Artifact-free recordings in human bidirectional brain–computer interfaces

Jeffrey M Weiss\textsuperscript{1,2,6}, Sharlene N Flesher\textsuperscript{2,3,6}, Robert Franklin\textsuperscript{4,6}, Jennifer L Collinger\textsuperscript{1,2,3,6} and Robert A Gaunt\textsuperscript{1,2,3,7}

\textsuperscript{1} Department of Physical Medicine and Rehabilitation, University of Pittsburgh, Pittsburgh, PA, United States of America
\textsuperscript{2} Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, United States of America
\textsuperscript{3} Center for the Neural Basis of Cognition, Pittsburgh, PA, United States of America
\textsuperscript{4} Blackrock Microsystems, Salt Lake City, UT, United States of America
\textsuperscript{5} VA Pittsburgh Healthcare System, Pittsburgh, PA, United States of America

E-mail: rag53@pitt.edu

Received 23 July 2018, revised 25 September 2018
Accepted for publication 10 October 2018
Published 16 November 2018

Abstract

Objective. Intracortical microstimulation has shown promise as a means of evoking somatosensory percepts as part of a bidirectional brain–computer interface (BCI). However, microstimulation generates large electrical artifacts that dominate the recordings necessary for BCI control. These artifacts must be eliminated from the signal in real-time to allow for uninterrupted BCI decoding. Approach. We present a simple, robust modification to an existing clinical BCI system to allow for simultaneous recording and stimulation using a combination of signal blanking and digital filtering, without needing to explicitly account for varying parameters such as electrode locations or amplitudes. We validated our artifact rejection scheme by recording from microelectrodes in primary motor cortex (M1) while stimulating in somatosensory cortex of a person with a spinal cord injury. Main results. M1 recordings were digitally blanked using a sample-and-hold circuit triggered just prior to stimulus onset and a first-order 750 Hz high-pass Butterworth filter was used to reduce distortion of the remaining artifact. This scheme enabled spike detection in M1 to resume as soon as 740 \(\mu\)s after each stimulus pulse. We demonstrated the effectiveness of the complete bidirectional BCI system by comparing functional performance during a 5 degree of freedom robotic arm control task, with and without stimulation. When stimulation was delivered without this artifact rejection scheme, the number of objects the subject was able to move across a table in 2 min under BCI control declined significantly compared to trials without stimulation \((p < 0.01)\). When artifact rejection was implemented, performance was no different than in trials that did not include stimulation \((p = 0.621)\). Significance. The proposed technique uses simple changes in filtering and digital signal blanking with FDA-cleared hardware and enables artifact-free recordings during bidirectional BCI control.

Keywords: ICMS, artifact rejection, brain–computer interface, bidirectional, electrical stimulation
1. Introduction

A major goal of neuroprosthetic research is to restore motor function to individuals with upper extremity paralysis. Recent work has shown that it is possible to achieve high degree of freedom control of a robotic arm using a brain–computer interface (BCI) [1, 2]. While performing simple object interactions with a BCI-controlled limb has been successful, dexterous control remains a significant challenge. One limitation of current BCI systems, and a possible contributor to the challenges in achieving dexterous hand control, is the absence of somatosensory feedback. This idea is highlighted by the observation that humans require somatosensory feedback to perform dexterous object manipulation [3, 4]. This information is unavailable in BCI systems, which provide only visual feedback.

While the concept of incorporating somatosensory feedback into a BCI system has been recognized as an important area for development [5–11], progress toward that goal has been relatively slow. Recently, however, we have shown that a person can perceive graded tactile percepts referred to multiple independent regions of the hand using intracortical microstimulation (ICMS) [12]. Using this microstimulation approach, sensors on a BCI-controlled end-effector can be used to modulate stimulus parameters in real-time and provide additional channels of information to the user. A significant technical hurdle to realizing this objective is the ability to record neural spiking activity while electrically stimulating in an adjacent brain area. Microstimulation generates voltages that are orders of magnitude greater than the microvolt-scale extracellular potentials used for control. When filtered using standard filters, these large, brief pulses can generate artifacts several milliseconds in duration. While the onset timing of the large pulses is precisely known, the filter response after each stimulus pulse is often unpredictable and of a long duration. These factors make reliably eliminating artifacts and detecting neural spikes difficult. A bidirectional BCI system must have a robust method for artifact rejection to decode uninterrupted motor commands during ICMS.

A number of artifact rejection solutions have been proposed previously, with varying degrees of simplicity and generalizability. One proposed solution is to alternate recording and stimulation trains in time, e.g. in 50 ms intervals. This method was used to instruct target selection via ICMS in a 2D BCI control task [13]. While effective at eliminating artifact-contaminated signal, only 50% of the neural data is preserved and stimulation can only occur in bursts. In order to increase the potential parameter space of the stimulation paradigms, we did not want to be constrained to bursting patterns. Rather, we desire to find and use stimulation schemes that will yield naturalistic percepts and can be driven continuously during object contact. Furthermore, we wanted to preserve as much neural data as possible for high degree of freedom BCI decoding, including control of a robotic arm and hand. An ideal solution would allow stimulation to occur continuously, minimize neural data loss and require minimal changes to a clinical BCI system. Other proposed solutions for artifact rejection include more sophisticated signal processing techniques such as template subtraction or modeling [14–18], regression-based referencing [19], or a combination of bipolar stimulation with adaptive-filtering [20]. Some of these alternative solutions allow spikes to be detected during the stimulus pulse itself, but only in experiments in which the amplifiers are never saturated. These methods sometimes require noncausal processing, resulting in additional challenges to rejecting artifacts in real-time [17, 19]. Additionally, many of these methods require a training dataset to generate stimulus artifact templates or model parameters. We sought to create a simple and generalizable method for artifact rejection that did not require training and could be implemented in real-time using existing hardware.

