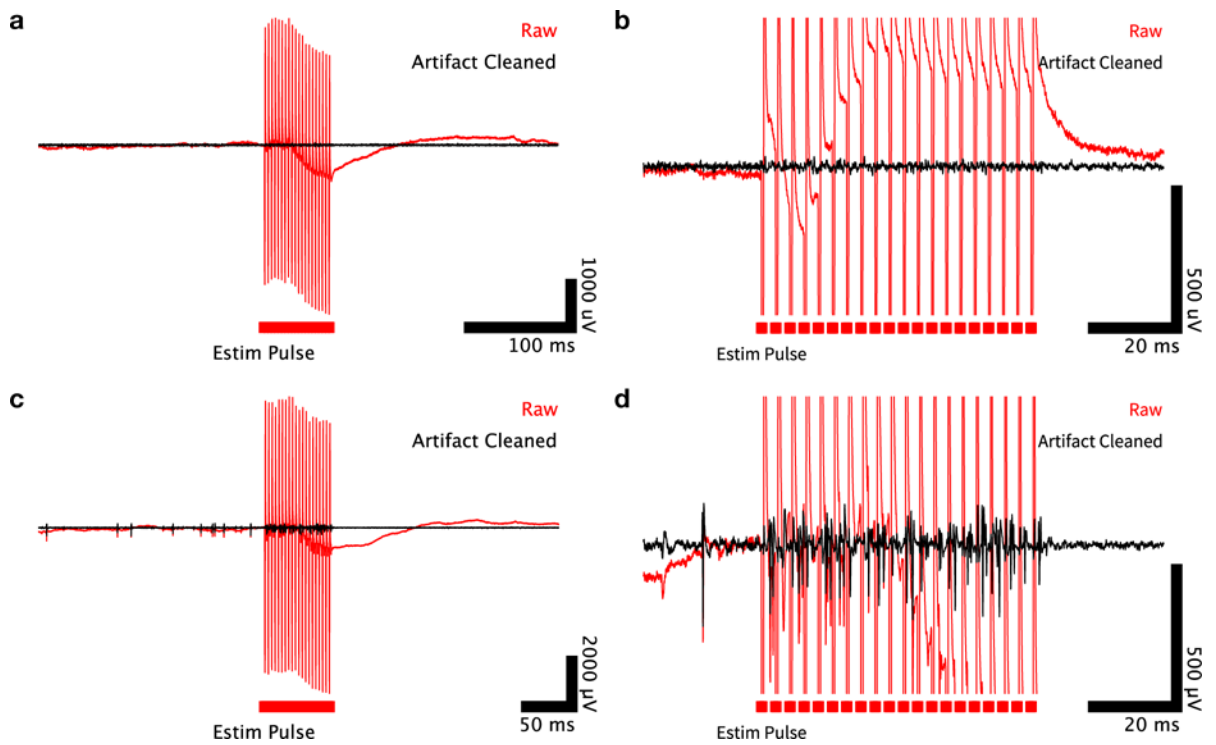


Figure 7. Comparison of raw and artifact cleaned signals. (a) Raw (red) and cleaned (black) voltage traces for a single trial on the most superficially located recording channel where no spontaneous activity was recorded, and no spiking activity is detected during stimulation. (b) Zoomed view of (a). (c) Same as (a) but for a channel located within cortex where spontaneous activity was present. (d) Zoomed view of (c) demonstrating evoked spikes visible during the stimulation period.

