

## **Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

---

Rachel Elizabeth Conn

---

Date

Coordination between motor units in an expiratory muscle in the Bengalese Finch  
across timescales

By

Rachel Elizabeth Conn  
Master of Science

Graduate Division of Biological and Biomedical Science

---

Sam Sober, Ph.D.  
Advisor

---

Gordon Burman, Ph.D.  
Committee Member

---

Ilya Nemenman, Ph.D.  
Committee Member

Accepted:

---

Kimberly Jacob Arriola, Ph.D.  
Dean of the James T. Laney School of Graduate Studies

---

Date

Coordination between motor units in an expiratory muscle in the Bengalese Finch  
across timescales

By

Rachel Elizabeth Conn  
B.S., Georgia Institute of Technology, 2017

Advisor: Sam Sober, Ph.D.

An abstract of  
A thesis submitted to the Faculty of the  
James T. Laney School of Graduate Studies of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science  
Graduate Division of Biological and Biomedical Science  
in Neuroscience  
2022

## Abstract

Coordination between motor units in an expiratory muscle in the Bengalese Finch  
across timescales

By Rachel Elizabeth Conn

Motor behaviors result from the impressive coordination of many neural signals and the effect of those signals on muscles in the body and its environment. Recent work implicates sub-behavioral timescale spike patterns for the control of motor behaviors. Whether sub-behavioral spike patterns are important for the coordination between motor units within a muscle remains an open question. We asked whether motor units within a songbird expiratory muscle exhibit coordination at various timescales during an anesthetized breathing behavior. We recorded from four motor units in the EXP of a Bengalese Finch while recording air sac pressure. We segmented spike times of all motor units into breath cycles defined by the air sac pressure waveform. We estimated the mutual information between the spike patterns of all pairwise combinations of motor units for three combinations of timescales: spike counts of each motor unit, the spike counts of one motor unit and the spike timings of another, and the spike timings of each motor unit. All pairs of motor units exhibited coordination in their spike counts, the spike timings of some motor units exhibited coordination with the spike counts of some other motor units. We were too data limited to draw conclusions about whether motor unit pairs were coordinated in their spike timings. We noticed that the spike patterns of each motor unit varied greatly throughout the data collection, so we split the data set into two consecutive halves and repeated spike count mutual information analysis for three example motor unit pairs for each consecutive half separately. We found that the coordination between each example motor unit pairs changed from the first half to the second half of the data collection. These results illustrate variability in the coordination between motor units across timescales, emphasize the importance of developing novel experimental techniques to increase data sample sizes and computational techniques to analyze under-sampled data, and highlights the propensity for neural data to exhibit nonstationarities even during a consistent behavior.

Coordination between motor units in an expiratory muscle in the Bengalese Finch  
across timescales

By

Rachel Elizabeth Conn  
B.S., Georgia Institute of Technology, 2017

Advisor: Sam Sober, Ph.D.

A thesis submitted to the Faculty of the  
James T. Laney School of Graduate Studies of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science  
Graduate Division of Biological and Biomedical Science  
in Neuroscience  
2022

## Acknowledgments

Many thanks to my advisor, Sam Sober for all of his support. He is truly an advocate for his students. Thank you to my mentors: Simon Sponberg for introducing me to my love of science and lab work and Kristen Frenzel for fostering a love of teaching and for many, many discussions filled with guidance. Thanks to my committee members, Gordon Berman and Ilya Nemenman for their willingness to answer questions and discuss difficult concepts. Thank you to my lab-mates, in particular Lyndie, Amanda, Bryce, Samir, Kyle, Raveena, and Andrea for engaging scientific discussions, guidance, math team-work, coding advice, and so much more. Thank you to my family- especially my husband, Nate, who has provided countless hours of moral support, encouragement, and belief in me as a human and to my son, Blaise, whose face continues to light up every time he sees me.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Quantitative theories of motor control . . . . .	2
1.1.1	Motor cortex: The representational view . . . . .	3
1.1.2	Motor cortex: The dynamical systems view . . . . .	6
1.2	Bridging motor control theories with what is known about biomechanics	8
1.2.1	Muscles: The size principle . . . . .	9
1.2.2	Muscles: Hill-type models . . . . .	11
1.3	Motor control via sub-behavioral, millisecond timescale changes in spike patterns . . . . .	13
1.4	Experimental Approach . . . . .	14
<b>2</b>	<b>Methods</b>	<b>16</b>
2.1	Surgical Procedures . . . . .	18
2.2	Data Processing . . . . .	19
2.2.1	EMG Data . . . . .	19
2.2.2	Pressure Data: Identifying Breath Cycle Onset Times . . . . .	20
2.2.3	Pressure Data: Omitting False Cycles and Outliers . . . . .	21
2.3	Data Analysis . . . . .	23
2.3.1	Overview of Mutual Information and the KSG Method . . . . .	23

2.3.2	Validating Stability, Estimating Error, and Parameter Selection for the MI Estimates . . . . .	28
2.3.3	Mutual Information to Identify Coordination Between the Ac- tivity of Pairs of Motor Units . . . . .	30
2.3.4	Error Propagation for Conditional MI Estimates . . . . .	32
<b>3</b>	<b>Results</b>	<b>36</b>
3.1	Motor units exhibit unique spiking properties that vary throughout the data collection. . . . .	37
3.2	Spike patterns of different motor units are coordinated in their spike counts. . . . .	38
3.3	The spike timings of some motor units may be coordination with the spike counts of others, but no coordination between the spike timings of pairs of motor units was detected. . . . .	40
3.4	The structure of motor unit coordination may change throughout the data collection. . . . .	43
<b>4</b>	<b>Discussion</b>	<b>49</b>
4.1	Motor unit coordination at sub-behavioral and mixed timescales . . .	49
4.1.1	Case 1: These motor units are coordinated at mixed and/or sub-behavioral timescales, but we were unable to detect enough coordination. . . . .	50
4.1.2	Case 2: Lack of coordination at mixed and subcycle timescales	56
4.2	Motor unit coordination across timescales . . . . .	58
4.3	A Challenge to the idea of a static code for motor control . . . . .	61
4.3.1	Nonstationarities in the coordination between motor units . .	62
4.3.2	Addressing non-stationarity concerns for motor control . . . .	64
4.4	Conclusions and future directions . . . . .	68

<b>Appendix A Proof of Upperbound on Error for Spike Timing MI Estimates</b>	<b>72</b>
A.1 Mixed Timescales: Timing and Count . . . . .	72
A.2 Subcycle Timescales: Timing and Timing . . . . .	77
<b>Bibliography</b>	<b>82</b>

# List of Figures

- 1.1 What is the timescale of motor unit coordination? **a** and **b**. Air sac pressure (**a**) and EMGs (**b**) were recorded during breathing while under anesthesia in Bengalese Finches. **b**. Simultaneous single unit activity was obtained using flexible 3D contact electrode arrays. **c**. The co-activity of the simultaneously recorded units were examined at small and large timescales. . . . . 4
- 1.2 Physiology of muscle contraction. **a**. According to the sliding filament theory, in the absence of  $Ca^{2+}$ , actin and tropomyosin are unable to bind to eachother. **b**. An influx of calcium permits myosin heads to bind to actin - cross bridge formation. This binding process repeats as the thick (myosin) and thin (actin) filaments slide past eachother, shortening the sarcomere and therefore the muscle fiber. **c**. Calcium concentration has a nonlinear relationship with muscle force. **d**. Membrane potential, calcium concentration, and tension curves vary in response to a current pulse indicated by the vertical dotted line. All subfigures were adapted from Eckert [16]. . . . . 9

2.1 Spike count and spike timings were calculated for each breath cycle to find the coordination at behavioral (spike count) and sub-behavioral (spike timing) timescales. Once breath cycle times were identified, the total number of spikes that occurred within each breath cycle were quantified for each motor unit. This spike count was representative of the neural activity at behavioral timescales. Additionally, the precise timing values of each spike within each breath cycle were measured. These values were representative of the neural activity at sub-behavioral timescales, even down to the millisecond. . . . . 17

2.2 Calibrated and filtered pressure waveforms were used to identify cycle onset times. **a.** A zero threshold was applied to the pressure data to identify breath cycle onset times (grey dotted line). The duration of a pressure cycle was calculated as the time from positive (blue diamond) zero threshold crossing to negative ( maroon diamond) zero threshold crossing. **b.** Some cycle onsets and offsets were artificially detected because of small amplitude/high frequency fluctuation in the pressure around the threshold (yellow circle). **c.** These falsely detected cycle times were omitted from the analysis by discounting all cycle times corresponding to cycles with durations less than 0.11s (yellow shaded region). . . . . 22

2.3 We examined the peak-to-peak amplitude of the pressure waveforms to ensure that we studied a consistent behavior in our analyses. **a.** For a subset of breath cycles, the bird suddenly began to breath very shallowly, and the air sac pressure cycles gradually increased in amplitude over the course of several breaths. Cycle onsets are marked by blue diamonds. The subset of cycles with small amplitudes and those surrounding were omitted from the mutual information analyses to make sure that coordination was studied under a consistent behavior for the entire analysis epoch. **b.** The peak-to-peak amplitude of all pressure cycles was calculated. **c.** The cycles associated with the sudden drops in amplitude were identified with a sliding threshold of  $3 \times IQR$  (top panel, cycles falling outside this threshold and preceding cycle marked in red). Once the first of each low amplitude cycle series was identified, the previous and next consecutive cycles were omitted until the peak-to-peak amplitudes fell within the sliding  $IQR$  (bottom panel, all omitted cycles marked in red). . . . .