This work was conducted as part of a BCI study with the ultimate goal of demonstrating the utility of ICMS to provide somatosensory feedback to improve high degree of freedom BCI control. This goal necessitates an electrophysiology system capable of recording neural activity in motor cortex while simultaneously stimulating in somatosensory cortex. Here we describe a simple, yet robust solution to artifact rejection that minimizes neural data loss, uses existing FDA-cleared clinical hardware, and avoids complicated signal processing techniques that may fail to generalize to previously untested stimulation scenarios.

2. Materials and methods

2.1. Array implantation

A 28 year old male participant with a C5 motor/C6 sensory ASIA B spinal cord injury sustained 10 years prior was enrolled in the clinical trial. Presurgical magnetoencephalography imaging, as described previously [12], was used to identify the hand areas of primary motor cortex (M1) and area 1 of somatosensory cortex (S1). Four Utah microelectrode arrays (Blackrock Microsystems, Salt Lake City, Utah) were implanted in the participant’s left sensorimotor cortex, as shown in figure 1. Two of the arrays were designed exclusively for recording and were implanted in the hand and shoulder areas of M1. The other two arrays were designed primarily for stimulation and were implanted in the hand area of S1. Each recording array was 4 mm × 4 mm in size with 88 functional electrodes with platinum-coated tips. Each stimulation array was 2.4 mm × 4 mm in size with 32 functional electrodes coated with a sputtered iridium oxide film, which increased the charge injection capacity for stimulation. The arrays were
wired to two percutaneous connectors attached to the skull. It was not possible to record from the S1 arrays during ICMS, and the goals of our study do not necessitate simultaneously stimulating and recording from the same array.

This study was performed under an Investigational Device Exemption granted by the US Food and Drug Administration. The study received approval from the Institutional Review Boards at the University of Pittsburgh and the Space and Naval Warfare Systems Center Pacific and is registered at ClinicalTrials.gov (NCT01894802). Informed consent was obtained prior to performing any experimental procedures.

2.2. Electrophysiology system

Neural recording was accomplished using two Neuroport clinical electrophysiology systems, each consisting of a neural signal processor (NSP) and Front-End Amplifier (Blackrock Microsystems, Salt Lake City, UT). All system components are shown in figure 2. The amplifier firmware was modified to add a digital sample-and-hold feature, which allowed the digital outputs on each channel of the amplifier to be fixed at their last sampled value when the voltage on a TTL input on the amplifier was set high. By setting this TTL input high during stimulus pulses, we could effectively ‘blank’ out the stimulus artifact prior to subsequent signal processing onboard the NSP, which was unmodified and could only perform standard signal processing operations for spike extraction. Spike extraction was performed on the NSP by high-pass filtering and thresholding the digitized and blanked signal to detect neural spikes.

ICMS was delivered using the Cerestim R96 microstimulator (Blackrock Microsystems). The stimulator has a ‘sync’ output, which produces a TTL pulse beginning 60 μs before each stimulus pulse. A monostable multivibrator (Texas Instruments CD74HC123, Dallas, TX) ‘blanking circuit’ was triggered by the rising edge of the sync signal and generated a TTL pulse of an arbitrary duration set by a potentiometer. This TTL blanking signal was then connected to the amplifier to blank the recordings during stimulation pulses. A timing diagram of this process is shown in figure 3. The Cerestim R96 and NSPs interfaced with a custom BCI software suite, developed in Matlab and C++, and which was configured to ignore registered spikes that occurred in the first sample after the offset of the sample-and-hold blanking period. These ‘spikes’ were the result of discontinuous sampling of the analog signal.

ICMS consisted of trains of charge-balanced, cathodic-leading asynchronous pulses delivered at 100 Hz. Each pulse had a 200 μs cathodal phase duration, 100 μs interphase duration, and 400 μs anodal phase duration. Cathodal amplitudes ranged from 2 to 100 μA; anodal amplitudes were always half of the cathodal amplitude. Up to twelve electrodes could be simultaneously stimulated, with a maximum total charge per phase of 144 nC.

The system presented above has two primary components that can be tuned to reject stimulation artifacts while preserving neural data. First, the digital filter used for spike extraction can be tuned to minimize the effect of stimulation artifacts. Second, the raw signals can be blanked during stimulation for an arbitrary duration. This blanking duration can be tuned to balance removing the artifact from the raw signal, while also preserving neural data. These features were tuned using raw voltage recordings collected from electrodes in M1 when ICMS was applied to S1.

2.3. Open-loop recording experiments

Electrode recordings from both M1 arrays were collected during trials in which ICMS was delivered to the S1 electrodes for the purpose of tuning and validating the artifact rejection system. Analog voltage recordings were band-pass filtered (0.3–7500 Hz) and digitized at 30 000 samples per second. Signal blanking and digital filters were not applied online. This allowed for the artifact rejection and spike detection features to be implemented and tuned offline on the raw, artifact-contaminated signals.