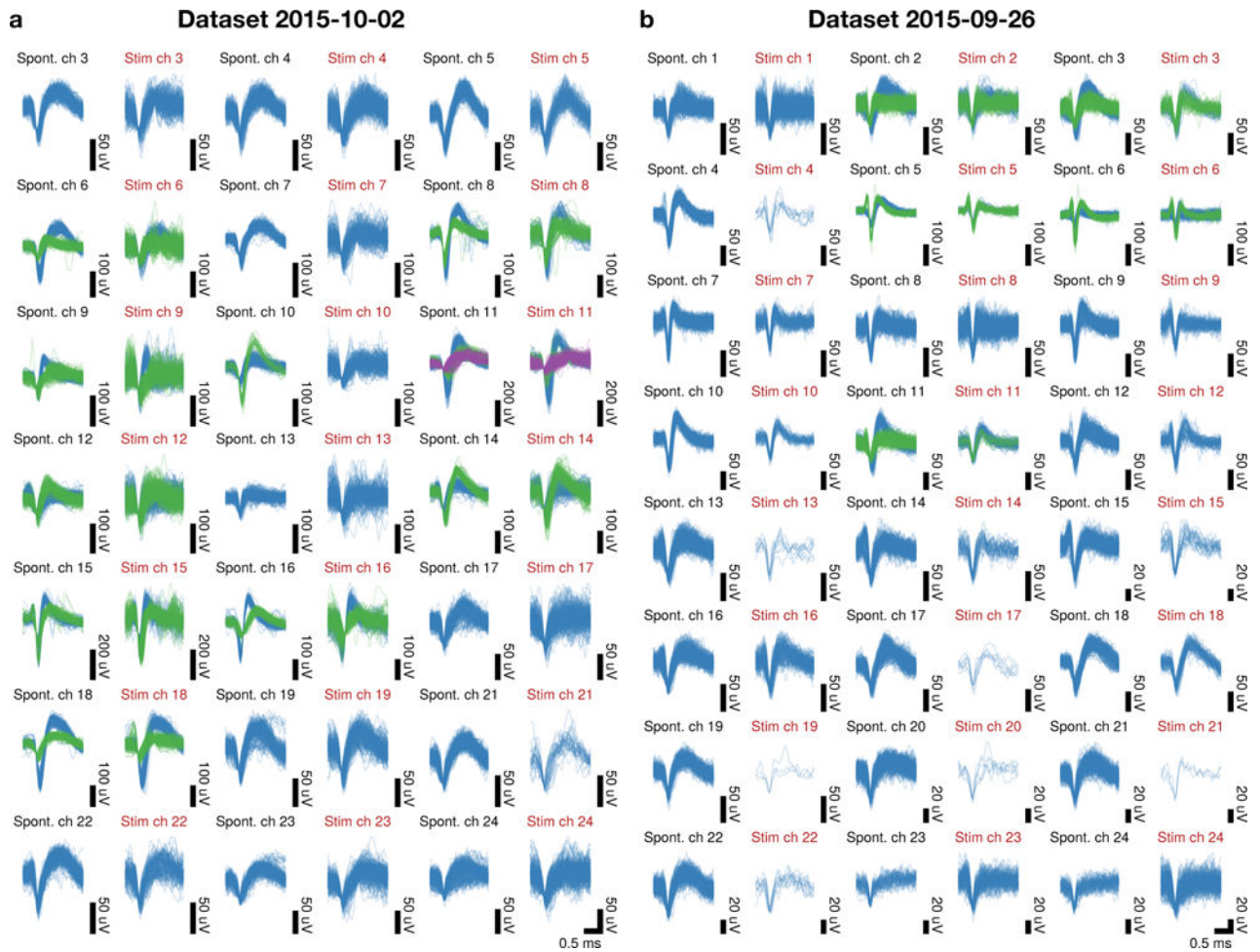


Figure 8. Recovered spiking waveforms detected during the stimulation period closely resemble spontaneously occurring waveforms recorded on the same channel. (a–b) Superimposed waveforms from each channel with spiking activity of two representative recording sessions. Each pair of columns compares spontaneously occurring (spont.) vs. peri-stimulus period (stim.) spiking waveforms. Waveform colors denote hand-sorted unit identity.

