A Semi-Autonomous Neuroprosthesis using Probabilistic Time-Series Inference

by

Guy Hotson

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

December, 2016

© Guy Hotson 2016

All rights reserved
Abstract

Upper limb neuroprosthesis is a rapidly progressing technology with the potential to restore function to victims of severe paralysis by enabling users to control robotic systems with their neural signals. Great strides have been made in directly mapping users’ cortical activity to control of the individual degrees of freedom of computer cursors and robotic end-effectors. After giving an overview of the field (chapter 1), the first half of this thesis (chapters 2-4) will focus on improvements and limitations to the level of real-time neural control attainable by human subjects with electrocorticography (ECoG). This entails offline analysis of the information available in the ipsilateral hemisphere of injured neural substrates (chapter 2), the first demonstration of independent, simultaneous real-time control over reaching and grasping using ECoG signals (chapter 3), and the first demonstration of real-time neural control of individual finger movements (chapter 4).

The second half of this thesis (chapters 5-7) will detail how environmental sensors can provide complimentary information to the neural recordings within the novel Hybrid Augmented Reality Multimodal Operation Neural Integration Environment
(HARMONIE) to yield robust neuroprostheses. HARMONIE fuses information extracted from RGB-D cameras and eye-tracking with control signals extracted from the user’s neural signals. The system was first demonstrated with neural signals serving as a ”go” signal for the autonomous limb movements (chapter 5). However, this deprives the user of flexibility to modify trajectories to accomplish tasks requiring user feedback. Chapter 6 outlines a framework for the continuous probabilistic fusion of neural signals with movement trajectories derived from prior knowledge of movement primitives. This was accomplished by using dynamic movement primitives to model the 3D endpoint trajectories generated by manipulating various objects, then using a switching unscented Kalman filter to continuously arbitrate between the 3D endpoint kinematics predicted by the dynamic movement primitives and control derived from neural signals. This framework was validated offline with non-human primate data (chapter 6), then implemented in a real-time control environment (chapter 7). This work provides the first demonstration of how patients can almost immediately gain control over reaching, grasping, and individual finger control without the need for long-term cortical adaptation. Further, it shows how this immediate neural control can be leveraged in a functionally useful manner by incorporating computer vision and intelligent robotics.

Primary Reader: Nitish Thakor

Secondary Reader: Brock Wester
Acknowledgments

I would like to thank everyone in Dr. Thakor’s BCI group for their input and support throughout my PhD. Geoffrey Newman for always making sure we got to coffee hour in time for the best snacks. Ryan Smith for not being afraid to give critical feedback. Luke Osborn for using his southern hospitality to shepherd undergrads and high school students away from messing with my stuff. Joseph Betthauser for working in lab at all hours of the night, scaring away any potential burglars. Heather Benz for practicing her judgmental mom look to make sure I never ducked out of lab early. I would also like to thank Dr. Crone’s lab. Matt Fifer for not letting his mid-PhD pessimism get in the way of being an amazing mentor and team-player. Griffin Milsap for his mishaps with soylent and his contagious excitement for new technologies. Kyle Rupp for his house parties and conversations about ketogenic diets. David McMullen for making debates about mundane terminology entertaining. Yujing Wang for only complaining a little bit whenever she had to sacrifice countless hours making sure we got good functional maps of ECoG activation. Tessy Lal for bearing the burden of becoming the sole motor ECoG tester when I’m gone. And Max Collard for all
ACKNOWLEDGMENTS

his highly caffeinated math and physics rants. Next, I’d like to thank all the people outside the Thakor/Crone lab ecosystem that helped make these studies happen. All the people involved at the APL who developed the MPL, various software components, and supported patient testing. In particular, Brock Wester for his great mentorship and enthusiasm, and Kapil Katyal for his long hours getting everything functioning. My thanks go out to Dr. Schieber’s lab for graciously letting me analyze some of their monkey data. Finally, I’d like to thank Molly Kim for all her emotional support and affection, my parents for supporting me in all my endeavors, and my brother for inspiring me to get my PhD in the first place. Because if he could do it, so could I.
Dedication

This thesis is dedicated to my sister, who taught me how to work hard and fight until the end, no matter what the world throws at you.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abstract</strong></td>
<td>ii</td>
</tr>
<tr>
<td><strong>Acknowledgments</strong></td>
<td>v</td>
</tr>
<tr>
<td><strong>List of Tables</strong></td>
<td>xvi</td>
</tr>
<tr>
<td><strong>List of Figures</strong></td>
<td>xvii</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Overview</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Motivation</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Upper Limb Biosignals</td>
<td>3</td>
</tr>
<tr>
<td>1.3.1 Electromyography (EMG)</td>
<td>4</td>
</tr>
<tr>
<td>1.3.2 Electroencephalography (EEG)</td>
<td>5</td>
</tr>
<tr>
<td>1.3.3 Microelectrode Arrays (MEAs)</td>
<td>6</td>
</tr>
<tr>
<td>1.3.4 Electrocorticography (ECoG)</td>
<td>7</td>
</tr>
<tr>
<td>1.4 Overview of Upper Limb Neuroprosthetic Control Strategies</td>
<td>10</td>
</tr>
</tbody>
</table>
CONTENTS

1.4.1 Direct Neural Control ............................................. 10
1.4.2 Hybrid Control ..................................................... 11
1.4.3 Shared Control ..................................................... 11
1.5 Dissertation Organization ........................................... 14

2 Coarse Electrocorticographic Decoding of Ipsilateral Reach in Patients with Brain Lesions 17

2.1 Overview ............................................................ 17
2.2 Materials and Methods .............................................. 21
  2.2.1 Data Acquisition ................................................. 21
  2.2.2 Experimental Paradigm ......................................... 24
  2.2.3 Neural Data Feature Extraction ............................... 26
  2.2.4 Decoding Model Construction and Evaluation ........... 28
  2.2.5 Principal Component Analysis ................................ 29
  2.2.6 Model Input Feature Selection ............................... 30
2.3 Results .............................................................. 31
  2.3.1 Feature Selection ................................................. 31
  2.3.2 Multiple Inputs .................................................. 33
2.4 Discussion .......................................................... 38

3 Intracranial EEG Control of Reaching and Grasping with the Modular Prosthetic Limb 45
CONTENTS

3.1 Overview ......................................................... 45

3.2 Materials and Methods ......................................... 47
  3.2.1 Subject Info .............................................. 47
  3.2.2 Neural Signal Acquisition ................................. 49
  3.2.3 Experimental Procedures ................................. 50
  3.2.4 iEEG Electrode Evaluation ............................... 52
  3.2.5 BMI Model Training ...................................... 53
  3.2.6 JHU/APL Modular Prosthetic Limb ....................... 55
  3.2.7 Online Testing ........................................... 56
  3.2.8 Quantitative Evaluation of Control ...................... 57

3.3 Results ......................................................... 58

3.4 Discussion ..................................................... 64

3.5 Study Acknowledgments ........................................ 68

4 Individual Finger Control of the Modular Prosthetic Limb using High-Density Electrocorticography in a Human Subject 70

  4.1 Overview ..................................................... 70

  4.2 Methods ....................................................... 73
  4.2.1 Subject Info ............................................... 73
  4.2.2 Offline Experimental Testing ............................. 74
  4.2.3 Vibrotactile Stimulation ................................ 75
  4.2.4 Finger Tapping ............................................ 76
CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.5 BMI Classifier Training</td>
<td>77</td>
</tr>
<tr>
<td>4.2.6 Online Finger BMI Testing</td>
<td>79</td>
</tr>
<tr>
<td>4.2.7 BMI Control Evaluation</td>
<td>79</td>
</tr>
<tr>
<td>4.2.8 Offline Analysis</td>
<td>81</td>
</tr>
<tr>
<td>4.3 Results</td>
<td>84</td>
</tr>
<tr>
<td>4.3.1 BMI Electrode Selection</td>
<td>84</td>
</tr>
<tr>
<td>4.3.2 Online Testing</td>
<td>85</td>
</tr>
<tr>
<td>4.3.3 Offline Analysis</td>
<td>89</td>
</tr>
<tr>
<td>4.4 Discussion</td>
<td>94</td>
</tr>
<tr>
<td>4.5 Conclusion</td>
<td>99</td>
</tr>
</tbody>
</table>

5 Demonstration of a Semi-Autonomous Hybrid Brain-Machine Interface using Human Intracranial EEG, Eye Tracking, and Computer Vision

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Overview</td>
<td>101</td>
</tr>
<tr>
<td>5.2 Materials and Methods</td>
<td>105</td>
</tr>
<tr>
<td>5.2.1 Subject Info</td>
<td>105</td>
</tr>
<tr>
<td>5.2.2 Neural Signal Acquisition</td>
<td>107</td>
</tr>
<tr>
<td>5.2.3 Training</td>
<td>109</td>
</tr>
<tr>
<td>5.2.4 Functional Localization and Neural Signal Analysis</td>
<td>109</td>
</tr>
<tr>
<td>5.2.5 BMI Model Training</td>
<td>110</td>
</tr>
<tr>
<td>5.2.6 Testing Paradigm</td>
<td>112</td>
</tr>
</tbody>
</table>
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.7</td>
<td>Computer Vision and Eye Tracking</td>
<td>113</td>
</tr>
<tr>
<td>5.2.8</td>
<td>Modular Prosthetic Limb</td>
<td>116</td>
</tr>
<tr>
<td>5.2.9</td>
<td>Online BMI Testing</td>
<td>118</td>
</tr>
<tr>
<td>5.3</td>
<td>Results</td>
<td>120</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Online Global Evaluation</td>
<td>121</td>
</tr>
<tr>
<td>5.3.2</td>
<td>iEEG-Based Decoding Evaluation</td>
<td>122</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Computer Vision and Eye Tracking</td>
<td>124</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Modular Prosthetic Limb Control</td>
<td>126</td>
</tr>
<tr>
<td>5.4</td>
<td>Discussion</td>
<td>127</td>
</tr>
<tr>
<td>5.5</td>
<td>Footnotes</td>
<td>135</td>
</tr>
</tbody>
</table>

## 6 High Precision Neural Decoding of Complex Movement Trajectories using Recursive Bayesian Estimation with Dynamic Movement Primitives

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>INTRODUCTION</td>
<td>136</td>
</tr>
<tr>
<td>6.2</td>
<td>MOTION PLANNING ARCHITECTURE</td>
<td>139</td>
</tr>
<tr>
<td>6.2.1</td>
<td>System Overview</td>
<td>139</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Dynamic Movement Primitives</td>
<td>142</td>
</tr>
<tr>
<td>6.2.3</td>
<td>HMM Temporal Update</td>
<td>143</td>
</tr>
<tr>
<td>6.2.4</td>
<td>HMM Measurement Update</td>
<td>144</td>
</tr>
<tr>
<td>6.2.5</td>
<td>UKF Temporal Update</td>
<td>146</td>
</tr>
<tr>
<td>6.2.6</td>
<td>Kalman Filter Measurement Update</td>
<td>148</td>
</tr>
</tbody>
</table>
## CONTENTS

6.3 EXPERIMENTAL VALIDATION .............................................. 149
  6.3.1 Experimental Details .............................................. 149
  6.3.2 Neural Modeling .................................................... 151
  6.3.3 Action Prediction Fitting ........................................ 151
  6.3.4 Dynamic Movement Primitive Fitting ......................... 153
  6.3.5 Action Prediction Results ....................................... 154
  6.3.6 Trajectory Prediction Results ................................... 156
  6.3.7 Global System Evaluation Details .............................. 156
  6.3.8 Global Performance Results ...................................... 158

6.4 DISCUSSION ............................................................. 159

6.5 CONCLUSION ............................................................. 161

7 Realtime Recursive Bayesian Inference for a Semi-Autonomous Neu-
roprosthetic ................................................................. 163
  7.1 Introduction ........................................................... 163
  7.2 Methods ............................................................... 166
    7.2.1 Experimental Setup ............................................. 166
      7.2.1.1 Training ....................................................... 166
    7.2.2 System Overview ............................................... 167
    7.2.3 EMG Recording ................................................... 167
    7.2.4 ECoG Recording .................................................. 169
    7.2.5 IMU Recording .................................................... 169
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Subject Demographic and Clinical Information</td>
<td>22</td>
</tr>
<tr>
<td>2.2</td>
<td>Correlation between actual and decoded kinematics with and without PCA</td>
<td>38</td>
</tr>
<tr>
<td>4.1</td>
<td>Experimental Overview</td>
<td>76</td>
</tr>
<tr>
<td>6.1</td>
<td>Variable Descriptions</td>
<td>142</td>
</tr>
<tr>
<td>6.2</td>
<td>Movement Trajectory Decoding Performance</td>
<td>156</td>
</tr>
<tr>
<td>6.3</td>
<td>Global Decoding Performance</td>
<td>159</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Example implantation of ECoG electrodes ................. 8

2.1 Reconstructions of electrode placements with ESM results. The grids shown are the subset of implanted electrodes that were recorded from during this study. The green highlighted areas correspond to regions of cortical lesions. The lesion in subject 3 could not be seen on the brain surface rendering because it was located beneath the surface of the brain. Colored rectangles joining electrodes imply that bipolar stimulation was applied to that pair of electrodes. In subject 1, several electrode pairs were further investigated by performing unipolar stimulation relative to a distant reference electrode. The results of this unipolar stimulation are shown with colored circles surrounding the electrodes. ................................................................. 23

2.2 Presurgical MRI and brain reconstructions. Reconstructions are shown for subject 1 (first row) subject 2 (bottom left and middle) and subject 3 (bottom right). The previous resection margins anterior to the pre-central gyrus in subject 1 are highlighted in green in the upper right. Superior oblique and top axial views of the reconstruction for subject 2 show the lesion from different viewpoints. Pre-surgical MRI of subject 3 reveals a lesion of posterior left insula also involving left internal capsule. ............................................................... 24

2.3 Subject Kinematics and Experimental Paradigm. (A) Shows the weights given to each of the three dimensions (height, depth, and lateral) when computing the first PC of the movement kinematics. Each point represents the PC weight calculated using a different cross-validation. (B-D) Show different views of the experimental paradigm and example kinematics from one of the subjects’ sessions. The red dashed line represents the first PC plotted in Cartesian space for that same session. 26
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4</td>
<td>Example Principal Components of Kinematics. A representative example of the first, second, and third principal components of the kinematics from each subject, overlaid with the original kinematics.</td>
<td>32</td>
</tr>
<tr>
<td>2.5</td>
<td>Anatomical patterns of average correlations between representative ECoG signal features and the first PC of the hand position. Correlations were calculated at each feature’s best lag (within 1 second) relative to the movement data. Color-coded circles corresponding to maximum correlations, averaged across experimental sessions, are shown for each subject.</td>
<td>33</td>
</tr>
<tr>
<td>2.6</td>
<td>Correlations between kinematics and neural features at different time lags relative to one another. Each row corresponds to one of eight extracted features from one electrode, resulting in 384, 352, and 360 rows for subjects 1, 2, and 3 respectively. The x-axis corresponds to the lag between the neural feature and the reach position. A negative lag represents changes in the neural features occurring before the corresponding kinematics, and a positive lag represents feature changes after the kinematics. Rows are ordered by the magnitude of their peak correlation. Correlations between the hand kinematics and the features were calculated every 50 ms between leads or lags of one second.</td>
<td>34</td>
</tr>
<tr>
<td>2.7</td>
<td>Relationship between number of model inputs (i.e., ECoG signal features) and model performance. Mean decoding accuracies of the first PC of movement across cross-validations and sessions are displayed for features chosen in decreasing order of correlation. Shading corresponds to the 95th percentile confidence interval for the mean.</td>
<td>35</td>
</tr>
<tr>
<td>2.8</td>
<td>Peak decoding accuracy for each ECoG signal feature type. Between 1 and 40 recording sites were selected for each signal feature type, and the means were calculated across cross-validations and sessions. The peak accuracies are displayed here with error bars corresponding to the standard error of the mean. Numbers over each bar indicate the minimal number of features required for statistical saturation.</td>
<td>37</td>
</tr>
<tr>
<td>2.9</td>
<td>Decoding model performance across sessions for the first PC of movement. Distributions are displayed for the five cross-validations of the performance for one, two, and nine input features. S1-1 stands for subject 1 session 1, S1-2 stands for subject 1 session 2, etc. The maximum of the 1024 chance decoding attempts for all five folds with shuffled neural data is shown with an asterisk.</td>
<td>39</td>
</tr>
</tbody>
</table>
3.1 Functional mapping of cue-averaged task-related high gamma activity in training set. (A) Reconstruction of the implanted grid location for Subject 1 is depicted; the electrode used for reaching (number 25) is highlighted in red and corresponds to the channel circled in red in the activation maps below, while the electrode used for grasping (number 11) is highlighted in blue and similarly corresponds to the electrode circled in blue below; the central sulcus is highlighted in green. (B) Reconstruction of the depth electrodes implanted in right hemisphere of Subject 2; electrodes used for reaching highlighted in red, electrodes used for grasping highlighted in blue (transparent medial view in inset).

3.2 Rasters of cue-averaged high gamma activity for Subject 1. The Reach and Grasp task maps display significant increases (red spectrum) or decreases (blue spectrum) in high gamma energy relative to baseline. Each row corresponds to a different iEEG electrode. Reaches vs Grasps is the result of a Wilcoxon test between two conditions for each (channel, time) pair with FDR correction for comparisons across multiple time points within each channel. Red corresponds to stronger activation for reach. Vertical lines are MO: movement onset; PT: pressed target; RT: released target; Home: return to home; Rest: released pressure bulb.

3.3 Schematics and photographs of experimental setup with MPL. (top) A schematic of the experimental setup is shown. Traces of the behavioral sensors, high gamma power, and MPL commands during a three trial segment are shown as an example. (A-C) The subject is seated on his hospital bed with his arm at rest on a pushbutton for reach offset detection. A laptop displayed a red bar indicating pressure exerted on the squeeze bulb. (A) In the background, the MPL is at its baseline state (rest posture). (B) The subject is executing a grasp movement, and (C) the subject is executing a reach movement.
3.4 The MPL software and hardware architecture consists of a distributed network of processors that include a Neural Fusion Unit (NFU), a Limb Controller (LC), 10 Small Motor Controllers (SMC), 4 Large Motor Controllers (LMC), and 3 Wrist Motor Controllers (WMC). The NFU is a processor capable of running on-board neural decoding and sensory stimulation algorithms for generation of limb motion commands. The LC is the main processor of the limb system and is responsible for receiving limb control commands, running high-level control algorithms, and coordinating the control of the individual motors in the system. The LMC, WMC, and SMC integrated software/hardware systems, which receive real-time data from temperature, torque, and position sensors located within each joint, are responsible for providing closed-loop position, velocity and torque control of the brushless DC motors in the limb system. In addition to internal polling, these joint sensors broadcast information to the user/experimenter along the MPL’s external communication CAN bus for data logging and external control processing.

3.5 Average power spectral density (PSD) relative to baseline, aligned to movement onset. (A) Reach and grasp electrodes are shown for Subject 1, and (B) two representative electrodes are shown for Subject 2. The first vertical dashed line corresponds to average audio cue onset. The solid line denotes movement onset (MO). In reach trials, the dashed lines after the solid line correspond to the average time of the reach completion (pressing target button, PT), release of the target button (RT), and return to home (resting on the home switch), from left to right. The dashed line in the grasp trials corresponds to average grasp completion time.

3.6 Limb performance accuracy metrics. (A, B) Accuracies are shown for reaching and grasping during trials where reach and grasp were executed simultaneously. (C, D) Reach and grasp accuracies are shown for reach and grasp only trials, respectively. The vertical dashed lines in A-D denote separate blocks. Distributions are summarized with boxplots of the peak sensitivities for grasps in Subject 1 (E), reaches in Subject 1 (F), grasps in Subject 2 (G), and reaches in Subject 2 (H). Bars above the boxplots with asterisks mark distributions with significantly different medians ($p < 0.05$, Wilcoxon test).

4.1 Reconstruction of high density ECoG electrode locations.

4.2 Flow diagram of the online BMI processing. Signals are rereferenced, high gamma is extracted and smoothed, and then a hierarchical classifier is used to determine if/which fingers should be moved.
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>Electrodes used for online BMI. Starred electrodes were selected for online BMI control. Sulci are accentuated in the inset for improved visibility, with the central sulcus highlighted in green. Post-hoc analysis showed the electrodes with gold stars contributed the most to decoding accuracy on the training set. Light blue electrodes showed significant activation during vibrotactile stimulation, and red electrodes were active during the motor task. The electrodes that were not available for the offline analysis are filled black. Purple outlines the interhemispheric fissure and red outlines the previously resected superior frontal gyrus.</td>
<td>85</td>
</tr>
<tr>
<td>4.4</td>
<td>Classification accuracies over time for movement and finger classification. Classifications were aggregated in 250 ms time bins, and accuracies were averaged across trials. The black dashed line depicts the average proportion of predictions that a window contained movement (i.e., versus rest). The solid black vertical line depicts movement onset.</td>
<td>86</td>
</tr>
<tr>
<td>4.5</td>
<td>Confusion matrices for finger classifications. The left column of matrices shows results for all five fingers, the right column shows results with pinky and ring fingers combined. The top row depicts the confusion matrices at the onset of the peak movement classification time, middle shows the average over the 1.5 seconds of peak movement detection time, and the bottom is any time within seven seconds of cue onset.</td>
<td>87</td>
</tr>
<tr>
<td>4.6</td>
<td>Average normalized MPL finger position of all five fingers during online BMI control (left), and with ring and pinky commands aggregated (orange, right). Solid lines depict the average position of the cued finger as a proportion of maximum flexion, where 1 corresponds to fully flexed and 0 corresponds to fully extended. The corresponding dashed lines show the average position of all the other fingers during the cued finger’s trials.</td>
<td>88</td>
</tr>
<tr>
<td>4.7</td>
<td>Somatotopic organization of the high gamma activation during the passive vibration task. Electrode color corresponds to the finger with the strongest high gamma activation. Intensity corresponds to the strength of the activation normalized relative to the average activation across the four other fingers. This provides an indication of how discriminable the respective finger is relative to all other fingers. Electrodes A and B were the two electrodes that contributed most to the decoding performance of the finger BMI based on training data.</td>
<td>90</td>
</tr>
</tbody>
</table>
4.8 Spectral activation during offline index finger tapping (top) and passive index finger vibration (bottom) for the two BMI electrodes that contributed most to performance with the training data (displayed in figure 4.7). For finger tapping (top row) the solid black line marks the onset (t=0) of any hand movement detected by the CyberGlove. The red curve is the average trace of the index finger sensors on the CyberGlove. For finger vibration (bottom row), the black line marks onset of the vibratory stimulus. Activation significantly exceeded baseline levels 160 ms and 48 ms before movement onset during the motor task (top row), and 48 ms after the onset of vibrotactile stimulation (bottom row).

4.9 Finger classification accuracy over time using CyberGlove and ECoG recordings during the motor task, and ECoG during vibrotactile stimulation. Motor task recordings are aligned to the onset of any movement by the subject’s hand (solid line), and vibrotactile stimulation is aligned to onset of the vibratory motor (solid line). The red, green, and blue dashed lines denote when decoding accuracy exceeded chance levels using motor task high gamma, vibrotactile stimulation high gamma, and motor task CyberGlove data. All decoding is causal, i.e. classification accuracy at time t is dependent only on samples collected at time t and earlier. Shading represents the 95% confidence interval of the mean.

5.1 HARMONIE combined BMI system block diagram. Components of the HARMONIE BMI system work together to allow the patient to perform a complex automated motor task. A patient implanted with iEEG electrodes is able to use eye tracking to visually select an object that has been segmented by computer vision. Patient initiates the task via their iEEG-BMI input. Task initiation and object information are used to control the MPL to perform the task.

5.2 Schematic of the HARMONIE experimental setup. Patient views a segmented scene sent from the computer vision device. Kinect sensor records both the depth information and RGB of the experimental workspace. Patient’s gaze is recorded by monitor-mounted eye tracking and movement intention is detected by iEEG-based BMI. MPL carries out the reach, grasp, and place motor task.
### Figure 5.3
Neural activation during reaches. Cue-averaged statistically significant high gamma modulations relative to pre-cue periods are plotted in a channel x time raster for Subject 1 with the median behavioral times marked for movement onset and target button press. Results are aligned relative to the audio cue, and show 1024 ms of the pre-cue period. The mean modulation across nine 16 ms time bins spanning 112 ms (centered on median movement onset time) is plotted on the patient’s brain reconstruction (B) for Subject 1, left hemisphere viewed from medial surface. Inter-hemispheric grid (IHG) ECoG electrodes 18, 19, and 26 (outlined in black boxes) displayed task-related modulations during this time period (B). Routine electrocortical stimulation caused movements in the right arm at IHG 18 and 26 and stiffening on the right side at IHG 19.

![Image](image.png)  
Observations:  
- Movement onset and target button press are marked.  
- Results are aligned with the audio cue.  
- Mean modulation across nine 16 ms time bins spanning 112 ms is plotted.  
- Electrode locations are marked on a brain reconstruction.  

### Figure 5.4
Four trial types of the HARMONIE system are shown. Blue lines represent the actual movement of Subject 1’s right hand, with an upward deflection marking the trial start (i.e., reach onset) when the subject lifted their hand off the home plate (zeroed when resting). Red stars demonstrate actual BMI prediction of a movement for Subject 1. The red filled areas represent the baseline window where no movement detections were expected, and the green areas represent time windows where we expect movement detections. (A) and (B) are representative of the successful trials. (C) shows an example of a trial where the baseline was discarded due to contamination from a previous movement. This baseline was not included in our accuracy analysis, but the uncontaminated active period surrounding movement onset was. (D) depicts a trial where both the baseline and the active period were classified incorrectly, resulting in a false positive and false negative.

### Figure 5.5
Offline demonstration of computer vision segmentation ability. Grey circles depict the system’s predicted size and location for each object. (A) shows that the object recognition software was able to simultaneously segment out eight spherical objects of varying sizes and materials. During online testing with the subject, only a single rubber basketball was used. (B) shows the potential for the system to segment out everyday lunch objects, such as an apple and a cup of yogurt.
6.1 Dynamic Bayesian network of motion prediction using computer vision and neural signals. Rectangles represent categorical (discrete) variables, circles are continuous. Blue is observed, and orange is inferred at each time step. Obj represents the type of selected object (possibly none), Act is the desired action to execute (e.g. rest or drink), Goal is the 3D endpoint the user is trying to reach, Kin is the current 3D position, velocity, and acceleration, and Neuro is the feature vector extracted from the neurological signals. Arrows denote conditional dependencies between variables.

6.2 Reach trajectories performed by the non-human primate recorded via optical tracking of the monkey’s wrist. The origin corresponds to the home location. Line colors correspond to the location of the target object. Shapes are centered where the monkey completed its trajectory.

6.3 Alternate angle of the optical recordings of the monkey’s reach trajectories.

6.4 State transition diagram for action prediction. Each object manipulation was modeled as occurring in two phases. The monkey began and ended each trial by resting its arm motionless.

6.5 Prediction of rest (grey), movement/object manipulation (light blue/green), and hold (orange/red) from neural data. Each row corresponds to a single trial (sorted by action type). Trials were aligned to movement onset (MO, solid line), and the median onset of the object hold period is shown with a vertical dashed line. Probabilities across action types (i.e. all blue/green and all red/orange labels) were subsequently summed together because each object only had one associated action. Green asterisks denote the 13 false positives transitioning from rest state to the first movement state, and black asterisks mark the 6 false positives transitioning from movement state 1 to 2, then back to state 1.

6.6 Flow diagram of the UKF-HMM algorithm. A prediction of the current action is first made based on the previous estimate of the action probabilities. This is then updated with the measurement of the neural signals, and the output is used to decide whether to move using the UKF, hold position, or rest. Dashed lines represent inputs from the previous time step.

6.7 Example trajectories (position relative to home location) for an example object-location pair (push). Line color corresponds to the three dimensions being decoded. The dashed lines are single trials, and thick solid lines are the averaged across all trials for that object-location pair.
LIST OF FIGURES

7.1 Features are extracted from the ECoG or EMG signals. These are used to predict the action and grasp. Computer vision with eye-tracking is used to determine the object type and location that the user intends to manipulate. This helps inform the HMM guiding the action/grasp predictions, as well as the DMP predictions. The DMP predictions are fused with the 3D velocity estimates derived from the ECoG features or IMU sensors. The movement information is then sent to the movement interface, which communicates with the MPL or virtual MPL. . . . . 168

7.2 Example augmented reality implementation of the system’s GUI. Top: a picture showing a user wearing a Hololens on the right looking at an apple. The video from the Hololens is depicted on the the screen to the left, with an apple identified through computer vision. Bottom: an image captured from the user’s view of the Hololens output. Here a user was viewing a cup, and it was labeled accordingly. . . . . . . . . . . . 179

8.1 Individual finger somatotopy during a passive vibration task. Color corresponds to the finger that showed the strongest response. The intensity signifies how much statistically stronger that finger responded than the pooled response of the rest of the fingers. . . . . . . . . . . . 181

8.2 Decoding results using LSTM’s for decoding. . . . . . . . . . . . . . 187
Chapter 1

Introduction

1.1 Overview

There is a great need for the restoration of upper limb function in populations with paralysis/amputation, yet the efficacy of current solutions is wanting. Within this thesis, I outline my work developing an intelligent upper limb neuroprosthetic. In Chapter 1, I give background of the field. In Chapters 2-4, I explore how electrocorticography (ECoG) can be leveraged to deliver state of the art control over a neuroprosthetic. In Chapters 5-7, I delve into how user control signals for neuroprosthetics can be combined with intelligent robotics to further improve system performance. Chapters 2-7 are all accompanied with separate abstracts, introductions, and conclusions to provide more specific overviews relevant to those chapters. Chapters 2-6 have all been developed into journal publications, as well as several
conference abstracts/papers [1, 2, 3, 4, 5, 6].