The final online artifact rejection implementation was validated by analyzing the distribution of spikes detected in the inter-stimulus interval. Spike times were detected and recorded online from both recording arrays while the participant performed a motor imagery task consisting of reaching and grasping movements that modulated neural activity in M1 [2]. While the participant performed this task, we collected three 60 s long trials under three conditions: no stimulation, stimulation without artifact rejection, and stimulation with artifact rejection. The participant was informed that the stimulation was irrelevant to his motor task. During trials with stimulation, three electrodes on the medial stimulation array delivered 60 μA pulses at 100 Hz in 500 ms bursts. Each burst was separated by 500 ms. The spikes during each stimulation train were binned into 0.5 ms increments within the 10 ms interval between stimulus pulses. A histogram of the mean binned spike counts with 95% confidence intervals was constructed for each stimulation condition. Stimulation times from the trials with artifact rejection were then applied to the...
this method and all sorted single units and multiunit signals from these channels were included in order to have a mixture of high- and low-SNR signals. A snippet detection threshold was set at \(-5.25\) times the root-mean-square voltage (\(V_{\text{RMS}}\)) of each filtered signal. The \(-5.25\) \(V_{\text{RMS}}\) multiplier has been used in previous studies (e.g., Collinger et al [1]) to extract units of varying SNR without being overly sensitive to low-amplitude noise. A biased estimate of \(V_{\text{RMS}}\) was calculated using the Blackrock Microsystems algorithm described in Christie et al [24] to reduce the effect of outliers. The extracted spike snippets were then sorted using principal component analysis and a Gaussian mixture model expectation-maximization algorithm [25]. Spike times were saved for each sorted unit. Units with less than five spikes were excluded from further analysis. A total of 167 sorted units across the 58 channels were analyzed further, which included both high-SNR single units and low-SNR multi-unit activity. The raw voltage signals were then re-filtered with a first-order high-pass Butterworth filter with varying cutoff frequencies. Sorted snippets were extracted for each filter using the previously determined spike times. The mean peak-to-peak voltage, noise estimate, and signal-to-noise ratio (SNR) were calculated for each unit and filter setting. The noise estimate was calculated for each channel as two times the standard deviation of the continuous filtered signal after removing each spike [23]. The SNR was calculated as the mean peak-to-peak voltage divided by the noise estimate. Unsorted threshold crossings were extracted for each cutoff frequency, using a threshold of \(-4.5\) \(V_{\text{RMS}}\), to determine the effect of changing the cutoff frequency on the number of threshold crossings, as used for online BCI control.

2.5. Bidirectional BCI system

The recording and stimulation electrophysiology system described above was used to develop a bidirectional BCI system that was based upon the high degree of freedom BCI system that our group has described previously [1, 2]. Multichannel spiking activity was recorded using the digital filters and thresholding techniques described above, but implemented in real-time on the NSPs. No spike sorting was applied; threshold crossings on each recording channel were treated as a single-unit. Firing rates were estimated using 20 ms bins of spike counts, which were then smoothed using an exponential function with a 440 ms sliding window. The two-phase calibration procedure previously described by Collinger et al [1] was used to train an optimal linear

### Figure 3. Timing diagram. Each biphasic stimulus pulse (red trace) consisted of a 200 \(\mu\)s cathodal period, followed by a 100 \(\mu\)s interphase period and a 400 \(\mu\)s anodal period. The stimulator outputs a TTL sync signal (orange trace) beginning 60 \(\mu\)s before the stimulus pulse and ending at 1000/f ms, where \(f\) is the commanded stimulation frequency. The rising edge of the sync signal triggered the blanking circuit to output a TTL blanking signal (green trace) of a fixed, experimenter-defined duration. This blanking signal triggered the amplifier’s sample-and-hold feature, resulting in each electrode recording (blue trace) to maintain a constant value from 60 \(\mu\)s prior to the stimulus pulse until the falling edge of the blanking signal.

### Figure 2. Bidirectional system diagram. Each recording array (blue) records neural signals, which are digitized, filtered, and decoded to control an end effector. Sensor data from the end effector are transformed to generate stimulation commands, which are delivered through the stimulating arrays (red). The recording signal is blanked during stimulus pulses to remove the majority of the artifact before filtering.
estimator (OLE) decoder to control the Modular Prosthetic Limb (Johns Hopkins University Applied Physics Lab). The resulting decoder predicts five degree of freedom (DOF) limb velocity by relating the square-root transformed firing rate, $f$, of each unit to 5D velocity, $v$, using equation (1).

$$f = b_0 + b_1 v_x + b_2 v_y + b_3 v_z + b_4 v_r + b_5 v_g.$$  
(1)

The 5D velocity includes the Cartesian endpoint components $x$, $y$, and $z$, an endpoint roll orientation component, $r$, and a grasp component, $g$. Optimal linear estimation with ridge regression was used to determine decoding weights from each $b$ coefficient to predict arm kinematics from the recorded firing rates. During testing, firing rates were scaled using equation (2) prior to decoding to account for global changes in firing rate [26]$. Individual unit firing rates, $f$, were scaled by the mean population firing rate during training, $F_{training}$, and divided by the mean population firing rate over the previous $300\text{ ms}$, $F_{300\text{ ms}}$, to produce a scaled firing rate, $f_{scaled}$.

$$f_{scaled} = f \cdot \frac{\text{mean}(F_{training})}{\text{mean}(F_{300\text{ ms}})}.$$  
(2)

To generate stimulation commands, torque sensors, located in the motors at the base of the digits of the Modular Prosthetic Limb, were mapped to electrodes in S1 [12]. Torque sensor values, $T$, were linearly converted to stimulation amplitude, $A$ (see equation (3)), such that electrodes began stimulating at a minimum amplitude, $A_{min}$, when the corresponding sensor value crossed a defined threshold, $T_{min}$. The stimulation amplitude then increased linearly with increasing torque until a maximum value, $A_{max}$, was reached at a defined maximum torque value, $T_{max}$. Stimulus frequency was fixed at $100\text{ Hz}$. Pulse phase durations were the same as those described previously. A $50\%$ duty cycle was applied to limit stimulation durations to $1500\text{ ms}$ within $30\text{ s}$, i.e. after $15\text{ s}$ of continuous $100\text{ Hz}$ stimulation electrodes were temporarily disabled for $15\text{ s}$.

$$A = \left( \frac{T - T_{min}}{T_{max} - T_{min}} \right) (A_{max} - A_{min}) + A_{min}; T_{min} \leq T \leq T_{max}.$$  
(3)