2.4 The KSG MI estimator uses a  $k$  nearest neighbors approach. The algorithm estimates the marginal entropies and joint entropies from the probabilities of finding data points within the distance to the  $k$ th nearest neighbor for each data point. **a.** A schematic of artificial data illustrates the  $6$ th nearest neighbor (red) from a point of interest (dark blue) along with the total number of data points found within the  $6$ th nearest neighbor distances (black dotted lines) in the x and y directions from the point of interest. **b.** A joint histogram of one example neuron pairs shows correlated structure in the x and y variables. **c.** For each MI estimate, the dependence of the MI on the data size and the selection of parameter  $k$  were examined. For the example distribution in **b**, an example of the data size dependence is plotted for  $k = 6$ . **d.** The  $k$  dependence for this same distribution is plotted for  $k = 1 - 9$ . Based on the data size dependence and the  $k$  stability,  $k = 6$  was selected for this example MI estimate. . . . . 27

3.1 Four different motor units exhibit distinct spiking characteristics throughout the data collection time. Unit D (blue) exhibited robust spiking throughout the beginning of the data collection, and then stopped. Unit G (green) exhibited sparse spiking throughout the entire data collection. Unit E (orange) exhibited sparse spiking specifically towards the end of the data collection. Finally, Unit B (yellow) exhibited robust tonic spiking throughout the entire data collection. . . . . 38

3.2 Motor units are coordinated at the timescale of the behavior. **a.** Histograms of spike count for each unit illustrate the spike count distributions for the entire data collection. Mutual information was found to be greater than zero for all pairs of units. **b.** Count-count MI values for all unit pairs are shown. Error bars show the estimated standard deviation the MI **c.** Joint histograms of spike counts for each example pair illustrate the joint distributions of spike count across the whole data collection. . . . . 40

3.3 Motor units may be coordinated at the sub-behavioral scale. **a.** Example histograms illustrating the distribution of spike timing from three example units are shown, each for one spike count condition. The distributions of spike timings for each unit for all breath cycles where Unit D spikes two times, where Unit G spikes once, and where Unit B spikes 22 times are shown. **b.** Joint distributions for mixed timescale coordination: spike timing and spike count are illustrated for one spike count condition for two example unit pairs: Units D and G and Units B and G. **c.** Additionally, the joint distributions for sub-behavioral scale coordination are illustrated for Units D and G for all breath cycles in which Unit D spiked twice and Unit G spiked once. The sub-behavioral timescales joint distribution is shown for Units B and G for all breath cycles in which Unit B spiked 23 times and Unit G spiked once. **d,e.** The weighted sum of mutual information for mixed timescales and for sub-behavioral timescales for all spike count conditions between Units D and G and Units B and G are shown. . . . . 44

- 3.4 The coordination between motor units changes between consecutive halves of the data collection. The data were split into two consecutive halves and the behavioral scale coordination analyses were repeated for each half. **a** and **b**. The marginal histograms for each unit are shown for the first half (**a**) and the second half (**b**) of the data set. In particular, the marginal distributions of Unit D, Unit E, and Unit B change from the first half to the second half. **c** and **d**. The mutual information estimates for the same example pairs of units as in Figure 2 are shown for each data half (**c** and **d**, respectively). Each example unit pair illustrates a different changes in coordination when the data set is split. **e** and **f**. Joint distributions of spike count for each example pair are shown for each half (**e** and **f**, respectively). . . . . 48
- 4.1 Our timing-count and timing-timing MI results appear to depend on the amount of data that was able to be included in the analysis. The top panel illustrates the dependence of the each timing-count and timing-timing MI on the percentage of data included in the probability term (all breath cycles with spike counts of one or higher for all timing units) and the percentage of breath cycles that yielded stable MI estimates (i.e. the percentage of breath cycles that was not omitted for any reason including zero spikes, dimensionality, data insufficiency, or instability across parameter values). The bottom panel illustrates the dependence of the error of each timing-count and timing-timing MI on the percentage of breath cycles included in the probability term and the percentage of breath cycles that yielded stable MI estimates. Timing-count unit pairs are shown in black, and timing-timing unit pairs are shown in red. . . . . 53

# Chapter 1

## Introduction

Movement requires the integration of many signals from the brain, the body, and the environment. The nervous system consists of millions of neurons that must work together to accomplish dexterous control. Understanding the principles of neural control may contribute to improved rehabilitative treatments for motor injuries/disorders [60]. In recent decades, there have been significant improvements in experimental methods [78] and analysis techniques [61]. A great deal of progress has been made identifying correlations between motor signals and motor behavior [27, 40, 72, 97, 18, 73] and identifying correlations between multiple motor signals [45, 12, 61, 76, 64].

The vast majority of theories of motor control and motor coordination from the central nervous system to the periphery impose a particular simplifying assumption about what signals are most relevant: that studying spike patterns at large timescales (i.e. around the timescale of the behavior) is sufficient to capture the most relevant control signal for motor control. Furthermore, although this assumption has been extremely useful for the field of motor control, it may fall short when the mechanisms of muscle contractile dynamics in more naturalistic, dynamic conditions are considered [16, 67, 85, 73]. For example, the same stimulation pattern to a cockroach leg

muscle produces different bodily responses during standing versus running and phase shifting that activation pattern during running causes the roach to turn [75]. There is abundant correlative evidence that supports the assumption that changes to spike counts within a behaviorally relevant time window are the primary signals for motor behavior [40, 27, 71, 43, 19, 12, 61] and some causal evidence in muscles in highly controlled conditions [18, 50]. However, alternatives to this assumption have not been investigated from the lens of motor control until a few recent studies. These recent studies mostly in single neurons or across different muscles indicate that theories of motor control have been overlooking what may be the most relevant neural signal - spike timings [64, 82, 77, 91, 73, 75]. Whether multiple motor units within a single muscle are coordinated via spike timing patterns remains an open question.

## 1.1 Quantitative theories of motor control

Neurons communicate via chemical and electrical signals. When excitatory postsynaptic potentials accumulate in a neuron to raise the membrane potential above a certain threshold, an action potential (an all or nothing spike event) is fired, triggering the cell to send an electrical signal towards neurons at its axonal connections. While there is a vast opportunity for subthreshold signals to play a role in neural communication, to account for all subthreshold potentials and other chemical signals would be technically challenging, computationally expensive, and unnecessary to answer many questions effectively. It is useful to make a reasonable, simplifying assumption: that motor systems of most animals from central brain regions to the periphery fire action potentials (spikes) as the primary communication signals for motor control and motor coordination. Even with the assumption of spikes as the primary communication signals, there are many different aspects of spikes that could serve as the most relevant signals for neural communication. For example, a neuron

may communicate information by altering the total number of spikes in a given behavioral time window (spike count) [27]. At the other extreme, the precise timings of spikes (spike timing) may serve as important neural signals [73]. Of course, neurons may communicate at some intermediate timescale (e.g. smoothed spike rate), and neurons may communicate using multiple timescales simultaneously [77]. Therefore, it is of particular interest to identify the timescale at which spike patterns transmit the most information (Figure 1.1). This topic has been debated and investigated in sensory systems in the field of neural coding [6, 15, 80, 11, 83, 46, 3, 5, 47, 55, 92]. From these studies, it is clear that sub-behavioral timescale changes in spike patterns (even down to a few milliseconds) are correlated with sensory stimuli, and therefore spike timing may serve as important neural signals for the sensory system. Despite the evidence that exists in sensory systems about the importance of spike timings as neural signals, the vast majority of theories of motor control and motor coordination have imposed the simplifying assumption that the most relevant control signals for motor control/coordination are contained in the behavioral scale spike counts or in the smoothed spike rates (often collectively referred to as rate coding). This assumption is pervasive across theories of motor control and coordination from motor cortex to the periphery [27, 40, 72, 97, 10, 32, 33].

### **1.1.1 Motor cortex: The representational view**

Two prevailing theories dominate the studies on motor cortex: "the representational view" and "the dynamical systems view". The "representational view" is one example of a theory of motor control that assumes that motor control is encoded via spike rates. Overall, the big question to answer according to the representational view of motor cortex is "what movement variables are represented by neural signals in motor cortex?". This view of motor cortex asserts that the spike patterns of neurons in motor cortex represent specific movement variables, such as limb position or muscle

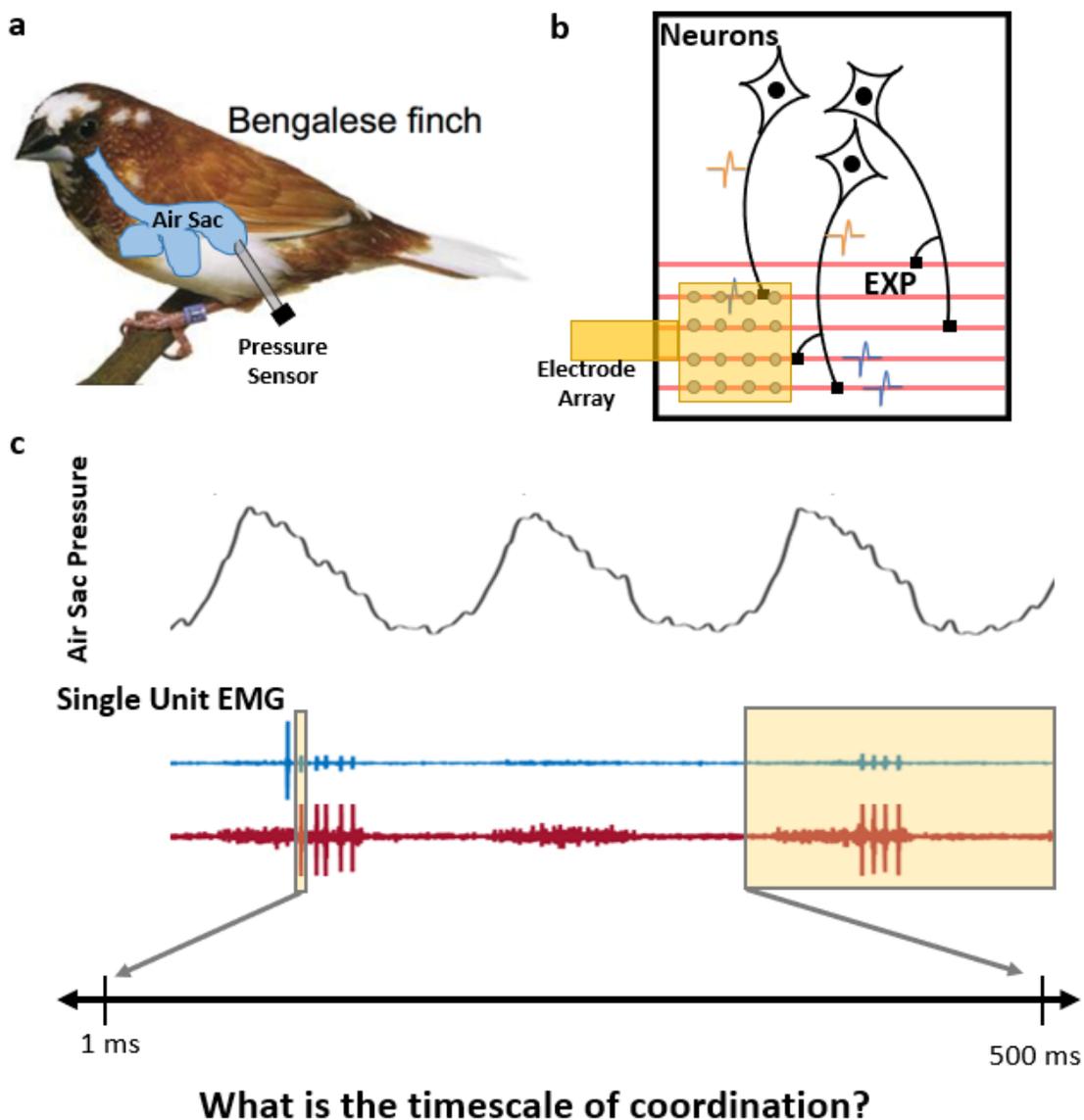


Figure 1.1: What is the timescale of motor unit coordination? **a** and **b**. Air sac pressure (**a**) and EMGs (**b**) were recorded during breathing while under anesthesia in Bengalese Finches. **b**. Simultaneous single unit activity was obtained using flexible 3D contact electrode arrays. **c**. The co-activity of the simultaneously recorded units were examined at small and large timescales.

force [40, 19]. This theory of motor control and motor coordination has been explored in individual neurons and at the population level [40, 39, 19, 2, 27, 71, 43]. Studies based on the representational view typically assume that the most relevant neural communication signals are spike patterns at the behavioral timescale. These spike rates are computed by counting the total number of spikes in behaviorally relevant time windows (on the order of hundreds of milliseconds) and dividing by the length of the time windows to compute average spike rate [19, 39, 2, 71], therefore these studies ignore all information that may be contained in the spike patterns at sub-behavioral timescales. Even still, there is a continued debate over what are the most relevant movement variables that are coded by these behavioral timescale spike patterns.