1.2 Motivation

In the United States alone, over 100,000 people are quadriplegic [7]. On a global scale, the reported incidence of spinal cord injury (SCI) has remained relatively constant between 10.4 and 83 per million inhabitants every year, with one out of three of these resulting in quadriplegia [8]. With a growing population and spinal cord injury rates refusing to flag, there is an increasing need to develop devices to enable individuals with profound motor disabilities to perform basic activities of daily living (ADL).

Many assistive devices have been designed to enable quadriplegics to accomplish various basic tasks. For example, electric wheelchairs can allow SCI patients to move autonomously, software applications such as Dasher facilitates the use of computers when conventional keyboards are not an option [9], eye-trackers such as the Tobii (Tobii Technology, Stockholm Sweden) can allow patients to control computer cursors, and speech synthesizers can allow patients without speech capabilities to communicate.

However, the control of such assistive devices is still very crude. The signals used by the disabled for control of devices like wheelchair are typically based on eye-tracking, speech recognition, head-tracking, or the rudimentary puff-sip or chin/mouth.
CHAPTER 1. INTRODUCTION

These signals all have their downsides. Overloading the eye with tasks such as cursor
control has many downfalls, such as the Midas touch problem of activating every-
thing you foveate upon [10], [11]. Speech recognition has limited accuracy, especially
in noisy environments differing from the training conditions [12], and would result in
patients being unable to converse with others while using their device. Furthermore,
control signals from head-tracking, puff-sip devices and the chin/mouth controller are
dependent on residual muscles being sufficiently intact, deliver only a small number
of independent control signals, and induce muscle fatigue [13]. Finally, none of these
signals are able to intuitively transfer cognitive intent or high-level goal into action.
In order to resolve the shortcomings of previous devices and potentially deliver a more
natural/intuitive assistive device, the field of neuroprosthetics has emerged.

1.3 Upper Limb Biosignals

Upper limb neuroprosthesis is an emerging technology beginning to see trans-
lational research for use with patients in need. In the early 20th century, it was
found that electrical activity from different muscle contractions could be recorded
with an oscilloscope [14]. Later, the investigation of cortical signals by neuroscien-
tists such as Vernon Mountcastle lead to the discovery of a topographic organization
in the brain where columns of neurons were tuned to different sensory-motor sig-
nals [15]. This somatotopic organization of the brain enables electrodes placed over
small populations of neurons to yield information about independent sensori-motor inputs/outputs. With the rapid advancement in recording technologies, neuroprosthetic control signals from the periphery and directly from cortex can now be streamed in realtime via various kinds of electrode arrays.

1.3.1 Electromyography (EMG)

When performing muscle contractions, muscle cells generate electrical signals that can be recorded by invasive needle electrodes and surface electrodes, resulting in electromyographic (EMG) signal recordings. The contraction of different muscle patterns generates unique signatures in the EMG recordings. These signatures can be detected and classified through pattern recognition software such as neural networks [16], fuzzy logic [17], and linear discriminant analysis [18, 19]. Trained pattern recognition classifiers can then be run in realtime in order to translate EMG signals into control signals for a neuroprosthetic [20, 21]. While victims of severe paralysis do not always have enough residual muscle activity for reliable use of EMG signals, upper limb amputees can potentially greatly benefit from EMG controlled neuroprosthetics. Even amputees missing a substantial amount of their upper limb can benefit from EMG control after undergoing targeted muscle reinnervation [22], a technique wherein nerves are surgically made to reinnervate alternative muscle groups, such as in the chest, to activate different muscle groups and generate the EMG control signal.
1.3.2 Electroencephalography (EEG)

One of the most studied recording modalities of cortical signals is electroencephalography (EEG), due to its minimal invasiveness with high temporal resolution and low financial cost. EEG has three primary features that yield useful control signals: the steady-state visually evoked potential (SSVEP), the P300 response, and event-related [de]synchronization (ERD/ERS). EEG signals show phase-locked modulation to visual stimuli flashing between about 5 to 75 Hz. This is referred to as the SSVEP, and has been used to enable users to select different options flashing on a screen for BCI control [23, 24]. The P300 refers to an increase in the smoothed signal amplitude 300 ms after observing an anticipated event. This has seen most use in P300 spellers, in which different letters light up on a screen, and a P300 is detected in response to the desired letter being shown [25]. EEG also shows spectral power increases and decreases, referred to as event-related synchronization (ERS) and event-related desynchronization (ERD) respectively, in response to various behavioral tasks. For example, the Mu power band (8-13Hz) over motor cortex shows ERD during movement. Groups lead by Wolpaw, Pfurtscheller, and Birbaumer investigated ERD and ERS in the EEG signals for BCI, and were able to deliver one-dimensional cursor control to subjects via their EEG signals [26, 27, 28]. Further advances have been made in order to deliver two-dimensional cursor control [29], which has been translated into control of external devices such as a virtual helicopter [30], a wheelchair [31], and a planar robotic limb [32].
CHAPTER 1. INTRODUCTION

1.3.3 Microelectrode Arrays (MEAs)

Due to the challenge of resolving signatures/features of interest from signals acquired from EEG, which has limited spatiotemporal resolution, similar studies have been conducted using more invasive electrode devices. MEAs directly penetrate through cortex with a small array of microelectrodes capable of recording from single neurons (e.g. 4 mm x 4 mm with 96 electrode shanks). The raw signal recorded by the electrodes corresponds to the local field potential (LFP), which can be decomposed into frequency bands that show ERD and ERS corresponding to different behavioral events. The spikes (i.e. action potentials) of individual neurons can also be extracted from the raw signal. The recorded LFP and/or spiking activity can be modeled and subsequently decoded in order to give paralyzed patients control of a robotic limb [33], with up to ten (DOF) independently controlled so far [34]. While MEAs currently provide the most impressive results, they have a number of drawbacks. These include (1) glial scarring which causes recordings to diminish over time [35], (2) neuronal dropout which necessitates daily retraining of the decoder [36], and (3) a limited recording area due to their size, which bottlenecks the potential throughput of information. One potential solution to these drawbacks is to supplement or replace the MEA recordings with electrocorticography (ECoG), which provide a middle ground between EEG and MEAs with electrode grids providing more detailed information than EEG, but over a wider area than MEAs.
1.3.4 Electroencephalography (ECoG)

ECoG signals are recorded from platinum disk electrodes placed on the surface of the brain, as shown in Figure [1.1]. They provide an compromise between the highly focused MEAs and low resolution wide area EEG recordings. ECoG grids come with varying electrode sizes and spacings, allowing for flexibility in trading off spatial resolution and wide-area coverage. They are regularly implanted in patients for clinical reasons. For example, they are often implanted for approximately one week in patients with intractable epilepsy for seizure localization. After implantation, the ECoG electrical activity can be streamed directly from the cerebral cortex to recording equipment, enabling clinicians to monitor cortical activity. Patients can then consent to participate in research experiments, allowing researchers to explore topics such as online decoding algorithms for translating ECoG signals into neuroprosthetic control signals.

Many features extracted from ECoG signals have been shown to yield information about tasks being performed. For example, the moving average of the signal (referred to as the local motor potential, or LMP) and low frequency content have been found to contain movement related information [37]. Unfortunately, the LMP typically requires a large degree of smoothing, which creates a group delay that may be too large for use in a real-time neuroprosthetic. The low and mid frequency power bands (mu and beta respectively, spanning 8-30 Hz), in motor areas typically show a decrease in power during movements, followed by a "rebound" in the beta band. However,
this power band desynchronization is seen across broad spatial regions of cortex for extended periods of time \cite{38}, restricting the amount of information it is able to yield about what specific events are occurring.

The high frequency high gamma band (about 70-150 Hz) offers more spatially and temporally precise information about the task being performed \cite{39}. The high gamma band has been shown to be a correlate of the local firing rate of constrained populations of neurons \cite{40}. The fact that the modulations occur in a higher frequency band enables power to be extracted with minimal group delay. All these factors combined make the high gamma band an ideal feature to use for neuroprosthetic
control. While additional features may yield some improvement to decoding accuracy, in this thesis the high gamma band extracted from the recording electrodes is the sole neural feature used for online human control of neuroprosthetics with ECoG.

ECoG has been widely used as a control signal for external devices. Cursor control has been previously demonstrated with one \cite{41, 42}, two \cite{37, 43, 44}, and three \cite{45} DOF. Offline decoding has been able to reconstruct the ballistics of reaching and grasping movements \cite{46, 47, 48, 49, 4}. Control of reaching and/or grasping with a robotic arm has been accomplished with up to three DOF \cite{50, 45, 51}.

However, these online neuroprosthetic efforts have largely been reliant on direct control schemes involving unnatural mappings. Control of more than one DOF has typically required an unintuitive mapping between intent and reality, such as tongue, hand, and/or shoulder movements to move a cursor \cite{41}, or attempted thumb/elbow/wrist movements to control a prosthetic limb in 3D \cite{45}. As more DOF are introduced into the systems, these control schemes will likely become untenable.
CHAPTER 1. INTRODUCTION

1.4 Overview of Upper Limb Neuroprosthetic Control Strategies

1.4.1 Direct Neural Control

The most straightforward approach to developing a neuroprosthetic is to give subjects direct neural control over the individual DOF of a device. This strategy involves mapping the instantaneous neural features directly to a specific DOF, such as vertical cursor velocity. This has been employed with ECoG using a linear SVM [50] and linear regression [45], and with MEAs using a Kalman filter [33] and linear regression [36]. This strategy has allowed patients to attain a rudimentary control of a robotic limb, and even perform basic tasks such as retrieving a cup for drinking [33] and a chocolate bar to eat [36].

The neural decoders were trained by having subjects view/attempt a movement (such as flexing their wrist or moving their arm up) and mapping the corresponding neural response to a low level command signal (such as moving the robotic arm up). Attaining more than two-three DOF takes a substantial amount of training time, with the most impressive (ten DOF) control to-date requiring over three months of training with a subject implanted with multiple MEAs [34].
1.4.2 Hybrid Control

In an attempt to overcome the limited throughput of information associated with using only a single extracted neural feature, various groups have recently proposed hybrid control setups leveraging a variety of neural and/or non-neural inputs. For example, Allison et al combined the ERD and SSVEP signals found in EEG to deliver two orthogonal axes of cursor control in [52]. In [53], Pfurtscheller et al describe several other ways of combining the SSVEP with ERD, such as using the SSVEP as a switch to reduce false positive rates when detecting ERD. Hybrid control can also entail combining neural features with non-neural inputs, such as EMG [54]. Eye-tracking can be used to select options in a GUI or provide information about potential goals [55, 56, 57].

1.4.3 Shared Control

The use of "shared control" can further improve neuroprosthetic performance in the face of limited signal fidelity. Shared control entails offloading some of the cognitive burden of control from the patient by automating portions of tasks with intelligent robotics. Shared control is exemplified in nature by the octopus, which contains about two-thirds of its neurons in its arms [58]. This potentially enables the octopus to give high level commands from its centralized brain, but relegate the low-level control details to its semi-autonomous arms [59]. Neuroprosthetics with shared
control take this same strategy; they avoid designating all the nuances of control to the user’s neural modulation, and instead accomplish tasks in a semi-autonomous fashion. This is akin to the strategy used by the fictional neuroprosthetic user Dr. Octopus (Otto Octavius) from the Spiderman franchise. Shared control helps alleviate the cognitive burden placed on users and can potentially increase accuracy and reliability of the system. It can be employed in a variety of manners, such as avoiding objects with a neurally controlled wheelchair [60] or stabilizing the neural control of reaching with a robotic limb [61].

The concept of assisting users by sharing control between direct user inputs and intelligent robotic planning algorithms has been greatly explored in the field of teleoperated robotics. It has been shown that environmental sensors (e.g. video+depth (RGB-D) cameras, light detection and ranging (LIDAR), and force sensors) can be leveraged to improve system performance [62]. Augmenting teleoperation with shared control schemes generally involves prediction of the user’s intent followed by arbitration between system autonomy and direct user control. When predicting user intent, systems typically attempt to infer the user’s goal or trajectory from a combination of external sensors and user inputs [63]. Prediction of user intent can include segmentation of user inputs with a hidden Markov model (HMM) [64]. This can enable the system to draw from a library of movement primitives that aid in task completion [65, 66].

The arbitration between user inputs and intelligent motion planning algorithms
is of great importance to user satisfaction; even if it leads to improvements in the ultimate system performance, users can become unsatisfied if too much of their autonomy is sacrificed [63]. How arbitration is implemented has been shown to be integral for the success of shared control with neuroprosthetics [67]. Most often, arbitration is performed by continuously blending a linear weighting of the user inputs and autonomous system predictions [68, 69, 63]. This strategy has delivered shared control of a neuroprosthetic that enabled paralyzed users to accomplish tasks they were unable to perform with neural control alone [70].

Probabilistic neuroprosthetic control strategies typically incorporate prior information about movement kinematics and/or inferences about the user’s goal in order to improve decoding. This can help overcome limitations in the different recording modalities, such as the poor representation of speed information in single unit recordings [71] and the poor representation of direction information in ECoG recordings [72]. Most commonly, decoding is performed using a Kalman filter with a state transition matrix built either from a library of sequences of movements [73], set manually according to assumptions of how the system should operate [74, 71]. Hidden Markov Models (HMMs) have been used to segment the stages of movements during reaching tasks based on the cortical signals [75, 76], and switching Kalman filters have been used to improve results [77, 78]. Performance with probabilistic neuroprosthetic control strategies can be further improved by leveraging information regarding the user’s potential goals [79, 80, 81]. Eye-tracking can be combined with the recorded biolog-
CHAPTER 1. INTRODUCTION

ical control signals to provide information about potential goals in order to further improve system performance [55]. Probabilities derived from neural and environmental sensors have been combined for discrete action selection (left, right, forward) for a wheelchair [31]. However, these systems have all relied on simple linear models of motion that are incapable of reproducing the complex movements necessary for many tasks. Within this thesis, I show how ECoG can be leveraged to immediately give a basic level of control over neuroprostheses after functional mapping of arm and hand regions of the brain. This cortical control can then be combined with an intelligent robotics system to provide a functionally useful system to patients without the need for extensive training. By using a probabilistic framework for the cortical features and the motion model, I further introduce a system that shares control between autonomous robotics and cortical inputs in a theoretically optimal manner. An outline of the thesis is found below.

1.5 Dissertation Organization

The organization of this thesis is as follows: Chapter 2. This chapter delineates an exploration of how different features of ECoG signals recorded from damaged motor systems are correlated with ipsilateral arm movements. This has strong implications for brain-machine interfaces put into practice for patients suffering from paralysis due to complications from issues like strokes. Chapter 3. Direct control
CHAPTER 1. INTRODUCTION

of reaching and grasping with the Modular Prosthetic Limb (MPL) was delivered to
two subjects using ECoG. This marked the first time human subjects were able to
use their neural signals recorded through ECoG to attain simultaneous, independent
control reaching and grasping without any need for operant conditioning. Chapter
4. Direct control of individual fingers with MPL. This built on the previous study
by showing that, not only does ECoG have the resolution necessary to distinguish
arm movements from hand movements, high density ECoG electrodes are capable
of providing online control of individual finger movements. This was the greatest
number of degrees of freedom controlled via ECoG to-date, pushing the bounds of
what is possible with direct ECoG control with patients’ natural underlying neural
correlates of movement. Chapter 5. Supervisory Control of the MPL for completing
a pick and place task with objects placed in a arbitrary locations within a workspace
in front of the user. This chapter describes how computer vision and eye-tracking can
be leveraged to enable users to accomplish complex sequences of movements without
the need for extensive training. Chapter 6. This chapter details the offline valida-
dation of a continuous shared control system using microelectrode array recordings
from a non-human primate performing reach and manipulate tasks. A switching un-
scented Kalman filter was used to continuously fuse non-linear models of motion with
estimates of the monkey’s kinematics attained through the neural signal recordings.
Chapter 7. The system detailed in Chapter 6 was implemented with the robot
operating system (ROS) and Caffe to enable real-time continuous shared control of
CHAPTER 1. INTRODUCTION

neuroprosthetics. Proof of concept testing was performed using electromyography (EMG) and inertial measurement unit (IMU) sensors. **Chapter 8.** This chapter consists of summary of the contributions of this thesis, followed by an overview of possible future directions.
Chapter 2

Coarse Electrocorticographic Decoding of Ipsilateral Reach in Patients with Brain Lesions

2.1 Overview

The brain-machine interface (BMI) is a tool to replace areas of human functionality lost due to trauma or degenerative disease. To date, BMIs have been used with humans to control cursors [26, 82, 43], move robotic limbs [51, 36, 33, 4], and allow patients with locked-in syndrome to communicate with the outside world [83]. Electrocorticography (ECoG) electrodes are implantable, non-penetrating, and occupy a unique space between the high fidelity of cortex-penetrating microelectrode
CHAPTER 2. ELECTROCORTICOGRAphIC DECODING OF IPSILATERAL REACH

arrays recordings and the low fidelity and bandwidth of noninvasive EEG. A long history of work with EEG and ECoG signals has highlighted mu (8–12 Hz), beta (14–30 Hz) [84, 38] and high gamma (>70 Hz) [85] band signals as indices of cortical motor processing. More recent work has shown that ECoG signals may also contain movement-related information in their smoothed amplitudes, or local motor potentials (LMPs) [86, 85, 87].

In most efforts to develop a brain-machine interface (BMI) to restore upper limb function, the design has assumed a normally functioning brain as the source of neural control signals, with the BMI serving to bypass lesions of neural pathways connecting the brain to its muscle effectors. However, a large patient population with hemiplegia from strokes and other cerebral lesions may also benefit from BMIs [88]. In these patients, a variety of approaches are already being used or developed to restore function. These include intensive neurorehabilitation [87, 89, 90], direct current stimulation of the motor cortex [91], [92], and neurobiological therapies, e.g. infusions of stem cells and growth factors, aimed at promoting neurogenesis in the damaged hemisphere [93, 94, 95]. It might also be feasible to restore function via neural signals from the undamaged hemisphere [96, 46, 97]. However, cerebral representations for upper limb control are primarily contralateral rather than ipsilateral, and the risks associated with surgical implantation of electrodes over healthy, functional cortex may not outweigh the potential benefits. To overcome these limitations, an alternative approach could be to utilize residual sensorimotor systems in the damaged hemisphere. This
approach would improve the risk-benefit ratio because subdural electrodes for an
ECoG-based BMI could be implanted along with other invasive experimental ther-
apies. Furthermore, training of the BMI, along with intensive neurorehabilitation,
could generate neural activity encouraging neurogenesis and plasticity in residual
brain tissue, potentially working synergistically towards functional recovery. How-
ever, in patients with total paralysis of the contralateral upper limb, arguably the
population to whom a BMI would be most attractive, it would be challenging to
train the BMI’s decoding algorithms. To overcome this limitation, subjects could be
trained by imagining their own arm moving [98], viewing arm movements [99, 100],
being passively moved [101], or even performing movements with the ipsilateral arm
[102]. The neurophysiological correlates of reaching movements in residual motor
systems of the ipsilateral cortex therefore warrant investigation.

The descending corticospinal tract is not solely comprised of axons that cross the
midline. A minority of the fibers also arise from ipsilateral cortex [103]. Ipsilateral
wrist movements may be impaired in hemiparetic subjects with unilateral brain lesions
[104]. A substantial portion of PMd and a nontrivial number of M1 cells display tuning
to ipsilateral arm movements in primates [105]. Indeed, the efficacy of rehabilitative
therapies in many patients with upper limb paralysis from unilateral brain lesions
[90] implies not only that functional reserve and potential for plasticity is present in
the lesioned hemisphere, but also that ipsilateral descending motor pathways can be
utilized for upper limb control. The degree to which arm movements can be decoded
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

from the neural activity of a damaged ipsilateral cortex, however, is not well known and is of particular interest for training an upper limb BMI in patients with total upper limb paralysis from a stroke or other unilateral cerebral lesion.

Reach trajectories have been decoded extensively in nonhuman primates performing center-out reaches with microelectrode arrays [106, 107, 108, 109], and more recently in humans and primates implanted with ECoG grids [110, 48, 111, 45], using neural signals contralateral to the reaching arm. Previous work has also shown the potential for extracting neural correlates of overt and imagined hand movements from intact ipsilateral cortex in humans using ECoG [102, 112] and EEG [113]. However, it has not previously been shown if neural signals from damaged cortical areas can retain any information about ipsilateral arm movements. In this study, we investigated the neural correlates of ipsilateral reaching movements in human subjects with lesions in the hemisphere recorded with ECoG. In two subjects, these lesions disrupted motor pathways and were associated with motor impairment of the contralateral arm. In one subject, the lesion involved superior parietal lobule, likely including the human homolog of parietal reach region, but there was no associated motor impairment.
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

2.2 Materials and Methods

2.2.1 Data Acquisition

Three subjects, two male and one female aged 24-36 years, were implanted with ECoG arrays prior to surgical resection of their seizure foci. Grid placement was determined solely on the basis of their clinical need without any influence from this study. Only subjects with motor cortex coverage were considered for this study. Table 2.1 summarizes the clinical and demographic information associated with each study participant. All subjects gave written informed consent and the study was conducted under a protocol approved by the Johns Hopkins Institutional Review Board (IRB). ECoG grids consisted of rectangular arrays of platinum electrodes (Adtech Medical Instrument Corp; Racine, WI) with 2.3 mm diameter and center-to-center spacing of 10 mm. Electrodes were embedded in a Silastic sheet and implanted on the brain in the subdural space. Arrays of nonpenetrating micro-ECoG leads (75-micron diameter, 0.66-mm spacing) were also implanted in each of the three subjects, but in subjects 1 and 2, these micro-ECoG leads were not over motor areas and therefore not included in our analysis. Of the implanted electrodes, 44-48 per subject were analyzed for this study. A 4x4 array of 16 micro-ECoG leads was included in the analysis for subject 3. These micro-ECoG leads were located in posterior middle frontal gyrus, anterior to electrodes where electrocortical stimulation mapping (ESM) affected motor function, and their signals did not produce notable decoding results. The tests described below
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

Table 2.1: Subject Demographic and Clinical Information.

<table>
<thead>
<tr>
<th>Subj</th>
<th>Sex</th>
<th>ECoG Coverage</th>
<th>Seizure focus/pathology</th>
<th>Neuro defecit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>Right frontal-parietal</td>
<td>Right frontal oligodendroglioma</td>
<td>Left hemiparesis</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>Right parietal-occipital</td>
<td>Right occipital</td>
<td>Left inferior quadrant visual defect</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>Left frontal-parietal</td>
<td>Stroke involving posterior limb of left internal capsule</td>
<td>Right spastic hemiplegia</td>
</tr>
</tbody>
</table>

were part of a larger battery of research testing. These tests were conducted between two and seven days after implantation, and removal of the grids occurred at the time of seizure focus resection, seven to eight days after implantation.

ECoG signals were referenced to an inactive intracranial electrode and recorded by a 64-channel Neuroscan Synamps2 (Compumedics; Charlotte, NC) for subject 1 and 2; ECoG signals from subject 3 were also referenced to an inactive intracranial electrode, but were recorded by a 128-channel NeuroPort system (Blackrock Microsystems; Salt Lake City, UT). Signals were digitized by the Neuroscan system at 1 kHz after being band-pass filtered between 0.15 and 200 Hz (subject 1 and 2). Signals recorded by the BlackRock system were band-pass filtered in analog with a low-pass cutoff of 7,500 Hz and a high-pass cutoff of 0.3 Hz, digitized at 30 kHz, then down-sampled to 1 kHz. Anatomical reconstructions of each subject’s electrode placement were generated by volumetrically co-registering post-implantation brain computerized tomography (CT) with high-resolution pre-implantation brain magnetic resonance imaging (MRI) images using BioImage [14]. Additional information from surgical photos, post-implantation skull X-rays in the antero-posterior and lateral axes, and
CHAPTER 2. ELECTROCORTICOGRAFHIC DECODING OF IPSILATERAL REACH

Figure 2.1: Reconstructions of electrode placements with ESM results. The grids shown are the subset of implanted electrodes that were recorded from during this study. The green highlighted areas correspond to regions of cortical lesions. The lesion in subject 3 could not be seen on the brain surface rendering because it was located beneath the surface of the brain. Colored rectangles joining electrodes imply that bipolar stimulation was applied to that pair of electrodes. In subject 1, several electrode pairs were further investigated by performing unipolar stimulation relative to a distant reference electrode. The results of this unipolar stimulation are shown with colored circles surrounding the electrodes.

ESM were used to verify electrode locations. Reconstructions of electrode locations, overlaid with the color-coded results of ESM, are shown in Figure 2.1. Figure 2.2 displays the locations of the cortical lesions for each subject.

The three-dimensional positions of the subject’s shoulder and arm were tracked optically using the Northern Digital Optotrak system with a sampling frequency of 100 Hz. Neural data and movement data were synchronized using parallel port
Figure 2.2: Presurgical MRI and brain reconstructions. Reconstructions are shown for subject 1 (first row) subject 2 (bottom left and middle) and subject 3 (bottom right). The previous resection margins anterior to the precentral gyrus in subject 1 are highlighted in green in the upper right. Superior oblique and top axial views of the reconstruction for subject 2 show the lesion from different viewpoints. Pre-surgical MRI of subject 3 reveals a lesion of posterior left insula also involving left internal capsule.

triggers sent simultaneously from the experimental computer over a split cable to the Neuroscan (subject 1 and 2) or NeuroPort (subject 3) amplifier’s external trigger inputs and the Optotrak Data Acquisition Unit (ODAU).

2.2.2 Experimental Paradigm

Subjects were instructed to make reaches to the tip of a long wooden dowel being moved by the experimenter in three-dimensional space. Subjects used the arm ipsilateral to the ECoG electrode implantation. Each reach was either terminated
by touching the dowel with the subject’s pointing index finger (subject 1, session 1; subject 2; subject 3) or with an index-thumb pinch (subject 1, session 2). The subject returned his or her hand to a comfortable resting position following completion of each reach to the target. The subject rested his or her hand in the home position for a variable amount of time (0-16.7 seconds), after which the target was moved to another point in three-dimensional space. The position of the target in three-dimensional space was determined by the experimenter in an attempt to probe the natural workspace of the subject’s upper limb as thoroughly as possible. The target tended to be placed in front of and above the subject’s rest position, but varied in the lateral (i.e., left or right) directions. The durations of the reaches performed by subjects 1, 2, and 3 were 3.2-6.2, 2.0-5.0, and 1.2-5.3 seconds, respectively, with median durations of 4.7, 3.0, and 2.7 seconds. A visual depiction of the workspace and experimental setup is shown in Figure 2.3, along with example subject kinematics. The subject was cued to begin each reach once the target had stopped moving to its new location. The cuing dowel was tracked using the same Optotrak system that was used simultaneously to track the subject’s shoulder and hand, with a sampling frequency of 100 Hz. Because the tracked target tip was occluded during manipulation by the subject, the position of the distal tip was estimated using proximal sensor locations and the known, fixed distance between the sensors and the end of the dowel.
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

Figure 2.3: Subject Kinematics and Experimental Paradigm. (A) Shows the weights given to each of the three dimensions (height, depth, and lateral) when computing the first PC of the movement kinematics. Each point represents the PC weight calculated using a different cross-validation. (B-D) Show different views of the experimental paradigm and example kinematics from one of the subjects’ sessions. The red dashed line represents the first PC plotted in Cartesian space for that same session.

2.2.3 Neural Data Feature Extraction

Recorded ECoG signals were imported into MATLAB (MathWorks, Natick, MA) and analyzed using the Signal Analysis Toolbox. These signals were re-referenced to a common average reference (CAR) to avoid spatial biases from varying distances between active and reference electrodes [115]. The CAR-filtered signals were then filtered forward and backward (to avoid phase distortions) using a Hamming window and a series of 400-order FIR filters with bandpass cutoff frequencies corresponding to
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

delta (0-4 Hz), theta (4-8 Hz), mu (7-13 Hz), beta (14-30 Hz), low gamma (30-50 Hz), high gamma 1 (70-110 Hz) and high gamma 2 (130-200 Hz) bands. The frequency limits for these bands were chosen to ensure inclusion of all relevant frequencies and to avoid 60 Hz line noise and its harmonics. The filtered signals were then down-sampled to 100 Hz to match the sampling frequency of the kinematic data. These signals were squared to yield power in each band, then passed through a smoothing integrator (1 second moving average, applied forward and backward), and log-transformed to approximate normal distributions. Transients in the neural data induced by the filtering operations at the beginning and end of each block of trials were not analyzed in this study.

In addition to frequency domain features, previous work indicated that smoothed time domain features contained information related to movement [37], [86]. These features, known as local motor potentials (LMPs) were extracted in this study using a moving average filter of two seconds, applied forward and backward to remove phase distortions. All feature extraction was done in a non-causal manner (using data from both the past and the future) to avoid introducing a group delay as a confound in our results.
2.2.4 Decoding Model Construction and Evaluation

The kinematics were modeled as a linear combination of the input features with a constant offset and a Gaussian noise term: \( y_k(t) = \beta_0 + \beta \cdot X(t) + \eta \), where \( y_k(t) \) denotes the predicted output of the kth dimension of the kinematic output, \( X(t) \) is the vector of feature values at time \( t \), \( \beta_0 \) is the constant offset term, \( \beta \) is the vector of linear weights applied to the features, and \( \eta \) is zero-mean Gaussian noise. The parameters of this model were estimated using the glmfit function in MATLAB.