### 2.6. Object transfer task

The performance of the bidirectional system was evaluated using an object transfer task. The goal of this experiment was to demonstrate that the artifact rejection system enabled bidirectional BCI performance that was functionally equivalent to BCI performance without stimulation. In each session, a five DOF velocity decoder was trained and the decoder’s functional performance was then evaluated in a task where the participant was required to use the Modular Prosthetic Limb to grasp a cylinder on the left side of a table, pick it up, and transfer it to the right side of the table. A $20\text{ cm}$ wide region in the center of the table was indicated by two lines of tape. The participant was required to hold the object above this region while transporting it, rather than simply drag the object across the surface of the table. If the object contacted the table at any point in this region, the participant was required to bring the object back to the left to reset the transfer. The task was scored based on the number of successful transfers in $2\text{ min}$. In each of two sessions, the system was configured with either a fourth-order high-pass Butterworth filter with a cutoff frequency of $250\text{ Hz}$ or a first-order high-pass Butterworth filter with a cutoff frequency of $750\text{ Hz}$. Filter settings, spike thresholds, and decoders were fixed throughout each session. Within each session, three task conditions were tested: no stimulation, stimulation with signal blanking, and stimulation without signal blanking. Here, signal blanking refers to the combination of blanking before digital filtering and thresholding, and software rejection to discard transient threshold crossings immediately after the blanking period. Each task condition was tested with both filter settings, for a total of six conditions. Each condition was tested in a block format with three repetitions per block. In all cases, coincidence detection, a standard feature of the NSP, was enabled to reject threshold crossings that occurred simultaneously on at least $45\text{ channels}$. During task conditions with ICMS, eight electrodes spanning both S1 arrays were mapped to the index finger torque sensor and simultaneously amplitude-modulated using the sensor-stimulus transform in equation (3). For this task, $T_{min} = .05\text{ Nm}$, $T_{max} = .25\text{ Nm}$, $A_{min} = 18\mu A$, and $A_{max} = 46\mu A$. These parameters were chosen with the intention of generating large ICMS artifacts on both M1 arrays during the task and were not necessarily intended to provide beneficial feedback for the task. The $50\%$ stimulation duty cycle, a required safety feature of our system, allowed for the possibility of intervals in which index finger torque did not generate stimulus pulses during ICMS trials. For example, if stimulation was delivered for $15\text{ continuous seconds}$, stimulation would automatically shut off for $15\text{ s}$. This could result in periods of grasp without associated ICMS pulses.

### 3. Results

#### 3.1. Unfiltered recordings during ICMS

The stimulation artifacts recorded in M1 during microstimulation of S1 were typically on the scale of millivolts, and often exceeded eight millivolts, saturating the amplifier. These artifacts were one to two orders of magnitude larger than most extracellular potentials, such as those shown in figure 4. However, even in the case of amplifier saturation, spikes could be observed in the raw voltage signal less than one millisecond after the offset of a microstimulation pulse. Therefore, we determined that it would be possible to discard artifacts and recover spikes a short time after each stimulus pulse. Our artifact rejection scheme is the result of applying a minimal number of changes to our existing clinical BCI system. Signal-blanking duration and digital filter parameters were tuned to allow us to reliably record artifact-free spikes between each stimulation pulse.

#### 3.2. Effect of digital filter parameters on spike extraction

A typical neural recording scheme involves recording a high bandwidth voltage signal (i.e. $0.3–7500\text{ Hz}$ bandpass-filtered
signal sampled at 30 000 samples per second), applying a digital ‘spike filter’, and then applying a voltage threshold to the filtered signal to detect spikes. A good spike filter is typically designed to remove low-frequency local field potentials (LFPs) while passing high-frequency action potentials. For example, in previous studies we used a fourth-order Butterworth high-pass filter with a 250 Hz cutoff frequency [1, 2]. However, ICMS introduces another source of high-frequency content that can cause these standard filters to produce undesirable output after the stimulus pulse is completed. We wanted to explore whether modified filters alone could minimize these effects.

Microstimulation generates high-amplitude stimulus artifacts that are impulse-like in nature. As a result, these high-amplitude artifacts generate filter output resembling the filter’s impulse response, distorting the artifact. For infinite impulse response (IIR) filters, this can cause ringing in the filter output, often resulting in spurious threshold crossings long after the initial stimulus artifact.

In many signal-processing applications, filter parameters are chosen to best approximate an ideal filter in the frequency domain. For this reason, higher-order Butterworth filters are often preferred for their short transition bands and flat passbands and stopbands. Here, preserving high frequency components and attenuating low frequency components of the signal are necessary to enable spike detection. However, the time domain properties of the filter are also critical because we wish to detect spikes as quickly as possible after each stimulation pulse. The post-stimulus filter ringing can be reduced or eliminated by choosing a filter with an impulse response exhibiting a fast settling time and few or no oscillations. The impulse responses to several high-pass Butterworth filter designs are shown in figure 5. Ultimately, we chose a first-order Butterworth filter design to eliminate oscillations in the impulse response. We further found that setting the high-pass cutoff frequency at higher values resulted in impulse responses with faster settling times.

In the example recording during microstimulation, Stimulation artifacts are several orders of magnitude larger than extracellular spike recordings and often saturate the amplifier. (a) Three stimulation artifacts are shown, along with several neural spikes. Each artifact saturates the amplifier during the cathodal phase. (b) Detailed view of the red boxed region in (a). Spikes are indicated by arrows at times 1.3, 3.5, and 6.2 ms. The first spike begins 0.6 ms after the offset of stimulation.

Figure 4.

Figure 5. Impulse responses for a high-pass Butterworth filter with different filter orders and cutoff frequencies. Lower order filters oscillate less in response to an impulse input and higher cutoff frequencies exhibit faster settling times.