For example, work by Evarts sought to determine whether pyramidal tract neurons in motor cortex code for limb displacement or code for muscle force during reaching tasks. Monkeys were trained to perform wrist flexion/extension movements in which opposing forces and timing constraints decoupled wrist displacement from the muscle forces that must be exerted to perform the task. Pyramidal neuron spikes were counted during the wrist displacement time window ( $500 - 600ms$ ) for each wrist movement direction and each load condition. Evarts found that at the timescale of the wrist movement behavior, the spike rates of pyramidal neurons better correlated with muscle force than with limb displacement [19]. However, many other examples of similar work exist in support of both stories [39, 2].

Additionally, Georgopoulos and colleagues explored the representational view in populations of neurons in motor cortex [27, 71, 43]. Monkeys performed targeted 3D arm movements while electrophysiological recordings were taken from arm-related cells in motor cortex. Multiple linear regression was used to determine the relationship between the average firing rate of each cell during a particular time window ( $200 - 1000ms$ ) and the direction of arm movement. Overall, they found that reach direction in monkeys can be decoded by summing the firing rates of a population of neurons and

taking their reach direction tuning curves into account. Moreover, neurons that are spatially nearby to one another and share a directional tuning are highly correlated in their spike rates [45]. Although the representational viewpoint has explored a variety of movement variables, and it would be plausible to incorporate spike timing into this theory, to our knowledge, every study that has investigated the representational view of motor cortex has imposed an assumption of rate coding.

### **1.1.2 Motor cortex: The dynamical systems view**

A framework for understanding motor cortex that has gained attention in recent years is the dynamical systems viewpoint. Studies of this view have typically assumed that smoothed spike rates (intermediate, sub-behavioral timescales) are the primary signals for neural communication. However, these studies are typically less interested in which motor variables are predicted by neural signals compared with studies of the representational view. Additionally, the dynamical systems viewpoint inherently focuses on a population of neurons rather than on individual neurons. Overall, the big picture question to answer according to the dynamical systems view of motor cortex is “how do the spike activities of populations of neurons in motor cortex that generate appropriate motor behavior arise from the network?” [72]. This viewpoint proposes to offer a principled explanation of why studies in support of the representational viewpoint seem to provide contradicting evidence that the spike rates of neurons in motor cortex code for different movement variables. If it is a dynamical system that governs the spike patterns of the network of neurons in motor cortex, then we might expect to see extensive variability in the spiking activity of individual neurons. Some neurons may code for motor variables at certain times, but ultimately the network must collectively generate the activity necessary to produce the desired motor output [72, 12, 13]. This view of motor cortex asserts that by identifying the dynamics (i.e. rules describing how things change over time) that underlie the

population activity, we can better decode the resulting movement variables. Studies based on the dynamical systems view have sought to predict motor variables from a relatively low-dimensional dynamical system where the activity of each individual neuron can be largely predicted by the state of the whole system [12, 72]. These studies typically consider intermediate, sub-behavioral timescale spike patterns, and ignore any information contained in millisecond scale spike timings.

For example, an application of the dynamical systems viewpoint has offered remarkable success at decoding reach trajectories from the spike patterns of populations of neurons in motor cortex at the single-trial level [61]. Monkeys were trained to perform a maze reaching task while neural data was collected from M1 and PMd. The spike patterns for 202 neurons were processed with latent factor analysis via dynamical systems (LFADS), a machine-learning method that infers the dynamics that govern a neural population and uses those dynamics to estimate de-noised spike rates for the recorded population for each trial. The LFADS bin size was set to 5ms for this data set, so sub-behavioral timescale spike patterns were considered, but not down to the millisecond scale spike timings. Overall, for the analysis on this data set, the de-noised spike rates that were outputted from LFADS could predict the monkey reach trajectories for individual trials much more effectively than representational viewpoint decoding methods [61].

Although both the representational and dynamical systems viewpoints have historically examined some form of smoothed spike rates in their investigations of motor control, it would be possible for both viewpoints to incorporate spike timings as potential signals governing motor control. The smaller the timescale at which spikes transmit information, the higher the information capacity of each neuron. For this reason, if the rate coding assumptions that have historically been implicit in the representational and the dynamical systems viewpoints of motor cortex are incorrect, and

motor behaviors are controlled via spike timings, then there is an even larger space for growth in our understanding of motor cortex [14].

## 1.2 Bridging motor control theories with what is known about biomechanics

Studies of motor control seek to understand how the brain controls the body. A particular advantage of focusing on the periphery to answer questions about motor control and coordination is that establishing causation is typically more accessible in the motor periphery because signals in the muscles directly cause a measurable behavioral output [16, 77, 74]. Motor control and coordination studies that are focused on the motor periphery intersect heavily with the fields of biomechanics and neuromechanics, which emphasize the importance of considering the effects of physical properties of muscles and the environment on the neuromuscular transform [85].

A great deal is known about muscle physiology and contractile dynamics of individual muscle fibers. When a muscle cell receives an action potential, an influx of calcium initiates cross-bridge formation, leading to a muscle contraction (Figure 1.2 a and b). Calcium plays an important role in determining the force of a muscle, and the relationship between calcium concentration and muscle force is nonlinear (Figure 1.2 c and d) [16]. Muscles exhibit activation, length, velocity, and history dependence — the activation curve from a single spike depends on which spike number in a series of spikes is being considered [67, 85]. When all of these dependencies are taken into account, it is clear that muscle force may depend on sub-behavioral timescale spike patterns, even down to the millisecond. However, because many different factors contribute to the contractile dynamics of each muscle, it can be challenging to study each one in isolation [67]. For this reason, muscles were initially studied under highly controlled conditions (e.g. isometric - constant length, isotonic - constant tension)

[50, 38]. Heavily based on these highly controlled studies, the hypotheses and tools that have dominated studies of motor control at the level of the muscles impose the assumption that behavioral or intermediate timescale spike patterns are the most relevant signal.

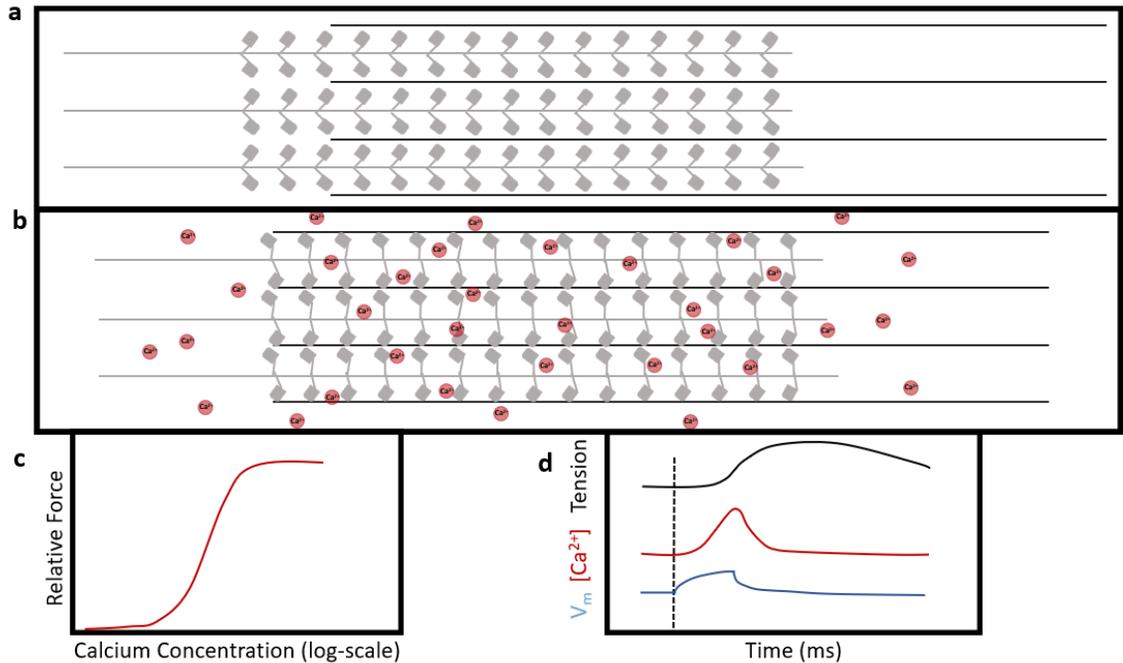


Figure 1.2: Physiology of muscle contraction. **a.** According to the sliding filament theory, in the absence of  $Ca^{2+}$ , actin and tropomyosin are unable to bind to each other. **b.** An influx of calcium permits myosin heads to bind to actin - cross bridge formation. This binding process repeats as the thick (myosin) and thin (actin) filaments slide past each other, shortening the sarcomere and therefore the muscle fiber. **c.** Calcium concentration has a nonlinear relationship with muscle force. **d.** Membrane potential, calcium concentration, and tension curves vary in response to a current pulse indicated by the vertical dotted line. All subfigures were adapted from Eckert [16].