All models were tested under fivefold cross-validation. Pearson’s correlation (\( r \)) between observed and predicted kinematics was used to quantify model accuracy. Feature selection and model training were performed on 80% of each block of trials, and the remaining 20% was used only for evaluation. Each training session contained at least 18 movements, and each testing session contained at least 4 movements. Chance decoding performance was calculated for each session by randomly shuffling the extracted and smoothed neural features 1024 times, similar to the random shuffle procedure done in [116]. These shuffled data were then used for testing and training with fivefold cross-validation. The distributions of the decoding accuracies with actual and shuffled data were compared with a Bonferroni-corrected Wilcoxon test.

Decoding accuracies were transformed into z-scores using the Fisher transformation, \( z = \frac{1}{2} \ln(\frac{1+r}{1-r}) \sqrt{N-3} \), where \( N \) is the number of points used to calculate the
correlation. These transformed accuracies were used as inputs to a two-way ANOVA
with the number of inputs and feature type (e.g., theta features only, all features)
as independent factors. A post-hoc t test was done using the Dunn-Sidak correction
for multiple comparisons to determine which numbers of inputs and/or which
feature types significantly differed, if any.

Both the neural features (after all preprocessing) and the kinematic data were
normalized by subtracting the means and dividing by the standard deviations. The
means and standard deviations were recomputed during every fold of cross-validation
using only the training data. This normalization was done to ensure the magnitude of
the features did not impact how heavily they were weighted in the decoding models.

2.2.5 Principal Component Analysis

While Cartesian coordinates for arm position (i.e., vertical, lateral, and depth) are
orthogonal, constraints on the positioning and capabilities of the subjects resulted in
an inability to fully probe these axes of the workspace independently. The kinematics
of the resulting movements in this space were correlated to the extent that presenting
decoding results in Cartesian axes would be somewhat misleading. We therefore used
principal component analysis (PCA) to decorrelate the kinematic dimensions by pro-
jecting the 3D arm position onto another set of orthogonal axes. These decorrelated
kinematics, or principal components (PCs), accounted for decreasing proportions of
the variance in the original signals. The first PC was therefore oriented in the di-
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

rection of largest variance (i.e. the vector along which the subject moved the most), while the second PC was oriented in the direction of the most remaining variance, etc.

The coefficients used for the PCA transform were recalculated for each fold of cross-validation using only the normalized training data. In all three subjects, decoding accuracy of the first PC was greater than the second and third PCs ($p < 0.05$, ANOVA with Dunn-Sidak post-hoc). All but one of the models built on the second and third PCs across all subjects were indistinguishable from chance. We therefore focus on the decoding accuracies of the first PC in this study.

2.2.6 Model Input Feature Selection

Models with increasing numbers of input features were trained; inputs were added in decreasing order of their correlation with the kinematic variables in the training set. Correlations were calculated for the kinematic signal lagged with respect to the neural features at time lags varying between 1 second at 50 ms intervals. A single spectral domain or time domain signal was only included once at its best lag. Results from model orders of up to 40 are reported below. Additionally, restricted models were trained from single feature types (e.g., high gamma 1 features only). Performance of these restricted models is also reported below in comparison to the full model performance for model orders one of up to 40. An analysis of the performance saturation of these models with increasing numbers of inputs was performed with
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

a Kruskal-Wallis non-parametric one-way analysis of variance. The least significant difference post-hoc test was used to determine the first number of inputs at which additional inputs did not statistically increase performance.

The large smoothing window applied to the extracted features meant it was possible that changes in the neural features could occur substantially before or after their associated movements but still yield a high correlation with the movements. To investigate whether the feature changes preceded or followed the associated movements, we calculated correlations every 50 ms between the extracted features and the first principal component of movement offset by at most 1000 ms relative to one another.

2.3 Results

2.3.1 Feature Selection

The extracted signal features had substantial correlations with the first PC of the ipsilateral hand position, and the maximally correlated feature types varied across subjects. The three-dimensional kinematics and their contributions to the first PC are shown in Figure 2.3. Examples of how each of the PCs relate to the original kinematics is shown in Figure 2.4. Figure 2.5 illustrates the correlations between different ECoG signal features (LMP and bandpass power) and the first PC of hand position in each subject. The magnitudes of these correlations are color-mapped onto the ECoG recording sites from which the features were derived. The lagged
correlations were calculated at 50 ms intervals between 1 and +1 seconds relative to the hand kinematics. The peak correlations (i.e., across all lags) were averaged across sessions for each electrode. Subject 3 contained micro-ECoG electrodes, but these are not depicted due to their lack of correlation.

The large smoothing window applied to the extracted features meant it was possible that changes in the neural features could occur substantially before or after their associated movements but still yield a high correlation with the movements. To investigate whether the feature changes preceded or followed the associated movements, we plotted the magnitudes of the correlations between the neural features and the hand kinematics at different time offsets relative to one another (Figure 2.6). Subject 1 typically had maximum correlations corresponding to neural features changing
after the hand kinematic changes, particularly in session 1. This may indicate the activations were primarily in response to sensory feedback. Subject 2 and 3 had peak correlations when the changes in the features occurred before the kinematics.

### 2.3.2 Multiple Inputs

Figure 2.7 displays the models’ mean correlations with the first PC of movement as a function of the number of model inputs. Input features were selected in order of decreasing correlation with the first PC of movement in the training set. Fivefold cross-validation was performed for each session, resulting in 10 folds for subject 1 and
CHAPTER 2. ELECTROCORTICOGRAphIC DECODING OF IPSILATERAL REACH

Figure 2.6: Correlations between kinematics and neural features at different time lags relative to one another. Each row corresponds to one of eight extracted features from one electrode, resulting in 384, 352, and 360 rows for subjects 1, 2, and 3 respectively. The x-axis corresponds to the lag between the neural feature and the reach position. A negative lag represents changes in the neural features occurring before the corresponding kinematics, and a positive lag represents feature changes after the kinematics. Rows are ordered by the magnitude of their peak correlation. Correlations between the hand kinematics and the features were calculated every 50 ms between leads or lags of one second.

3 and 15 for subject 2. For subject 1 and 3, the addition of more inputs to the model did not significantly improve model performance. In subject 2, models using a single input performed significantly worse \((p < 0.05, \text{ANOVA with Dunn-Sidak post-hoc})\) than models with the best performing number of inputs.

The input features in Figure 2.7 were selected from all signal feature types. This analysis was repeated with inputs from different ECoG recording sites restricted to a single signal feature type (e.g. LMP, delta band power, etc.). Figure 2.8 shows the
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

Figure 2.7: Relationship between number of model inputs (i.e., ECoG signal features) and model performance. Mean decoding accuracies of the first PC of movement across cross-validations and sessions are displayed for features chosen in decreasing order of correlation. Shading corresponds to the 95th percentile confidence interval for the mean.

decoding performance for models restricted to one signal feature type, using the number of inputs that saturated performance. Most of these restricted models saturated with only a single electrode, and all but one saturated with three or fewer electrodes. Decoding performance for each signal feature type was fairly unique to each subject. In all subjects, models trained with all features significantly outperformed all single feature models ($p < 0.05$). Smaller ranges of time lags (e.g. 100 to 0, 200 to 0, 100 to +100, 200 to +200, and 0 to +200 ms) resulted only in small reductions in performance. The low frequency features (i.e., delta, theta, mu, and beta) and LMP
performed significantly better than the high frequency features (low gamma through high gamma 2) for subject 1 \((p < 0.05)\). In subject 2, there was no significant difference between models trained with high gamma 1 or high gamma 2, and these sets of models outperformed all other single signal feature models \((p < 0.05)\). In subject 3, models trained with delta, low gamma, and high gamma 2 performed statistically worse than all other single signal feature models, which had no significant difference between each other \((p < 0.05)\). All p-values reported between feature types were obtained as a part of the two-way ANOVA with Dunn-Sidak post-hoc test described in the Methods.

The decoding performance of the first PC of movement for each session is shown in Figure 2.9. Results are shown for the two sessions for subject 1 and 3 and the three sessions for subject 2. Five folds of cross-validation using one, two, and nine input features are plotted for each session. During post-hoc analysis, we found that the decoding accuracy did not significantly improve \((p < 0.05\), one-way ANOVA with least significant difference correction) if more signal features and/or recording sites were added to the one best model input for subject 1 or 3, or to the two best model inputs for subject 2. Models trained with nine inputs were chosen a posteriori as a proxy for the peak decoding accuracy across sessions and subjects, as it gave consistently high results. Decoding accuracy was found to be significantly higher than chance \((p < 0.05\), Bonferroni-corrected Wilcoxon). Table 2.2 displays results for both raw kinematics and all PC’s.
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

Figure 2.8: Peak decoding accuracy for each ECoG signal feature type. Between 1 and 40 recording sites were selected for each signal feature type, and the means were calculated across cross-validations and sessions. The peak accuracies are displayed here with error bars corresponding to the standard error of the mean. Numbers over each bar indicate the minimal number of features required for statistical saturation.

Videos S1, S2, and S3 show a virtual depiction of the actual and decoded kinematics for representative folds of cross-validation for each subject. The actual and decoded kinematics were scaled and translated by the same amount to fit in the virtual workspace. The actual and decoded positions were then sent to a virtual modular prosthetic limb [118, 119, 120] through VulcanX limb control software, as described in an online setting in [4, 5]. The videos do not account for any delays which would occur in an online control scenario because of the non-causal methods we employed.
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

Table 2.2: Correlation between actual and decoded kinematics with and without PCA

<table>
<thead>
<tr>
<th>Subject, Session</th>
<th>Median correlation (Pearson’s r)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Depth</td>
</tr>
<tr>
<td>S1,1</td>
<td>0.75</td>
</tr>
<tr>
<td>S1,2</td>
<td>0.75</td>
</tr>
<tr>
<td>S2,1</td>
<td>0.65</td>
</tr>
<tr>
<td>S2,2</td>
<td>0.64</td>
</tr>
<tr>
<td>S2,3</td>
<td>0.73</td>
</tr>
<tr>
<td>S3,1</td>
<td>0.49</td>
</tr>
<tr>
<td>S3,2</td>
<td>0.69</td>
</tr>
</tbody>
</table>

2.4 Discussion

Our findings in three different subjects show that it is possible to decode a low-dimensional representation of natural reaches from ipsilateral ECoG electrodes in the presence of damaged cortical motor systems. This decoding attained a median Pearson’s correlation (r) between actual and predicted reach kinematics of 0.77, 0.73, and 0.66 for subjects 1-3. This accuracy was accomplished using as little as 133 seconds, and no more than eight minutes, of training data in each session. Furthermore, two of the subjects in this study (subject 1 and 3) had severe upper limb weakness contralateral to ECoG recording grids, arising from lesions of brain structures critical for motor control. Although subject 2 did not have contralateral limb weakness, he had a lesion of parietal lobe structures that participate in visually guided reaching and grasping movements of the upper limb [121, 122, 123]. Our findings indicate that the first PC of movement is robustly represented in the ipsilateral hemisphere, even in the face of damaged sensorimotor systems.
We found that high frequency signal features such as high gamma power tended to have a positive correlation with the first PC of movement, while low frequency features such as delta generally had a negative correlation. This result agrees with previous studies [38, 39], which have found that functional activation of cortex, including sensorimotor cortex, is accompanied by an event-related increase in power in the gamma band (>30 Hz) and an event-related decrease in power in lower frequencies, especially alpha (8-13 Hz) and beta (15-25 Hz) bands. The presence of this phenomenon in ipsilateral cortex agrees with results found in [112]. While typically not as robust as theta or mu, the delta band displayed notable power suppression during movements. The filters we employed were of an exceedingly high order (400),
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

providing a roll-off steep enough to discount the possibility of bleed over from other low frequency bands. The LMP, however, did not consistently have a positive or negative correlation with the reaching movements.

There was substantial variability as to which spectral components of the ECoG signals were most predictive of the movement trajectory. For example, in subject 1 the mu band models achieved statistically higher decoding accuracies for the first PC of movement than did those constructed from high gamma 1 features; the opposite was true in subject 2, and both model types performed equivalently in subject 3. This variability in the performance of different ECoG spectral features across subjects suggests that an agnostic approach may be necessary when evaluating and selecting signal features for modeling ipsilateral movement kinematics, particularly in the face of different lesion locations and types. Much of the variability we observed was likely due to substantial differences in lesion size and location across our three subjects. Although patients with motor impairments do not frequently undergo intracranial monitoring, future studies might accumulate enough subjects to determine the spectral features that yield the best performance across subjects. Better yet, this might make it possible to better characterize the relationship between lesion characteristics (e.g. extracted from MRI via voxel based morphometry) and the performance of different spectral features. For example, more intact motor and premotor cortices in subjects 2 and 3 may have allowed for better decoding with high gamma power.

It is also possible that compensatory movements with the contralateral hemibody
presented a confound [102, 105]. Indeed, reaching movements impart forces which are often accompanied by postural changes throughout the body, presenting a potential confound in any study of this sort. Our results should therefore be interpreted as the result of a movement involving complex dynamics in systems throughout the body.

The results depicted in Figure 2.6 show that neural feature changes in subject 1 occurred following the onset of movement. This delayed change suggests that these neural responses arose chiefly from processing sensory feedback, perhaps because ipsilateral motor pathways were lesioned. Specifically, cortical lesions from prior resections in subject 1 probably involved hand and arm motor areas of the primary motor cortex. In contrast, subject 2 did not have motor deficits and the motor deficits in subject 3 were associated with a lesion in the left internal capsule. Unlike neural feature changes in subject 1, neural feature changes in subjects 2 and 3 preceded movement onset.

We found that decoding performance saturated with a small number of model inputs. That is, a few signal features at a few recording sites yielded the best performance, and the addition of more signal features and/or recording sites did not significantly improve performance. This finding could reflect a very coarse neural representation for ipsilateral limb movement that is widely distributed across sensorimotor systems typically specialized for contralateral limb movement. Alternatively, it could reflect transcallosal activation of these systems by homologous systems in the contralateral hemisphere. In this way, activity in sensorimotor cortex contralateral
to movement (e.g. left hemisphere controlling right hand in subject 1 and 2) could modulate the activity of sensorimotor cortex ipsilateral to movement (e.g. in right hemisphere in subject 1 and 2). The encoding of this modulation during reaching movements might also be coarse, serving primarily to coordinate the movements of the two arms during reaching. Such explanations are necessarily highly speculative. The spatial resolution and anatomical distribution of our ECoG recordings were not sufficient to test these hypotheses. Nevertheless, our findings indicate that only a few electrodes are needed to achieve reach decoding from ECoG features, and that these electrodes qualitatively correspond to sensorimotor cortex as identified by ESM (Figure 2.1).

We found that decoding of the second and third PCs of the movement kinematics was significantly worse than decoding of the first PC ($p < 0.05$, ANOVA with Dunn-Sidak post-hoc), and that all but one of the models built for the second and third PCs were indistinguishable from chance. One possible explanation is that the second PC captured bi-directional movements in the lateral dimension, i.e. to the left or to the right (Figure 2.4). The third PC qualitatively appeared to contain the most noise, and also often contained some bidirectional movements. These PCs stand in contrast to the first PC, in which movements only resulted in unidirectional deflections. One possible explanation for the poor accuracy with the second PC is that it is not the direction of movement, but instead its magnitude that is being decoded. It is possible that the correlation between movement effort and neural activity captured by
CHAPTER 2. ELECTROCORTICOGRAPHIC DECODING OF IPSILATERAL REACH

ECoG is highly robust, such that it can even be captured in ipsilateral sensorimotor systems, even when these brain systems have been damaged. The directionality of movement, however, may present a greater challenge for decoding, especially under the circumstances of the subjects in this study. Whether the directionality of naturalistic reaching movements can be decoded from ECoG under more optimal conditions remains to be seen. Previous efforts [102, 113] have shown it is possible to decode hand movements from intact ipsilateral cortex in a constrained one- or two-dimensional experimental environment. However, decoding the directionality of movements may be more challenging in the setting of natural 3D movements, especially when performed by subjects with lesions of sensorimotor systems. While a preliminary analysis with ANN’s did not improve decoding accuracy, there may be non-linearities within the data that could allow for directionality to be better decoded.

The ability to decode motor commands from ipsilateral sensorimotor systems, even when lesioned, may facilitate the development of brain-machine interfaces (BMIs) for a wider population of patients with upper limb paralysis from brain lesions such as stroke, trauma, or surgical resection of a tumor. For example, upper limb function could be partially restored by low dimensional control of a neuroprosthetic device in which ECoG electrodes are implanted over healthy ipsilateral cortex. Alternatively, electrodes could be implanted in and around cortical sensorimotor areas that have been damaged, in order to control the affected, contralateral upper limb. An important advantage of this approach is that it avoids the risks of surgically implanting
chronic indwelling electrodes in healthy cortical areas controlling the patient’s only functioning upper limb. In addition, implantation of ECoG electrodes in the lesioned hemisphere can be done in conjunction with other invasive interventions, such as infusions of stem cells and growth factors encouraging neurogenesis and neuroplasticity, thereby sharing the risk with other experimental therapies.

Our results suggest that it may also be possible to bootstrap a BMI system controlling contralateral arm movements by first using it to control ipsilateral arm movements. For example, the movements of the ipsilateral limb under BMI control could be mirrored in the contralateral arm using robotic assistance or functional electrical stimulation. This could provide an opportunity for development of contralateral arm control through functional reorganization and even neurogenesis. The neural activity generated while training the BMI could potentially facilitate neurobiological interventions through activity-dependent neuroplasticity, allowing the patient to eventually wean off ipsilateral arm control as control of the contralateral arm is restored. This might expedite the training process for BMIs by exploiting the more intact movements of the patient’s unaffected limb.

This chapter has supplementary downloadable material available at

http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0115236
Chapter 3

Intracranial EEG Control of Reaching and Grasping with the Modular Prosthetic Limb

3.1 Overview

Reaching to and grasping objects is an important capability that forms the basis for many activities of daily living (ADLs). It is thus an important target for brain-machine interfaces (BMIs) being developed for patients with impaired limb function due to neurological lesions of motor pathways (e.g., spinal cord injury, amyotrophic lateral sclerosis, stroke, etc.). Recent work has demonstrated that grasp types \cite{47}, grasp timing \cite{124}, hand postures \cite{49}, and reach parameters \cite{110} can be decoded
CHAPTER 3. REACHING AND GRASPING

from spectral changes in human intracranial electroencephalograhic (iEEG) signals, and that movement-related spectral modulation of iEEG can be used for online control of BMIs, for example during dexterous grasping [125], when selecting between grasp types and elbow movement [50], or for three dimensional cursor control [45]. We therefore sought to determine whether human iEEG could be used to provide simultaneous and independent online control of reaching and grasping movements, thus demonstrating segregation of these two movement types at the spatial scale of iEEG macroelectrodes. iEEG is an attractive platform for the development of BMIs because of the potential for better long-term signal stability than multi-unit recordings, as well as the availability of subjects who have accepted the risks of electrode implantation for the mapping of their seizure onset zones prior to epileptic resection surgery [126].

Relative to scalp EEG, iEEG provides better spatial resolution and better signal quality for high gamma activity [127, 128]. There is substantial empirical evidence from local field potential studies in humans [129] and nonhuman primates [40] that this high gamma activity closely tracks population firing rates. The degree of control that can be achieved with the large-scale population activity recorded with iEEG [130, 43] is unknown, however, especially with chronic training beyond the time constraints of seizure monitoring. To ensure that the risk of long-term electrode implantation is offset by the benefit of stable long-term BMI use, it would be advantageous to confirm at least basic control of the intended prosthetic at the time of implantation.
CHAPTER 3. REACHING AND GRASPING

Previous work in scalp EEG and iEEG has demonstrated two- and three- dimensional cursor control where at least one dimension is controlled by behavior unrelated to the task at hand (e.g., vocalization or tongue movement) [130, 43, 26, 82]. Although it has been demonstrated that training and operant conditioning can be used to learn BMI control on the time-scale of months [131], it is unclear to what extent an unnatural mapping will scale up to more complex tasks in more complex environments. We therefore sought to determine whether the command signals for forward reaching and grasping of the Johns Hopkins University Applied Physics Lab (JHU/APL) Modular Prosthetic Limb (MPL) could be derived from high gamma (70-110 Hz) neural population activity associated with naturalistic reaching and grasping movements, respectively. These commands were interpreted by the hardware in the MPL and converted to multi-axial anthropomorphic movements spanning two controllable joints for forward reaching and 10 controllable joints for grasping.

3.2 Materials and Methods

3.2.1 Subject Info

The subjects for this study were 55 year old (Subject 1) and 30 year old (Subject 2) right-handed males implanted with intracranial electrodes to map the ictal onset zone of medically resistant seizures prior to surgical resection. In Subject 1, an 8x8 grid of subdural platinum-iridium electrodes (Adtech, Racine, WI, 2.3 mm diameter
exposed surface, 1-cm spacing) were surgically implanted over right frontal-parietal regions (see Figure 3.1, in addition to a 4x5 electrode grid over right lateral occipital cortex and a 1x8 electrode strip stretching from right mid-temporal regions to dorsolateral prefrontal cortex (both not shown). Subject 2 was implanted with a 1x8 electrode strip (Adtech; Racine, WI; as above) across right frontoparietal cortex, six depth electrodes with eight platinum macrocontacts each (Adtech; 2.41-mm long, 6.5 mm center-to-center spacing) placed medially from the right premotor area to the posterior parietal lobe, and one hybrid depth with eight platinum macro contacts (Adtech; 1.57 mm long, 5 mm center-to-center) and sixteen microcontacts (Adtech; 75 micron diameter). Neuronavigation via the Cranial Navigation Application (BrainLab; Westchester, IL), was used during placement of the depth electrodes in Subject 2. Anatomical reconstructions of the subjects’ brains with the location of implanted electrodes were generated by volumetrically co-registering the pre-surgical MRI with a post-surgical CT using BioImage (Figure 3.1) [114]. Subject 1’s seizures began after a bout of viral encephalitis with coma at 33 years of age. His complex partial seizures were typically preceded by a somatosensory aura in his left hand with spread to the face and subsequent shaking of the left hand, and were sometimes followed by secondary generalization. Subject 2 had previously undergone chronic recording with partial resection of his right post-central gyrus and superior parietal lobule. Both patients gave informed consent for research testing, which was done in accordance with a protocol approved by the Institutional Review Board of the Johns Hopkins
CHAPTER 3. REACHING AND GRASPING

Figure 3.1: Functional mapping of cue-averaged task-related high gamma activity in training set. (A) Reconstruction of the implanted grid location for Subject 1 is depicted; the electrode used for reaching (number 25) is highlighted in red and corresponds to the channel circled in red in the activation maps below, while the electrode used for grasping (number 11) is highlighted in blue and similarly corresponds to the electrode circled in blue below; the central sulcus is highlighted in green. (B) Reconstruction of the depth electrodes implanted in right hemisphere of Subject 2; electrodes used for reaching highlighted in red, electrodes used for grasping highlighted in blue (transparent medial view in inset).

3.2.2 Neural Signal Acquisition

Using a NeuroPort system (BlackRock Microsystems; Salt Lake City, UT), iEEG signals were initially sampled at 30 KHz with an analog bandpass filter with cutoffs of 0.3 Hz and 7500 Hz. The NeuroPort system then applied a digital 4th order Butterworth lowpass filter with a 250 Hz cutoff and downsampling to 1000 Hz. Artifactual channels were visually identified and excluded from all further analysis. Acquired iEEG signals were broadcast with the universal datagram protocol (UDP) to an experimental workstation, where they could be accessed for online spectral feature
3.2.3 Experimental Procedures

Short offline data sets of 30 (Subject 1) or 50 (Subject 2) auditorily cued trials were each collected for forward reaches and grasping movements. Audio cues of "reach" or "grasp" were delivered via external speakers by E-Prime software (PST,
Inc.; Sharpsburg, PA). For Subject 2 only, the reach and grasp trials were interspersed with 50 "Reach and Grasp" trials. The onset of each trial was manually initiated by the experimenter to ensure that the preceding reach was completed and an additional varying delay had passed before a cue was given. Behavioral states were detected using analog sensors sampled at 1000 Hz on the same hardware as the neural data: 1) the onset and offset of each reaching movement were detected using a pushbutton embedded in a wooden lap desk, 2) the termination of each reach on a distal target was detected using a pushbutton, and 3) the onset and offset of each grasp were detected using a pneumatic squeeze bulb connected via flexible tubing to an electronic pressure sensor. A detailed schematic of the experimental setup is included in Figure 3.3, and a video of the training procedure is shown for Subject 1 in (Supplemental Video 1). Reaches by Subject 1 ranged in duration from 1.3 to 1.8 seconds (median = 1.4 seconds) with response latencies ranging from 330 to 500 ms (median = 410 ms), while grasps (i.e., as detected by the squeeze bulb) ranged in duration from 0.6 to 2.7 seconds (median = 0.9 seconds) with response latencies ranging from 380 to 930 ms (median = 460 ms). Reaches by Subject 2 ranged in duration from 1.9 to 4.8 seconds (median = 2.7 seconds), with response latencies ranging from 450 to 1450 ms (median = 790 ms), while grasps ranged in duration from 0.8 to 3.5 seconds (median = 2.0 seconds) with response latencies ranging from 640 to 2070 ms (median = 1010 ms).
3.2.4 iEEG Electrode Evaluation

Following collection of the reach and grasp datasets, event-related high gamma activations were analyzed. The audio cue played to the subject was split and fed into the BlackRock NeuroPort system; the beginning of this cue was detected and used as a stimulus onset (SO) marker. The 1024 ms prior to SO was pooled into baseline distributions for each channel, while the 3072 ms following the onset of the audio cue
was used as a post-stimulus epoch. The 1024 ms prior to SO and 3072 ms following SO were segmented into 128 ms windows with 112 ms overlap. Each window was reduced to a single estimate of the high gamma analytic amplitude in a 16 ms bin using a Hilbert transform with an embedded, flat-top Gaussian bandpass filter with cutoffs of 72 and 110 Hz. Separate distributions were created for each post-stimulus 16 ms time bin and channel and referenced to the channel baseline distributions using two-sample t tests with significance threshold $p < 0.05$. The thresholds for p-value significance of these tests were corrected for multiple comparisons within each channel using the false discovery rate (FDR) correction [132]. Any resulting significant p-values were then log10 transformed, and any significant modulation was labeled as an increase or a decrease. This resulting matrix of statistical significance measures therefore contained timing information about activation that was used to exclude channels which displayed modulation in response to the audio cue. This entire analysis was performed with custom MATLAB (MathWorks, Inc.; Natick, MA) software, from which the results were available within the experimental session (see Figure 3.2).

### 3.2.5 BMI Model Training

For Subject 1, a final training set was recorded in which the verbal commands "reach," "grasp," and "reach and grasp" were pseudo-randomly chosen and played to the subject via external speakers with E-Prime; this training set contained 46 trials and lasted approximately five minutes. For Subject 2, the 150 trials spanning ap-
proximately sixteen minutes collected for electrode evaluation were used as a training set. Also for Subject 2, the initially trained model was used to drive a virtual version of the MPL [118, 119, 120] as visual feedback during an additional 120 trials (i.e., 40 each of "reach," "grasp," and "reach and grasp"). The iEEG and behavioral data recorded during this block were used as the training set for online testing.

Signals in each training set were first spatially filtered with a common average reference [115] of all channels not excluded by visual inspection because of artifact or noise. Autoregressive power was extracted from the streamed signals using the Burg algorithm with model order 16 on a 400 ms window. The logarithm of the spectral power from components between 72.5 and 110 Hz were then averaged to yield an estimate of the broadband high gamma power. In offline data collection for model training purposes, feature extraction windows were overlapped by 300 ms.

In Subject 1, one electrode each was chosen for reach and grasp using information from the functional maps of post-stimulus activation. The high gamma log-powers during movement and rest movement were compared to manually establish a threshold for movement classification. In Subject 2, four channels each were selected as model inputs to separate binary linear discriminant analysis (LDA) classifiers for reach and grasp. In addition, transition probabilities were adjusted manually before the testing session to smooth the output from the classifier. For this study, we used a probability of 0.95 for the probability of a rest classification if currently at rest (i.e., 0.05 for a movement classification), and 0.8 for the probability of a movement classification if
CHAPTER 3. REACHING AND GRASPING

currently in the movement state (i.e., 0.2 for a rest classification).