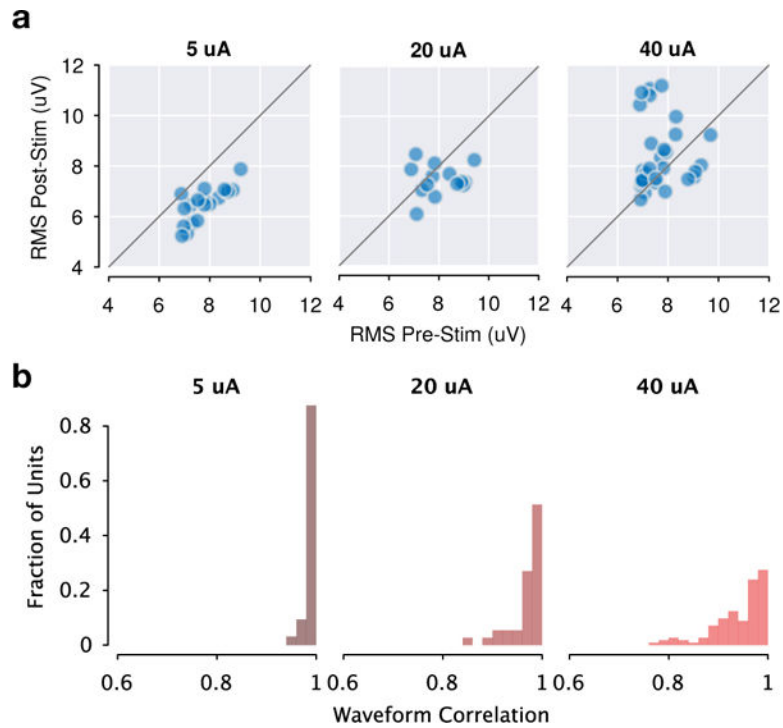


Figure 9. Summary of artifact removal quality across recording sessions. (a) RMS voltage on electrode channels which displayed no spontaneous spiking activity, across all recording sessions. Post-cleaning peri-stimulation RMS voltage is similar to spontaneous pre-stim RMS on each channel. (b) Summary of correlation coefficients between spontaneous spiking waveforms and recovered waveforms during stimulation, across all recovered units in all recording sessions.

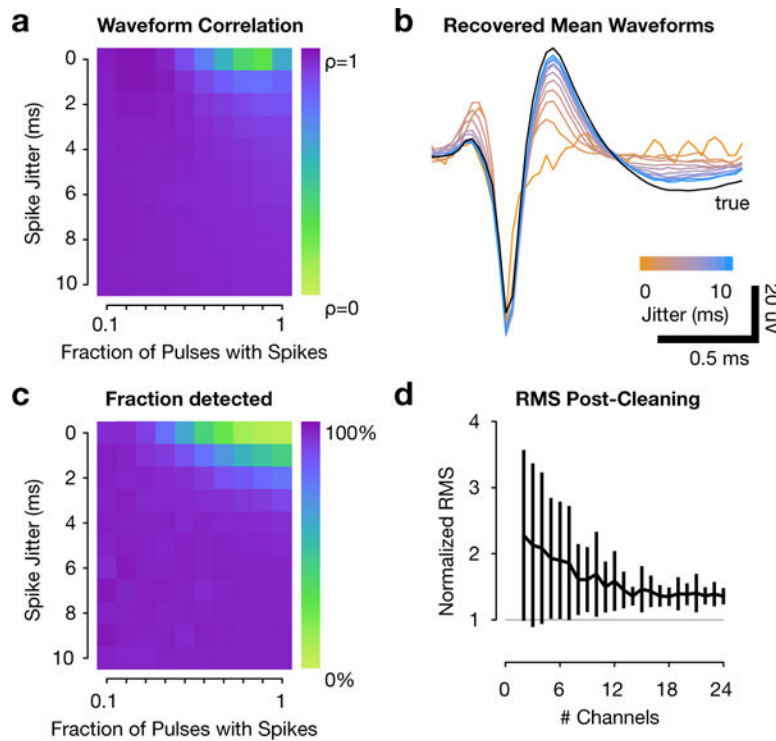


Figure 10.