Our final spike filter implementation consisted of a first-order Butterworth high-pass filter with a 750 Hz cutoff frequency as it minimized filter ringing and exhibited a fast settling time after stimulation. The effects of this filter, and the more common fourth-order, 250 Hz Butterworth filter, on an example stimulus artifact are demonstrated in figure 6. The effects can particularly be seen in panels 6(c) and (d). Ultimately, this modified filter reduces, but does not eliminate, the distortion of the stimulus artifact. Further increases in cutoff frequency, while resulting in faster settling times, would attenuate much of the signal power in the spike band, commonly cited as approximately 300–6000 Hz [27, 28], with much of the information at about 1000 Hz [29]. While the 750 Hz cutoff frequency is within this band, most of the spike signal is preserved because of the gradual roll-off of the first-order filter, as shown in figure 7.
3.3. Signal blanking

Stimulation waveforms are distorted by IIR spike filters, resulting in long duration artifacts. The first-order 750 Hz high-pass filter reduces, but does not eliminate, this distortion after the end of the stimulus pulse. We therefore implemented a method of blanking the recorded signal during pulse delivery to eliminate the bulk of the stimulus artifact itself, thus avoiding contaminating the filter input with a large impulse-like response. This was achieved using a digital sample-and-hold function that was implemented in the amplifier prior to digital filtering on the NSP.

While signal blanking was effective at removing stimulus artifacts, additional artifacts were caused by discontinuities at the offset of blanking, as highlighted in figure 8(a). These discontinuities, while much smaller in amplitude than the stimulus artifact, often resemble a step input. The resulting step response drives the filtered signal near or past the spike threshold, resulting in spurious artifacts after blanking. The ringing resulting from this step response can be seen in figure 8(a) between approximately 1.5 and 4 ms. Because filter
ranging persists in the signal after blanking, the criteria used to tune the digital filter in response to stimulus artifacts still apply after blanking the signal during stimulation. In figure 8(b), the use of a first-order 750 Hz high-pass Butterworth filter effectively eliminates oscillations in response to the discontinuity at the offset of the blanking period and does not introduce additional artifacts after blanking. The combination of signal blanking and a digital filter with a fast settling time after impulse or step inputs can eliminate stimulus artifacts and resume spike detection shortly after each stimulus pulse.

In our final implementation, the recorded voltage signals were blanked during each stimulation pulse using a sample-and-hold period of 1467 µs (44 samples at 30 kHz). This includes 60 µs before each stimulation pulse, 700 µs while the stimulator is delivering current, and an additional 707 µs as the voltage begins to recover from the anodal phase, which in some cases saturates the amplifier. One additional sample was blanked in software, after the application of the digital filter and threshold, to discard occasional false spikes caused by the discontinuity at the offset of the blanking period. This complete implementation allowed for reliable spike detection as soon as 740 µs after the offset of each stimulus pulse and a total signal-blanking duration of 1500 µs.

To quantify the ability of the system to detect spikes over the inter-stimulus interval, we measured the occurrences of spikes in 0.5 ms time bins between successive stimulus pulses. Spikes were detected and recorded online from both recording arrays while the participant performed an imagined reaching and grasping task. Task-relevant stimulation was delivered during bursts on a group of electrodes with and without artifact rejection. The histogram in figure 9(a) demonstrates that artifacts dominate the neural recordings when a fourth-order Butterworth 250 Hz high-pass filter is used without signal blanking. Most detected ‘spikes’ occur between 0–0.5 ms and 3.5–4.5 ms, which can be explained by the initial stimulus pulse and filter ringing, as seen in figure 6(c). When the first-order 750 Hz high-pass filter and signal blanking are implemented, spikes can be recorded throughout the interval between the blanking period and subsequent stimulus pulse (figure 9(b)). While a slightly elevated spike rate occurs from 1.5 to 5 ms after each stimulus pulse, the flat distribution from 5 to 10 ms is not significantly different from the distribution seen in the absence of stimulation (two-sample Kolmogorov–Smirnov test, \( p = 0.058 \)).

3.4. Signal quality

As discussed previously, the spike filter was tuned to quickly return to steady state after perturbations related to stimulation. However, the filter parameters also have an effect on overall signal quality. To investigate the effects of cutoff frequency on signal quality, we performed an analysis based on the filter analysis in Lempka et al [23] using neural recordings in the absence of ICMS. The relationships of cutoff frequency on peak-to-peak voltage, noise level, SNR, and unsorted threshold crossings are shown in figure 10. We validated that increasing high-pass cutoff frequency decreased in peak-to-peak voltage, while increasing SNR. These results agree with the trends presented in Lempka et al. Increasing the first-order high-pass cutoff frequency from 250 Hz to 750 Hz resulted in a 20.0 µV decrease in median peak-to-peak voltage (−13.8% change, figure 10(a)), a 3.6 µV decrease in the median standard deviation of the noise (−25.6% change, figure 10(b)), and an increase in SNR of 0.67 (13.0% change, figure 10(c)). The change in cutoff frequency had a minimal effect on the median number of detected unsorted threshold crossings (−1.8% change, figure 10(d)).

3.5. Bidirectional BCI control

The complete artifact rejection scheme was validated online in a bidirectional BCI control task where the participant was required to transfer an object across a table as many times as possible in 2 min using a BCI-controlled robot arm. Microstimulation delivered through a group of eight electrodes was triggered by the reaction torque measured at the index finger motor as the robot hand contacted the object.
Stimulation electrodes and amplitudes were chosen to generate large artifacts and were not necessarily selected to provide feedback that would lead to improvements in the task. Task performance for each condition (no stimulation, stimulation without blanking, and stimulation with blanking for both the fourth-order 250 Hz high-pass filter and first-order 750 Hz high-pass filter) is summarized in figure 11, and representative trials with and without artifact rejection can be seen in supplementary video 1. A one-way analysis of variance revealed a significant difference in number of transfers between task conditions ($F(5,12) = 75.64$, $p < 0.001$). Post-hoc Tukey tests revealed that the type of filter caused no significant difference in performance without stimulation ($p = 0.462$). Further, when the full artifact rejection scheme was implemented, including the first-order filter, task performance with stimulation was not significantly different than both the no stimulation fourth-order filter condition ($p = 0.761$) and the no stimulation first-order filter condition ($p = 0.994$). When artifact rejection was not included, stimulation significantly impaired performance compared to all other conditions ($p < 0.001$) to the point that the task was essentially impossible. A mean ($\pm$ standard deviation) of 3766 $\pm$ 322 pulses (i.e. 31% of the trial) were delivered during each trial with artifact rejection. 3411 $\pm$ 1619 pulses (i.e. 28% of the trial) were delivered during all other stimulation trials.