3.2.6 JHU/APL Modular Prosthetic Limb

Developed by JHU/APL under the Defense Advanced Research Project Agency (DARPA) Revolutionizing Prosthetics Program, the MPL (Figure 3.4) is an advanced upper-body extremity prosthetic and human rehabilitation device [133]. The MPL has 17 controllable degrees of freedom (DoF) and 26 articulating DoF in total (Figure 3.4). To facilitate control from neural decoded motion intent, the MPL has a custom software interface, VulcanX, that receives movement/motion commands locally and sends them over a controller area network (CAN) bus to a limb controller (LC) board in the hand of the MPL [134]. Three types of high-level control commands, passed through VulcanX, are fused together to form individual actuator commands by the LC: 1) Degree of Motion Control (DOM) commands, which allow each degree of motion to be controlled individually with position and/or velocity commands; 2) Endpoint Control (EP) commands, which allow the hand’s position and orientation to be controlled in Cartesian space using a Jacobian-based algorithm for computing inverse kinematics; and, 3) Reduced Order Control (ROC) commands, which allow pre-programmed hand grasp patterns to be actuated in a coordinated fashion as a single degree of freedom [135]. EP velocity and ROC commands were utilized to control reach and grasp, respectively, in this study.
CHAPTER 3. REACHING AND GRASPING

Figure 3.4: The MPL software and hardware architecture consists of a distributed network of processors that include a Neural Fusion Unit (NFU), a Limb Controller (LC), 10 Small Motor Controllers (SMC), 4 Large Motor Controllers (LMC), and 3 Wrist Motor Controllers (WMC). The NFU is a processor capable of running on-board neural decoding and sensory stimulation algorithms for generation of limb motion commands. The LC is the main processor of the limb system and is responsible for receiving limb control commands, running high-level control algorithms, and coordinating the control of the individual motors in the system. The LMC, WMC, and SMC integrated software/hardware systems, which receive real-time data from temperature, torque, and position sensors located within each joint, are responsible for providing closed-loop position, velocity and torque control of the brushless DC motors in the limb system. In addition to internal polling, these joint sensors broadcast information to the user/experimenter along the MPL’s external communication CAN bus for data logging and external control processing.

3.2.7 Online Testing

Once the high gamma thresholds for movement were established, classification outputs from the trained models of reach and grasp movements were simultaneously used to actuate the MPL via the VulcanX interface. Whenever the classifiers predicted the subject was reaching and/or grasping, the MPL was commanded to reach and/or grasp, respectively, at a set rate. If the subject was predicted to be resting, the limb was commanded to return to its rest posture at an equal rate. For Subject 2 only, the return rate for reaching was adjusted to be 50% higher than the forward rate.
CHAPTER 3. REACHING AND GRASPING

High gamma log-power calculations were performed in 400 ms windows (i.e., as in training) computed as quickly as possible on the streaming iEEG signals to provide inputs to the trained model (i.e., 11 ms for Subject 1, slowed to 32 ms for Subject 2 purposefully to avoid inundation of the MPL). Both subjects completed three blocks of online trials by performing the same overt movements with their native limbs as during the training set. In Subject 1 only, the second and third blocks were separated by a battery of physical and imagined movements that were not analyzed as a part of this study.

3.2.8 Quantitative Evaluation of Control

The physical movement blocks lasted approximately 4, 11, and 13 minutes for Subject 1 and 11, 15, and 10 minutes for Subject 2 (respectively). The MPL VulcanX control software created a log of commands sent to the limb with timestamps, which was compared offline to the timestamps of salient cues and behavioral events recorded by the BlackRock system (e.g., subject leaves the home switch, subject grasps the squeeze bulb, etc.). Trials were designated as starting 500 ms prior to the earliest of the reach and/or grasp onsets and ending 500 ms prior to the onset of the next trial. For each trial, we recorded the proportion of correct commands (e.g., the percentage of grasp’ commands with a positive velocity when a physical grasp was performed) in a window of equal length to the corresponding physical movement duration for that trial. To account for variable response latencies by the subject and
CHAPTER 3. REACHING AND GRASPING

an inconsistent system latency, the start of the window relative to the onset of the trial was selected individually for each trial to maximize the accuracy. For reach-and-grasp trials, durations and latencies were selected separately for the reach and grasp components. As a control, a window whose length equaled the average duration of the reaches or grasps was used to compute the peak reach or grasp command accuracy in grasp-only and reach-only trials, respectively. Accuracy for each trial was computed as the average of the single trial sensitivity (i.e., proportion of reach or grasp commands within the selected movement window) and the single trial specificity (i.e., proportion of rest commands outside of the selected movement window). The median reach command accuracies for reach-only vs. grasp-only and reach-and-grasp vs. grasp-only and the grasp command accuracies for grasp-only vs. reach-only and grasp-only vs. reach-and-grasp were compared using a nonparametric two-sided Wilcoxon rank sum test.

3.3 Results

Both subjects were able to attain a high degree of subjective control over reaching and grasping with the MPL across the experimental session with no model adaptation while moving their native limbs (Supplemental Video 2 and 3). Furthermore, both subjects were able to achieve a level of performance throughout the experimental session that was qualitatively similar to the first block.
CHAPTER 3. REACHING AND GRASPING

We investigated the spectrogram of modulation time-locked to salient stimuli and behavioral events to validate our choice of the high gamma band for online control. As shown in the functional mapping results (Figure 3.2), the electrodes used for control of the MPL exhibited robust high gamma modulation. Figure 3.5 shows the time-frequency response of the reaching electrode during reach-only and reach-and-grasp trials of the online task, as well as the grasping electrode during grasp-only and reach-and-grasp trials. High gamma modulation in the reaching electrode occurred within the frequency range of 80-160 Hz for Subject 1, in contrast with the more spectrally restricted 60-120 Hz modulation in the grasping electrode. Subject 2 displayed activation at a lower frequency range, centralized around 40-90 Hz. These frequency ranges of power modulation show that while our choice of 72.5-110 Hz for control may not have exactly matched the neurophysiological response to the task, it did capture a substantial amount of the power modulation for both tasks. The temporal envelope of activation was relatively restricted in the reaching electrode for both subjects, with mean power modulation peaking roughly 200 ms before the onset of movement. Subject 2 had similarly tight timing in grasp-related cortical activation. In contrast, in Subject 1 power modulation in the grasping electrode began an average of 300 ms prior to movement onset and peaked more than 300 ms after movement onset. The reach-related high gamma power modulation also differed from grasp-related power modulation in the presence of two distinct temporal peaks, time-locked to outward reach and the subsequent return to rest. Figure 3.5 (bottom row) provides
CHAPTER 3. REACHING AND GRASPING

Figure 3.5: Average power spectral density (PSD) relative to baseline, aligned to movement onset. (A) Reach and grasp electrodes are shown for Subject 1, and (B) two representative electrodes are shown for Subject 2. The first vertical dashed line corresponds to average audio cue onset. The solid line denotes movement onset (MO). In reach trials, the dashed lines after the solid line correspond to the average time of the reach completion (pressing target button, PT), release of the target button (RT), and return to home (resting on the home switch), from left to right. The dashed line in the grasp trials corresponds to average grasp completion time.

verification that gamma power modulations in the grasp and reach electrodes were markedly lower during execution of reach and grasp, respectively.

During online control in Subject 1, we observed that control of grasping was less reliable for reach-and-grasp trials than grasp-only trials. Figure 3.5 (middle row) shows that high gamma modulation in the reaching and grasping electrodes during reach-and-grasp trials was qualitatively reduced relative to reach-only and grasp-only trials. To evaluate this effect, log high gamma power was extracted in 300
ms around the onset of movement. Statistical analysis revealed that log power in the grasping electrode around the onset of grasp was significantly higher in grasp-only than in reach-and-grasp trials ($p < 0.05$, Wilcoxon test); the log power in the reaching electrode around the onset of reach was not significantly different in reach-only and reach-and-grasp trials, however. Identical analyses performed in Subject 2 did not reveal any significant differences in movement-related power modulation between reach-and-grasp trials and either the reach-only or grasp-only trials in any of the electrodes used for control ($p > 0.05$, Wilcoxon test, Bonferroni-corrected).

To evaluate the high gamma power modulation associated with movement state in reach-only trials, log power was also extracted in time windows around the onset of stable hold and the onset of return, in addition to a baseline window preceding the cue. For Subject 1, log-power in the reach electrode was significantly higher in the reach window and return window than in the hold window, all of which were significantly higher in the baseline window ($p < 0.05$, one-way ANOVA with Tukey’s honestly significant difference post-hoc). In all four electrodes used for reaching control in Subject 2, median hold activity was lower than median reaching and returning activity; the difference was significant in three out of four electrodes ($p < 0.05$, one-way ANOVA, with Tukey’s honestly significant difference post-hoc). Reaching, returning, and intermediate hold windows similarly exhibited higher levels of high gamma activity than in baseline windows in Subject 2.

Classification accuracy for both reaching and grasping started and remained high
throughout all three blocks of the online task. The mean reach classification accuracy across all trials was 86% (Subject 1) and 82% (Subject 2) for reach-only trials; the reach accuracy across reach-and-grasp trials was 83% (Subject 1) and 89% (Subject 2). The mean grasp classification accuracy across all grasp-only trials was 81% (Subject 1) and 96% (Subject 2); the grasp accuracy across reach-and-grasp trials was 55% (Subject 1) and 88% (Subject 2). The evolution of classification accuracies showed no significant effect of block ($p > 0.05$, one-way ANOVA) in either subject. The trial-by-trial reach and grasp accuracies are depicted in Figure 3.6. Reach accuracies were significantly higher than chance for both reach-only trials and reach-and-grasp trials ($p < 0.05$, Wilcoxon test with Bonferroni correction), while grasp accuracies were significantly higher than chance for grasp-only trials ($p < 0.05$, Wilcoxon test with Bonferroni correction), but not reach-and-grasp trials in Subject 1 only ($p = 0.078$, Wilcoxon test). Grasp accuracies were significantly higher in grasp-only trials than in reach-and-grasp trials in both subjects ($p < 0.05$, Wilcoxon test). Reach accuracies were not significantly higher in reach-only trials than in reach-and-grasp trials for Subject 1 ($p > 0.05$, Wilcoxon test), although reach accuracies were higher in reach-and-grasp trials than in reach-only trials for Subject 2 ($p > 0.05$, Wilcoxon test).

To investigate whether reaching and grasping were indeed independent, sham sensitivities were calculated as a control; reach sensitivities were calculated during grasp-only trials and grasp sensitivities were calculated during reach-only trials (Figure 3.6). Since no physical reaches took place in grasp-only trials, nor physical grasps
Figure 3.6: Limb performance accuracy metrics. (A, B) Accuracies are shown for reaching and grasping during trials where reach and grasp were executed simultaneously. (C, D) Reach and grasp accuracies are shown for reach and grasp only trials, respectively. The vertical dashed lines in A-D denote separate blocks. Distributions are summarized with boxplots of the peak sensitivities for grasps in Subject 1 (E), reaches in Subject 1 (F), grasps in Subject 2 (G), and reaches in Subject 2 (H). Bars above the boxplots with asterisks mark distributions with significantly different medians ($p < 0.05$, Wilcoxon test).

during reach-only trials, the average reach and grasp durations were used as surrogates. Peak reach sensitivities were significantly higher in cued reach-only and reach-and-grasp trials than in cued grasp-only trials for both Subjects ($p < 0.05$, Wilcoxon test); reach sensitivities were significantly higher in reach-only trials than in reach-and-grasp trials for Subject 2 ($p < 0.05$, Wilcoxon test) but the difference was not significant in Subject 1 ($p = 0.16$, Wilcoxon test). Peak grasp sensitivities were higher in cued grasp-only and reach-and-grasp trials than in cued reach-only trials for
both subjects \( (p < 0.05, \text{Wilcoxon test}) \); grasp sensitivities were significantly higher in grasp-only trials than in reach-and-grasp trials for Subject 1 \( (p < 0.05, \text{Wilcoxon test}) \) but the difference was not significant in Subject 2 \( (p = 0.15, \text{Wilcoxon test}) \).

### 3.4 Discussion

We enabled two human subjects to control the MPL using a control scheme that exploited individual functional anatomy, i.e., the population responses in cortical regions used for control of each subject’s native arm. This allowed our subjects to achieve control without extensive training. To identify iEEG control sites and characterize their response selectivity, we used iEEG functional mapping during reaching and grasping. By using electrodes over cortical areas that were differentially activated during reaching and/or grasping, we were able to afford the patient independent control over the reaching and grasping functionalities of the MPL arm. We showed that these two movements, when executed individually, elicited cortical responses in the high gamma band (72-110 Hz) that generalized to their simultaneous execution, although the same responses occurred with a reduced magnitude.

Additionally, the subject’s control over the arm did not wane over the course of three separate blocks using thresholds derived from a short training block. Models were equally effective across blocks with no adaptation or re-training, providing evidence that control was achieved by accurately detecting the naturalistic circuits for
reaching and grasping, not via adaptation or operant conditioning. Reach and grasp commands were controlled independently, suggesting functional segregation of these movements at the spatial scale of clinically routine iEEG electrodes. There is abundant evidence from experiments in non-human primates that reaching and grasping engage different networks of cortical areas \[136\]. As in non-human primates, human premotor cortices engaged by reaching are likely dorsal to those engaged by grasping \[137\]. As expected, the iEEG site activated by and used for control of reaching was dorsal to the site activated by and used for control of grasping.

Our BMI used event-related high gamma power modulation as an index of task-related neural activity during physical movements. This choice was based on a body of literature which demonstrates that high gamma band modulation is an index of cortical processing in humans \[38, 39, 138, 139\] and recent experimental evidence that high gamma power changes are strongly and positively correlated with the firing rates of neuronal populations in close proximity to recording electrodes \[129, 140, 40, 141\]. Our findings are consistent with empirical evidence that compared with power changes in other frequencies, high gamma power augmentation has high spatial selectivity with respect to task-related cortical activation, such that adjacent iEEG electrodes can yield signals with greater independence at higher frequencies \[142, 143\]. High gamma responses are also robust enough to be detected in single trials, a necessary requirement for BMI applications. Furthermore, several studies have shown that high gamma features extracted from human iEEG outperform corresponding lower
CHAPTER 3. REACHING AND GRASPING

frequency features for offline motor decoding \[47\] \[85\] and online BMI control \[50\] \[43\] \[51\].

This study focused on the control of reach and grasp in the MPL since they are fundamental to upper limb use and provides a proof of concept for the systems-level integration groundwork necessary for more complicated and dexterous tasks. Reaching and grasping movements were decoded for actuation of the MPL with high accuracy and stability; furthermore, this was achieved in a clinical epilepsy monitoring setting under time constraints that did not allow for long-term training or testing. Although this prohibited testing the long-term stability of MPL control, it did demonstrate the feasibility of obtaining MPL control within a compressed timeframe, which could have important clinical benefits. Specifically, it would be highly advantageous to demonstrate acceptable brain control of a neuroprosthetic at the time of surgical implantation, in order to verify the placement of electrodes and troubleshoot any technical difficulties at the time of the operation. Non-invasive methods of functional mapping (e.g., fMRI) can be used to perform gross surgical planning, but intraoperative verification of control with iEEG would be extremely useful to refine the final implantation site. This would help to avoid the need for re-implantation because the patient is unable to control the neuroprosthetic. This would be both costly and increase surgical risk. Although the total time for our experiment was longer than that of an awake craniotomy, most of this time was due to experimental setup and troubleshooting, and thus could be reduced with additional practice.
CHAPTER 3. REACHING AND GRASPING

We observed during online testing with the subject that it was fairly common for the MPL to exhibit a secondary reach or grasp as the subject returned to the resting position. This corresponded to a burst of high gamma activity as the subject initiated return of his limb to the home switch or as the subject relaxed his hand after squeezing the bulb. This was best demonstrated in the reaching trials by post-hoc offline analysis of the high gamma power in windows associated with reaching to, holding at, and returning from the distal reach target, which demonstrated a higher degree of modulation for reaching and returning than for the intermediate holding in the reaching electrode for Subject 1 and for a subset of the reaching electrodes in Subject 2.

This report provides additional evidence for the potential utility of iEEG as a source of control signals for BMIs. Although the participants in this study did not suffer from upper limb paralysis, we believe that the technique of rapid trial-averaged spatiotemporal mapping of high gamma modulation can be used to identify sites that are activated when subjects with motor impairments attempt to perform movements. These patients often have residual motor function and could attempt to move with assistance, or could be moved passively, or could observe upper limb movements in a trial-based framework.

A large amount of decoding and BMI success has been achieved using command signals derived from iEEG [50, 45, 130, 43, 51]. Although iEEG macroelectrodes [130, 43], iEEG microelectrodes [144], and multi-electrode arrays [33, 36] have all
been used to demonstrate effective BMI control in small populations, no large-scale longitudinal studies have compared the tissue response and control performance between these classes of implants. Much previous work has illustrated a significant redundancy of motor encoding at the single neuron level [145], suggesting that population activity could be useful for prosthetic control. Nevertheless, there is evidence from studies in motor, perceptual, and cognitive systems that the richness of encoding increases with improvements in spatial resolution (i.e., iEEG macroelectrodes exhibit coarser encoding than iEEG microelectrodes, and iEEG microelectrodes exhibit coarser encoding than local field potentials from multi-electrode arrays) [146]. It is possible that as the spatial resolution of iEEG implants improves and more comparative studies are done between iEEG and multi-electrode arrays, that iEEG implants for BMI control will be an attractive option for some patients [147, 148]. In the meantime, iEEG recordings in patients undergoing epilepsy surgery will continue to serve as a platform for demonstrating the degree of useful control that can be achieved without extensive training, prior to chronic implantation of iEEG electrodes for BMIs.

3.5 Study Acknowledgments

The authors wish to thank Heather Benz, Anna Korzeniewska, Zachary Huff, and Griffin Milsap for lab meeting discussions of our approach and analysis, Howard
CHAPTER 3. REACHING AND GRASPING

Conner for building the stand used to hold the MPL, and Charles Schuman for constructing the mounting interface.

This chapter has supplementary downloadable material available at http://ieeexplore.ieee.org/xpl/abstractMultimedia.jsp?arnumber=6646313.
Chapter 4

Individual Finger Control of the Modular Prosthetic Limb using High-Density Electroctographicography in a Human Subject

4.1 Overview

Brain-machine interfaces (BMIs) offer a promising approach for restoring function to patients with severe paralysis. By delivering direct cortical control over robotic prosthetic devices, BMIs could enable spinal cord injury (SCI) patients to perform activities of daily living (ADLs) necessary for self-sufficiency. However, many of
CHAPTER 4. INDIVIDUAL FINGER CONTROL

these activities, such as preparing food or taking medications, require a level of hand
dexterity that has yet to be achieved by BMIs. This dexterity can only be achieved by
means of complex hand movements based on control of individual fingers. Decoding
the neural correlates of finger control has been explored \cite{85, 149, 150, 151, 152, 153},
but researchers have so far been unable to demonstrate online independent control of
individual fingers.

To date, online cortical control of finger movements has only been achieved in the
context of coordinated movements of multiple fingers. Electrocorticography (ECoG)
signals recorded from sensorimotor areas have been used offline to reconstruct hand
aperture \cite{86} and classify different hand gestures \cite{47, 49, 154}. ECoG has also been
used online to continuously control grasping movements in parallel with arm move-
ments \cite{50, 4} and to classify different grasp types \cite{49}. Neuronal firing rates from
microelectrode arrays (MEAs) have also been used to achieve online cortical control
over grasping in both nonhuman primates \cite{155, 61, 156, 157} and humans \cite{33, 36, 34}.

Offline decoding of single finger movements has been achieved with MEAs in non-
human primates \cite{149, 158}, but MEA recordings in humans are rare and suffer from
attrition of reliable single units over time \cite{159}. Moreover, the restricted cortical areas
covered by MEAs may not sufficiently sample the cortical networks of potential use
in BMIs. Finger movements are represented in a larger area of sensorimotor cortex
in humans than monkeys, and although there may be significant overlap in motor
representations for individual fingers \cite{160, 161}, it is inherently difficult for MEA’s to
leverage what somatotopic organization is present.

In contrast, ECoG is able to deliver stable large-area coverage of sensorimotor areas \[144\], delivering control from multiple anatomical sites. For example, coverage of both arm and hand areas with a single ECoG grid enabled a subject to simultaneously control the ability to reach and grasp with a prosthetic limb \[4\]. This and other studies have chiefly used ECoG high gamma responses for online control. These spectral responses are robust enough to be detected in single trials, but because they reflect firing rate changes in populations of cortical neurons, their utility depends on the degree of functional segregation among these populations.

In spite of the evidence of limited finger somatotopy in motor cortex \[160, 162, 161\], both ECoG and fMRI studies have found a degree of separability in the peak population responses for different fingers in the precentral gyrus \[163, 164\]. These differential finger responses have been used to perform offline classification and reconstruction of finger movements from ECoG recordings over sensorimotor areas \[85, 150\]. Offline classification between rest and four fingers (ignoring pinky) was near perfect in one subject in \[49\], though the system’s performance dropped drastically when translated to an online grasp classification system synchronized to cue presentation. Additionally, several groups have performed offline regression of single finger movements \[85, 151, 152, 153\]. However, until now these achievements in offline classification and regression have not been translated into asynchronous (i.e., without knowledge of the cue) online control of individual fingers.
CHAPTER 4. INDIVIDUAL FINGER CONTROL

Here we show that high-density electrocorticography (ECoG) electrodes over sensorimotor areas can not only discriminate individual fingers offline, but that they can be used to asynchronously detect and classify finger movements online. A human subject demonstrated online individual finger control of the highly dexterous Modular Prosthetic Limb (MPL) [133]. The BMI relied only on the subject’s cortical responses during movements of his corresponding native fingers, without extensive training or arbitrary mapping of user inputs to finger movements. Offline analysis indicated control was not solely from somatosensory feedback, suggesting this control can be utilized by deafferented paralyzed patients. Our approach of using the native functional anatomy of sensorimotor cortex obviates the need for operant conditioning, potentially providing immediate intuitive control to patients that can be expanded to a large number of degrees of freedom without placing significant cognitive burden on the patient. These results demonstrate, for the first time, online control of individual fingers of a prosthetic hand, as well as offline analyses indicating a role for pre-movement signals in achieving this control.

4.2 Methods

4.2.1 Subject Info

The patient, a 20-year old right-handed male, underwent a left craniotomy for implantation of intracranial electrodes to localize the brain regions responsible for
his seizures to guide resective surgery. These included a high-density 8 x 16 array of subdural electrodes (PMT Corp., Chanhassen, MN; 3mm center-to-center spacing, 1mm diameter) over sensorimotor regions. ECoG signals were recorded using distant standard subdural electrodes (2.3 mm diameter exposed) for ground and reference.

The subject gave informed consent for testing to be done according to a protocol approved by the Institutional Review Board of the Johns Hopkins Medical Institutions. Electrodes were localized with respect to cortical surface anatomy by volumetric co-registration of the subject’s pre-implantation MRI with their post-surgical CT using the BioImage Suite \[114\]. This was used to help guide our electrode selection process for the BMI. After running the BMI, the reconstruction was checked against intraoperative photos from the implantation and explantation of the high density grid. The electrode locations on a two-dimensional snapshot of the reconstruction were manually adjusted relative to the underlying cortex via rotation, scaling, and translation of the grid to optimize the alignment between the grid and prominent gyral and sulcal landmarks present in both the 3D reconstruction and the intraoperative photos. Figure 4.1 displays the corrected reconstruction.

4.2.2 Offline Experimental Testing

A finger tapping task was performed to collect training data for the online finger decoder. Online BMI testing was performed the subsequent day. A passive vibration task and an additional finger tapping task were performed for offline analysis. Table
4.1 outlines the experiments performed. The tasks are described in more detail below.

4.2.3 Vibrotactile Stimulation

To investigate activation due to somatosensory feedback, we ran a vibrotactile stimulation experiment. While activation of post-central gyrus has been observed in paralyzed individuals [165] and able-bodied individuals during motor-imagery (without overt movement) [166, 167, 168, 169, 42], deafferented subjects lack cortical activation directly induced by somatosensory feedback. To address this, we used a vibrotactile stimulation experiment as a control. Vibrational motors were placed on
CHAPTER 4. INDIVIDUAL FINGER CONTROL

Table 4.1: Experimental Overview

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vibrotactile Stimulation</td>
<td>Control for sensory feedback</td>
</tr>
<tr>
<td>Finger Tapping</td>
<td>Two sessions: one to provide training data for our decoder, and one as an independent dataset for offline analysis</td>
</tr>
<tr>
<td>Online Finger BMI Testing</td>
<td>Online test of the decoder built on finger tapping data</td>
</tr>
</tbody>
</table>

the fingertips of the subject’s right hand. The motors were turned on for 0.5 seconds every two seconds in a pseudorandom order by an Arduino microcontroller (Sparkfun Electronics, Boulder, CO) with two Adafruit v2.3 motor control shields (Adafruit Industries, New York City, NY). Each finger was vibrated 50 times, for a total of 250 trials.

4.2.4 Finger Tapping

The subject performed a visually cued finger tapping experiment. Trials were initiated by an experimenter on a presentation computer running psychtoolbox-3 in MATLAB (MathWorks, Inc., Natick, MA, USA). After a trial was initiated, the system paused for two seconds, then displayed an image of a hand with one finger pseudo-randomly highlighted. The patient then alternately flexed and extended the corresponding finger on his right hand 5 times. Each finger was cued 25 times, for a total of 125 trials. Finger movements were recorded with a CyberGlove II
4.2.5 BMI Classifier Training

To perform online classifications of finger movements, we trained a hierarchical classifier on the offline finger tapping data. For any given time point, a binary linear discriminant analysis (LDA) classification was first used to determine if a movement was occurring. This allowed us to decode finger movements without synchronization to any cues (i.e. asynchronously). If a movement was detected, a classification from a second LDA classifier was used to determine which of the five fingers moved. The binary movement classifier was given a first-order Markovian transition probability; the prior for a prediction at time $t$ was determined by the prediction at time $t-1$. These priors were calibrated online just before the online testing session to find a good tradeoff between false positives and false negatives.

To extract features for training the classifiers, we first re-referenced the high-density grid channels to a common average reference (CAR) with bad channels removed. The high gamma power was then extracted from each channel using the Hilbert transform with a flat-top Gaussian bandpass filter with cutoffs of 72 and 110 Hz using a 256 ms window and 128 ms step size. The average of the high gamma powers extracted from an 896 ms period before cue onset of each trial was used to gauge baseline activity for each electrode. The average of the high gamma power in an 896 ms period during finger movement of each trial was used to estimate activation...
CHAPTER 4. INDIVIDUAL FINGER CONTROL

Figure 4.2: Flow diagram of the online BMI processing. Signals are rereferenced, high gamma is extracted and smoothed, and then a hierarchical classifier is used to determine if/which fingers should be moved.

due to the finger movement. This resulted in 125 feature vectors (one for each trial) for both baseline and activation. These feature vectors were then used to train the movement detection and finger classification LDA classifiers.

When using the classifiers online, we used the same feature extraction workflow we used for training. To ensure smooth classifications, the features from the previous 50 samples were averaged at each time point. Due to slower computation speeds than anticipated, this resulted in a moving average of approximately 2.3 seconds of data to be used for classifications.
4.2.6 Online Finger BMI Testing

When a finger movement was detected from the neural data, the corresponding MPL finger was commanded to flex at a fixed velocity. If a finger was not classified as moving, it was commanded to extend back towards the resting position at a fixed velocity. Before the online testing session began, the subject was given open use time and a practice session to acclimate to controlling the MPL fingers. During this acclimation period, the false positive rate of the movement classifier was tuned by adjusting the priors on the Markovian transition probabilities. The speed at which the fingers would flex/extend was also adjusted at this time, resulting in an extension velocity set to 40% of the flexion velocity. Figure 4.2 displays a flow diagram for the online BMI.

During the online testing session, the experimenter typically waited for all the MPL fingers to return to rest before initiating the next trial. There was a two second system pause before the appearance of the visual cue after trial initiation. This was used as a baseline (rest) period during post-hoc analysis.

4.2.7 BMI Control Evaluation

To evaluate the performance of the BMI control, we inspected the accuracy of the binary movement detection classifications and the 5-way individual finger classifications. Movement times relative to trial onsets were obtained from video analysis.
CHAPTER 4. INDIVIDUAL FINGER CONTROL

for trials where the subject’s hand was visible. The proportion of movement vs non-
movement predictions was binned in 250 ms bins relative to movement onset and
averaged across trials. The proportion of movement predictions over time was used
to evaluate the performance of the movement detection classifier.

The individual finger classifier performance was evaluated similarly by binning the
single finger classifications that occurred when a movement was predicted. Because
it was observed that the ring and little finger movements were often linked, the
classification accuracy was calculated both considering them as the same and as
separate fingers. Finger classification confusion matrices were calculated at the onset
of peak movement detection, and during the time period where movements were most
consistently predicted. We also evaluated the global asynchronous finger classification
performance. This was accomplished by calculating the accuracy of the individual
finger classifier at any time a movement was predicted within seven seconds of cue
onset. We also calculated the percentage of trials where the cued finger, and not any
of the others, was fully flexed by the MPL. To find the balanced accuracy of flexing
the cued finger, we averaged the accuracy of fully flexing during the correct trial with
the accuracy of not fully flexing during the incorrect trial.

To approximate the relative contributions of the electrodes to classification accu-
racies, we examined the cross-validated regularized LDA performance on its training
data with varying numbers of electrodes.
CHAPTER 4. INDIVIDUAL FINGER CONTROL

4.2.8 Offline Analysis

Offline analysis was performed to determine the expected range of decoding performance with an optimized electrode selection during time periods preceding somatosensory feedback. While a lack of somatosensory feedback does not preclude activation of S1 [171, 172, 173], direct cortical activation caused by somatosensory feedback is not be present in deafferented patients in need of BMIs. We attempted to control for this by comparing the time course of finger-specific activation during the finger tapping and finger vibration tasks. We expect similar patterns of activation between paralyzed and able-bodied patients during the time period preceding somatosensory feedback, as suggested by the spatial patterns of activation seen in paralyzed patients during attempted movements [165]. The signals during this pre-sensory time period are likely representative of the minimum of what could be expected from paralyzed individuals for BMI control.