### 1.2.1 Muscles: The size principle

For example, studies of motor control have largely been dominated by the size principle and Hill-type models, in which muscle force is determined by the differential recruitment of motor unit types and the smoothed spike rates of those units, respectively [32, 33, 97]. Classically taught in introductory neuroscience courses [41] and

initially studied in cat stretch reflexes [32, 33], the size principle states that within a behavior that requires force production, motor neurons are activated in order of excitability — smallest and most excitable neurons are activated first, and largest, least excitable neurons are activated last. The size of the motor neuron generally correlates with force produced by the muscle fiber(s) it innervates [41]. Conceptually, this ordering of motor unit excitability presents a reasonable recruitment strategy for efficient and effective motor control. The slow fibers which exert the least amount of force and fatigue the most slowly will be recruited first. Once slow fibers have been activated, if more force is necessary, fast, strong, fatigable fibers will be recruited [32, 33, 35]. Whole muscle force may be controlled by varying the number of active motor units [18]. Additionally, it is thought that changes in motor unit recruitment work in tandem with spike rate to affect the force output of muscles differently in different mechanical contexts [18, 50].

Although motor coding via recruitment and rate coding have been historically useful ideas, when the dynamics of natural behaviors are taken into account, these principles may be violated. It is thought that for the control of faster, more naturalistic behaviors that require rapid force production, it may be advantageous for organisms to recruit fast fibers before slow fibers, thereby violating the size principle of recruitment order [48]. The spinal cord and other central structures have mechanisms in place which may affect the coordination between motor units and recruitment order. For example, Renshaw cells, which are well known to execute recurrent inhibition in the spinal cord, have been shown to differentially inhibit motor neurons depending on fiber type [24], and stimulating motor cortex can affect recruitment order [48]. Similarly, in behaviors that require rapid force production, it might be advantageous for organisms to leverage spike timing sensitivities to impact motor behavior. For example, the catch-like property of skeletal muscle describes a possible mechanism by which organisms may leverage spike timing to affect muscle

force: beginning low-frequency muscle stimulation with a few high-frequency spikes can augment the force output [8].

### 1.2.2 Muscles: Hill-type models

While the size principle is a hypothesis that is thought to govern the order of recruitment of muscle fiber types, Hill-type models of contractile dynamics are commonly used to estimate the force of muscles in studies of motor control and coordination [97, 10]. These models rely on the simplifying assumption that a muscle fiber’s force output depends only on its activation, length, and velocity [97]. A typical Hill-type model consists of a contractile element and an elastic element in series. The elastic element may be modeled as a spring to represent the passive force properties of the muscle and/or tendon [97, 62]. The contractile element typically depends on three variables: muscle activation, muscle length, and muscle velocity. An example of equations governing the Hill-type model used by Perreault et al. are shown below [62]:

$$F_{Hill} = F_{CE} + F_{EE}F_{CE} = F_{max}A(t)F_{LT}(L)F_{FV}(V) \quad (1.1)$$

Here,  $F_{Hill}$  refers to the total force of the muscle as calculated by the model.  $F_{CE}$  refers to the contractile element of the force, and  $F_{EE}$  refers to the elastic element of the force.  $F_{max}$  refers to the maximum force elicited by a tetanus at the velocity and length corresponding to the peak force.  $A(t)$  refers to the activation term.  $F_{LT}(L)$  refers to the length-tension relationship, and  $F_{FV}(V)$  refers to the force-velocity relationship (see [62] for more details). The activation term in a Hill-type model is often related to calcium dynamics, or some smoothed function of spike rate [97]. Therefore, the force of a Hill-type muscle model typically depends on some variation of rate coding. Biomechanics literature supports that muscles exhibit a spike rate dependence — multiple spikes sent to a muscle increase the tension of the muscle [50]. The three

inputs to a Hill-type model are assumed to impact muscle force independently from one another. Therefore, Hill-type models are liable to miss nonlinear muscle force properties that depend on interactions between these terms [38].

Although Hill-type models are useful tools to estimate force output, they have been shown to exhibit the largest errors in the most naturalistic conditions when compared to experimentally measured force output [62]. The errors identified by Perreault et al. were thought to be attributed to the failure of Hill-type models to consider the interactions between the primary determinants of muscle force [62, 38]. Naturalistic behaviors, which rarely resemble the highly controlled conditions on which the Hill-type model is based, may be particularly sensitive to muscle history dependence and the other nonlinearities in muscle responses, contributing to the errors in Hill-type models. Naturalistic behaviors may also exhibit biomechanical sensitivities that impact the overall limb dynamics in a way that Hill-type models cannot predict [73, 85]. In addition to the nonlinearities and biomechanical sensitivities, Hill-type models may not account for the catch-like property of skeletal muscle mentioned above [8]. Each of these properties offer potential mechanisms by which millisecond scale spike timings may affect motor behavior [73].

When all the factors that contribute to muscle force are taken into account, there are many reasons to suspect that sub-behavioral timescale spike patterns may play an important role for motor control. Many natural behaviors such as running or flying involve cyclic muscle length changes [75, 74]. Due to the length and velocity dependencies of muscle force, the time of muscle activation within each length-change cycle may have major implications on the power output of the muscle. Therefore the nervous system may modulate the timing of one or more spikes within each cycle of muscle-length change in order to affect behavior. This spike timing modulation has been shown to serve as a mechanism of hawk moth flight control [74]. Further-

more, there is evidence of millisecond scale spike timings affecting the motor output in a songbird expiratory muscle [77]. If researchers reexamine the leading theories of motor control by considering sub-behavioral scale spike patterns as relevant neural signals for motor behavior, it would help to merge what is known about muscle contractile dynamics and the biomechanics of complex, naturalistic behaviors with the existing theories of motor control and coordination, which would provide much stronger and more integrated hypotheses about how the brain controls the body.

### **1.3 Motor control via sub-behavioral, millisecond timescale changes in spike patterns**

Although the current theories of motor control and coordination have made important contributions to our understanding of the motor system and have resulted in remarkably successful rehabilitative innovations [60], if the rate coding assumptions for motor control are incorrect, then our understanding of the motor system and rehabilitation techniques could be greatly improved by identifying whether sub-behavioral timescale spike patterns serve as important signals for motor control. As stated above, recent evidence from motor coding experiments supports what may naturally flow forth from well-known details about muscle contractile dynamics and biomechanics (see 1.2 above) [73]. This evidence indicates that motor behavior is not only encoded by behavioral or intermediate timescale spike patterns (rate coding), but that spike timings are important encoders of motor behavior [82, 77, 64, 74, 91].

We know that the spike timings of single neurons and the spike timings of motor units exhibit correlation with motor behaviors [82, 64, 77], and we know that small shifts in the spike timings of individual muscles have causal implications for motor behavior [74, 77]. This evidence in support of spike timing's central role for motor coding opens the door to new questions about the role that spike timing may play for

coordination between motor signals. It is unknown how multiple motor signals within a muscle correlate with each other at different timescales. Different motor signals within a muscle may not be coordinated, or they may exhibit some combination of the following types of coordination: coordination at the sub-behavioral timescale (coordination between the spike timings of different motor units), coordination at the timescale of the behavior (coordination between spike counts of different motor units), and coordination at mixed timescales (where the spike counts of one motor unit coordinates with the spike timings of another motor unit). As described below, we turned to a songbird expiratory muscle group [77] to study how motor units within a muscle are coordinated with one another across different timescales to control respiration.

## 1.4 Experimental Approach

Assessing the coordination between simultaneously recorded motor units with a high degree of temporal resolution is technically challenging. It requires stably recording well-isolated single units during a stereotyped behavior for hundreds or thousands of trials. We overcame this challenge by combining our novel flexible electrode technology with the many advantages of the songbird respiratory musculature. We recorded from multiple single units in the expiratory (EXP) muscle group of the Bengalese Finch during breathing while under anesthesia [77]. This behavioral paradigm serves as an excellent model to study coordination between motor units because we are able to stably record simultaneous single units for thousands of breath cycles without signal corruption from movement artifact.

Songbirds are capable of remarkably precise respiratory control. Songbirds have both lungs and air sacs: lungs for gas exchange and air sacs for ventilation [69]. In order for a bird to sing its highly stereotyped song, it must actively regulate

the pressure in its air sacs [68]. Air sac pressure is directly related to respiratory muscle activity, and both inspiration and expiration in birds requires neural drive [69]. Respiratory muscles have been shown to exhibit highly stereotyped, temporally precise activation patterns during singing [96]. The abdominal expiratory muscles can respond to changes in somatosensory feedback within tens of milliseconds [81]. Additionally, a small shift in the spike timing patterns produced by single motor units in the EXP group can cause changes to the air sac pressure during anesthetized breathing [77]. Therefore, not only do the respiratory muscles exhibit precise control during singing, but also during breathing while under anesthesia. Yet, we do not know whether and at what timescale multiple motor units in the EXP coordinate with one another. In this work, we used an information theoretic approach to answer the question: To what extent and at what timescale are motor units within the EXP group coordinated with each other during anesthetized breathing?

# Chapter 2

## Methods

In order to identify whether and at what timescale the activity of single motor units coordinate with one another, we recorded single unit electromyographic data (EMG data) from the expiratory muscle group (EXP) in addition to the air sac pressure in the Bengalese Finch during breathing while under anesthesia (Figure 1.1 a and b). We identified all spike times of the motor units we recorded. We cleaned and processed the airsac pressure data, using positive threshold crossing to identify breath cycle onset times. Next, we segmented the spike data into individual breath cycles - notating the spike count (the total number of spikes) in each breath cycle and the spike timings in each breath cycle (the time of each spike relative to the onset of that breath cycle) for each motor unit. We used techniques from information theory to quantify how the activity of motor units coordinate with each other at different timescales (Figure 2.1 c). We estimated the mutual information (MI) between the spike counts (behavioral timescale spike patterns) of motor units, the spike timings (sub-behavioral timescale spike patterns) of motor units, and the spike counts of one unit and the spike timings of another unit for all pairwise combinations of motor units to determine whether and at what timescale motor units within the EXP coordinate with one another.