Figure 9. Distribution of spikes across both recording arrays during the inter-stimulus interval, compared against spiking in the absence of stimulation. Bars denote mean and 95% confidence interval (error bars) during stimulation. Dotted line (mean) and shaded region (95% confidence interval) show spiking distribution during trials without stimulation. (a) Fourth-order 750 Hz filter without signal blanking. The distribution is dominated by artifact. Inset highlights that the distribution remains corrupted by filter ringing between 5 and 9 ms. (b) First-order 250 Hz filter with signal blanking. Spikes are absent during the blanking period but occur regularly in the time between stimulation pulses.

Figure 10. Relationship between first-order high-pass cutoff frequency and signal quality. Black lines and gray areas indicate median and interquartile range of units studied. Blue asterisks highlight 750 Hz. Red x and error bars denote median and interquartile range for a fourth-order Butterworth 250 Hz high-pass filter. (a) Peak-to-peak voltage decreases with increasing cutoff frequency. (b) Noise estimate decreases with increasing cutoff frequency. (c) Overall SNR increases with cutoff frequency. (d) Uns sorted threshold crossings remain stable between 250 Hz and 750 Hz.
During trials without artifact rejection, the participant typically had no volitional control of the arm during ICMS. In some of these trials, the hand continued to grasp the object, resulting in stimulation pulses that continued for the maximum 15 s. After this, the electrodes were temporarily, but automatically disabled for the subsequent 15 s due to a 50% duty cycle enforced to ensure a safe ICMS protocol. During these 15 s periods, the task appeared to revert to the no stimulation condition, allowing the participant to complete a small number of transfers, as seen in the two ICMS conditions without signal blanking in figure 11. During the fourth-order filter condition without signal blanking, the artifact-corrupted signal frequently resulted in a decoded velocity towards the top-right of the workspace. While the participant reported that he had little-to-no volitional control during stimulation, the object would sometimes drop out of the hand in the correct side of the workspace, or would be held over the correct side of the workspace until the 15 s duty cycle elapsed and he could drop the object in the correct location without ICMS.

This strategy permitted an average of five transfers per trial during this condition.

4. Discussion

We sought to implement a scheme for ICMS artifact rejection that would enable us to develop a bidirectional brain computer interface. We were able to accomplish this goal by making only minor modifications to an FDA-cleared clinical system. By reducing the high-pass Butterworth digital spike filter order from fourth-order to first-order, we eliminated filter ringing that occurred in response to the large stimulus artifact. Furthermore, we decreased the settling time of the filter output after stimulation by increasing the cutoff frequency from 250 Hz to 750 Hz. By blanking the signal prior to digital filtering, we were able to reduce the size of the perturbation seen by the filters and further reduce the magnitude of the filter’s response to stimulation. Ultimately, the goal of stimulus artifact reduction is to enable spike recordings for BCI control. Our system allows for spike detection as soon as 740 µs after the offset of each stimulus pulse. The change in filter design also resulted in an overall increase in the SNR. As a final functional test of this system, we demonstrated that the introduction of ICMS did not impair BCI performance with the artifact rejection scheme in place.

The proposed artifact rejection scheme is simple to implement with standard clinical BCI hardware and has few parameters, allowing it to be easily tuned for different participants and experiments. The blanking duration can be varied to favor preserving neural data or eliminating artifact from the raw signal. The high-pass filter cutoff frequency can also be tuned to balance peak-to-peak voltage with fast settling times and increased SNR, as demonstrated in figure 10. While the use of blanking places upper limits on the stimulation rate, the stimulation parameter space is largely unconstrained. The technique allows many pulse timing patterns to be used, unlike past approaches that limited stimulation to bursting patterns [13]. The scheme also works with monopolar microstimulation and makes no assumptions about amplifier saturation, unlike proposals that necessitate the use of bipolar stimulation to localize the magnitude and volume affected by artifact [20]. Alternative approaches to artifact rejection often rely on complex modeling and signal processing to estimate and subtract templates, or even noncausal algorithms [17]. These methods may be useful tools for offline analysis, but may be difficult or impossible to adapt to a robust real-time system where the exact configuration of electrodes and stimulus amplitudes might vary moment-by-moment. When continuous high-bandwidth raw data is collected during stimulation, such as in the development phase of this study, a variety of methods can be applied for offline analysis that would be challenging to implement in real-time due to limitations related to hardware and the need for fast and causal signal processing.

There are several limitations of our proposed scheme. Notably, the use of signal blanking results in a loss of neural data. When stimulating at 100 Hz, a total blanking duration of 1500 µs yields a 15% loss of neural data. In some cases, it may be possible to reduce the signal blanking duration and recover additional spikes. For example, using shorter, symmetric stimulus pulses is likely to be an effective method of reducing the blanking duration. From a practical BCI perspective, we have found that this loss of data has a negligible impact on performance and the method is very effective as long as synchronous stimulus pulses are used with a sufficiently low pulse rate. A model presented by Young et al. [19] suggests that decoder SNR is expected to drop off with the square root of the fraction of blanked data (SNR_{blanked} = SNR (\sqrt{1 - b}), b = fraction of blanked data). This model predicts that our decoder experiences a 7.8% decrease in SNR during 100 Hz stimulation, but this theoretical decline in decoder performance did not manifest as a measurable decrease in task performance. This theoretical reduction in SNR requires that stimulus pulses be delivered synchronously on all channels. If stimulation were to occur in an asynchronous fashion across multiple channels, as might occur with biomimetic pulse trains [30], all of the available signal would be effectively blanked out, eliminating all recordings. In this case, another solution such as amplifiers with a higher input range that can fully sample both artifacts and spike signals, or different modalities of recording and stimulation, such as optical methods, may be required.