To view the time-course of activation relative to movement onset, the ECoG and CyberGlove data were segmented using the CyberGlove recordings. However, it was observed that the subject often initiated trials with a gross relaxation/extension movement of his hand and fingers before flexing the finger of interest. We therefore aligned the recordings separately to both the onset of any detected movement and the onset of flexion by the finger of interest. We then excluded any trial with more than a 400 ms disparity between the two alignment points, or excessive hand movements before onset of the finger movement; 19 trials were excluded in this way. The selection of
CHAPTER 4. INDIVIDUAL FINGER CONTROL

onset times was verified by performing cross-validated LDA to predict which finger was moving using only the CyberGlove data (i.e., finger kinematics, not neural data). Chance levels were computed via permutation testing, and time points significantly above chance were detected via the Wilcoxon rank-sum test with a Bonferroni correction for multiple comparisons. We found the CyberGlove tracings of hand movement had differentiable features for different fingers during the preliminary hand movement. We therefore aligned to our more conservative estimate of movement onset for all our subsequent offline analyses.

The conservative movement onset marker was used to compare the time course of activation during the finger tapping and vibration experiments. Spectrograms were created using the multitaper method from Chronux toolbox (http://chronux.org) with a window size of 128 ms, step size of 10 ms, passband of 16-160 Hz, cutoff frequencies of 10 Hz and 170 Hz, a time-bandwidth product of three, and five tapers.

We then investigated the time course of high gamma activation by using the same high gamma feature extraction method used for our online BMI. We removed visibly bad electrodes, CAR filtered the high-density grid electrodes, and used the Hilbert transform with a 72 to 110 Hz bandpass filter to compute the high gamma band power with a window of 128 ms and a 16 ms step size. For every electrode in the motor and vibrotactile stimulation experiments, we only included activity during trials for the finger that evoked the strongest high gamma response. This allowed us to detect significant trial-averaged activations even if an electrode was only modulated
CHAPTER 4. INDIVIDUAL FINGER CONTROL

by one finger. We then performed a Bonferroni corrected Wilcoxon rank-sum test for significant activation of high gamma relative to baseline at every point between 256 ms before and 240 ms after the onset of movement (or vibration, in the sensory trials). This allowed us to identify the earliest time points of cortical activation in each electrode for the motor and sensory task.

Finally, we investigated the time course of discriminable information between individual fingers during the motor and vibrotactile experiments. This was done to examine the decoding accuracy of finger movements during the period of time before finger-specific information was received through somatosensory feedback, which would not be present in deafferented subjects. The extracted high gamma features were smoothed using a 15th order moving average filter (window centers spanning 224 ms). We then performed 10-fold cross-validated LDA classification on the smoothed features; ten disjoint testing sets (each containing 10% of the trials) were each evaluated by (during each fold, separately) using the other 90% of the trials for feature selection and model training. The feature selection was performed using cross-validated regularization \cite{174}. This regularization phase helped inform us as to what electrodes most consistently contributed to decoding performance. The number of times each electrode was chosen across the 10 folds of cross-validation was used as an index of its importance. We used extracted features in a causal manner, meaning that for a given time point $t$, features pertaining to that time point were extracted from windows which did not extend past $t$. This alignment was used to ensure our classifications
CHAPTER 4. INDIVIDUAL FINGER CONTROL

were not dependent on future information, and therefore approximate the earliest time points when sensory feedback could have been included in the classifications. Feature extraction was repeated in a non-causal manner, aligned to the peak flexion time with a 30th order moving average filter (window centers spanning 464 ms) to estimate the upper limit on expected online decoding performance.

4.3 Results

4.3.1 BMI Electrode Selection

Preliminary inspection of high gamma activation during the motor task and vibrotactile stimulation, along with a preliminary reconstruction of the electrode locations, informed our feature selection process for the online decoder. Spectral features were used from a subset of electrodes that was chosen to avoid postcentral somatosensory cortex, though an error in the preliminary electrode localization led to the inclusion of two electrodes over postcentral gyrus. In training data from this subset of electrodes, cross-validated classification achieved 80% accuracy. Figure 4.3 depicts the locations of the electrodes chosen for the training set and for online BMI testing (below), as well as which electrodes showed significant activation during the motor and sensory tasks.
Figure 4.3: Electrodes used for online BMI. Starred electrodes were selected for online BMI control. Sulci are accentuated in the inset for improved visibility, with the central sulcus highlighted in green. Post-hoc analysis showed the electrodes with gold stars contributed the most to decoding accuracy on the training set. Light blue electrodes showed significant activation during vibrotactile stimulation, and red electrodes were active during the motor task. The electrodes that were not available for the offline analysis are filled black. Purple outlines the interhemispheric fissure and red outlines the previously resected superior frontal gyrus.

4.3.2 Online Testing

The subject performed 39 visually cued trials during online control of the MPL fingers (example performance, Supplementary Video 1). The subject began moving the prompted finger, on average, 1.43 seconds after cue onset.

Figure 4.4 shows the accuracy of the binary movement classifier and the individual finger classifier as a function of time relative to cue onset. Movement detection accuracy reached a peak of 97%, sustained between 1.6 and 3.1 seconds after movement onset. Classifier accuracy during the baseline period was 87%, resulting in a balanced
Figure 4.4: Classification accuracies over time for movement and finger classification. Classifications were aggregated in 250 ms time bins, and accuracies were averaged across trials. The black dashed line depicts the average proportion of predictions that a window contained movement (i.e., versus rest). The solid black vertical line depicts movement onset.

accuracy of 92%. Individual finger prediction accuracy peaked at 81% (chance 20%) when predicting which of all five fingers was moving, and 94% when classification for the ring and little finger was coupled (chance 25%).

At the onset of the peak movement detection period (1.6 seconds post-movement-onset), classification among all five fingers had an accuracy of 76%. When combining the ring and pinky fingers (they often moved together), accuracy was 88% (Figure 4.5). The average classification accuracy during the peak movement detection period
Figure 4.5: Confusion matrices for finger classifications. The left column of matrices shows results for all five fingers, the right column shows results with pinky and ring fingers combined. The top row depicts the confusion matrices at the onset of the peak movement classification time, middle shows the average over the 1.5 seconds of peak movement detection time, and the bottom is any time within seven seconds of cue onset.

(1.6-3.1 seconds post-onset) was 67% for all five fingers, and 79% when combining ring and pinky fingers. The average accuracy at any time a movement was detected within seven seconds of trial onset (cue displayed) was 55% and 68% for five and four fingers, respectively. Given that the accuracy in the baseline period was 87%, this equates to 60% and 72% accuracy for 6 and 5 class classifiers that include rest (chance levels of 17% and 20%, respectively).

The balanced accuracy of flexing the cued MPL finger (within seven seconds of trial onset) was 64%, and 77% when pinky and ring positions were summed together. The cued finger was the only finger fully flexed by the MPL during 43% of the trials, and 62% when pinky and ring commands were combined. The average traces of the
Figure 4.6: Average normalized MPL finger position of all five fingers during online BMI control (left), and with ring and pinky commands aggregated (orange, right). Solid lines depict the average position of the cued finger as a proportion of maximum flexion, where 1 corresponds to fully flexed and 0 corresponds to fully extended. The corresponding dashed lines show the average position of all the other fingers during the cued finger’s trials.

normalized MPL finger positions during their respectively cued trials are shown in Figure 4.6. The dashed lines represent the average position of the other four fingers during that trial (e.g. the solid blue line depicts the thumb, while the dashed blue line depicts the mean position of the other fingers during thumb trials).

A cross-validated, regularized parameter search limited to the electrodes selected for the online BMI revealed that adding five electrodes (shown in Figure 4.3) led to the greatest reductions in error on the training data, and the addition of other electrodes
typically increased error. In spite of excluding electrodes identified by preliminary electrode localization as being over the postcentral gyrus, the four electrodes which contributed the most to decoding accuracy all showed significant activation during the vibration task. We therefore performed offline analyses to investigate the motor decoding results when controlling for the contributions of somatosensory cortex activity.

4.3.3 Offline Analysis

To investigate the relative contributions of somatosensory and motor cortex to the BMI, we compared the spatiotemporal patterns of activation during offline passive vibrotactile stimulation vs. finger movement. The somatotopic organization of activation is shown in Figure 4.7. Figure 4.8 shows spectral activation during passive vibration and offline finger movement in the two electrodes that contributed most to online decoding performance. Intensities are saturated to better see the earliest activation times. A total of 33 electrodes displayed significant task-related spectral modulation during finger movement in the 512 ms centered on movement onset. These electrodes crossed significance thresholds a median of 16 ms before onset of movement. The 25 electrodes with significant modulation in the 512 ms centered on vibrotactile stimulus onset reached significance a median of 80 ms after vibration onset. This suggests that a good deal of the activation during finger movements was not solely due to somatosensory feedback.
Figure 4.7: Somatotopic organization of the high gamma activation during the passive vibration task. Electrode color corresponds to the finger with the strongest high gamma activation. Intensity corresponds to the strength of the activation normalized relative to the average activation across the four other fingers. This provides an indication of how discriminable the respective finger is relative to all other fingers. Electrodes A and B were the two electrodes that contributed most to the decoding performance of the finger BMI based on training data.

Spectral modulations during the motor task significantly exceeded baseline levels in four of the five electrodes that contributed most to the classifier performance. In these electrodes, activation first became significant during feature extraction windows (256 ms long) centered 16, 48, 128, and 160 ms before the onset of movement (one did not display significant activation). In the sensory task, activation first reached significance 48 ms after the onset of vibration in three of the five electrodes, 64 ms in one electrode, and never in one electrode.
Mean classification accuracy is shown for motor task and vibrotactile stimulation in Figure 4.9. Time points were referenced to the onset of any detected movement (via CyberGlove) or the onset of vibration. At each time point, the mean of the previous 15 feature extraction windows was used for cross-validated classification. This causal decoding scheme made use of the 352 ms of data preceding each time point. Classification accuracy during the finger tapping task was significantly higher than the corresponding accuracy during vibrotactile stimulation between -32 ms to
CHAPTER 4. INDIVIDUAL FINGER CONTROL

224 ms ($p < 0.05$, Wilcoxon rank sum, Bonferroni corrected for 30 comparisons). The classification accuracy for finger vibration exceeded chance 192 ms after stimulation onset. The mean classification accuracy during the motor task reached 70.7% at the corresponding time of 192 ms after movement onset. Decoding based on the ECoG signals showed similar accuracy over time relative to using the CyberGlove recordings for finger decoding. This was in spite of the fact that the ECoG features were associated with a group delay of 184 ms, whereas CyberGlove relied only on the instantaneous time value.

While peak performance without somatosensory feedback cannot be reliably ascertained from this data, the time course of classification accuracies indicate that accuracy using signals only related to motor function (without accuracy inflated by sensory feedback not present in paralyzed subjects) would likely achieve at least 70.7% accuracy in discriminating which finger was being moved. When data was aligned to the peak flexion time and the smoothing window was increased to 480 ms, we obtained a maximum accuracy of 96.5% and a mean accuracy of 90.2% between 190 ms before to 1250 ms after the peak flexion (after correcting for group delays). Based on these findings, we estimate that a high-density ECoG array over sensorimotor areas would yield individual finger classification accuracies in the range of 70.7-96.5% using motor signals in the deafferented patient population that would benefit most from invasive BMI systems.
**Figure 4.9:** Finger classification accuracy over time using CyberGlove and ECoG recordings during the motor task, and ECoG during vibrotactile stimulation. Motor task recordings are aligned to the onset of any movement by the subject’s hand (solid line), and vibrotactile stimulation is aligned to onset of the vibratory motor (solid line). The red, green, and blue dashed lines denote when decoding accuracy exceeded chance levels using motor task high gamma, vibrotactile stimulation high gamma, and motor task CyberGlove data. All decoding is causal, i.e. classification accuracy at time t is dependent only on samples collected at time t and earlier. Shading represents the 95% confidence interval of the mean.
CHAPTER 4. INDIVIDUAL FINGER CONTROL

4.4 Discussion

We have demonstrated, for the first time in humans, online neural decoding of individual finger movements to control a dexterous modular prosthetic limb. We utilized high gamma power extracted from a high-density ECoG grid situated over a subject’s sensorimotor cortex to produce intuitive BMI finger control. This was accomplished without arbitrary mappings of user inputs or extensive training by using the native cortical representations of finger movements. During online control, we utilized all available neural information in a subset of electrodes to produce highly accurate control of fingers. For further offline analysis, we limited our analysis to time periods preceding discriminable sensory afferent information to more closely mimic SCI patients’ neural activity. We found this still yielded discriminable information useful for finger control.

We were able to accurately decode finger flexions online using the native sensorimotor representations of fingers in the brain. The classification accuracy for the online BMI at the onset of the most reliable time period for detecting movements (1.6 seconds post-cue) was 76% (88% with ring and pinky combined). The balanced accuracy of correctly flexing the cued MPL finger was 64% (77% with ring/pinky combined). Our asynchronous binary prediction of whether the subject was moving any finger yielded a balanced accuracy of 92%. While previous groups have decoded finger movements offline [86], to our knowledge, this is the first online demonstration by a human subject of neural control of individual finger flexion with a robotic...
prosthetic limb. We further believe that the control of five individual fingers in this study marks the greatest number of distinct degrees of freedom controlled online with ECoG signals. By accomplishing this without the need for operant conditioning, we have shown that high density ECoG can leverage the native somatotopy of individual fingers to provide patients with immediate, more intuitive finger control.

Our online results contribute to the growing literature of online ECoG control over robotic prosthetics. Subjects have used ECoG to asynchronously modulate the aperture of two grasp types (77% of trials correct to completion) while simultaneously controlling elbow movements (34% of trials correct to completion) with a prosthetic limb [50]. The ability to simultaneously control reach and grasp independently with a prosthetic limb was demonstrated with mean accuracies of 84% and 81% for reach and grasp respectively [4]. ECoG control signals have been further integrated with intelligent robotics to pick up and move objects with a mean success rate of 70% [5]. Yanagisawa et al. was able to asynchronously detected hand movements (61% detected within one second) then decode which of three different hand movements was performed (69.2% accuracy) [51]. Chestek et al. showed 97% accuracy in classifying between resting and grasping at a fixed time relative to cue in two subjects [49]. Their accuracy was 57% in one subject when expanded to two different grasp types vs rest, and 55% accuracy in another subject when classifying between rest and four grasp types. ECoG signals from hand somatosensory areas were also used in [45], where attempted thumb, wrist, and elbow movements were used as control signals for 3D
CHAPTER 4. INDIVIDUAL FINGER CONTROL

arm position. Their subject achieved 80% accuracy in a 3D cursor task leading up to their MPL endpoint control, demonstrating the ability of postcentral hand area ECoG signals recorded from a paralyzed subject to control a BMI. Our online asynchronous classification accuracies of 64% and 77% for rest vs five and four fingers respectively build on the previous literature to show, for the first time, that neural recordings can be used to provide immediate online control of individual fingers without the need for operant conditioning.

In order to more closely mimic the neural activity of an SCI patient, we performed further offline analyses where we limited neural activity to time periods preceding sensory feedback. Slight pre-trial movements of our subject’s hands led to loss of data from conservative estimates of movement onset determination. While cortical reorganization occurs after SCI/amputation, multiple groups have demonstrated continued sensorimotor activity during imagined or triggered phantom hand movements [175, 176, 177, 178]. We believe SCI patients could harness this continued neural activity to control a robotic limb BMI. Indeed, groups working with a SCI patient implanted with ECoG have demonstrated SCI patients’ ability to modulate this rhythms to control a BMI [45]. In spite of the very conservative estimate of movement onset, spectrograms of high gamma activation revealed significant pre-movement activity (Figure 4.8). We estimate a lower bound of 70.7% decoding accuracy without sensory feedback. Higher single finger decoding accuracies at movement onset was previously found [85], but this was achieved with a non-causal methodology that did
CHAPTER 4. INDIVIDUAL FINGER CONTROL

not account for sensory feedback. Our peak accuracy during offline analysis of the motor task reached 96.5%, with a mean accuracy of over 90% during a time period lasting over a second. This establishes an upper bound of over 90% for expected online single finger decoding accuracy using high-definition ECoG over sensorimotor areas. A comparable drop in performance was seen previously [49] when predictions were only made using data from before movement onset. This is consistent with our estimate of the expected lower bound accuracy.

Our analysis revealed widespread pre-movement activity across both precentral and, interestingly, postcentral gyri during the finger tapping task. Much of our decoding accuracy was derived from electrodes that also showed activation during vibrotactile stimulation. This may have been due in part to the high-density grid having better coverage of the finger representation on the postcentral gyrus than the precentral gyrus. Another possible reason the somatosensory-modulated electrodes were essential to decoding can be traced to more distinct somatotopy of individual fingers in the postcentral gyrus than in the precentral gyrus [160, 162, 161, 179, 180]. A large portion of the post-movement activity detected in postcentral gyrus was no doubt the result of afferent cutaneous and proprioceptive information. However, it is possible that a subset of the pre-movement activity in this region was the result of an efference copy sent from premotor and associated motor planning cortices [167, 181, 182]. These results are similar to a human ECoG study that showed pre-movement signals in sensory areas that were led by signals in task-related areas in premotor cortices
CHAPTER 4. INDIVIDUAL FINGER CONTROL

This movement-related activation, observed in both pre- and postcentral cortex, has been noted in paralyzed subjects [171, 172, 173]. Furthermore, postcentral activity was recently used to provide control over a three dimensional brain-machine interface to a paralyzed subject [45].

Further advancements in ours and other systems are necessary before BMIs will be able to meet the clinical needs of SCI patients. Although we were able to demonstrate immediate robust online decoding and actuation of individual fingers, future users will need more accurate systems before they can perform the ADLs necessary for full autonomy. However, while the raw decoding accuracy found in this study would not meet the performance requirements of a functionally useful prosthetic device, cortical adaptation has been shown to increase BMI performance over time [102, 184, 131]. This would likely improve results for patients that can work with the BMI system longer than the brief testing session performed in this study. Additionally, decoding from more comprehensive coverage of precentral areas would likely further improve system robustness [85, 49]. This could potentially be accomplished with the use of large arrays of flexible high density micro ECoG arrays, which could provide higher spatial resolution without sacrificing wide area coverage [185, 186]. Finally, other avenues of research may also create more functionally useful BMIs. Incorporating intelligent robotics into neural-only BMIs may allow users to perform a wide variety of tasks [5, 70]. By using environmental sensors to inform the system of which fingers the user likely wants to flex, the system could incorporate prior knowledge of grasp...
trajectories to accomplish advanced tasks in the presence of noisy cortical control signals.

4.5 Conclusion

We demonstrated the ability of a patient implanted with high density ECoG over sensorimotor areas to asynchronously control individual finger movements intuitively in an online BMI, without the need for online training sessions. Offline analysis showed a large amount of the decoding accuracy in our coverage area was derived from postcentral gyrus. While our online control was likely highly influenced by sensory feedback, we found that the signals in somatosensory areas showed significant activation before movement onset, which might still occur in patients lacking afferent somatosensory inputs. Furthermore, decoding accuracy over the course of the movement mirrored the separability of hand positions—the more the CyberGlove recordings became identifiable as a specific finger movement, the more the finger movement could be decoded from the preceding window of neural data. Our findings support previous studies suggesting an efference copy activating somatosensory cortex [167, 181, 182]. Here we demonstrated that this information can be used to control individual fingers in an able-bodied subject. Future studies will need to determine whether the same or similar signals can be used in paralyzed patient populations, but recent work has demonstrated promising results showing sensorimotor activation in deafferented
CHAPTER 4. INDIVIDUAL FINGER CONTROL

patients [171, 172, 173, 45].

We have demonstrated, for the first time, online control of all five individual fingers of a robotic prosthetic hand independently and hope that insights from this study can one day be applied to paralyzed patient populations to enable them to perform dexterous ADLs.
Chapter 5

Demonstration of a Semi-Autonomous Hybrid Brain-Machine Interface using Human Intracranial EEG, Eye Tracking, and Computer Vision

5.1 Overview

Current brain-machine interfaces (BMIs) lack widespread clinical use due to their inability to provide paralyzed patients with reliable control of prosthetic devices to
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

perform everyday tasks. The more than 100,000 Americans with complete quadriplegia [7] suffer a loss of autonomy that affects their general quality of life and the general health care with lifetime costs near five million dollars [187]. Surveys of patients with motor disabilities reinforce the limited state of functional rehabilitation that BMIs and other assistive technologies currently provide [188]. These studies show that patients who use assistive technology are most dissatisfied with their manipulation ability and they rank restoration of their ability to perform activities of daily living as highly important for future devices.

The shortcomings of current BMIs are no longer due to a lack of robotic ability. The modular prosthetic limb (MPL) developed by the Johns Hopkins Applied Physics Laboratory (JHU/APL) is an upper limb neuroprosthetic with 17 controllable degree of freedom (DOF) capable of performing a full range of activities of daily living [133]. While these capabilities are far greater than those offered by traditional prosthetics, they have not yet been fully utilized in a BMI. This is due, in part, to the current training paradigm of mapping independent DOFs to neural activity generated during volitional or imagined movements. Groups using this training paradigm have demonstrated multi-dimensional control in humans implanted with invasive multi-electrode arrays, electrocorticography (ECoG), and depth electrodes [4, 36, 45, 33, 189, 41, 50, 44]. While recent work demonstrating 7-DOF neural control with microelectrode arrays (MEAs) [36] is impressive, this still represents a fraction of potential robotic capabilities and requires sustained selective attention and
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

mental effort by the patient to exert the low-level control necessary for performing basic tasks.

Limitations of current BMI technology have led researchers to develop novel methods of controlling robotic actuators. Three strategies have been explored to fully utilize robotic ability: hybrid BMIs, supervisory control, and intelligent robotics [190]. Hybrid BMIs combine a traditional input signal from the brain with an additional signal, which can also be from the brain (e.g., combining electroencephalography (EEG)-based SSVEP and P300 potentials for BMI) or from physiological signals such as electromyography (EMG) or electrooculography (EOG) [190, 53, 191, 192, 193]. Input from other assistive devices, such as eye tracking, can also be incorporated into a hybrid BMI design to make use of intuitive input controls [56, 194, 195]. One of the major benefits of hybrid BMIs is their ability to limit the number of BMI false positives since errors from two distinct sources would be needed to lead to a misclassification [53]. In the context of assistive motor BMIs, limiting the number of false positives is essential to real world applications where inadvertent task initiation could lead to a time intensive robotic motor action.

Supervisory control allows sharing of control between a human patient and a robotic actuator, with varying degrees of robot autonomy traded with the patient [196]-[33]. Current BMI research has largely focused on low-level (i.e., direct manipulation or process oriented) strategies where each DOF is continuously controlled, requiring sustained attention from the patient [36, 45, 33, 189]. On the other
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

hand, semi-autonomous strategies allow patients to concentrate on high-level (i.e.,
goal-oriented) control while allowing an intelligent robotic controller to calculate and
execute the prosthetic limb’s low-level kinematics [195, 197, 198, 199]. Some studies
have shown that high-level control schemes can be used to make BMI tasks quicker
and more accurate [200, 201]. Intelligent robotics technology incorporates sensors to
provide feedback from the patient’s environment to the robot, for example, recognizing
objects in a workspace, planning obstacle avoidance, and optimizing pathways
[56, 202, 203].

The Hybrid Augmented Reality Multimodal Operation Neural Integration Envi-
ronment (HARMONIE) system reported herein utilizes an intelligent semi-autonomous
hybrid control strategy to allow patients to perform functional tasks via brain control
(Figure 5.1). Computer vision is used to detect objects in 3-dimensional space, while
eye tracking allows the patient to select one of these objects to be acted upon by
the MPL (Figure 5.2). Intracranial EEG (iEEG) signal changes correlated with the
patient’s natural reaching movements are used to initiate a sequential reach-grasp-
and-drop of an object by the MPL. The HARMONIE assistive rehabilitation robotic
system is being developed to allow patients to use natural eye movements and motor
intention in a hybrid BMI to control a semi-autonomous robotic limb. In this paper,
we report results from a pilot study of the HARMONIE system in two human sub-
jects undergoing intracranial recordings for epilepsy surgery. Results obtained during
testing of the system in the first subject led to improvements of the system prior to
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

Figure 5.1: HARMONIE combined BMI system block diagram. Components of the HARMONIE BMI system work together to allow the patient to perform a complex automated motor task. A patient implanted with iEEG electrodes is able to use eye tracking to visually select an object that has been segmented by computer vision. Patient initiates the task via their iEEG-BMI input. Task initiation and object information are used to control the MPL to perform the task.

subsequent testing in the second subject. Subject 1 used the system to perform the reach-grasp-and-drop task with just one object, while Subject 2 used it to perform the task with three objects.

5.2 Materials and Methods

5.2.1 Subject Info

Two subjects were implanted with subdural ECoG arrays and/or penetrating depth electrodes (Adtech; Racine, WI) for recording seizures prior to surgical resection
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

Figure 5.2: Schematic of the HARMONIE experimental setup. Patient views a segmented scene sent from the computer vision device. Kinect sensor records both the depth information and RGB of the experimental workspace. Patient’s gaze is recorded by monitor-mounted eye tracking and movement intention is detected by iEEG-based BMI. MPL carries out the reach, grasp, and place motor task.

to treat medically resistant epilepsy with placement based on clinical needs. The ECoG electrodes had a 2.3-millimeter (mm) diameter exposed recording surface with 10-mm inter-electrode (center-to-center) spacing. The depth electrodes had eight 2.41-mm length (6.5-mm center-to-center spacing) platinum contacts. A macro-micro depth electrode had eight macroelectrodes contacts with 1.57-mm length platinum contacts (5-mm center-to-center spacing), as well as four microelectrodes interposed between each of the five most distal contacts.

Our first subject, Subject 1, was a 42 year old right-handed female with a lesion in
her left anterior cingulate gyrus. She was implanted with ECoG electrodes including a double-sided 2 x 8 interhemispheric grid over her left mesial frontal cortex in the interhemispheric fissure (Figure 5.3B), as well as three 1 x 8 electrode strips over her left dorsolateral prefrontal cortex, extending to the Sylvain fissure. Our second subject, Subject 2, was a 30 year old right-handed male who had previously undergone partial resection of his right post-central gyrus and superior parietal lobule. He was implanted with seven depth electrodes (one being a macro-micro depth electrode) placed medially in his right hemisphere from anterior to posterior from the pre-motor area back into the motor cortex in pre-central gyrus, as well as sensory areas in post-central gyrus and parietal lobe (Figure 5.3D). Neuronavigation, via the Cranial Navigation Application (BrainLab; Westchester, IL), was used during placement of the depth electrodes. The patient also had a 1 x 8 ECoG strip placed over motor and sensory cortices laterally. Both patients gave informed consent for testing to be done according to a protocol approved by the Institutional Review Board of the Johns Hopkins Medical Institutions. Electrode locations were confirmed by volumetric co-registration of the subject’s pre-implantation MRI with their post-surgical CT using the BioImage Suite [114] (Figure 5.3B and 5.3D).

5.2.2 Neural Signal Acquisition

A 128-channel NeuroPort system (BlackRock Microsystems; Salt Lake City, UT) was used for iEEG signal acquisition in parallel with clinical long-term EEG moni-
Figure 5.3: Neural activation during reaches. Cue-averaged statistically significant high gamma modulations relative to pre-cue periods are plotted in a channel x time raster for Subject 1 with the median behavioral times marked for movement onset and target button press. Results are aligned relative to the audio cue, and show 1024 ms of the pre-cue period. The mean modulation across nine 16 ms time bins spanning 112 ms (centered on median movement onset time) is plotted on the patient’s brain reconstruction (B) for Subject 1, left hemisphere viewed from medial surface. Inter-hemispheric grid (IHG) ECoG electrodes 18, 19, and 26 (outlined in black boxes) displayed task-related modulations during this time period (B). Routine electrocortical stimulation caused movements in the right arm at IHG 18 and 26 and stiffening on the right side at IHG 19.

During training, the NeuroPort system used a sampling rate of 30 KHz with an analog third-order Butterworth anti-aliasing filter with cutoffs of 0.3 Hz and 7,500 Hz. The NeuroPort system applied a digital low-pass filter with a 250 Hz cutoff (4th order Butterworth) and downsampled data to 1 kHz for streaming over UDP to an experimental computer. Any channels with substantial noise or artifacts, as identified by a clinical neurophysiologist by visual inspection, were excluded from subsequent analyses.
5.2.3 Training

The subjects performed reaching and grasping movements (30 each for Subject 1, 50 each for Subject 2) with their right (Subject 1) or left (Subject 2) arms, contralateral to their implants. The subjects were instructed to loosely hold a pneumatic squeeze bulb while resting their hand on a home plate sensor in their lap. When a sound clip instructing Reach was played, the subjects reached forward and pushed a target button before returning to the home plate. When the Grasp auditory cue was played, the subjects tightly grasped the squeeze bulb. Cues were presented pseudorandomly. Signals from the audio cue, home plate, target button, and pneumatic air pressure sensor were fed into the analog ports of the NeuroPort system for synchronized recording with the iEEG signals. Subject 2 went through training twice, once for functional localization and a second session after which the BMI model was trained.

5.2.4 Functional Localization and Neural Signal Analysis

Electrodes that displayed significant task-related high gamma modulation for reach-related activity were identified using a custom analysis script in MATLAB (Mathworks; Natick, MA, USA). The Hilbert transform was calculated on common average referenced (CAR) iEEG data in 128 ms windows (112 ms overlap); this sig-
nal was binned in the time domain by averaging 16 adjacent samples at 1000 Hz for an effective time resolution of 16 ms. The Hilbert transform was augmented with a multiplication of the frequency spectrum by a flat-top Gaussian spanning 70-110 Hz to yield an estimate of the high gamma analytic amplitude.