(a–c) Recovery of spiking waveforms from synthetic datasets in which one channel was replaced by a generated signal with known spike waveforms and times. (a) Correlation of mean recovered waveform after application of ERAASR with the true waveform added to the signal, as a function of the fraction of pulses containing a spike and the jitter in the pulse to spike latency. Recovered waveforms displayed distortion as the spikes are evoked very reliably (higher fraction of pulses) and with very tight latency (lower jitter), as they become less distinguishable from artifact. (b) Mean waveforms recovered as a function of increasing jitter when all pulses contained a spike. (c) Fraction of spiking waveforms detected as threshold crossings. (d) Quality of artifact removal when ERAASR was applied to subsets of channels of varying cardinality on the multielectrode array. RMS voltages are calculated only for channels which exhibited no spontaneous or evoked spiking activity and are normalized to each channel's RMS voltage during non-stimulated trials. Error bars denote standard deviation over channels.

Algorithm 1

ERAASR algorithm pseudocode

Input : Raw data tensor $\mathbf{X}_{\text{raw}}(c, t, p, r)$
with size C channels by T timepoints per pulse by P pulses by R trials

Output : Cleaned data tensor $\mathbf{X}_{\text{cleaned}}(c, t, p, r)$ with same size

Parameters: K_C, K_P, K_R - number of principal components to describe artifact structure over channels, pulses, trials
 $\lambda_C, \lambda_P, \lambda_R$ - the number of adjacent channels, pulses, trials excluded as regressors during artifact reconstruction
 $\beta_P, \beta_C \in \{\text{false}, \text{true}\}$ - Whether to perform cleaning over pulses separately on each channel. Below we assume β_P is false and β_C is true for clarity.

$\mathbf{X} \leftarrow \mathbf{X}_{\text{raw}}$

Clean across channels

- Unfold X into \mathbf{M}^C with size RPT by C .
- $\mathbf{W}^C \leftarrow$ projection along leading K_C principal components of \mathbf{M}^C
- for** $c \in 1, \dots, C$ **do**
 - $\mathbf{A}_c \leftarrow \mathbf{M}^C \mathbf{W}_c^C$ with weights zeroed for channel c and channels within λ_C
 - $\mathbf{X}(c, :, :, :) \leftarrow \mathbf{X}(c, :, :, :) -$ projection of $\mathbf{X}(c, :, :, :)$ onto \mathbf{A}_c
- end**

Clean across pulses

- Unfold X into \mathbf{M}^P with size TRC by P .
- $\mathbf{W}^P \leftarrow$ projection along leading K_P principal components of \mathbf{M}^P
- for** $p \in 1, \dots, P$ **do**
 - $\mathbf{A}_p \leftarrow \mathbf{M}^P \mathbf{W}_p^P$ with weights zeroed for pulse p and pulses within λ_P
 - $\mathbf{X}(:, :, p, :) \leftarrow \mathbf{X}(:, :, p, :) -$ projection of $\mathbf{X}(:, :, p, :)$ onto \mathbf{A}_p
- end**

Clean across trials, separately for each channel

- for** $c \in 1, \dots, C$ **do**
 - Unfold $X(c, :, :, :)$ into \mathbf{M}_c^R for with size TP by R .
 - $\mathbf{W}_c^R \leftarrow$ projection along leading K_R principal components of \mathbf{M}_c^R
 - for** $r \in 1, \dots, R$ **do**
 - $\mathbf{A}_{c,r} \leftarrow \mathbf{M}_c^R \mathbf{W}_{c,r}^R$ with weights zeroed for trial r and trials within λ_R
 - $\mathbf{X}(c, :, :, r) \leftarrow \mathbf{X}(c, :, :, r) -$ projection of $\mathbf{X}(c, :, :, r)$ onto $\mathbf{A}_{c,r}$
 - end**
- end**

return $\mathbf{X}_{\text{cleaned}} \leftarrow \mathbf{X}$