Another limitation of the proposed approach is that the parameters must be tuned for worst-case scenarios since the implementation is global and blanks all channels for an equal duration for each stimulation pulse, regardless of artifact amplitude. When stimulation currents are low, the artifact amplitude is smaller and spikes can be recovered much closer to the offset of stimulation. A system that automatically adjusts blanking time based on stimulation amplitude or electrode location could theoretically perform better, but at the expense of added complexity and less generalizability. A simpler modification that may improve performance would be to linearly interpolate during the blanking period rather than holding a constant value. This computation would require a processing delay equal to at least the blanking duration and could not be implemented sample-by-sample in real-time. However, this approach would eliminate discontinuities at the
offset of blanking. In this case the blanking duration could theoretically be reduced to the duration of the stimulus pulse plus any transients caused by the analog filters. This method cannot be easily implemented with our hardware, but could be implemented by processing the raw digital signals for spike extraction in software. Our method takes advantage of the capabilities of the NSPs, but this limited our ability to modify the default signal processing scheme. Processing the raw signals in software would permit much more flexibility to preprocess the signals before applying a digital filter, much like when raw data is processed offline.

Finally, first-order filters exhibit undesirable characteristics in the frequency domain. The filter’s frequency response has a wide transition band with gradual roll-off, and thus a large range of frequencies below the cutoff frequency are only partially attenuated. This can be seen in the filter output during stimulation; low-frequency responses to stimulation are not entirely filtered out and thus low frequency drift can be observed in the mean filtered signal during stimulation. While the amplitude of this variation is small, it does bring the mean signal closer to or further from the threshold at various time points, which may affect the sensitivity of detecting low SNR peri-threshold spikes in a time-varying manner. For example, a slight increase in spike rate can be observed from 1.5 to 3.0 ms in figure 9(b). Further analysis is required to identify the source of this increase, which could be a neural response to the stimulation, or could be an increased sensitivity to noise due to the first-order filter response after the blanking period. Despite the undesirable frequency domain characteristics of the first-order filter, we chose this filter due to its favorable time-domain characteristics, specifically the absence of ringing in the impulse response. Because the increase in high-pass cutoff frequency resulted in a faster settling time, it is possible that some ringing can be tolerated from a higher order Butterworth filter if the cutoff frequency is increased to yield a sufficiently fast settling time. This would result in a steeper roll-off and better attenuation of low frequencies, but at the expense of attenuating lower frequencies within the spike power band.

Several unanswered questions and future steps may yield further improvements to the system. First, it is unclear what effect the changes in filter order and cutoff frequency have on unit discriminability. While SNR increases with high-pass cutoff frequency, peak-to-peak voltage decreases. This decrease in peak-to-peak voltage may affect the ability to distinguish multiple units. In recent years, spike sorting has fallen out of favor in the BCI community [21, 24, 31–33], but this may be of concern to groups interested in studying single unit activity. Second, only IIR Butterworth filters were investigated in this study. FIR filters, by definition, can be defined to have a short impulse response and may outperform the first-order IIR filter used here. However, FIR filters typically require many coefficients to perform well in the frequency domain. Finally, modifying the scheme to linearly interpolate during the blanking period, while adding complexity to the system, is likely to yield significant performance gains without introducing more parameters requiring tuning.

5. Conclusion

The presented system demonstrates that neural spike signals can be recorded during ICMS of an adjacent brain area by performing digital blanking and filtering operations in real-time. This simple modification is a robust method to reject all stimulus artifacts while preserving enough neural data to maintain BCI performance during ICMS. This artifact rejection system is a critical component of a bidirectional BCI system, and will allow for future work investigating the use of ICMS to provide somatosensory feedback to BCI users.

Acknowledgments

This work was supported by the Defense Advanced Research Projects Agency (DARPA) and Space and Naval Warfare Systems Center Pacific (SSC Pacific) under Contract No. N66001-16-C-4051 and the Revolutionizing Prosthetics program under Contract No. N66001-10-C-4056. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA, SSC Pacific or the United States Government.

The authors would like to thank Nathan Copeland for his continued dedication to this study. The authors would also like to thank W Max Maguire, Elizabeth Tyler-Kabara, Michael Boninger, John Downey, Cody Wabiszewski, Angelica Herrera, and Christopher Hughes for their assistance with data collection and technical discussions. The authors would also like to thank Scott Lempka for providing data from Lempka et al [23] that aided in interpretation of our results, and Patrick Loughlin for discussions and advice about digital signal processing. The authors have confirmed that any identifiable participants in this study have given their consent for publication. RF is an employee of Blackrock Microsystems LLC.
ORCID iDs
Jeffrey M Weiss https://orcid.org/0000-0003-1332-674X
Sharlene N Flesher https://orcid.org/0000-0003-0740-0018
Robert Franklin https://orcid.org/0000-0001-8968-9390
Jennifer L Collinger https://orcid.org/0000-0002-4517-5395
Robert A Gaunt https://orcid.org/0000-0001-6202-5818