A baseline distribution of high gamma amplitudes was created for each channel by pooling amplitude measurements from the 1024 ms prior to the audio cue. For each channel, separate distributions were created for each time bin after the audio cue. Significant modulations of high gamma amplitude at each channel were found by applying a two-sample t test between each post-stimulus distribution (i.e., one for each time and channel pair) and the baseline distribution. A significance threshold of alpha formula was used, with corrections for multiple comparisons carried out using the false discovery rate (FDR) correction within each channel. Significant modulations were plotted on a channel versus time raster and also on the patient’s brain reconstruction to visually determine which electrode locations displayed the greatest task-related activity [Figure 5.3(a) and (c)].

5.2.5 BMI Model Training

The results from the aforementioned reaching trial analyses were used to select the electrodes and temporal windows to be used for training the BMI model. The electrodes were selected using a system optimized for rapid event-related feature extraction and statistical testing in order to capture complete spatial-temporal maps of
cortical activation. These methods were not optimal for the real-time control of the HARMONIE system. Instead, the Burg algorithm was used to estimate the spectral power of the recorded iEEG signal with a sixteenth-order autoregressive model \[204\]; an estimate of the high gamma log-power was obtained for each window by computing the mean of the log-transformed spectral power in bins between 72.5 and 110 Hz for each electrode. This frequency band was chosen to capture high gamma modulation while avoiding 60 and 120 Hz noise.

Distributions of high gamma log-power during the active (i.e., related to moving) and baseline periods were built using 400 ms windows spaced every 100 ms. For Subject 1, windows centered at times between 1800 ms and 1200 ms before onset of the audio cue were used as model inputs representing baseline. The active distribution was built from the 600 ms surrounding (i.e., 300 ms before to 300 ms after) the onset of movement, as detected by the release of the home-plate. This time period was chosen as it displayed robust high-gamma activation during preliminary analysis of the training data used to train the decoder. For Subject 2, the baseline distribution was built from windows whose centers fell in the 1024 ms prior to the audio cue of each trial. The active distribution was built from windows whose centers were between movement onset and movement offset, as detected by the home plate. This change in the baseline and active distributions used was implemented to increase consistency with our functional localization methods (see Section II-D). For Subject 2 only, the trained reach classifier was also augmented with additional baseline points
from periods where only grasping movements were being made, in an attempt to
isolate features responsible for reaching only.

The high gamma log-power distributions during baseline and active periods were
then used as signal features to train a binary linear discriminant analysis (LDA)
classifier, which was employed online to detect reaching movements. In Subject 2,
transition probabilities were modulated heuristically to help the time course of clas-
sifications approach that of the true movements. This was done a priori (before the
testing sessions), with values of 0.95 for the probability of staying at rest if currently
at rest (0.05 for entering an active period), and 0.8 for the probability of remaining
active if currently active (0.2 for returning to rest).

5.2.6 Testing Paradigm

A video monitor mounted with a portable Tobii PCEye (Tobii Technology; Stock-
holm, Sweden) eye tracking sensor was placed in front of each subject (Figure 5.2).
This monitor displayed video from a workspace adjacent to the subject that consisted
of objects (one ball for Subject 1, three for Subject 2) placed on a table positioned
in front of the MPL and computer vision system. Subjects were able to control, via
eye tracking, a cursor displayed on the monitor in order to select objects in the video.
For Subject 1, the reach trials were not cued and the subject decided when to initiate
reach movements after the ball was placed in the workspace. For Subject 2, trials
were cued by the experimenter. The experimenter would wait for Subject 2 to focus
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

on a home area of the screen before indicating which of the three colored balls he should select. At the onset of each trial, the subject was instructed to visually fixate on the ball displayed on the monitor and then perform a reach toward the ball on the monitor (though it was possible to reach without proper fixation). When neural signals of reaching, concurrent to visual fixation, were detected, the MPL initiated a complex sequence of actions consisting of reaching, grasping, picking up, moving, and dropping the ball under computer vision guidance. Subject 1 had two blocks of online testing while Subject 2 had one block.

Performance of the HARMONIE system was determined, in part, by video analysis of the testing session. Trials were deemed successful if the patient performed a reaching movement that subsequently initiated the MPL’s motor task (see Online BMI Testing for timing). Additionally, Subject 2 had to select the correct ball from three possibilities for the trial to be considered a success. For Subject 1, each distinct reaching movement toward the monitor by the patient was counted as a separate trial since she was not cued.

5.2.7 Computer Vision and Eye Tracking

Computer vision and eye tracking were used for identification and selection of targets in the workspace for MPL action. For computer vision, a Kinect (Microsoft; Redmond, WA, USA) infrared range sensor system array was used to capture 3-D point cloud depth data and RGB video from the subject’s workspace (Figure 5.2).
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

The Kinect was placed on a base stand above the MPL, approximating the location of the human head relative to the right upper limb. The Kinect also captured RGB video of the workspace to be viewed by the subject on the computer monitor. The Kinect streamed, via USB 2.0, to a Linux computer running the Robot Operating System (ROS) Fuerte platform (Willow Garage; Menlo Park, CA, USA) (Figure 5.1).

Computer vision was used to segment spherical objects. The Point Cloud Library (PCL) was primarily used for 3-D processing. A pass-through filter was first applied to remove 3-D points outside of the workspace of the MPL. Secondly, a sample consensus algorithm was used to find and segment the largest planar surface in the scene corresponding to the tabletop. A Euclidean clustering extraction algorithm was applied to the 3-D points on the tabletop to assign groups of points to an object in an unsupervised fashion \[205\]. A sample consensus algorithm was then applied to each of the objects to determine if the object corresponded to a sphere or cylinder. The output from this classification also included geometric properties of the object, including the radius and orientation. These object properties assisted in selecting optimal grasp types for interacting with the object. The RGB video and object parameters (shape and position relative to the MPL) identified by the ROS system were streamed over UDP from the Linux machine to a computer controlling the monitor in front of the subject.

The RGB video from the Kinect was displayed on the monitor in front of the subject utilizing a MATLAB graphical user interface (GUI). Spherical objects successfully
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

identified through segmentation of the point cloud were outlined in the RGB visualization on the monitor by a light grey circle. The Tobii PCEye eye tracker mounted at the base of the monitor allowed the user to control the mouse cursor on the monitor with their gaze position alone.

Visual feedback was provided to the patient through the GUI to indicate successful eye tracking and iEEG-based initiation by changing the color of the highlighting circle surrounding the segmented object. When the eye-tracker-controlled cursor was not over the object, the circle outlining the object was grey. For Subject 1, when the cursor was within a distance of twice the object’s radius from its center, the circle changed to red to indicate successful fixation. This feedback was modified for Subject 2 due to possible overlap of extending the radius with multiple objects in the workspace. For Subject 2, when the cursor was within the actual radius of the object, the circle changed from grey to blue to indicate successful fixation.

If an intended reach signal was detected by iEEG while the circle was fixated upon, the system entered the MPL actuation state and the circle turned green (for both subjects). Upon entering the actuation state the MPL began its action sequence to grasp, pick up, move the selected object, and drop it into a container. Upon completion of the action sequence, the MPL returned to its home position. For Subject 1, the monitor froze on the display frame during the MPL action while for Subject 2 it constantly streamed video of the MPL action. If an intended reach was detected from the iEEG and the subject’s gaze was not yet over an object, the subject
had 2 s to look at an object before the command was ignored.

We performed offline testing of the computer vision module after system improvements were made for Subject 2. We recorded 200 video frames with different types of objects while varying the number and placement of the objects in the workspace. We investigated the number of frames each object was identified in, as well as the variability in their measured sizes and positions. This was done in addition to tracking computer vision errors during online testing.

5.2.8 Modular Prosthetic Limb

The MPL is an anthropomorphic robotic limb with 17 controllable DOF and 26 articulating DOF \[133\]. The MPL software architecture and control system allows for high-level commands, e.g., those translated from neural signals \[4\], to actuate a coordinated sequence of movements \[135\]. Endpoint control (EP) and reduced order control (ROC) commands allow developers to specify 3-D hand end-point positions, velocities, and grasp patterns. The sequence of EP and ROC commands given can be determined by the shape and position of the objects in the workspace. A Jacobian-based algorithm is then employed to find the inverse-kinematics needed to execute the specified motor task.

A right-handed MPL was attached on the right side of the same stand and base used to mount the Kinect (Figure 5.2). This corresponded to the limb controlled by the contralateral (left-sided) supplementary motor area recorded with ECoG for
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

Subject 1. The right-handed MPL was ipsilateral to the patient’s right-sided implant in Subject 2; therefore the arm was positioned facing the subject in such a way to mirror his left hand. The same object was used repeatedly for Subject 1 while several objects of varying consistency and texture were used with Subject 2.

The object position and shape derived from computer vision and selected with eye tracking were translated into MPL EP and ROC commands (Figure 5.1). These commands were sent to the MPL controller and inverse kinematics were used to determine MPL motion commands. The object used with Subject 1 was placed on a pedestal (small roll of tape) to stabilize and elevate the ball to help the MPL pick it up (similar to [36]) while no pedestals were used with the multiple objects for Subject 2.

The path-planning module calculated the final hand position, orientation, and grasp type based on the object type and orientation in the workspace. Standard inverse kinematic techniques were then used to calculate target joint angles for the arm and wrist to achieve the target hand position and orientation, with the additional degree of freedom in the arm used to minimize the elbow position. Additional waypoints were added to the target trajectory to control the direction of approach, and their joint angle representations were also estimated using inverse kinematics. The target waypoints, converted to joint space, were used to generate and stream more continuous intermediate joint commands to the MPL via linear interpolation.
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

5.2.9 Online BMI Testing

The HARMONIE system was tested online with the patient two days (Subject 1) and one day (Subject 2) after the training session with the patient. For online BMI control, signals were streamed and processed and the high gamma activation in the selected electrodes was calculated with the same parameters as the training data. High gamma log-power calculations were done as in the training set on 400 ms windows in real time with a median time of 12 ms between iterations for Subject 1 and 31 ms for Subject 2. Each power calculation was used as an input to the LDA classifier constructed from the training data. To reduce spurious classifications, the LDA model had to return a positive classification for each time point within a 500 ms window. Once a movement was detected, the system was flagged for 2 s to initiate a reach if an object was visually selected in this timeframe and the MPL was not already moving. Classifications by the LDA then resumed after these 2 s, and the arm could be flagged to move again once another 500 ms of data was collected and processed.

Offline analysis of the HARMONIE system was performed by synchronizing the recorded NeuroPort data and clinical monitoring video. To assess the sensitivity and specificity of the BMI movement detection, active and baseline periods based on movement off of the home plate (i.e., reach onset) were analyzed. Trials in which the system predicted a movement between 500 ms before and 3000 ms after the subject’s hand left the home plate were denoted as a true positive [Figure 5.4(a)]. This
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

Figure 5.4: Four trial types of the HARMONIE system are shown. Blue lines represent the actual movement of Subject 1’s right hand, with an upward deflection marking the trial start (i.e., reach onset) when the subject lifted their hand off the home plate (zeroed when resting). Red stars demonstrate actual BMI prediction of a movement for Subject 1. The red filled areas represent the baseline window where no movement detections were expected, and the green areas represent time windows where we expect movement detections. (A) and (B) are representative of the successful trials. (C) shows an example of a trial where the baseline was discarded due to contamination from a previous movement. This baseline was not included in our accuracy analysis, but the uncontaminated active period surrounding movement onset was. (D) depicts a trial where both the baseline and the active period were classified incorrectly, resulting in a false positive and false negative.

A time window was selected to reflect pre-cue activity and the median reach duration observed in the training set (Figure 5.4). The movement prediction sensitivity was defined as the percent of trials detected as true positives. Movement of the MPL was initiated immediately upon an iEEG movement detection; the true positive window was used only for offline accuracy analysis.

The true negative rate was found by looking at a baseline period of 3500 ms (i.e., equal in duration to the active period) ending 500 ms before movement onset. If
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

no classifications were made during the baseline period, it was counted as a true negative. The percent of baseline periods which were counted as true negatives was then defined as the movement prediction specificity. Baseline periods were excluded from consideration if the subject moved within 1000 ms before the baseline began. We then verified that none of the remaining baselines had a movement classification within 2500 ms before its start. This was done since a classification event would cause the system to pause detection for 2 s, in addition to the 500 ms of positive classifications required for a detection.

The false positive rate was determined by analyzing all periods of time during testing where the subject’s arms were at rest for at least one minute. This included time during which the MPL was moving and the subject interacted with experimenters. The false positive rate was calculated as the total number of movement classifications during the rest blocks divided by the total duration of the rest blocks.

5.3 Results

The HARMONIE system allowed both subjects to perform complex object manipulations with the MPL using only eye gaze and iEEG signals corresponding to natural reaching movements (Supplemental Video 1). Subject 2 was able to choose among three balls to successfully power the HARMONIE with improvement in task completion time (Supplemental Video 2). Control of movement initiation was activated
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

by contralateral reaches but not by an ipsilateral reach when tested with Subject 1 (Supplemental Video 3).

5.3.1 Online Global Evaluation

The subjects attempted 28 self-paced trials (Subject 1) and 31 cued trials (Subject 2) over the course of an hour long testing session (for each subject). Individual trials were deemed global system successes if combined eye-tracking and reach initiated the MPL to complete the entire reach-grasp-and-drop task with one ball (Subject 1) or three balls (Subject 2).

For Subject 1, the testing session was divided into two blocks: a block of 18 reaches where the subject’s movements were recorded via the home plate, and a block of 10 reaches with no home plate recordings. Trials of the BMI system initiated by movement toward the monitor were successful during the first block in 14 of the 18 attempts (77.8% system success rate). The success rate of the system decreased to 20 of 28 (71.4%) attempts by the subject when trials from the second block were added. For the 28 trials, four of the eight system failures were traced to failure to classify the iEEG signal, and three of the errors were traced back to eye tracking. Identification of false positives in the second block was not possible due to a lack of movement timing information from the home plate (it was not in use). In one case, the patient attempted an identical reach with her ipsilateral arm (neurologically inappropriate for implant site) and failed to trigger the system. The patient subsequently reached
with her contralateral arm and successfully triggered the HARMONIE system.

For Subject 2, 21 of the 31 trials led to successful completion of the entire system (67.7% system success rate). Of the 10 global system failures, nine were traced back to failure of the MPL and computer vision system to pick up the ball. Failure of one trial was traced back to an inability classify movement from the iEEG signal. The patient had trouble remaining in the visual home area after being cued to ball type but before the Go cue was given. For the 11 trials in which he adequately waited for the cue and the data could be synced, there was a median 3.55 s delay between cue and system initiation.

5.3.2 iEEG-Based Decoding Evaluation

Initiation of movement was determined by the iEEG portion of the hybrid BMI using an LDA decoding model based on a training session two days prior to the testing session for Subject 1 and one day prior for Subject 2. During the training session, three electrodes located over the supplementary motor area (SMA) of Subject 1 [Figure 5.3(b)] showed robust high gamma activation for a substantial duration near the onset of reaches. Subject 2 also demonstrated similar patterns of activity in four contacts in three depth electrodes located within hand motor and sensory cortices, as well as the inferior parietal lobe [Figure 5.3(d)]. Only trials where the subject’s movement was adequately recorded via the home plate were used for iEEG analysis.

For Subject 1, of the 18 trials considered, 16 were detected by the system, resulting
in a movement prediction sensitivity of 88.9% [Figure 5.4(a)]. 15 of these trials had baselines uncontaminated by previous movements. A false positive classification was made during one of these baselines, yielding a specificity of 93.3%. Therefore, the balanced accuracy of the iEEG-based movement prediction module (i.e., the mean of the specificity and sensitivity, see Online BMI Testing) was 91.1%. We then calculated the chance prediction accuracy while accounting for the pacing of our movement predictions. We shuffled the intervals between movement predictions 10 000 times and computed the balanced accuracies on these shuffled predictions with trial timing held constant. The 95th percentile of these shuffled accuracies was chosen as the Formula significance threshold. The shuffled dataset had a median accuracy of 51.7% and a significance threshold of 61.1%. Therefore, the original balanced classification accuracy of 91.1% was significantly higher than chance Formula. In trials where the movement was accurately predicted, the median latency of the prediction relative to the detection of their movement onset was 200 ms. In order to quantify the rate of false positives made by the iEEG module, we identified nine contiguous blocks of time where the subject rested their arms for at least one minute. During the 13.2 min of rest time, the iEEG module made 14 false positive classifications, resulting in an average rate of 1.06 iEEG-only based false positives per minute.

For Subject 2, 28 of the 31 trials were considered, with three dropped due to technical failure of the home plate. Of the 28 trials considered, 24 were detected by the system, resulting in a movement prediction sensitivity of 85.7%. The baseline
intervals were uncontaminated by previous movements in 27 of the 28 trials. No false positive classifications were made during these baseline intervals, yielding a specificity of 100%. This yielded a balanced accuracy of 92.9%, significant with a threshold shuffled accuracy of 57.1% and a median shuffled accuracy of 51.4%. The median time between movements and their iEEG prediction was 273 ms. The false positive rates were calculated as above and demonstrated three false positives over 14.7 min of rest time, resulting in 0.20 false positives per minute for the iEEG module.

### 5.3.3 Computer Vision and Eye Tracking

During online testing, the computer vision reliably segmented spheres (one object for Subject 1, three for Subject 2) with radii of 4.7 centimeters (cm) throughout a workspace measuring 66 45 cm. When computer vision was tested offline with the ball used for Subject 1, the ball was correctly identified in all 200 recorded frames. The range of predictions of the ball’s radius ranged between 4.51 and 4.60 cm. The object’s 3-D position varied at most by 0.14 cm.

We also investigated how closely two objects could be positioned next to each other. We found that two tennis balls with radii of 3.35 cm could be reliably detected (both objects accurately detected in 196/200 frames) when separated by just 2.5 cm. Their detected radii varied between 3.52 and 3.99 cm and their positions varied by at most 0.17 cm.
CHAPTER 5.  A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

Figure 5.5: Offline demonstration of computer vision segmentation ability. Grey circles depict the system’s predicted size and location for each object. (A) shows that the object recognition software was able to simultaneously segment out eight spherical objects of varying sizes and materials. During online testing with the subject, only a single rubber basketball was used. (B) shows the potential for the system to segment out everyday lunch objects, such as an apple and a cup of yogurt.

The computer vision reliably detected a large number of objects simultaneously. When three tennis balls were well spaced in the workspace (21.59-24.28 cm between each object as measured by the system; similar to the set up for Subject 2), the system accurately detected all the objects in 196/200 frames. When eight objects of various size and material were placed in the workspace, the system simultaneously detected all eight of the objects in 192/200 frames [Figure 5.5(a)]. Of the objects tested, the system detected balls of radii ranging 3.35-12 cm.
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

Subject 1 used eye tracking to select the objects with a moderate amount of difficulty. To minimize subject eye strain, the minimal selectable area for hovering the cursor to perform a selection was doubled from the original radius size. For Subject 1, three of the seven total system errors were due to eye tracking difficulties causing the object to not be selected.

Subject 2 demonstrated robust control over the eye tracking system (without subsequent improvements). He was able to visually select the correct ball (indicated by color) out of three objects 100% of the time. No global system errors were attributed to the eye tracking in this subject. For the 13 trials in which the subject maintained fixation within the home area until the go cue was given, the subject was consistently able to select the correct object within approximately 1-2 s of the cue.

5.3.4 Modular Prosthetic Limb Control

Once initiated by the combined iEEG and eye control-based BMI, the MPL performed the complex reach-grasp-and-drop task 100% of the time for Subject 1 and 70% of the time for Subject 2. As mentioned above, Subject 2 used three balls with varying textures and, unlike Subject 1, the objects were not placed on pedestals. For Subject 1, the mean time from task initiation to ball drop was 22.3 s with a range not appreciable at our monitor’s frame rate (less than 0.1 s). Improvements to the MPL control system trajectory planning algorithms led to shorter completion times for Subject 2, with a mean completion time of approximately 12.2 s and a similar
 CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

range of less than 0.1 s (an improvement of 45.3%).

5.4 Discussion

The present study demonstrates that the final version of the HARMONIE system reliably and accurately combined eye tracking, computer vision, iEEG BMI, and advanced robotics cohesively to perform complex interactions between the MPL and a real object in cued and uncued trials. To our knowledge, this is the first demonstration of a semi-autonomous hybrid BMI system using intracranial signals from humans. The success of this demonstration is due to the combination of supervisory control, hybrid BMI, and intelligent robotics.

The HARMONIE system uses a supervisory control strategy to allow patients to perform a complex motor task with the MPL. Supervisory control strategies range from high-level (goal oriented) control to low-level (direct manipulation or process oriented) control. We focused on the patient’s intention to initiate the task (high-level) with a reach. The HARMONIE system performed well using supervisory control and successfully performed the complex motor task 71.4% and 67.7% of the time in two subjects. The median iEEG movement detection delays were only 200 ms and 273 ms. An important benefit of high-level control is that once initiated, the robotic actuator can proceed autonomously, thus freeing the patient from the sustained attention necessary to continuously control the many DOF of the MPL during a complex motor
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

task.

In the online demonstration, the HARMONIE system was able to integrate information from both the patient and environment to successfully determine the unique path the MPL took for each trial. The robotic limb performed the entire complex task 100% (Subject 1) and 70% (Subject 2) of the time after initiation with a mean completion time of 22.3 s and 12.2 s respectively, both with a range of less than 0.1 s. The decrease in MPL success for the second subject may be attributed to more challenging and realistic conditions, i.e., adding two additional objects and not using a pedestal to elevate and stabilize the objects. The semi-autonomous control framework enabled rapid implementation of modular improvements. For example, improvements in MPL movement transitions resulted in a reduction in completion time during subsequent testing with Subject 2 to 12.2 s, a 45.3% improvement (Supplemental Video 2). Our results demonstrate how high-level control enables the patient to complete a complex motor task with semi-autonomous components (the MPL) thereby freeing the patient to concentrate on other matters while the highly reliably prosthetic performs the action.

High-level control stands in contrast to low-level control, or direct manipulation. Work done in humans has demonstrated high-DOF control of robotic actuators with intracortical MEAs [36, 33]. While this strategy requires sustained attention by the patient, the degree of control is generalizable to multiple tasks. Recent work done with MEAs implanted in a paralyzed patient demonstrated control of the MPL to
complete a task similar to our experimental setup of reaching to a spherical object and placing it \[30\]. In the Collinger et al. study, the average time to task completion was 15.4 s (quicker than our Subject 1, slower than Subject 2). However, the range of completion times was 18.4 s for their direct control, while ours was less than 0.1 s (a benefit of the robotic control in semi-autonomous planning). The ability of the patient implanted with an MEA to use the same BMI approach for several other tasks demonstrates the generalizability of direct manipulation. However, one of the main benefits of mixed supervisory and shared control systems is the flexibility by which they can incorporate both high-level and low-level control. Future development of the HARMONIE system is expected to incorporate both levels of control into its system architecture. This will allow patients to choose when to directly manipulate an object versus choosing high-level control with less cognitive burden. High-level control can also be used in this scenario to immediately provide a patient with provisional BMI control of common, functionally useful tasks while they engage in training with the BMI to achieve and maintain direct manipulation.

In addition to supervisory control strategies, the use of a hybrid BMI is integral to the success of the HARMONIE system. Two goals of hybrid BMIs are to decrease the number of false positives and to utilize a second input signal for natural control that complements the primary input modality. Implementing assistive BMIs in the real world requires minimizing the number of false positives since each false positive can initiate complex and long tasks. The hybrid portion of the HARMONIE system
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

prevented the acceptable [53] (but still high) baseline iEEG-based false positives from Subjects 1 and 2 (1.06 and 0.2 false positives a minute, respectively) from initiating the system. Only one iEEG-BMI false positive led to a full system initiation once eye-tracking was added to the hybrid BMI. This reduction in false positives is an important benefit of the hybrid BMI approach and of importance when implementing BMIs in real world situations.

An important consideration when using hybrid BMIs is the choice of primary and secondary input modalities. The primary input modality should be a brain signal that can be reliably measured and classified. Our study illustrates the ease and stability of using an iEEG-based BMI since our subjects trained only once with a limited number of trials. Similar to other recent ECoG-based BMI work [45], we were able to use the same model reliably (91.1% and 92.9% balanced accuracy, Formula compared with chance) one and two days later, respectively, indicating stability of the ECoG and depth electrode signals over time. This is in contrast to MEA studies where retraining occurs daily to account for unit dropout over time [36, 98], therefore requiring subsequent recalibration of the BMI system [36]. A possible solution, mentioned above, is to combine MEA and iEEG recording modalities. In this setup, a stable iEEG signal could be decoded for high-level control whenever the MEA-based BMI requires retraining.

One possible drawback of the study is the use of actual, instead of imagined, reaching movements by the patient. Several EEG-based combined assistive systems
use combinations of SSVEP, P300 or repetitive, arbitrarily-mapped motor imagery to power their BMIs [191, 199, 202]. While these modalities have been demonstrated to work with paralyzed patients, they can be associated with cognitive fatigue and eye strain [200]. Also, relative to paralyzed patients, patients with intact motor function have decreased neural activity during motor imagery [207, 208], perhaps due to the need to inhibit movement of the native limb. We believe the actual reaches performed by our patients served as a reasonably good analogue for paralyzed patients attempting to move. We focused on movement-induced iEEG responses in the supplementary motor area (Subject 1) and motor and pre-motor cortices (Subject 2), making it highly unlikely that our BMI was reliant on sensory feedback that would not be present in the intended patient population. Future studies could focus on pre-movement activation as another means of controlling for sensory interference in active movements [209]. Further research must be done to quantify the accuracy of iEEG motor intention-based BMIs in appropriate patient populations to control the HARMONIE system.

The purpose of secondary inputs into a hybrid BMI is to complement the primary brain-derived signal. Our study corroborates recent reports showing that the combination of eye tracking with brain control in a hybrid BMI approach improves upon traditional eye tracking assistive devices [53, 191]. Using eye tracking as part of the physiological hybrid BMI leverages patients’ natural ability to use eye gaze to indicate the location of an object relative to the patient. However, eye tracking is also subject
to false positives that can interfere with everyday use. Indeed, this is an important 
motivation for using iEEG and not eye-tracking or P-300-based object selection alone. 
When visual fixation is used alone, there is a risk of the Midas Touch problem, in 
which fixation on an object (without intent to manipulate it) inadvertently triggers 
its selection and manipulation [53]. Traditional eye tracking systems have used dwell 
times for object selection, but users have indicated a preference for BMI object selec-
tion [56]. The eye tracking, object recognition, and motor BMI-LDA based selection 
solution that we used addresses the false positive Midas Touch problem. Our patients 
demonstrated adequate control of the eye tracking system with three and zero of the 
total system errors due to a lack of object fixation before reach initiation. Subject 2 
demonstrated robust control of the eye tracking, selecting from the correct of three 
objects in 100% of trials within approximately 1-2 s of the cue being given. Subject 
1 may have had more difficulty with the eye tracking due to her use of glasses and 
light eye color. Our 71.4% and 67.7% global system success rate is similar to other 
hybrid BMI and eye control studies that demonstrated 74.5% correct selection [194] 
and demonstrate the patients’ natural control of the eye tracking and BMI selection 
scheme.

The hybrid BMI was able to perform a complex motor task with the aid of intelli-
gent robotics. Objects were identified from a segmented point cloud using computer 
vision [195] and acted upon by the MPL. Each successful trial required the HARM-
ONIE system to integrate the object’s endpoint position and shape determined by
computer vision into a sequence of low-level command necessary to reach, grasp, and drop. Testing with Subject 1 was done with one object while testing with Subject 2 used three objects, with offline testing demonstrating capability of up to eight objects [Figure 5.5(a)]. One to three objects could be segmented in 196/200 frames that only decreased to 192/200 frames when eight objects of various sizes (range of 3.35-12 cm) were added. A reliable and robust computer vision system further reduces the possibility of false positives by limiting selectable objects in the workspace to only those that should or could be acted upon (as opposed to every object on a monitor). While these results demonstrate the capability of the HARMONIE system, more robust computer vision is needed for the HARMONIE system to perform in the real world, outside of carefully controlled laboratory settings [Figure 5.5(b)]. Future work incorporating a second Kinect or other IR depth sensors could allow for more robust identification of a wider variety of objects.

The present study offers a proof of principle of the HARMONIE system, a first time demonstration of a semi-autonomous, iEEG-based, hybrid BMI in human patients. The unique combination of these modalities allowed two patients to intuitively control a complex motor task consisting of reaching, grasping, and dropping an object. Future work will focus on testing the system in a larger pool of patients, both healthy and paralyzed. Modular improvements in the system are expected to achieve quicker and more reliable control.

A major limitation of our pilot system is that the current HARMONIE system
CHAPTER 5. A SEMI-AUTONOMOUS HYBRID BRAIN-MACHINE INTERFACE

shares only a limited amount of control with the user via eye tracking and iEEG. Currently, the system is semi-autonomous and once the iEEG initiation step is completed, the MPL proceeds autonomously with the task. Future work will aim to incorporate true shared control as demonstrated in wheelchair guidance [60] and in a continuous manner [61], building on previous efforts from our team demonstrating online control of continuous reach and grasp of the MPL [4].