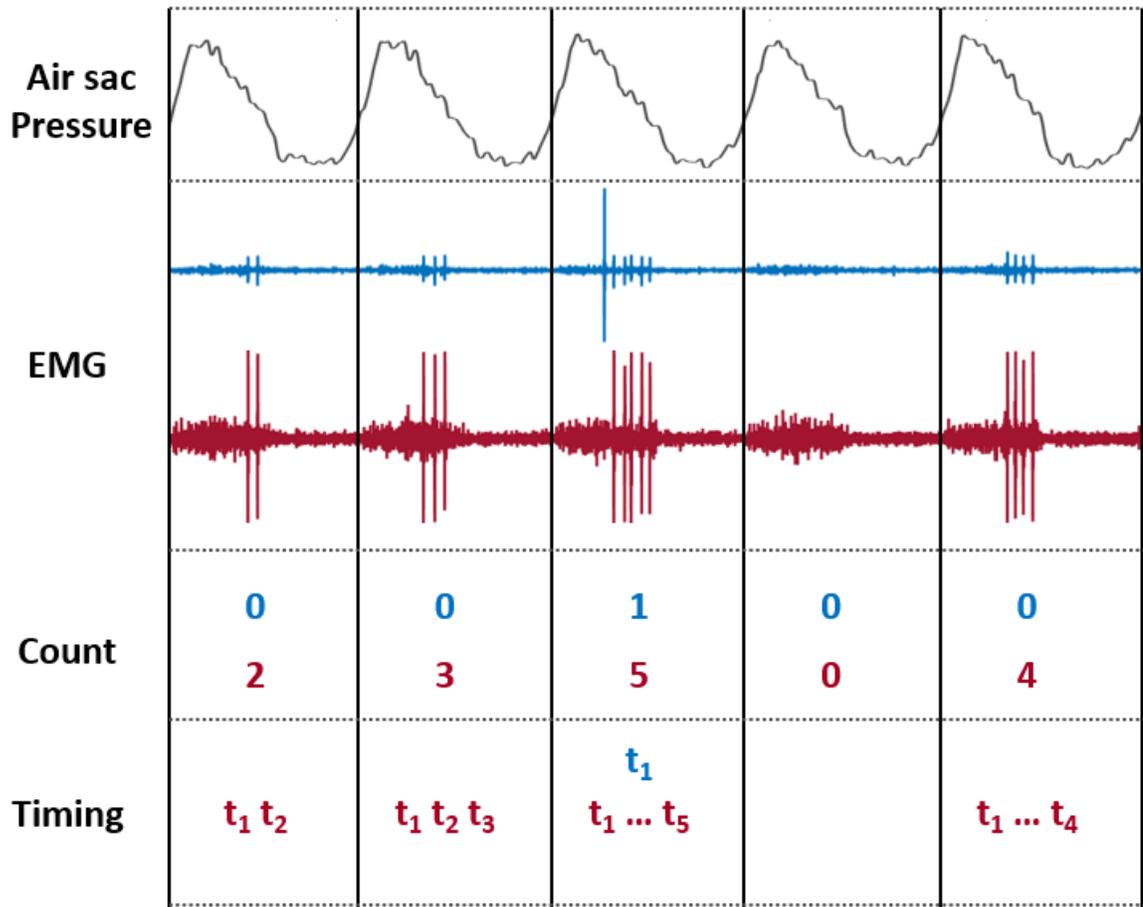


Figure 2.1: Spike count and spike timings were calculated for each breath cycle to find the coordination at behavioral (spike count) and sub-behavioral (spike timing) timescales. Once breath cycle times were identified, the total number of spikes that occurred within each breath cycle were quantified for each motor unit. This spike count was representative of the neural activity at behavioral timescales. Additionally, the precise timing values of each spike within each breath cycle were measured. These values were representative of the neural activity at sub-behavioral timescales, even down to the millisecond.

## 2.1 Surgical Procedures

Following similar procedures to Srivastava et al. we conducted simultaneous electromyographic (EMG) recordings from multiple single units of the expiratory muscle group (EXP) in addition to air sac pressure in the Bengalese Finch during anesthetized breathing [77]. All of the procedures were approved by the Emory University Institutional Animal Care and Use Committee. We note that all of the experiments and spike sorting were conducted by Bryce Chung, while Rachel Conn led the data analysis and interpretation.

To simultaneously record multiple single units, we used a flexible electrode array system with 16 3D contacts. We developed these electrode arrays with our collaborators at Georgia Tech [99, 98]. The 3D contacts were made out of gold, and held together with  $20\mu m$  thick polyimide. The contacts were  $125 - 200\mu m$  in diameter and spaced  $75 - 100\mu m$  apart in a  $4 \times 4$  grid (Figure 1.1 b).

To secure electrode arrays to EXP for acute recordings, we made an incision dorsal to the leg attachment and rostral to the pubic bone on one or both sides of the animal to target the EXP group on one or both sides of the animal. The electrode array connected to a custom Omnetics adapter, which allowed us to connect to a digital amplifier (RHD2132; Intan Technologies). We sampled the voltage signal from each electrode contact at 30kHz.

We recorded air sac pressure in parallel with electromyographic signals from the EXP. We followed procedures from Srivastava et al. and others [77, 96, 28, 4]. We inserted a  $6mm$  Silastic Tube (Dow Corning) through an incision inferior to the rib cage. The tube was beveled on one end to make insertion easier. We attached a pressure sensor 20INCH-D-4V (All Sensors) to the other end of the tube, and recorded the voltage read out from the pressure sensor at 30kHz using the same Intan system (Figure 1.1 a).

## 2.2 Data Processing

We processed the EMG and the pressure data prior to analysis. We identified spike times from the EMG voltage traces, and we identified the onset times of each breath cycle from the pressure waveforms.

### 2.2.1 EMG Data

To accurately identify spike times from multiple single motor units, we used a spike sorting algorithm ("Myosort") developed by Najja Marshall and Mark Churchland [48]. This algorithm first filtered and normalized the EMG traces. Next it detected and aligned spike waveforms using a threshold and peak detection. Similar waveforms within each channel were identified by clustering using the first three principle components of the waveforms. From all of the identified clusters, the user next manually selected the clusters that best represent the waveforms that correspond to units truly present in the EMG data. Finally, myosort employed Bayesian optimal template matching to identify each time in the EMG trace where a waveform from one of the template clusters occurred [23].

Because our EMG recordings took place over long periods of time, we needed to verify that the units we recorded exhibited consistent firing properties and waveform shapes throughout the experiment. We therefore validated the spike times detected via myosort using DataView [30, 31]. We looked at changes to specific parameters throughout the data collection. The parameters that we used include: the peak to peak amplitude of the action potential waveforms, the instantaneous firing frequency, and the phase of spikes within a behavior cycle. If each of these parameters remained consistent throughout the data collection, we considered the spike sorting output to accurately reflect the spike times of single units. Additionally, we randomly selected a subset of spike bursts to estimate a false positive and false negative error rate in

our spike detection.

### 2.2.2 Pressure Data: Identifying Breath Cycle Onset Times

In addition to processing the EMG data, we calibrated and processed the air sac pressure data so that we could consider the coordination between spike patterns within each breath cycle. We first calibrated the pressure recording from a voltage signal into units of  $kPa$  by applying a known pressure value to the sensor. We calibrated the pressure according to the following equation:

$$P = am(v/\bar{v})\bar{v}_1 - P_0 \quad (2.1)$$

Where  $a$  is a constant to ensure that we converted into units of  $kPa$  (the device we used for the calibration was in units of  $inH_2O$ ),  $m$  is the slope of the calibration curve,  $v$  is the raw voltage,  $\bar{v}$  is the mean voltage across the experiment,  $\bar{v}_1$  is the mean of the first 90 seconds of the experiment, and  $P_0$  is the offset of the voltage recording from zero when the pressure is at atmospheric pressure.

Since we recorded from the EXP group, we were interested in the spike patterns corresponding to the portion of the breath cycle when the pressure was larger than zero (corresponding to expiration). We considered the onset of a pressure cycle to be the time when the air pressure changed from negative to positive (the onset of expiration, Figure 2.2 a). In order to robustly identify these cycle onsets, we used a bandpass filter with a finite impulse response, Hanning window. The average frequency of the breath cycles was about  $1.8Hz$ . We filtered out the frequency components of the signal that were above  $50Hz$  and below  $1Hz$ . We then identified all time points where the pressure signal crossed the zero threshold in the positive direction (Figure 2.2 a). After some additional data cleaning (see 2.2.3 below), each positive threshold crossings served as the onset time of a breath cycle and each next consecutive positive

threshold crossing served as the offset time of that breath cycle. Each spike time from each neuron was sorted into its respective breath cycle. The breath cycle onset times were subtracted from the spike times corresponding to that breath cycle to output the spike timing values for each breath cycle for each motor unit. The total number of spikes within each breath cycle were summed to output a spike count variable for each breath cycle for each motor unit (Figure 2.1).

### 2.2.3 Pressure Data: Omitting False Cycles and Outliers

Prior to segmenting the EMG data into spike timings and spike counts for each breath cycle, we cleaned our breath cycle times by omitting falsely detected cycle times and outlier breath cycles. Due to low amplitude noise in the pressure signal, we detected some threshold crossings that were not true breath cycles. To ensure that we did not include artificially short pressure cycles that were detected because of small fluctuations near the threshold, we characterized the duration of a cycle by finding the length from positive threshold crossing to negative threshold crossing (Figure 2.2 b). We then omitted all cycles that had a duration less than a minimum value (Figure 2.2 c). We verified that we had successfully omitted only and all cycle times that were falsely detected by small fluctuations around the threshold by visually auditing all cycle waveforms.

Throughout the data collection, there were periodic cases where the pressure waveforms suddenly dropped in amplitude and rapidly increased over the course of a few cycles (Figure 2.3 a). We chose to omit these cycles from our analysis in order to focus on a consistent behavior. Additionally, a major assumption of our primary analysis method is that all of the data comes from a stationary probability distribution (e.g. the distribution of the data does not change over time). Omitting these cycles helped us to better satisfy the stationarity assumption. We identified these regions by finding cycles in which the peak-to-peak amplitude (Figure 2.3 b) dropped below a threshold