References
[1] Collinger J L, Woodliger B, Downey J E, Wang W, Tyler-Kabara E C, Weber D J, McMorland A J C, Velliste M, Boninger M L and Schwartz A B 2013 High-performance neuroprosthetic control by an individual with tetraplegia Lancet 381 557–64
[2] Woodliger B, Downey J E, Tyler-Kabara E C, Schwartz A B, Boninger M L and Collinger J L 2014 Ten-dimensional anthropomorphic arm control in a human brain-machine interface: difficulties, solutions, and limitations J. Neural Eng. 12 016011
[3] Johansson R S, Häger C and Bäckström L 1992 Somatosensory control of precision grip during unpredictable pulling loads Exp. Brain Res. 89 204–13
[4] Monzée J, Lamarre Y and Smith A M 2003 The effects of digital anesthesia on force control using a precision grip J. Neurophysiol. 89 672–83
[5] Abbott A 2006 Neuroprosthetics: in search of the sixth sense Nature 442 125–7
[6] Bensmaia S J and Miller L E 2014 Restoring sensorimotor function through intracortical interfaces: progress and looming challenges Nat. Rev. Neurosci. 15 313–25
[7] Fitzsimmons N A, Drake W, Hanson T L, Lebedev M A and Nicolelis M A L 2007 Primate reaching cued by multichannel spatiotemporal cortical microstimulation J. Neurosci. 27 5593–602
[8] Lebedev M A and Nicolelis M A L 2006 Brain–machine interfaces: past, present and future Trends Neurosci. 29 536–46
[9] Mussa-Ivaldi F 2003 Brain–machine interfaces: computational demands and clinical needs meet basic neuroscience Trends Neurosci. 26 329–34
[10] Venkatraman S, Elkabany K, Long J D, Yao Y and Carmena J M 2009 A system for neural recording and closed-loop intracortical microstimulation in awake rodents IEEE Trans. Biomed. Eng. 56 15–22
[11] Weber D J, Friesen R and Miller L E 2012 Interfacing the somatosensory system to restore touch and proprioception: essential considerations J. Mot. Behav. 44 403–18
[12] Flesher S N, Collinger J L, Foldes S T, Weiss J M, Downey J E, Tyler-Kabara E C, Bensmaia S J, Schwartz A B, Boninger M L and Gaunt R A 2016 Intracortical microstimulation of human somatosensory cortex Sci. Transl. Med. 8 361ra141
[13] O’Doherty J E, Lebedev M A, Ilft P J, Zhuang K Z, Shokur S, Bleuler H and Nicolelis M A L 2011 Active tactile exploration using a brain–machine–brain interface Nature 479 228–31
[14] Klaes C, Shi Y, Kellis S, Minxha J, Revechshk B and Andersen R A 2014 A cognitive neuroprosthetic that uses cortical stimulation for somatosensory feedback J. Neural Eng. 11 056024
[15] Limnuson K, Lu H, Chiel H J and Mohseni P 2014 Real-time stimulus artifact rejection via template subtraction IEEE Trans. Biomed. Circuits Syst. 8 391–400
[16] Mena G E, Grosberg L E, Madugula S, Hottonoy P, Litke A, Cunningham J, Chichilnisky E J and Paninski L 2017 Electrical stimulus artifact cancellation and neural spike detection on large multi-electrode arrays PLoS Comput. Biol. 13 e1005842
[17] O’Shea D J and Shenoy K V 2018 ERAASR: an algorithm for removing electrical stimulation artifacts from multielectrode array recordings J. Neural Eng. 15 026020
[18] Wogenaar D A and Potter S M 2002 Real-time multi-channel stimulus artifact suppression by local curve fitting J. Neurosci. Methods 120 113–20
[19] Young D, Willet F, Membarg W D, Murphy B, Walter B, Sweet J A, Miller J, Hochberg L, Kirsch R F and Ajiboye A B 2018 Signal processing methods for reducing artifacts in microelectrode brain recordings caused by functional electrical stimulation J. Neural Eng. 15 0–21
[20] O’Doherty J E and Sabes P N 2016 Towards artificial proprioception for brain-machine interfaces 2016 Neuroscience Meeting Planner (San Diego, CA: Society for Neuroscience) p 288.08 (www.abstractsonline.com/pp8/#!/4071/presentation/32952)
[21] Fraser G W, Chase S M, Whitford A and Schwartz A B 2009 Control of a brain–computer interface without spike sorting J. Neural Eng. 6 055004
[22] Simeral J D, Kim S P, Black M J, Donoghue J P and Hochberg L R 2011 Neural control of cursor trajectory and click by a human with tetraplegia 1000 days after implant of an intracortical microelectrode array J. Neural Eng. 8 025027
[23] Lempka S F, Johnson M D, Moffitt M A, Otto K J, Kipke D R and McIntyre C C 2011 Theoretical analysis of intracortical microelectrode recordings J. Neural Eng. 8 045006
[24] Christie B P, Tat D M, Irwin Z T, Gilja V, Nuyujukian P, Foster J D, Ryu S I, Shenoy K V, Thompson D E and Chestruk C A 2015 Comparison of spike sorting and thresholding of voltage waveforms for intracortical brain–machine interface performance J. Neural Eng. 12 016009
[25] Downey J E, Schwed N, Chase S M, Schwartz A B and Collinger J L 2018 Intracortical recording stability in human brain–computer interface users J. Neural Eng. 15 046016
[26] Downey J E, Brane L, Gaunt R A, Tyler-Kabara E C, Boninger M L and Collinger J L 2017 Motor cortical activity changes during neuroprosthetic-controlled object interaction Sci. Rep. 7 16947
[27] Stark E and Abeles M 2007 Predicting movement from multitunit activity J. Neurosci. 27 8387–94
[28] Quian Quiroga R 2009 What is the real shape of extracellular spikes? J. Neurosci. Methods 177 194–8
[29] Irwin Z T et al 2016 Enabling low-power, multi-modal neural interfaces through a common, low-bandwidth feature space IEEE Trans. Neural Syst. Rehabil. Eng. 24 521–31
[30] Saal H P and Bensmaia S J 2015 Biomimetic approaches to bionic touch through a peripheral nerve interface Neuropsychologia 79 344–53
[31] Chestek C A et al 2011 Long-term stability of neural prosthetic control signals from silicon cortical arrays in rhesus macaque motor cortex J. Neural Eng. 8 045005
[32] Oby E R, Perel S, Sudalter P T, Ruff D A, Mischel J L, Montez D F, Cohen M R, Batista P A and Chase S M 2016 Extracellular voltage threshold settings can be tuned for optimal encoding of movement and stimulus parameters J. Neural Eng. 13 036009
[33] Trautmann E M, Stavisky S D, Lahiri S, Ames K C, Kaufman M T, Ryu S I, Ganguli S and Shenoy K V 2017 Accurate estimation of neural population dynamics without spike sorting (bioRxiv:229252)