In its current form other simple communication interfaces, such as eye tracking alone or P300 or SSVEP-based EEG systems, could initiate the system as well. However, the demonstration of a working system using iEEG is integral to future development. For example, it has been shown that three dimensions of continuous control can be decoded from iEEG [48]. Incorporating this kind of control into our intelligent robotic system will eventually allow for useful continuous shared control of the robotic tasks, allowing the user to control the progression of the arm forward and back through a trajectory constantly updated from information from the environment. Continuous monitoring of error-related neural responses found in iEEG signals could also be used to modify or reset arm positions as necessary [211].

Elucidating a larger repertoire of neural responses will allow for a wider variety of high-level commands to be implemented, leading to a greater number of functional tasks that can be performed with objects. The current system could also be modified to allow the user to select the desired location to place the object by using eye tracking to indicate location and iEEG initiation to begin the placement. As with reaching
and grasping in the current system, it is likely that a hybrid combination of iEEG
and eye-tracking will provide users with a high degree of flexible yet accurate control
over object manipulation and task execution.

The HARMONIE system, using the aforementioned and future improvements in
relevant technologies, will allow patients to achieve useful and reliable control of an
advanced multi-DOF powered robotic upper limb without extensive training and with
minimal cognitive load.

5.5 Footnotes

Data analysis and online control were supported in part by the National Institute
of Neurological Disorders and Stroke under Grant 3R01NS0405956-09S1, in part by
DARPA under contract 19GM-1088724, and in part by the National Institute of
Biomedical Imaging and Bioengineering under Grant 5T32EB003383-08. The MPL
was developed with funds from DARPA under Contract N66001-10-C-4056.

This chapter has supplementary downloadable material available at
Chapter 6

High Precision Neural Decoding of Complex Movement Trajectories using Recursive Bayesian Estimation with Dynamic Movement Primitives

6.1 INTRODUCTION

Neural control over robotic systems may soon enable victims of severe paralysis to regain the autonomy necessary to perform many essential activities of daily living.
By directly tapping into patients’ cortical signals, brain-machine interfaces (BMI) can deliver a basic level of control of prosthetic devices to quadriplegic users [34]. However, many daily tasks, such as using utensils or retrieving an object from a cluttered workspace, require complex trajectories with a degree of precision that has yet to be obtained from direct neural control. To achieve the level of robust neural control needed for widespread clinical use, it is likely that BMIs will need to incorporate shared control strategies that intelligently capitalize on information obtained from environmental sensors.

Environmental sensors (e.g. RGB-D cameras) and intelligent robotics have been shown to improve system performance in teleoperated robotics [62]. Augmenting teleoperation with shared control can be separated into two core components: prediction of the user’s intent, and arbitration between system autonomy and direct user control. Prediction often focuses on inference of a goal location and/or trajectory prediction [63]. It can also involve segmentation of user inputs with a hidden Markov model (HMM) [64] in order to draw from a library of movement primitives that aid in task completion [65, 66]. Arbitration between the system predictions and user commands in teleoperated robotics is often accomplished with a continuous linear blending of user inputs with system predictions [68, 69, 63], which has been used to deliver shared control of a BMI [70]. Intelligent arbitration is paramount to user satisfaction [63] with teleoperated robotics, and is integral for the success of shared control with BMIs [67].
Probabilistic BMI control strategies have been employed with and without information about the user’s goal, which can potentially be obtained from external sensors. Most commonly, recursive Bayesian estimation is employed with a state transition matrix built from a large corpus of kinematic data [73] or set manually according to system assumptions [74, 71]. Performance with recursive Bayesian estimation can be increased by including information about potential goals [79, 80, 55, 81]. Static canonical trajectories have been created for the goal locations and traversed through with neural data [212]. HMMs have been used to segment primates’ cognitive states during reaching tasks [75, 76], and switching Kalman filters have been used to improve results [77, 78]. Combining cortical signals with eye-tracking and/or computer vision can further improve system performance [5, 55, 213, 61]. However, these systems have all relied on simple linear models of motion that are incapable of reproducing the complex movements necessary for many tasks.

Here we present a probabilistic robotics framework for arbitrating between direct neural control and non-linear predictions of user intent, allowing for robust movement decoding of complex trajectories to novel locations. The type of task the user is engaged in (e.g. resting or pushing a button) is predicted through a hidden Markov model. The inferred task is then used to inform a non-linear prediction of the user’s desired kinematics using dynamic movement primitives (DMPs) [214]. This prediction is fused with neural sensor measurements via unscented Kalman filtering in order to localize the user’s intended position. We demonstrate the utility of this framework
by reconstructing the complex trajectories taken by a non-human primate performing four different actions on objects placed in various locations. By intelligently leveraging information from our motion models and potential environmental sensors, we dramatically improve offline 3D trajectory decoding over using neural signals alone.

6.2 MOTION PLANNING ARCHITECTURE

6.2.1 System Overview

We formalize the problem of tracking the user’s desired movements by using a dynamic Bayesian network (DBN), shown in Figure 6.1. We observe features extracted from the neural signals (Neuro). We also assume we observe the type (Obj) and location (Goal) the user intends to reach to. This can be accomplished with eye-tracking and computer vision [5]. The continuous kinematics variable, $Kin$, is a 9-dimensional vector tracking the position, velocity, and acceleration of the user’s desired 3D endpoint. $Kin$ is assumed to have Gaussian noise and follows switching Kalman filter dynamics [215]. The switching variable, $Act$, represents the category/state of the current action the user wants to take. $Act$ is a discrete-valued latent variable with first-order Markov properties.

In general, the value of $Act$ can represent simple sub-component of a movement
(e.g. moving to the object) or sophisticated sequences of movements (e.g. moving to then turning a handle). The transitions of \( \text{Act} \) represent the onset/offset of movements, which determine when the DMPs start/stop their predictions. The value of \( \text{Act} \) can also determine the type of DMP selected (e.g. a DMP for a drinking motion vs. a pushing motion). Within this study, however, we make the simplifying assumption that each object has only one possible sequence of actions. During our global system evaluation, the estimated value of \( \text{Act} \) is only used for detecting the onset and offset of movement, and the type of DMP used for motion prediction is determined by the object type.

The value of \( \text{Kin} \) is predicted from \( \text{Kin}_{t-1} \) using the current DMP. The parameters of the DMP are determined by the values of \( \text{Act}_{t-1} \) (the movement type) and \( \text{Goal}_{t-1} \) (the location of the object). Details of the DMP updates are outlined in section 6.2.2. \( \text{Act}_t \) is predicted from \( \text{Act}_{t-1} \) using an \text{Obj} dependent state transition matrix. For example, if the identified object is a cup, \( \text{Act} \) is more likely to transition from rest to drinking than from rest to poking.

After \( \text{Act}_t \) and \( \text{Kin}_t \) are predicted from the previous time step, they are updated using the measurement of \( \text{Neuro}_t \). The value of \( \text{Act}_t \) is updated using the HMM forward algorithm, and \( \text{Kin}_t \) is updated using the Kalman filter update. The 3D position estimate from \( \text{Kin}_t \) is then used as the system output. Table 6.1 describes each of the tracked variables.
CHAPTER 6. HIGH PRECISION NEURAL DECODING OF COMPLEX
MOVEMENT TRAJECTORIES USING RECURSIVE BAYESIAN ESTIMATION
WITH DYNAMIC MOVEMENT PRIMITIVES

Figure 6.1: Dynamic Bayesian network of motion prediction using computer vision
and neural signals. Rectangles represent categorical (discrete) variables, circles are
continuous. Blue is observed, and orange is inferred at each time step. Obj represents
the type of selected object (possibly none), Act is the desired action to execute (e.g.
rest or drink), Goal is the 3D endpoint the user is trying to reach, Kin is the current
3D position, velocity, and acceleration, and Neuro is the feature vector extracted from
the neurological signals. Arrows denote conditional dependencies between variables.
CHAPTER 6. HIGH PRECISION NEURAL DECODING OF COMPLEX MOVEMENT TRAJECTORIES USING RECURSIVE BAYESIAN ESTIMATION WITH DYNAMIC MOVEMENT PRIMITIVES

Table 6.1: Variable Descriptions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Values in this Study</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obj</td>
<td>Discrete</td>
<td>Sphere, Push Button, Handle, Mallet</td>
<td>Yes</td>
</tr>
<tr>
<td>Act</td>
<td>Discrete</td>
<td>Rest, and Two Movement States Per Object</td>
<td>No</td>
</tr>
<tr>
<td>Goal</td>
<td>Continuous</td>
<td>8 possible locations</td>
<td>Yes</td>
</tr>
<tr>
<td>Kin</td>
<td>Continuous</td>
<td>Position, Velocity, Acceleration</td>
<td>No</td>
</tr>
<tr>
<td>Neuro</td>
<td>Continuous</td>
<td>Spike Rate (Hz) of the Neurons</td>
<td>Yes</td>
</tr>
</tbody>
</table>

6.2.2 Dynamic Movement Primitives

BMIs often improve predictions of kinematics by combining a neural observation model with a state transition model with recursive Bayesian estimation. When modeling a user’s kinematics with BMIs, previous efforts have typically assumed the user’s kinematics evolve linearly over time [74][77]. However, linear models are often unable to capture many of the dynamics associated with more sophisticated trajectories. We therefore used dynamic movement primitives [214] to develop a library of motion models capable of capturing the nonlinearities involved in complex movements without sacrificing the flexibility needed for navigating dynamic environments. We used these models to predict the kinematic state transitions during Kalman filtering (see section 6.2.5).

When modeling the non-rhythmic movement of a variable $y$ to a goal location
CHAPTER 6. HIGH PRECISION NEURAL DECODING OF COMPLEX MOVEMENT TRAJECTORIES USING RECURSIVE BAYESIAN ESTIMATION WITH DYNAMIC MOVEMENT PRIMITIVES

$g$, DMPs combine a simple point attractor system (a damped spring model) with a nonlinear forcing function $f$:

$$
\tau \ddot{y} = \alpha_y (\beta_y (g - y) - \dot{y}) + f
$$

where $\tau$ is a time constant, and $\alpha_y$ and $\beta_y$ are positive constants. The point attractor dynamics ensure that the system will end at $g$, while the forcing function $f$ allows arbitrary trajectories to be taken while approaching $g$. The forcing function is composed of $N$ weighted Gaussian kernels $\psi_i$ indexed by a variable $x$ that exponentially decays to zero over time:

$$
f(x) = \sum_{i=1}^{N} \frac{\psi_i(x)w_i}{\sum_{i=1}^{N} \psi_i(x)} (g - y_0)x
$$

$$
\dot{x} = -\alpha_x x
$$

where $w_i$ is the weight of kernel $i$ and $y_0$ is the initial position. As $x$ decays, it causes $f(x)$ to go to zero, leaving only the stable point attractor that approaches the goal. Readers are referred to [216] for a comprehensive overview of the theoretical framework of dynamic movement primitives.

6.2.3 HMM Temporal Update

To track the state of our switching variable, we apply a hidden Markov model approach. At every time step, the probability distribution over actions, $\text{Act}$, is first estimated using the estimate of $\text{Act}$ from the previous time step and a state-transition model. We make a first-order Markov assumption that $\text{Act}_t$ is conditionally inde-
pendent of everything in the past given $Act_{t-1}$ and $Obj_{t-1}$. To construct an initial estimate the state distribution, we multiply the posterior estimate of state from the previous time step by the object-dependent state transition matrix. The possible combinations of prior states and transitions are combined by marginalizing over the estimate of $Act$ from the previous time step:

$$P(Act_t|Obj_{t-1}) = \sum_{Act_{t-1}} P(Act_t|Act_{t-1}, Obj_{t-1})P(Act_{t-1}|Neuro_{t-1}, Obj_{t-2})$$

where the action state transition probabilities, $P(Act_t|Act_{t-1}, Obj_{t-1})$, can either be learned from example data or intelligently set by an operator.

### 6.2.4 HMM Measurement Update

After obtaining $P(Act_t|Obj_{t-1})$, we update our estimates of the action by using the most recent observations of the neural signals. The action is updated following the forward algorithm under the naive assumption that each neural feature is independent:

$$P(Act_t|Neuro_t, Obj_{t-1}) = \frac{P(Neuro_t, Act_t, Obj_{t-1})}{P(Neuro_t, Obj_{t-1})}$$

$$= \frac{P(Neuro_t|Act_t, Obj_{t-1})P(Act_t, Obj_{t-1})}{P(Neuro_t|Obj_{t-1})P(Obj_{t-1})}$$

$$= \frac{P(Neuro_t|Act_t, Obj_{t-1})P(Act_t|Obj_{t-1})}{P(Neuro_t|Obj_{t-1})}$$

$$\propto P(Neuro_t|Act_t, Obj_{t-1})P(Act_t|Obj_{t-1})$$
where $P(\text{Neuro}_i|\text{Act}_t)$ is the Gaussian distribution of the $i^{th}$ neural feature given current action, with a mean and covariance estimated from training data. Equation $v$ assumes $\text{Neuro}_i$ is conditionally independent of $\text{Obj}_{t-1}$ given $\text{Act}_t$. Equation $v$ also incorporates the naive assumption of conditional independence of the neural features. The denominator of equation $iii$ is a normalization constant, because $P(\text{Neuro}_i|\text{Obj}_{t-1})$ is not influenced by the $\text{Act}_t$ being estimated.

The estimated probabilities of the actions being performed over time can be used to inform the predictions of the continuous dynamics. However, tracking every possible sequence of trajectories with a switching dynamical system is intractable \cite{215}. In our experimental validation, we make the simplifying assumption that each object has only one possible sequence of actions associated with it (rest, movement state 1, and movement state 2), which is consistent with the experiment performed. For example, when manipulating the mallet, the movement began at rest (Rest), moved to the mallet (Mallet 1), transitioned to holding the mallet in a constant position (Mallet 2), then rested at that position (Rest). We update the predictions of the kinematics using only the most likely action at any given time.
6.2.5 UKF Temporal Update

The Kalman filter is a recursive Bayesian method for alternating between predicting the continuously valued state of a series, and updating the prediction with a noisy measurement of the true value. At each time step, both the estimate and the certainty (i.e. the covariance) of the value are tracked. Whenever a prediction is made, noise is introduced and the certainty decreases. Every time a measurement is made, the certainty of the estimate increases. When measurements have higher certainty, they correct the predictions more heavily.

To make a prediction with the standard Kalman filter, a linear model with zero mean offset is assumed. This can be represented with a matrix, $F$. The mean and covariance of the tracked state $X$ can then be updated as follows:

$$\mathbb{E}[X_t] = \mathbb{E}[FX_{t-1}] = F \mathbb{E}[X_{t-1}]$$

$$\text{Cov}(X_t, X_t) = \text{Cov}(FX_{t-1}, FX_{t-1}) + Q$$

$$= F\text{Cov}(X_{t-1}, X_{t-1})F + Q$$

where $Q$ is the covariance of the noise introduced by the prediction. However, in this study, the predictions are performed using nonlinear DMP functions that cannot be represented with a linear matrix multiplication.

Two methods are commonly employed with Kalman filters to account for the impact of a nonlinear transformation on the distribution of the random variable being tracked. The extended Kalman filter uses a first order, local linear approximation of
the nonlinear transform. This linear approximation can then be used to update the estimates of the expected value and covariance. However, in situations where there is a large degree of uncertainty in the estimates, this local approximation can have an undesirable degree of inaccuracy that accumulates over time.

The other common method for utilizing a nonlinear function within the Kalman filter is to employ the unscented transform [217]. The unscented transform is a method of approximating the mean and covariance of a random variable after a nonlinear transformation. This is accomplished by propagating samples called sigma vectors through the nonlinear function. The first sigma vector is set to the mean of the distribution, and the rest of the sigma vectors are offset along the main axes of the random variable’s covariance. After the nonlinear function is applied to the sigma vectors, they are used to calculate an empirical estimate of the new mean and covariance. Readers are referred to [217] for an in-depth analysis of the method.

In this study, the random variable undergoing a nonlinear transformation is $Kin$, the 3D position, velocity, and acceleration. The nonlinear functions applied to $Kin$ are the DMPs. The unscented transform is used to calculate $\tilde{Kin}_{t|t-1}$ and $P_{t|t-1}$, the \textit{a priori} estimates of the mean and covariance of $Kin_t$ after the DMP is applied to the previous timestep. The \textit{a priori} estimates are then updated with the measurements from the neural signals to get the \textit{a posteriori} estimates, denoted as $\tilde{Kin}_{t|t}$ and $P_{t|t}$ respectively.
6.2.6 Kalman Filter Measurement Update

While the prediction step is nonlinear, the mapping between $\text{Kin}$ and $\text{Neuro}$, represented with a matrix $H$, can be represented as a set of linear equations [34]. This enables the standard Kalman filter measurement update to be employed:

Measurement Residual: $\tilde{z}_t = \text{Neuro}_t - HK\hat{\text{Kin}}_{t|t-1}$

Residual Covariance: $S_t = HP_{t|t-1}H^T + R$

Kalman Gain: $K_t = P_{t|t-1}H^T S_t^{-1}$

A Posteriori State Estimate: $\hat{\text{Kin}}_{t|t} = \hat{\text{Kin}}_{t|t-1} + K_t \tilde{z}_t$

A Posteriori Cov Estimate: $P_{t|t} = (I - K_t H)P_{t|t-1}$

The residual covariance, $S_t$, dictates how much weight to place on the measurement. As the certainty of the measurement decreases and $R$ goes to infinity, $K$ will go to 0 and the weight will all be placed on the a priori estimate. As the measurement certainty increases and $R$ goes to 0, the Kalman gain will approach $H^{-1}$, thereby fully correcting the a priori estimate with the measurement residual, resulting in all the weight being placed on the measurement at every timestep.
6.3 EXPERIMENTAL VALIDATION

6.3.1 Experimental Details

A male rhesus monkey was trained to perform a center-out reach, grasp, and manipulate task. All procedures were approved by the University Committee on Animal Resources at the University of Rochester. The monkey was implanted with floating microelectrode arrays (FMAs) containing 16 microelectrodes each. Eight FMAs were implanted in the left motor and premotor areas. Signals were sorted into spike trains using software from Plexon (Plexon, Dallas, TX). We excluded neurons with low mean firing rates, leaving 80 units of 104 recorded.

![Figure 6.2: Reach trajectories performed by the non-human primate recorded via optical tracking of the monkey’s wrist. The origin corresponds to the home location. Line colors correspond to the location of the target object. Shapes are centered where the monkey completed its trajectory.](image-url)
A detailed description of the experimental setup can be found in [218]. In brief, the monkey was seated in front of an experimental apparatus consisting of a central home object, a sphere, push button, coaxial cylinder (pull), and perpendicular cylinder (mallet). The objects were spaced with 45° intervals in a circular arc centered on the home object with radius of 13 cm. Target objects were indicated by an LED shortly after the monkey pulled on the home object. The monkey then rotated the sphere, pulled one of the two cylinders, or pressed the button. The final hold state was then maintained for one second for the monkey to get a reward. The apparatus was pseudo-randomly rotated in 22.5° intervals after the monkey performed a block of trials to an object/location pair. This resulted in eight potential locations for the four objects. The reaching kinematics of the monkey’s wrist were recorded with a Vicon optical motion capture system (Vicon Motion Systems, Oxford, UK), shown in Figure 6.2 and 6.3.

Figure 6.3: Alternate angle of the optical recordings of the monkey’s reach trajectories.
6.3.2 Neural Modeling

We modeled the smoothed multi-unit firing rates as a function of the kinematics and the action being taken. Firing rate was calculated by convolving spike times with a truncated exponential with a time constant of 20 ms, then smoothing with a moving average filter of 100 ms. We used these features with additional tap delays of 0, 20, and 60 ms to allow for transduction delays from cortical activity to arm kinematics (consistent with previous works [73]). All firing rates and kinematics were standardized using the mean and standard deviation from the training data.

The mapping, $H$, between the neural features and the kinematics was modeled using ordinary least squares (OLS) regression, and the error covariance was estimated from the residuals. The neural only decoding model was likewise built using OLS. When performing decoding using neural signals alone, we decoded the trajectory continuously without the HMM. Cross-validation was performed by creating a hold-out set of all trials for a particular object/location pair and training on all remaining trials. The fitted model was then evaluated on the hold-out set. This process was repeated for all object/location pairs.

6.3.3 Action Prediction Fitting

The switching dynamics of the actions were used to predict the onset and offset of movement. Trials were segmented into rest, movement, and hold periods by
thresholding the movement speed of the monkey’s hand. An HMM with nine states (shown in figure 6.4) was built to track the value of Act: one for rest, and two for each of the object/action types (push, pull, mallet, sphere). Two movement states were established for each movement type due to a substantial shift in neural activity as the monkey transitioned from moving towards the object to a constant grasp of the object. The Act variable tracked this shift in activity to provide an estimate of when the end effector should maintain its current position. Transition probabilities between the movement states for different actions were set to zero.

Figure 6.4: State transition diagram for action prediction. Each object manipulation was modeled as occurring in two phases. The monkey began and ended each trial by resting its arm motionless.

The neural features were modeled as conditionally independent Gaussians given Act. For simplicity, we performed supervised learning of the HMM parameters. To do so, the early movement state was defined as the first 200 ms of the movement (the movement to/initial manipulation of the object), and the second state was the
transition to hold. Knowledge of the duration of the states was not included in the testing phase. The HMM was assumed to start at rest, then progress to movement state 1, movement state 2, then rest again. The transition from rest to movement state 1 marked movement onset, the beginning of the continuous decoding of the trajectory.

6.3.4 Dynamic Movement Primitive Fitting

The DMPs were fit on the unnormalized kinematics using locally weighted regression [216] with 35 Gaussian kernels. It was noted that the monkey’s movement strategy could substantially change in different regions of space. This indicated the need for a mixture of DMPs with a dependency on object location, similar to [219]. For each object-location pair, we trained one DMP on all the trajectories taken to the object-location pair to the left, and another DMP on all the trajectories taken to the object-location pair to the right. The weights for these two DMPs were then averaged together. For example, to predict the trajectory taken for the sphere at location 3, we fit a DMP on all the trials to sphere location 2 and averaged it with a DMP fit on all trials to sphere location 4. To help ensure generalization to novel locations, we did not use any of the trajectories taken to the object-location pair in the test set when training the DMPs. Object locations that did not have examples to both the left and the right of the target location were excluded from testing to avoid the need for extrapolation, leaving 408 trials total across our testing sets.
We also noted that the final endpoint of the hold position ended at an offset from the object locations. The final goal location of the monkey’s wrist was not the center of the object being manipulated. The offset from the center of the object was dependent on the object type, because different objects are grasped in different locations. We therefore averaged the offset from the two adjacent locations within the training set to approximate the desired offset at the test location.

The error covariance for the DMP models was estimated from the error when predicting $K_{n_t}$ based on $K_{n_{t-1}}$.

### 6.3.5 Action Prediction Results

Prediction of movement transitions based on the neural signals was highly reliable, and is shown in Figure 6.5. With one exception among 408 trials, the predictions occurred between 94 ms before and 55 ms after movement onset. There was one trial where a prediction occurred 289 ms after movement, resulting in an exceptionally delayed movement onset prediction. Predictions occurred a median of 14 ms before movement onset, with a standard deviation of 26 ms. There were 13 false positives (ie predictions during the rest or hold periods) across the 20.8 minutes of data used in the testing sets. The hold period was detected in all the trials within 398 ms, with a median offset of 43.5 ms before hold onset, and a standard deviation of 77.6 ms. There were six false positives incurred by transitioning from state 1 to state 2 then back to state 1. Transitions to state 2 directly from rest were ignored.
While the predictions of movement onset and offset were highly reliable, the HMM was relatively unreliable in determining which object manipulation was being performed. Although the transition probabilities between actions were set to 0, the action with the highest probability over time would often switch. This was due to the states all having nonzero probabilities, and measurements indicating different states being the most likely over time. The correct object manipulation was predicted an average of 64% and 94% of the time respectively when the first and second movement states were being predicted.

**Figure 6.5:** Prediction of rest (grey), movement/object manipulation (light blue/green), and hold (orange/red) from neural data. Each row corresponds to a single trial (sorted by action type). Trials were aligned to movement onset (MO, solid line), and the median onset of the object hold period is shown with a vertical dashed line. Probabilities across action types (i.e. all blue/green and all red/orange labels) were subsequently summed together because each object only had one associated action. Green asterisks denote the 13 false positives transitioning from rest state to the first movement state, and black asterisks mark the 6 false positives transitioning from movement state 1 to 2, then back to state 1.
CHAPTER 6. HIGH PRECISION NEURAL DECODING OF COMPLEX
MOVEMENT TRAJECTORIES USING RECURSIVE BAYESIAN ESTIMATION
WITH DYNAMIC MOVEMENT PRIMITIVES

6.3.6 Trajectory Prediction Results

We first evaluated results of using the DMPs to augment motion prediction from
neural signals assuming the correct movement ($Act$). This was done to test the accu-
cracy of the trajectory decoding without any errors induced by the HMM. We excluded
the initial and final hold periods of each trial, and assumed the movement onset time
was known. Table 6.2 shows the median distance between the actual and predicted
trajectories, the mean correlation across the three dimensions (i.e. the average cor-
relation for the lateral, vertical, and forward dimensions), and the dynamic time
warping (DTW) distance (to test for trajectory shape ignoring time). We compared
performance using the neural signals alone, DMPs alone, and the unscented Kalman
filter (UKF) combining the two.

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
<th>Distance</th>
<th>DTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Alone</td>
<td>0.80</td>
<td>31.1</td>
<td>27.7</td>
</tr>
<tr>
<td>DMP</td>
<td>0.98</td>
<td>9.7</td>
<td>6.5</td>
</tr>
<tr>
<td>UKF</td>
<td>0.98</td>
<td>9.9</td>
<td>6.4</td>
</tr>
</tbody>
</table>

6.3.7 Global System Evaluation Details

The global performance (i.e. including the rest and hold periods) was evaluated
by using the HMM dynamics to detect movement onsets for the DMPs (DMP-HMM)
and UKF (UKF-HMM) algorithms. The DMP-HMM relies only on the DMP for
continuous decoding of the kinematics after the HMM has detected the movement from the neural signals. In contrast, the UKF-HMM fuses the DMP predictions of the kinematics with the neural measurements, as outlined in the Section 6.2.5 and 6.2.6.

Figure 6.6 depicts the flow diagram used for system evaluation. At rest, the DMP-HMM and UKF-HMM both remained at the home location. Because the object type was known and every object was manipulated in only one way, we added up the HMM probabilities across all possible object types for movement state 1 and movement state 2 during global system evaluation. We then used the Act with the highest probability to determine whether the monkey was at rest, in movement state 1, or in movement state 2. When a movement transitioned from rest to movement state 1, the DMP-HMM and UKF-HMM began decoding the continuous trajectory. If the predicted Act with the highest probability reverted back to rest from movement state 1, a point attractor pulled the position back to the home location and the DMPs were reset. If the Act predicted with the highest probability transitioned from movement state 1 to movement state 2, then the predicted hand position was held constant after a 300 ms delay. This delay is used to allow the monkey’s hand to settle into its final hold position.

When movement onset was predicted by the HMM, the DMP corresponding to the object being manipulated began decoding the continuous trajectory. The beginning of the second movement state marked a transition to hold, when the DMP was stopped.
and the predicted position was held constant (after a 300 ms delay to allow the monkey’s hand to stabilize).

Figure 6.6: Flow diagram of the UKF-HMM algorithm. A prediction of the current action is first made based on the previous estimate of the action probabilities. This is then updated with the measurement of the neural signals, and the output is used to decide whether to move using the UKF, hold position, or rest. Dashed lines represent inputs from the previous time step.

6.3.8 Global Performance Results

The average resulting traces are shown in Figure 6.7. Quantitative results are outlined in Table 6.3. The use of the DMP with cortical signals predicting onset and offset of movement substantially improved results. Using the UKF-HMM, however, did not improve over using the DMP-HMM. The supplemental video shows the decoded trajectories at each object/location combination used for testing.
CHAPTER 6. HIGH PRECISION NEURAL DECODING OF COMPLEX MOVEMENT TRAJECTORIES USING RECURSIVE BAYESIAN ESTIMATION WITH DYNAMIC MOVEMENT PRIMITIVES

Table 6.3: Global Decoding Performance

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
<th>Distance</th>
<th>DTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Alone</td>
<td>0.80</td>
<td>31.1</td>
<td>29.5</td>
</tr>
<tr>
<td>DMP-HMM</td>
<td>0.99</td>
<td>6.4</td>
<td>6.4</td>
</tr>
<tr>
<td>UKF-HMM</td>
<td>0.99</td>
<td>7.0</td>
<td>6.7</td>
</tr>
</tbody>
</table>

Figure 6.7: Example trajectories (position relative to home location) for an example object-location pair (push). Line color corresponds to the three dimensions being decoded. The dashed lines are single trials, and thick solid lines are the averaged across all trials for that object-location pair.

6.4 DISCUSSION

Here we show that complex movement trajectories can be reconstructed from the Bayesian fusion of neural measurements and movement primitives by leveraging information attainable from environmental sensors. By only testing on object-location pairs not seen in the training data, we demonstrate that our strategy has the ability to generalize to new locations. In spite of the difficulty of decoding the non-linear kinematics with multiple points of inflection present in the monkey’s movements, our
system is able to maintain high performance.