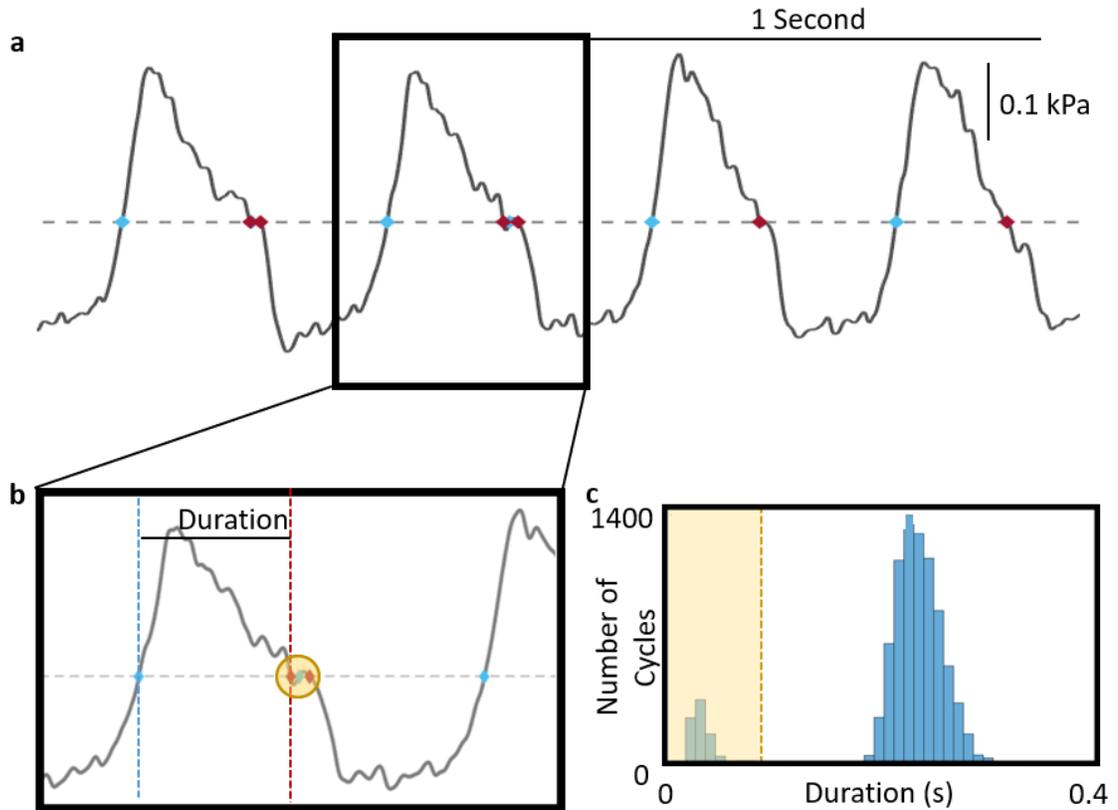


Figure 2.2: Calibrated and filtered pressure waveforms were used to identify cycle onset times. **a.** A zero threshold was applied to the pressure data to identify breath cycle onset times (grey dotted line). The duration of a pressure cycle was calculated as the time from positive (blue diamond) zero threshold crossing to negative (maroon diamond) zero threshold crossing. **b.** Some cycle onsets and offsets were artificially detected because of small amplitude/high frequency fluctuation in the pressure around the threshold (yellow circle). **c.** These falsely detected cycle times were omitted from the analysis by discounting all cycle times corresponding to cycles with durations less than 0.11s (yellow shaded region).

of  $\mu_{1/2} - 3 * IQR$ , where  $\mu_{1/2}$  is the median for a sliding window of 405 pressure cycles, and  $IQR$  is the inter-quartile range of the same sliding window (Figure 2.3 c). If a cycle had a peak-to-peak amplitude that dropped below this threshold, we omitted the preceding cycle and omitted the next consecutive cycles until the peak-to-peak amplitude fell within the  $IQR$  around the median (Figure 2.3 c).

## 2.3 Data Analysis

Overall, our goal was to understand how the activity of multiple motor units within EXP coordinated with each other during anesthetized breathing. We assessed the correlation between spike patterns at different timescales within the time course of one breath cycle. Similarly to previous work from our lab [77], in this manuscript, we focused on spike patterns at two extreme timescales- the behavioral scale, by finding the total number of spikes within each complete breath cycle (about 550 ms, spike count) and sub-behavioral timescale, by finding the millisecond scale spike timings within each complete breath cycle (Figure 2.1, spike timing). We turned to mutual information, a technique from information theory, to examine the coordination between units at these extreme timescales separately.

### 2.3.1 Overview of Mutual Information and the KSG Method

After identifying the spike times of simultaneously recorded single units and the breath cycle onset and offset times, we estimated mutual information (MI) to assess the correlation between the activity of motor units at different timescales. Although there are many mathematical techniques to measure the correlation between two variables, most techniques make assumptions about the underlying distributions of the variables (e.g. parametric techniques) or require prior knowledge about the shape of the relationship between two variables (e.g. linear or other regressions). The benefits

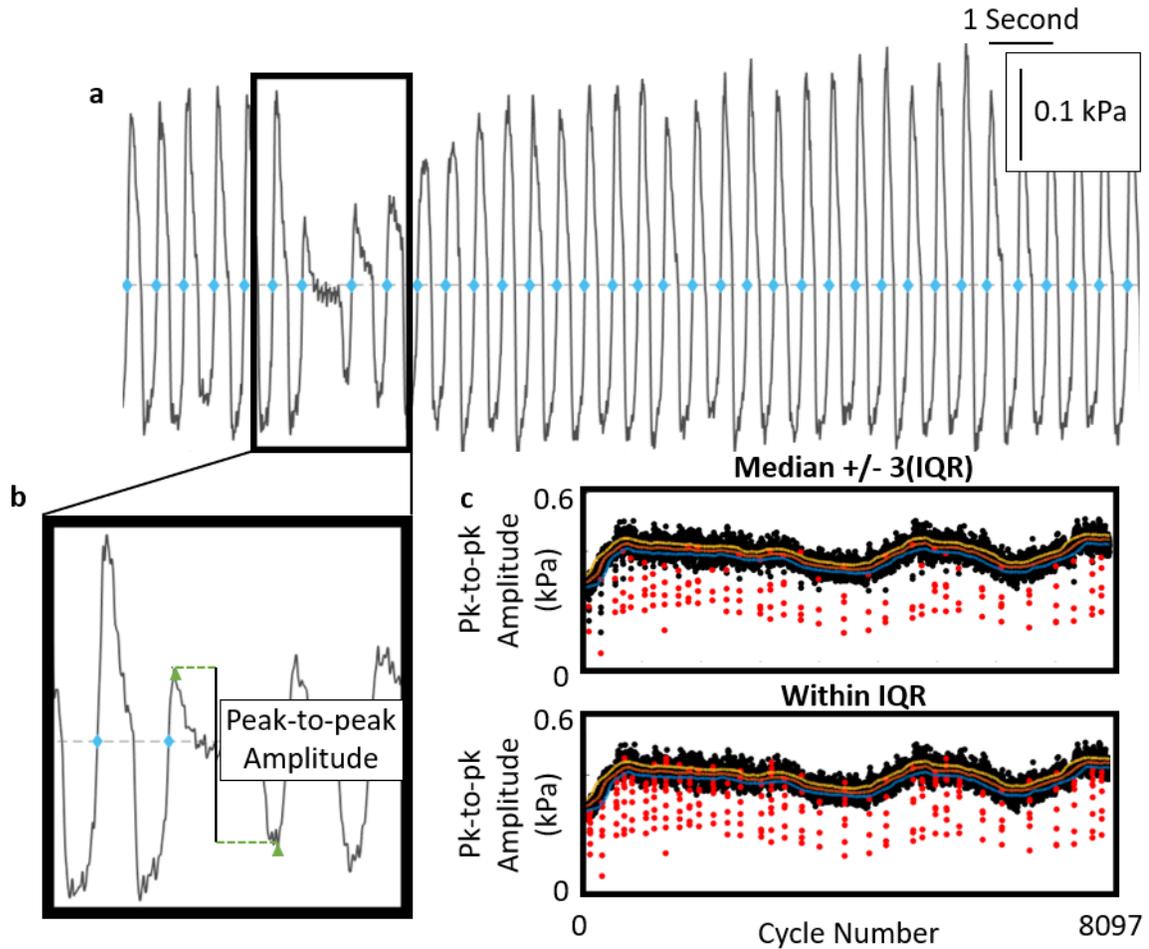


Figure 2.3: We examined the peak-to-peak amplitude of the pressure waveforms to ensure that we studied a consistent behavior in our analyses. **a.** For a subset of breath cycles, the bird suddenly began to breath very shallowly, and the air sac pressure cycles gradually increased in amplitude over the course of several breaths. Cycle onsets are marked by blue diamonds. The subset of cycles with small amplitudes and those surrounding were omitted from the mutual information analyses to make sure that coordination was studied under a consistent behavior for the entire analysis epoch. **b.** The peak-to-peak amplitude of all pressure cycles was calculated. **c.** The cycles associated with the sudden drops in amplitude were identified with a sliding threshold of  $3 \times IQR$  (top panel, cycles falling outside this threshold and preceding cycle marked in red). Once the first of each low amplitude cycle series was identified, the previous and next consecutive cycles were omitted until the peak-to-peak amplitudes fell within the sliding  $IQR$  (bottom panel, all omitted cycles marked in red).

of MI are (1) that it is model-independent - it describes the linear and nonlinear correlation between two random variables (e.g.  $X$  and  $Y$  below) and (2) that it is agnostic to the distributions of the variables. We used the Kraskov, Stoegbauer, Grassberger (KSG) k-nearest-neighbors estimator of MI for continuous variables [44]. To quantify the bias of our estimates and to ensure that the estimates were reliable for each dataset, we followed the procedures established by Holmes and Nemenman [36].

When the true marginal and joint distributions of both random variables are known, MI can be calculated exactly as follows:

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \quad (2.2)$$

However, in experimental neuroscience, we do not have the luxury of knowing the underlying probability distributions of spiking neurons due to limited datasets, etc. For this reason, MI must be estimated based on finite data samples. In particular it is not uncommon for electrophysiological data to be severely under-sampled (meaning that the ratio of the number of samples to the number of possible outcomes is much less than one) due to the technical limitations of recordings. In these cases, without carefully selecting an estimator that is designed for under-sampled regimes, it is easy for MI estimates to become inflated because the joint probability ( $p(x, y)$ ) distribution is more severely under-sampled than the marginals ( $p(x)$  and  $p(y)$ ). We chose to use the KSG estimator of mutual information for two primary reasons: (1) It is designed for continuous variables, so we can use this estimator on precise spike times. (2) It is robust to under-sampled data with a variety of marginal distributions.

The KSG estimator estimates the mutual information between two continuous variables by examining distances between data points. Because the KSG estimator is a continuous estimator, it eliminates the need to bin the spike timing data to make discrete data points. Discrete estimators, such as the NSB estimator require

binning [82, 53, 54], which limits the resolution with which we can study spike timing patterns to the size of the time bins. With estimators that require binning, the bin size must be one millisecond in order to capture millisecond-scale variability in spike patterns. Therefore, if the behavior of interest happens on the order of hundreds of milliseconds, the dimensionality of the binned spike data can become unwieldy for discrete estimators. The KSG estimator works for continuous data by using a nearest-neighbors approach. This estimator iterates through each data point and finds the distance to its  $k$ th nearest neighbor. This distance is then used to find the probability of finding data points within the  $k$ th neighbor distance in each marginal (e.g. in the  $X$  and  $Y$  direction(s), Figure 2.4 a). These probabilities are used to estimate the marginal entropies ( $H(X)$  and  $H(Y)$ ) and the joint entropy ( $H(X, Y)$ ), which can then be used to estimate the mutual information:

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (2.3)$$

Where  $H(X)$ ,  $H(Y)$ , and  $H(X, Y)$  are defined as:

$$H(X) = - \sum_{x \in X} p(x) \log_2 p(x) \quad (2.4)$$

$$H(Y) = - \sum_{y \in Y} p(y) \log_2 p(y) \quad (2.5)$$

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 p(x, y) \quad (2.6)$$

Note that the KSG estimator does not analytically calculate these quantities, but estimates them based on the distances between nearest neighbor data points in the marginal and joint distributions [44].