The neural signals were essential for estimating the HMM movement state. The UKF-HMM also enabled the neural signals to continuously alter the DMP trajectories with performance comparable to the DMP-HMM. While the UKF-HMM did not improve performance over the DMP-HMM in this study, the quality of neural control dramatically improves with closed loop practice [220]. Re-evaluation of model covariances would allow users to gain more autonomy as their neural control improves. Better cortical coverage would further improve decoding performance. Finally, it may be worth a small degradation in performance to enable users to continuously alter the automated trajectories via the UKF-HMM.

Upon detecting movements from the neural signals, the DMPs displayed remarkable performance in predicting the trajectories taken by the monkey. This may be in part due to the highly trained monkey performing very stereotyped movements across trials. Humans may move with more variability between trials, causing the deterministic DMPs to degrade in performance. The input of neural signals would be essential for enabling the system to account for inter-trial variability of the movements.

While the system was highly accurate in predicting movement onset and offset (rest, phase 1, and phase 2 of the movements), the prediction of the specific object manipulation being performed (e.g. push vs pull) was unstable and often incorrect. In this study, this was resolved by limiting objects to only one possible sequence of actions. During global evaluation, the HMM predicted onset and offset of movement
and the object type determined the type of manipulation being performed. However, in general, a single object can be manipulated in multiple ways. Future work will investigate how many actions performed on a single object can be reliably decoded with high accuracy. Techniques such as interacting multiple models (IMM) [221] can be employed to simultaneously track the most likely sequences of actions by employing parallel Kalman filters. The kinematics associated with multiple possible actions could also be tracked using particle filters [222], which can often improve accuracy. However, particle filters can quickly become computationally prohibitive to implement in an online setting.

The immense utility of incorporating continuous shared control with environmental sensors has been shown online with human subjects [70, 213]. By using environmental sensors for object localization, the subjects in [70] were able to complete tasks consistently that could not be completed as well with neural control alone. Here we built on this by incorporating DMPs with Bayesian inference to track the user’s current action and handle the challenge of arbitration between user commands and the nonlinear dynamics of autonomous robot control.

6.5 CONCLUSION

The probabilistic robotics framework formalized here establishes a novel approach to arbitrate between system automation and neural control of prosthetics for individ-
CHAPTER 6. HIGH PRECISION NEURAL DECODING OF COMPLEX MOVEMENT TRAJECTORIES USING RECURSIVE BAYESIAN ESTIMATION WITH DYNAMIC MOVEMENT PRIMITIVES

...uals with severe motor impairment. Our system allows for cortical signals to determine a user’s desired action, and uses DMPs to assist with the neural control over the end effector’s movement trajectories. Our experimental validation demonstrates the system’s potential to allow BMI users to gain robust control of the movement of neuroprosthetics, enabling them to achieve an unprecedented level of autonomy.

This chapter has supplementary downloadable material available at

http://ieeexplore.ieee.org/document/7378310/media
Chapter 7

Realtime Recursive Bayesian Inference for a Semi-Autonomous Neuroprosthetic

7.1 Introduction

Neuroprosthetics hold the potential to restore upper-limb function to amputees and victims of severe paralysis by bypassing their damaged motor systems. Previous groups have shown cortical signals can yield varying levels of control over robotic arms [45, 50, 34, 4]. EMG signals from residual muscles in amputees have been used to provide control over grasping with prosthetics [21, 20, 22]. However, while modern robotic prosthetic limbs are now approaching human levels of dexterity [133],
neuroprosthetic control with cortical signals or EMG has yet to provide patients with a level of control needed for widespread clinical use. To avoid restricting performance to what is possible using patient control signals alone, neuroprosthetics may need to intelligently share control with autonomous motion planning systems.

One common approach to improving neuroprosthetic performance is to incorporate prior knowledge of movement dynamics via recursive Bayesian estimation. Knowledge of movement dynamics is typically incorporated using a linear state transition matrix fixed according to prior knowledge \[74, 71\] or fit on a large set of recordings of movement kinematics \[73\]. Gating classifiers have been used to improve performance during offline decoding by assuming the hand is stationary during non-movement phases \[78, 223, 224\]. Estimation of the user’s goal, either through neural decoding \[80\] or eye-tracking \[55\], can be incorporated to further improve performance. In \[212\], it was shown in simulation that system performance could be improved by using neural signals to traverse static canonical trajectories for given goal locations.

Sharing control between intelligent robotic algorithms leveraging both environmental sensors (e.g. RGB-D cameras) and user inputs can improve system performance with teleoperated robotics \[62\] and, more recently, with cortically controlled neuroprosthetics \[5, 70, 225\]. The task of implementing a shared control framework can be divided into two steps. First, predictions of the user’s intent are made. This can involve predicting the user’s goal and/or the user’s desired low-level kinematics \[63\]. For example, user inputs can be segmented with a hidden Markov model (HMM)
which can enable the system to draw from a library of movement primitives to assist in the task. Next, the shared control system must blend/arbitrate between the user inputs and prediction of the user’s intent. It is critical that this be implemented well for user satisfaction with teleoperated robotics and neuroprosthetics. The most common approach is to use a continuous linear blending of user inputs and system predictions. This approach has been used to deliver shared control of a neuroprosthetic by human users, enabling them to accomplish tasks they were unable to perform with their cortical signals alone.

In, we presented a probabilistic robotics framework for arbitrating between user inputs and non-linear predictions of user intent. This was validated with offline analysis of single-unit neuron recordings in a non-human primate performing complex movement trajectories. Here we present a realtime implementation for use with electrocorticography (ECoG) or electromyography (EMG) with inertial measurement units (IMU). Computer vision is used to detect potential objects in the workspace. Eye-tracking is then used to enable users to select the object they wish to manipulate. Given the selected object, an HMM is used to segment tasks into sequences of movement primitives. The user’s desired kinematics associated with each movement primitive are predicted using dynamic movement primitives (DMPs). Finally, an unscented Kalman filter is used to arbitrate between user inputs and predictions derived from the DMPs. We demonstrate the utility of this framework by comparing success in an object retrieval task and a drinking task implemented in a virtual environment.
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A
SEMI-AUTONOMOUS NEUROPROSTHETIC environment [118, 119, 120]. By intelligently leveraging information from user in-
puts, prior knowledge of motion models, and an environmental sensor, we are able to
dramatically improve user performance in completing functionally useful tasks.

7.2 Methods

7.2.1 Experimental Setup

7.2.1.1 Training

During training, subjects follow 3D trajectories within a virtual environment by
controlling a virtual implementation of the Modular Prosthetic Limb (MPL). The
3D path to follow was shown in black. The path turns blue at a target speed it is
intended to be followed. When the path is followed sufficiently close, it turns green
and a score counter is increased.

This training phase is performed just once when using IMU control with EMG.
When using ECoG, the virtual arm first follows the intended path in an automated
path while the subject attempts to follow it without input to the system. This can
be used to form an initial mapping between the user’s neural signals and the velocity
of the virtual MPL (vMPL). They then repeat the training scenario with this initial
mapping for closed loop control. A refined mapping of their neural signals to the
intended velocity can subsequently be computed using the Re-Fit algorithm [74].
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A
SEMI-AUTONOMOUS NEUROPROSTHETIC

After performing the training scenario, lasso regression is used to form a mapping between the user inputs and the target velocity.

7.2.2 System Overview

A Microsoft Kinect for Windows V2 (Microsoft, Redmond WA) was set up to view a table that encompassed the potential workspace. A monitor placed in front of the user displayed the RGB video stream of the Kinect with bounding boxes around objects detected in the scene. A second monitor displayed the virtual limb with virtual instantiations of the detected objects. As augmented reality (AR) continues to improve, we anticipate porting the displays to a portable AR device. Figure 7.1 shows a flow diagram of the overall system for ECoG or EMG+IMU.

7.2.3 EMG Recording

EMG data was streamed into the Robot Operating System (ROS) \cite{226} from a Myo armband (Thalmic labs, ON, Canada) at 50 Hz. The Myo band was placed on the forearm. The system was validated with a user resting their hand, flexing their wrist, or extending their wrist. During training, the user repeatedly made the three movements and pressed keys to label/record data while the movements were maintained. The only features used were the eight channels of filtered EMG amplitude streamed from the Myo. The user made these movements in multiple locations in front
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC

Figure 7.1: Features are extracted from the ECoG or EMG signals. These are used to predict the action and grasp. Computer vision with eye-tracking is used to determine the object type and location that the user intends to manipulate. This helps inform the HMM guiding the action/grasp predictions, as well as the DMP predictions. The DMP predictions are fused with the 3D velocity estimates derived from the ECoG features or IMU sensors. The movement information is then sent to the movement interface, which communicates with the MPL or virtual MPL.

of them to ensure robustness. A nearest neighbors classifier was continuously updated as new labeled data was streamed in. A GUI provided a bar graph for visual feedback of the previous 25 classifications made by the nearest neighbors classifier. The user continued to train until they were satisfied the three class types were distinguishable in multiple locations, requiring approximately two minutes.


CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC

7.2.4 ECoG Recording

ECoG data was streamed into ROS from a Neuroport data acquisition system (Blackrock Microsystems, Salt Lake City UT). ECoG signals were sampled at 1000 Hz. The high gamma power was extracted at 50 Hz using autoregressive spectral analysis with an order of 16 and a window size of 128.

The high gamma power of ECoG recordings has been shown to be significantly linearly correlated with reaching kinematics [46]. While speed is the most reliably represented kinematic variable in ECoG recordings [72], it is limited as a control signal for 3D movements with a neuroprosthetic. However, there is still significant tuning to 3D reaching position and velocity [224]. Velocity has been shown to be a more useful control signal for neuroprosthetics than position [227], and has been used to deliver 3D cursor control to a paralyzed individual [45]. We therefore use a linear mapping between high gamma power and reach velocity to estimate the desired user kinematics.

While no ECoG subjects were recruited in this study, the ROS ECoG collection system was validated using analog input ports of the Blackrock system.

7.2.5 IMU Recording

Quaternions were streamed into ROS at 40 Hz from three MPU9150 MotionTracking devices (InvenSense, San Jose CA) after sampling with a Teensy 2.0 microcon-
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC CONTROLLER (PJRC, Sherwood, OR), as done in [228]. These quaternions were converted to a 3D Cartesian point representing the estimated hand location in the shoulder frame.

7.2.6 Computer Vision and Eye-tracking

A monitor displaying the RGB video stream from a Kinect was placed in front of the user. A Tobii EyeX Controller (Tobii, Stockholm Sweden) was used to enable users to control a computer cursor with their eye gaze.

Faster RCNN [229] was implemented on a Amazon Web Services (AWS) g2.2xlarge server (Amazon, Seattle WA) using Caffe [230] with a model pretrained on the MS COCO object detection data set [231]. The RGB images from the Kinect stream were compressed and sent via TCP to the AWS. The AWS continuously performed object detection on the most recent RGB frame and returned the object types and their bounding boxes. For simplification within this study, each detected object was recategorized as being a cylinder, sphere, or ignored.

Bounding boxes for the most recently detected cylinders and spheres were asynchronously updated on the video stream. Using the Kinect SDK, the center pixels of the bounding boxes were mapped to their nearest corresponding depth points, which were converted to 3D Cartesian points in camera space. These were then mapped from the camera frame to the virtual world frame by assuming a 45 degree rotation for the Kinect and a fixed offset from the virtual limb. The object types and their corresponding locations were sent over UDP to the virtual world. The objects were then

170
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC

instantiated as cylinders and spheres of fixed sizes/shapes. When the gaze-controlled cursor was located within one of the bounding boxes, the bounding box color was changed and the object type and location in the virtual world frame were also sent through UDP communication to the motion planning system.

7.2.7 Action and Grasp Classification

Each object had a sequence of possible actions associated with it (table 1). We used HMMs with inference performed using the forward algorithm to detect transitions between actions and between grasp open/close:

\[
p(x_t, y_{1:t}) = p(y_t|x_t) \sum_{x_{t-1}} p(x_t|x_{t-1})p(x_{t-1}, y_{1:t-1})
\]  

(7.1)

where \(x_t\) is the predicted class at time step \(t\) (e.g. binary transition or don’t transition between actions/grasps), and \(y_t\) is the vector of observed features at time \(t\) (i.e. high gamma or EMG amplitude). We assumed the features were conditionally Gaussian distributed:

\[
y_t|x_t \sim \mathcal{N}(\mu, \sigma^2)
\]

where \(\mu\) and \(\sigma\) are estimated from training data.

If an HMM detected a transition out of rest while no object was selected via eye-tracking, it was ignored. Once a transition out of rest was detected, the computer vision software was paused and the object type/location were assumed to be constant.
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC

<table>
<thead>
<tr>
<th>Selected Object (Eye-Tracking)</th>
<th>Action Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Rest</td>
</tr>
<tr>
<td>Sphere</td>
<td>Rest, Outward Reach, Return Reach, Rest</td>
</tr>
<tr>
<td>Cylinder</td>
<td>Rest, Outward Reach, Move to Mouth, Return to Table, Return Reach, Rest</td>
</tr>
</tbody>
</table>

The grasp conformation (i.e. cylinder vs spherical grasp) and wrist orientation were determined by the object type.

7.2.7.1 EMG Classification

For EMG classification, we used an HMM to classify between rest, wrist extension, and wrist flexion, meaning $x_t$ in equation 7.1 corresponded to these three conditions. We calculated the empirical mean and covariance of the EMG channels during each condition. The state transition matrix was set to not allow transitions between wrist flexion and extension, requiring a return to rest first. If a wrist extension was detected, the hand was opened if it was currently closed, or closed if it was currently open. The type of grasp used was determined by the most recently selected object. If a wrist flexion was detected, then the arm proceeded to the next action, which was solely dependent on the type of object selected. If no object was selected while the user was detected as resting, then the action would not transition to anything.
7.2.7.2 ECoG Classification

Due to the somatotopic organization of independent reach and grasp areas of the brain, it is possible to simultaneously classify reach and grasping movements in parallel [4]. We therefore use two independent HMM classifiers to simultaneously detect transitions in grasping and transitions between actions. Both the action and grasp HMMs make binary classifications, where $x_t$ from equation 7.1 corresponded to whether or not a transition was occurring at time $t$. The $\mu$ and $\sigma$ values can be calculated for $y_t|x_t$ for grasp transitions during the time period surrounding grasp onset.

7.2.8 3D Reach Prediction

The user’s desired 3D reach position was calculated recursively at each timestep. This was done by first predicting the desired position given the previous timestep, then updating the prediction given the user’s velocity measured through either the neural signals or the IMUs.

7.2.8.1 Dynamic Movement Primitives

Dynamic movement primitives (DMP) are a biologically inspired non-linear dynamical system capable of modeling complex movement kinematics that evolve over time [214]. These can be used to build a library of motion models that can be drawn
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC

on to form arbitrarily complex movements. For example, they have been used to enable robotic limbs to play table tennis [219], pour water [232, 233], and throw balls [234].

When modeling the dynamics of a variable $y$ moving towards a target location $g$, DMPs combine a point attractor (a damped spring model) with a nonlinear forcing function $f$:

$$\tau \ddot{y} = \alpha_y (\beta_y (g - y) - \dot{y}) + f$$

where $\tau$ is a time constant, and $\alpha_y$ and $\beta_y$ are positive constants. Because $f$ goes to zero over time, the point attractor dynamics ensure a stable system that approaches $g$ as $t$ goes to infinity. The forcing function $f$ changes the shape of the trajectories taken by $y$ as it approaches $g$. The forcing function is composed of $N$ weighted Gaussian kernels $\psi_i$ with centers $c_i$ and variances $h_i$.

$$\psi_i = \exp(-h_i(x - c_i)^2)$$

These kernels $\psi_i$ are a function of a variable $x$ that exponentially decays to zero over time:

$$f(x) = \frac{\sum_{i=1}^{N} \psi_i(x)w_i}{\sum_{i=1}^{N} \psi_i(x)} (g - y_0)x$$

$$\dot{x} = -\alpha_x x$$

where $w_i$ is the weight of kernel $i$ and $y_0$ is the initial position. The forcing function $f(x)$ goes to zero as $x$ decays.

Readers are referred to [216] for a more comprehensive overview of dynamic move-
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A
SEMI-AUTONOMOUS NEUROPROSTHETIC

ment primitives.

In this study, we focused on an object retrieval task and a drinking task. Object
retrieval was decomposed into two movement primitives: the outward reach, and the
return reach. Drinking was decomposed into four: outward reach, bringing the cup
to the mouth, returning the cup to the table, and returning to a resting position.
We developed 50 location dependent movement primitives for both segments of the
object retrieval task, and 30 for each of the four segments of the drinking task. When
performing online testing, the current DMP selected was determined by the DMP
with the most similar offset between initial and target positions for that segment of
the movement.

7.2.8.2 Unscented Kalman Filter

The details and offline results on our use of unscented Kalman filtering for neural
decoding with DMPs are presented in [6]. In brief, at each time step, we make
a prediction using the dynamics of the selected DMP, then update using a measurement
of the 3D velocity based on the IMU sensor or the ECoG signals.

Whenever a prediction is made, the transition of the state (3D position, velocity,
and acceleration) alters the uncertainty about the state. If multiplying the state by
a linear state transition matrix $F$, the covariance $P$ will become $FPF$. Additionally,
process noise $Q_k$ is introduced with the prediction, increasing the uncertainty of
the current state. With a standard Kalman filter, the predicted (a priori) estimate
The covariance update from timestep $k - 1$ to $k$ can therefore be written as:

$$P_{k|k-1} = F_k P_{k-1|k-1} F^T_k + Q_k$$

where the subscript $k|k-1$ denotes that the variable has not yet been updated with the observation at timestep $k$.

However, DMPs are a nonlinear system. To estimate the impact of the nonlinear DMP prediction on the uncertainty of the current state, we use the unscented transform. Every time a prediction is made, samples of the current state are propagated through the current DMP function. The new uncertainty is then estimated from these transformed samples.

While the DMP provides a new estimate of the position, velocity, and acceleration, we only estimate the 3D velocity from the IMU or ECoG features at each time step. We use a linear mapping between the state estimate of the velocity and the IMU or ECoG features. This allows the standard Kalman filter measurement update equation to be used, where the prediction and measurement are linearly fused with a weighting determined by their uncertainties.
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC

7.3 System Runtime Statistics

7.3.1 Computer Vision

The total round trip time (from sending a compressed high definition Kinect frame to AWS to receiving the object detections) took approximately 700 ms. About 650 ms of this was taken by the object detection, with the remainder taken by the TCP communication. This resulted in a refresh rate for object detections of about 1.4 Hz.

7.3.2 ROS System Performance

Processing was performed on a Lenovo Y50 laptop with an i7 processor. Predictions from the unscented Kalman filter and the HMM were able to be reliably made at the set frequency of 50Hz. EMG and IMU signals were acquired approximately every 40Hz. Additional CyberGlove II (CyberGlove Systems LLC, San Jose, CA) recordings were also able to be logged at approximately 40Hz.

Packets of the 1000Hz recordings of the ECoG signals were sampled at 50Hz, with each channel potentially having a different number of samples. This resulted in CAR filtering potentially being off by one sample in some channels, with negligible impact on the extracted ECoG features. These samples were buffered and processed with a custom C implementation of the autoregressive spectral analysis Burg algorithm.
CHAPTER 7. REALTIME RECURSIVE BAYESIAN INFERENCE FOR A SEMI-AUTONOMOUS NEUROPROSTHETIC

7.4 Discussion

Here we show an example instantiation of a real-time neuroprosthetic control system for shared control over reaching and grasping. The computer vision module is used to detect objects within the workspace. The depth stream of the Kinect One is used to find the approximate 3D position of the objects relative to the fixed shoulder position of the limb. Eyetracking is used to identify the desired object to manipulate. Either ECoG or EMG could be used to cycle through actions and grasps. Either ECoG or IMU data could be used to share control over the limb’s 3D endpoint with prior knowledge of the reach trajectories incorporated via dynamic movement primitives. By utilizing an unscented Kalman filter with covariances estimated through recorded data, the weighting of influence from the dynamic movement primitives and sampled inputs would be balanced in a theoretically optimal manner.

The system could be improved by incorporating object selection and GUI display into a wearable augmented reality visor such as the Hololens. Figure 7.2 displays an example implementation of this. Here the single shot multibox detector was running on AWS. The Hololens sent the acquired RGB images to the AWS server, then received the bounding boxes of all detected objects. The objects with bounding boxes overlapping with the center of the video were displayed in front of the user.
Figure 7.2: Example augmented reality implementation of the system’s GUI. Top: a picture showing a user wearing a Hololens on the right looking at an apple. The video from the Hololens is depicted on the the screen to the left, with an apple identified through computer vision. Bottom: an image captured from the user’s view of the Hololens output. Here a user was viewing a cup, and it was labeled accordingly.
Chapter 8

Conclusion

8.1 Summary of Results

8.1.1 ECoG Decoding

ECoG high gamma activation provides robust, localized activation that can be used for control of BMI with minimal training. Independent neural control over reaching and grasping with the MPL was given to two subjects without the need for any subject training. Additionally, by using high density ECoG coverage, we were able to uncover individual finger somatotopy during sensory (Figure 8.1) and motor tasks with unprecedented detail. This was leveraged to deliver online control of all five individual fingers, yielding an ECoG BMI that controls the highest number of degrees of freedom to date.
Figure 8.1: Individual finger somatotopy during a passive vibration task. Color corresponds to the finger that showed the strongest response. The intensity signifies how much statistically stronger that finger responded than the pooled response of the rest of the fingers.
CHAPTER 8. CONCLUSION

However, while high gamma ECoG signals provide robust task-related activation selective to different body parts, it remains to be seen how well this can be tuned to the directionality of the movements. With macro ECoG electrodes, correlation with the velocity of 3D movements is somewhat modest [224]. High density ECoG electrodes over arm motor areas of the brain would likely yield higher performance than what has been historically seen with macro electrodes. However, the results and simulations seen in [72] point towards the need to de-emphasize full 3D velocity decoding from ECoG, and focus more heavily on shared control for enabling paralyzed patients to accomplish activities of daily living.

8.1.2 Shared Control

We developed two proof-of-concept shared control systems to overcome the limitations of cortical signals such as ECoG. The first system uses the principle of supervisory control to enable users to initiate an automated sequence of actions using their cortical signals. This was implemented online using ECoG control, providing the first demonstration of a shared BMI controlled by human subjects.

A crucial question with any shared control system is how to balance between the intelligent autonomous control and the direct user inputs. This balance is imperative for both system performance and user satisfaction [63]. We therefore set out to build a system that achieves an optimal balance between the two. While eye-tracking with computer vision gives a robust estimate of the object type and location, using eye-
tracking for selection suffers from the "Midas touch" problem of mistakenly activating anything that is dwelled on for too long. By tapping directly into the motor cortex of the brain, our system is able to selectively reach for objects when actual movement intention is detected. This delivers an intuitive system with minimal training and very few unintentional movements from false positives.

Our second shared control system is capable of enabling users to continuously alter the path of the autonomous motion planning. The noise associated with the user inputs and the motion planning algorithm was estimated from training data. These noise estimates were then used to fuse the user inputs with the intelligent pathing algorithm in an approximately optimal manner. In our offline analyses, this resulted in control that greatly exceeded the capabilities of user inputs alone, while allowing the users to modify the autonomous movement. A proof-of-concept online control system was put together using EMG and IMU sensors, capable of robustly grasping objects within a virtual environment \cite{118, 119, 120} while enabling users to alter the 3D movement trajectories.

\section*{8.2 Future Directions}

\subsection*{8.2.1 Computer Vision}

The most recent iteration of the computer vision system documented here utilizes the faster R-CNN algorithm \cite{229} to detect objects from the RGB image. The center
of the detected object is then localized from the depth image and fed into the shared control system. While this works in many cases, there are many objects that are not best to grasp from the center. For example, a cup with a handle should be grasped from the handle. This would require 3D pose estimation, something which could be done by fitting 3D object models to the corresponding point clouds.

### 8.2.2 Grasping

In the current shared control system, automated object grasping does not succeed as reliably as would be necessary in a clinical system. This is largely because the autonomous pathing is all done open loop without feedback, meaning small errors can accumulate and lead to poor grasping of the object. This could potentially be resolved using visuo-servoing once the user initiates a grasp within the vicinity of an object. Another strategy would be to incorporate closed loop touch feedback, making use of force/contact sensors to enable the limb to reflexively grasp objects [235].

Finally, user inputs could be incorporated to further improve grasping capabilities. Rather than restricting the system to one grasp type per object, the system could be extended to have a prior on what grasps are likely to be used, then perform object dependent grasp prediction based on user inputs. Individual finger control could also be incorporated as in [1]. Just like the current system for continuous control of the endpoint, finger control could be implemented by inferring a latent action state
CHAPTER 8. CONCLUSION

8.2.3 Decoding

8.2.3.1 Graphical Model Inference

The current implementation for performing inference assumes that the action type and object type are predicted with 100% certainty. This could be modified easily for object type by marginalizing across all object types when inferring the action type:

\[
P(Act_t|...) = \sum_{Obj_{t-1}} \sum_{Act_{t-1}} P(Act_t|Act_{t-1}, Obj_{t-1})P(Act_{t-1}|...)P(Obj_{t-1})
\]

Where \(Act_t\) is the action at time \(t\), \(Obj_t\) is the predicted object, and \(...\) represents the variables previously observed. The value of \(P(Obj_{t-1})\) could be estimated by the softmax outputs of the neural network. This would effectively weight all the object dependent action transition probabilities at each timestep by the likelihood of that object given the RGB image.

However, the probabilistic nature of the inferred latent action state is more difficult to incorporate. The switching motion dynamics for different action state causes the system to be a switching non-linear dynamical system. Exact inference here is intractable, as there is an exponential explosion in the possible number of state transitions over time \([215]\). However, this can be approximated using techniques such as interacting multiple models \([221]\) or multiple hypothesis tracking \([236]\). Alternatively, the system could be modeled with a particle filter. Each particle would be
updated by a motion model, where the motion model used is selected with a probability corresponding to the probability of its inferred action state. With a sufficiently high number of particles, this would effectively marginalize out the latent action via Monte Carlo sampling. However, as the dimensionality of the kinematic state variables and the number of possible motion models increase, the number of required particles would likely exceed what is practical in an online system. This would be particularly problematic if predicting the kinematics of each joint angle, rather than the 3D endpoint.

8.2.3.2 Deep Learning

Rather than impose prior beliefs about how the kinematics and neural signals relate to each other and evolve over time, it may be beneficial to learn these dynamics from scratch. One of the most promising current techniques for modeling relationships in time-series data is to use recurrent neural networks (RNN). However, vanishing gradients have typically restricted long-term relationships from being uncovered in the data [237]. This is a large obstacle for neural decoding, where pre-movement activity can be informative of latent variables (such as the action being performed) that have downstream influences on the kinematics.

To resolve this problem, methods such as long-short term memory (LSTM) [238] and gated recurrent unit (GRU) [239] RNNs have been developed. These rely on gates that can effectively disable activations, enabling gradients to essentially short-circuit
Figure 8.2: Decoding results using LSTM's for decoding.
CHAPTER 8. CONCLUSION

through the network during backpropagation. Figure 8.2 displays some preliminary results of using an LSTM, which uses memory cells to capture long-term dynamics. Inputs for the LSTM consisted of the normalized neural signals, one-hot encoding of the object type, and normalized 3D target location. Outputs consisted of normalized 3D position, velocity, and acceleration, with a mean-squared error loss function. Correlation (Pearson’s r) between the predicted and actual test data was 0.70, 0.90, and 0.92 for the 3 dimensions. While this greatly improved upon the use of a linear model alone (correlation of 0.23, 0.50, and 0.56 when using linear regression on each time point independently), it falls short of the type of performance seen with the graphical model. However, optimization of the hyper-parameters and fine-tuning of the approach with a larger dataset could yield significant improvements. Furthermore, it may be possible to blend the approach of deep learning with prior knowledge of the system constraints to increase the performance of shared control of brain-machine interfaces beyond what is capable of either framework alone.
Bibliography


BIBLIOGRAPHY


BIBLIOGRAPHY


[33] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel,
BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


“A braincomputer interface using electrocorticographic signals in humans,”
Available: http://iopscience.iop.org/1741-2552/1/2/001


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY

Sundberg, and others, “Phantom limb imaginary fingertapping causes primary
Abstract/1996/12200/Phantom_limb_imaginary_fingertapping-causes.42.aspx

[178] H. Alkadhi, P. Brugger, S. H. Boendermaker, G. Crelier, A. Curt,
about Motor Imagery: Evidence from Paraplegic Patients,” Cerebral
http://cercor.oxfordjournals.org/content/15/2/131

central gyrus in humans,” Cerebral Cortex (New York, N.Y.: 1991), vol. 18,

[180] T. P. Powell and V. B. Mountcastle, “Some aspects of the functional organi-
zation of the cortex of the postcentral gyrus of the monkey: a correlation of
findings obtained in a single unit analysis with cytoarchitecture,” Bulletin of

[181] T. B. Crapse and M. A. Sommer, “Corollary discharge across the animal


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


239
Vita

Guy Hotson received a B.S. degree in Computer Engineering from Santa Clara University in 2011, and enrolled in the Electrical and Computer Engineering Ph.D. program at Johns Hopkins University in 2011. He received the joint Johns Hopkins University (JHU) JHU Applied Physics Lab (JHU/APL) Graduate Fellowship in 2014. His research focuses on real-time, probabilistic control of neuroprosthetic devices, and his paper on individual finger control won the 2015 international BCI Award.

Starting in January 2017, Guy will work at Ford’s Research and Innovation Center in Palo Alto, where he will work on machine learning for autonomous vehicles.