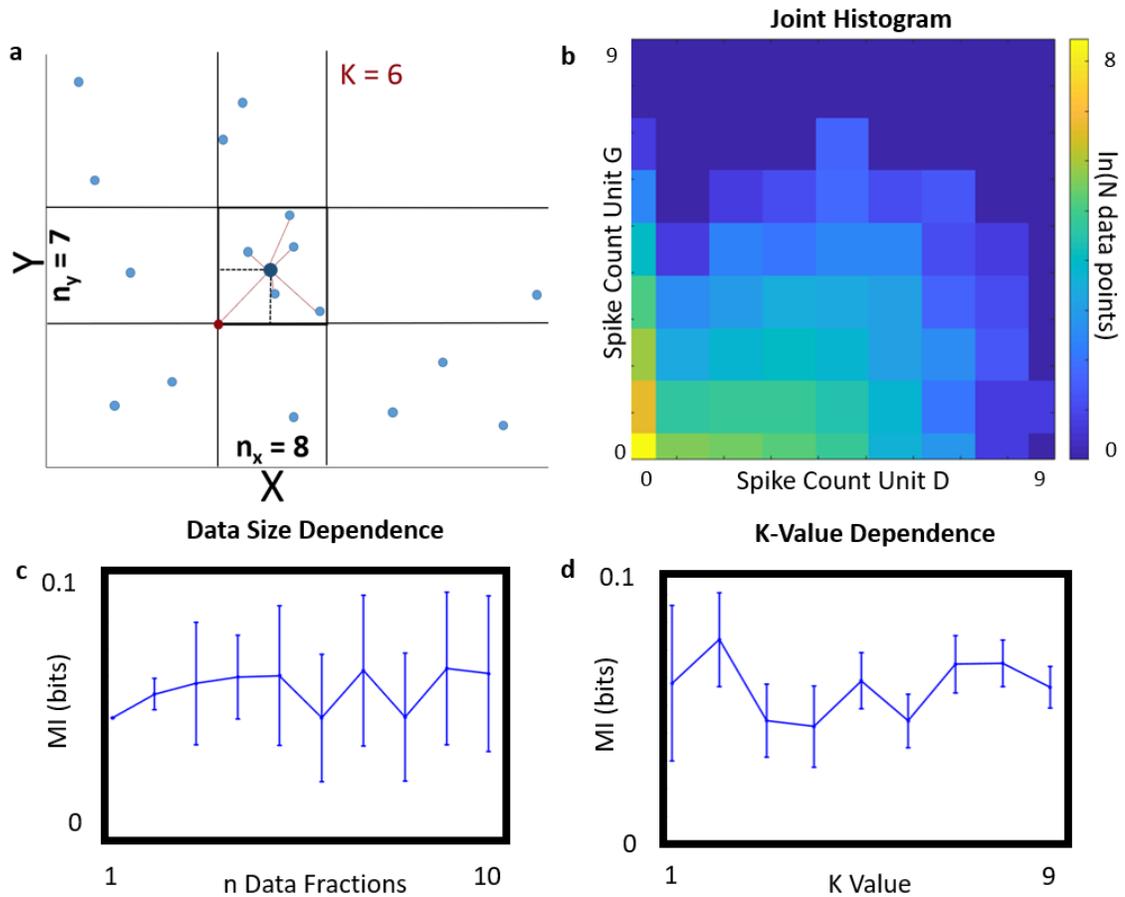


Figure 2.4: The KSG MI estimator uses a  $k$  nearest neighbors approach. The algorithm estimates the marginal entropies and joint entropies from the probabilities of finding data points within the distance to the  $k$ th nearest neighbor for each data point. **a.** A schematic of artificial data illustrates the 6th nearest neighbor (red) from a point of interest (dark blue) along with the total number of data points found within the 6th nearest neighbor distances (black dotted lines) in the x and y directions from the point of interest. **b.** A joint histogram of one example neuron pairs shows correlated structure in the x and y variables. **c.** For each MI estimate, the dependence of the MI on the data size and the selection of parameter  $k$  were examined. For the example distribution in **b**, an example of the data size dependence is plotted for  $k = 6$ . **d.** The  $k$  dependence for this same distribution is plotted for  $k = 1 - 9$ . Based on the data size dependence and the  $k$  stability,  $k = 6$  was selected for this example MI estimate.

### 2.3.2 Validating Stability, Estimating Error, and Parameter Selection for the MI Estimates

To validate the stability, select the parameter,  $k$ , and to estimate overall error for a single MI estimate, we followed the procedure outlined in Holmes and Nemenman [36]. Because the variance of the MI estimate should depend on the sample size and the particular distributions being sampled, we used the equation:

$$\sigma_{kSG}^2(N) = B/N \quad (2.7)$$

Where  $N$  refers to the total number of samples, and  $B$  is a constant that is dependent on the distributions. We randomly subsampled the data into many non-overlapping, equally sized fractions. For example, we split the data into two halves, three thirds, etc. We repeated the MI estimate for each subsample. We then averaged across all the equally sized subsamples to produce one estimate for each data fraction size (Figure 2.4 c). We considered the MI estimate to be stable if the data fractions with subsamples that contained 1/4 of the dataset or more were all within one standard deviation of each other. We also calculated the variance for all data fraction sizes with more than one subsample. We used these quantities along with the amount of data in each subsample to estimate the value of  $B$ . From this value, we were able to estimate the variance of the MI for the full data set.

The KSG estimator relies on the parameter  $k$ , the number of nearest neighbors that are used to estimate the entropies and MI. We developed a semi-automated process for selecting a  $k$ -value for each MI estimate. This procedure selected a value for  $k$  by repeating the subsampling procedure with estimates corresponding to different values of  $k$ , ranging from one to nine. The algorithm first identified all values of  $k$  which elicited a stable MI estimate (criteria for stability based on data fractions described above). Once these  $k$  values were identified, there was a check that the

MI estimates elicited by the identified  $k$  values were within one standard deviation of each other (Figure 2.4 d). If the MI values corresponding to the identified values of  $k$ , the algorithm selected the value of  $k$  that satisfied the data fraction stability criteria, maximized the stability across smaller data sizes, and that minimized the overall error estimate. The semi-automated procedure effectively selected a value for about 20% of the MI estimates without requiring user input. We manually followed the same criteria to select  $k$ -values for all MI estimates for which the algorithm was not effective. In order to do this, we constructed data fraction plots for  $k = 1 - 9$  (Figure 2.4 c). We then visually examined the plots for data fraction stability for up to four data fractions. We notated which values of  $k$  elicited MI estimates with data fraction stability. Then we constructed  $k$ -value plots (Figure 2.4 d). We determined whether the estimates corresponding to  $k$  values with data fraction stability were within one standard deviation of each other (by visually comparing error bars). To make sure our MI estimate was conservative, we selected a  $k$ -value that minimized the MI value (across the data stable  $k$ s that were identified) and minimized the error. For a small portion of MI estimates, we were unable to select a  $k$  value due to instability across data fractions or insufficient data. These estimates were small subgroups of data from our timing-timing or timing-count MI estimates (See Methods 2.3.3 below, Equation 2.9 and 2.10). In these cases, we artificially set MI to zero, assuming that these subgroups did not contribute to the coordination between the activity of motor units. Additionally, we manually selected  $k$ -values for a subset of MI estimates and compared them with those selected by the algorithm to validate our semi-automated algorithm. To validate our manual selection protocol, all lab members who contributed to the  $k$ -selection process manually selected  $k$ -values for a full analysis independently (count-count, count-timing, timing-count, and timing-timing) for the same pair of neurons, and we compared our selected  $k$ -values.

### 2.3.3 Mutual Information to Identify Coordination Between the Activity of Pairs of Motor Units

To identify how simultaneously recorded motor unit activity is coordinated, we found the MI between the spike trains from each unit. We separated our MI estimate into four components of different timescales:

$$\begin{aligned}
 I(U_1; U_2) = & I(U_{1,c}; U_{2,c}) + \langle I(U_{1,t}; U_{2,c} | U_{1,c}) \rangle \\
 & + \langle I(U_{1,c}; U_{2,t} | U_{2,c}) \rangle + \langle I(U_{1,t}; U_{2,t} | U_{1,c}, U_{2,c}) \rangle
 \end{aligned}
 \tag{2.8}$$

Here,  $U_1$  and  $U_2$  denote two different motor units, and  $c$  and  $t$  denote spike count (behavioral scale) or timing (sub-behavioral scale, Figure 2.1). To find the spike count variable for a given motor unit, we counted the total number of spikes within each breath cycle (each breath cycle was assumed to be an independent sample). Each  $U_c$  term was a  $(1 \times N)$  vector, where  $N$  is the number of breath cycles in the data set. Note that since the KSG algorithm is a continuous estimator, we added low amplitude jitter ( $\ll 1$ ) to all values to break degeneracies in spike counts. We do not expect this jitter to affect our estimates.  $I(U_{1,c}; U_{2,c})$  was therefore straightforward to estimate.

To find the spike timing variable for a given motor unit, we found the time of each spike that takes place within a breath cycle relative to the onset of that cycle. Because the spike count per cycle varied, the size of the timing variable varied in dimensionality from cycle to cycle. To accurately estimate MI, we needed the dimensionality of the timing variable to be consistent. For this reason and in order to properly separate information in spike timing from information in spike count, we separated all MI estimates that include a timing variable into subgroups with consistent count values for the timing variable(s). We found the average conditional

mutual information in spike timing given spike count. For the MI between the timing of one unit and the count of another, we used the following equation:

$$I(U_{1,t}; U_{2,c}) = \sum_{i=1}^{C_{1,max}} I(U_{1,t}; U_{2,c} | U_{1,c} = i) p'(U_{1,c} = i) \quad (2.9)$$

Where  $p'(U_{1,c} = i)$  represents the re-weighted probability that Unit 1 has a spike count of  $i$ . Since there are no spike timings for any breath cycle where a motor unit did not spike ( $i = 0$ ), we excluded all  $i = 0$  from the probability term - re-weighting the probabilities so that  $\sum_{i=1}^{C_{1,max}} p'(U_{1,c} = i) = 1$ . We therefore estimated the average MI between the spike timing of Unit 1 and the spike count of Unit 2 for all breath cycles where Unit 1 spiked at least one time.  $I(U_{1,t}; U_{2,c} | U_{1,c} = i)$  is the conditional MI between the spike timing of Unit 1 and the spike count of Unit 2 given the spike count of Unit 1. Each  $U_t$  term was an  $(i \times n_i)$  matrix, where  $i$  is the spike count and  $n_i$  is the total number of breath cycles in the data set with a spike count of  $i$ .  $C_{1,max}$  represents the maximum number of spikes within a cycle for Unit 1.

For the MI between the timing of two units, we used the following equation:

$$I(U_{1,t}; U_{2,t}) = \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} I(U_{1,t}; U_{2,t} | U_{1,c} = i, U_{2,c} = j) p'(U_{1,c} = i, U_{2,c} = j) \quad (2.10)$$

Where  $p'(U_{1,c} = i, U_{2,c} = j)$  represents the re-weighted joint probability that Unit 1 has a spike count of  $i$  and that Unit 2 has a spike count of  $j$ . Similarly to above, since there are no spike timings for any breath cycle where a motor unit did not spike ( $i = 0$  or  $j = 0$ ), we excluded all  $i = 0$  and all  $j = 0$  from the probability term - re-weighting the probabilities so that  $\sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} p'(U_{1,c} = i, U_{2,c} = j) = 1$ . We therefore estimated the average MI between the spike timing of Unit 1 and the spike timing of Unit 2 for all breath cycles where both units spiked at least one time.  $I(U_{1,t}; U_{2,t} | U_{1,c} = i, U_{2,c} = j)$  is the conditional MI between the spike timing of Unit

1 and the spike timing of Unit 2 given the spike count of Unit 1 and the spike count Unit 2. Each  $U_{1,t}$  term is an  $(i \times n_{i,j})$  matrix, where  $i$  is the spike count for Unit 1, and  $n_{i,j}$  is the total number of breath cycles in the data set where Unit 1 has a spike count of  $i$  and Unit 2 has a spike count of  $j$ . Each  $U_{2,t}$  term is a  $(j \times n_{i,j})$  matrix, where  $j$  is the spike count for Unit 2 and  $n_{i,j}$  is the total number of breath cycles where Unit 1 has a spike count of  $i$  and Unit 2 has a spike count of  $j$ .  $C_{1,max}$  and  $C_{2,max}$  represent the maximum number of spikes within a cycle for Unit 1 and Unit 2, respectively.

### 2.3.4 Error Propagation for Conditional MI Estimates

In order to propagate the error on the mutual information estimate for each weighted sum for  $I(U_{1,t}; U_{2,c})$  and  $I(U_{1,t}; U_{2,t})$  (Equation 2.9 and 2.10), we used Gaussian error propagation identities combined with maximum likelihood estimation. To find the MI, we summed the product of each conditional MI ( $I(U_{1,t}; U_{2,c} | U_{1,c} = i)$  or  $I(U_{1,t}; U_{2,t} | U_{1,c} = i, U_{2,c} = j)$ ) and each probability ( $p'(U_{1,c} = i)$  or  $p'(U_{1,c} = i, U_{2,c} = j)$ ), see Equations 2.9 and 2.10). We assumed that the MI estimate and estimate of the probability of each spike count condition were Gaussian distributed and independent. This assumption allowed us to use the Gaussian identities for error propagation (e.g.  $\sigma_{A+B} = \sigma_A + \sigma_B$ ) for the sum and product of two distributions to find the error of the overall MI estimate for mixed timescale coordination:

$$\sigma_{t,c}^2 = \sum_{i=1}^{C_{1,max}} \left( \frac{\sigma_{I_i}^2}{I_i^2} + \frac{\sigma_{p_i'}^2}{p_i'^2} \right) I_i^2 p_i'^2 \quad (2.11)$$

Here,  $\sigma_{t,c}^2$  is the variance of the mixed timescales MI estimate.  $i$  denotes each spike count condition for the unit that serves as the spike timing variable, and  $C_{1,max}$  denotes the maximum spike count for that motor unit.  $I_i$  is the conditional MI estimate for spike count  $i$ .  $\sigma_{I_i}^2$  is the variance of the conditional MI estimate for

spike count  $i$ .  $p'_i$  is the re-weighted probability of spike count  $i$  (Equation 2.13).  $\sigma_{p'_i}^2$  is the variance of the re-weighted probability, which we estimated using maximum likelihood estimation (Equation 2.14).

We followed the same procedure to find the error of the overall MI estimate for small timescale coordination:

$$\sigma_{t,t}^2 = \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} \left( \frac{\sigma_{I_{i,j}}^2}{I_{i,j}^2} + \frac{\sigma_{p'_{i,j}}^2}{p'_{i,j}{}^2} \right) I_{i,j}^2 p'_{i,j}{}^2 \quad (2.12)$$

Similarly to above,  $\sigma_{t,t}^2$  is the variance of the small timescales MI estimate.  $i$  and  $j$  denote the spike count condition for each of the two motor units in the estimate.  $C_{1,max}$  and  $C_{2,max}$  denote the maximum spike count condition for each motor unit, respectively.  $I_{i,j}$  is the conditional MI estimate for spike counts  $i$  and  $j$ .  $\sigma_{I_{i,j}}^2$  is the variance of the MI estimate for spike counts  $i$  and  $j$ .  $p'_{i,j}$  is the re-weighted joint probability of spike counts  $i$  and  $j$  (Equation 2.15).  $\sigma_{p'_{i,j}}^2$  is the variance of the re-weighted joint probability which we estimated using maximum likelihood estimation (Equation 2.16).

To find the probability terms and the variance of the probabilities, we used maximum likelihood estimation. The maximum likelihood estimate of the probability was very straightforward. For the mixed timescales estimate:

$$p'_i = \frac{n_i}{N'} \quad (2.13)$$

Here,  $p'_i$  is the probability that the motor unit that serves as the spike timing variable in the mixed timescales MI estimate exhibits a spike count of  $i$ . We note that we excluded all breath cycles with zero spikes from the mixed and small timescales analysis, so we re-weighted the probability term such that  $\sum_{i=1}^{C_{1,max}} p'_i = 1$ .  $n_i$  is the total number of breath cycles in the data set where this motor unit exhibits a spike count of  $i$ .  $N'$  is the total number of breath cycles where the unit of interest spiked at least

one time in the experiment.

From this estimate, we were able to estimate the variance of the re-weighted probability as:

$$\frac{\sigma_{p_i}^2}{p_i'^2} = \frac{(1 - p_i')}{n_i p_i'} \quad (2.14)$$

Here  $\sigma_{p_i}^2$  is the variance of the probability that the motor unit serving as the spike timing variable exhibits a spike count of  $i$ . The other variables are the same as above.

Similarly, the small timescales re-weighted joint probability term was estimated as:

$$p'_{i,j} = \frac{n_{i,j}}{N'} \quad (2.15)$$

Here,  $p'_{i,j}$  is the re-weighted joint probability that the each motor unit has a spike count of  $i$  and  $j$ , respectively. Again, we note that we excluded all breath cycles where the spike timing units exhibited zero spikes from the mixed and small timescales analysis. Therefore, we re-weighted the probability terms such that  $\sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} p'_{i,j} = 1$ .  $n_{i,j}$  is the total number of breath cycles in the data set where each motor unit exhibits a spike count of  $i$  and  $j$  respectively.  $N'$  is the total number of breath cycles in the experiment in which each motor unit spikes at least one time.

From this estimate, we were able to estimate the variance of the probability as:

$$\frac{\sigma_{p_{i,j}}^2}{p_{i,j}'^2} = \frac{(1 - p'_{i,j})}{n_{i,j} p'_{i,j}} \quad (2.16)$$

Here  $\sigma_{p_{i,j}}^2$  is the variance of the probability that the motor units exhibit spike counts of  $i$  and  $j$ , respectively. The other variables are the same as above. Although re-weighting the probability term did affect the estimate of the error, since the individual probabilities increased from re-weighting them, the error estimate increased as well. Therefore, we estimated an upperbound on the error. Because of this, we

are confident that we did not artificially inflate the significance of our results. For additional details see the Appendix.

# Chapter 3

## Results

We simultaneously recorded single unit EMG from multiple motor units from the EXP muscle group in addition to air sac pressure during breathing in an anesthetized Bengalese Finch. We isolated spike waveforms from single units, and segmented the EMG data into breath cycles. We considered how the spike patterns of different motor units coordinated with each other- at the timescale of the behavior (spike counts), at sub-behavioral timescales (spike timings), and at a mixed timescales (spike timing and spike count) both cumulatively across the entire data collection and within smaller time epochs. The spiking activity of each of our motor units varied greatly throughout the data collection. We found that all pairs of motor unit exhibited coordination between spike counts. Some pairs of motor units exhibited coordination between spike timings and spike counts. None of the motor units exhibited coordination between their spike timings. Finally, we examined how the coordination between spike counts of motor units changed from the first half of the data collection to the second half, and we found that the coordination between the example pairs of units exhibited differences between the first and second halves of the data collection.

We used a quantity called mutual information, a measure of correlation that considers both linear and non-linear correlation to study the coordination between motor

units across timescales. We segmented the pressure data into cycles, which were defined as the time from one positive zero crossing in the pressure waveform to the next positive zero crossing. We created a spike count variable for each motor unit: the total number of spikes within each breath cycle. We also identified the precise spike timings of each motor unit relative to the onset of each breath cycle. We found the mutual information between the spike counts of all pairwise combinations of motor units, between the spike counts of one motor unit and the spike timings of another motor unit for all pairwise combinations of motor units, and between the spike timings of all pairwise combinations of motor units. Additionally, we repeated the count-count MI analysis for a three example pairs of motor units for each consecutive half of the data collection separately to examine the stationarity of the data in addition to examine how the coordination between units changed between halves.

### **3.1 Motor units exhibit unique spiking properties that vary throughout the data collection.**

Throughout our data collection, we identified four motor units with distinct firing characteristics (Figure 3.1). Three units exhibited sparse firing at different epochs throughout the data collection: Unit D, Unit G, and Unit E. Unit D fired zero to nine spikes per cycle during the first third of the data collection. Unit G fired zero to seven spikes per cycle on and off throughout the entire data collection. Unit E fired zero to 18 spikes per cycle at the end of the data collection. In addition to the sparse units, we also recorded from one tonic unit, which fired many spikes per cycle throughout the data collection: Unit B. This unit fired zero to 44 spikes per cycle throughout the entire data collection (Figure 3.1).

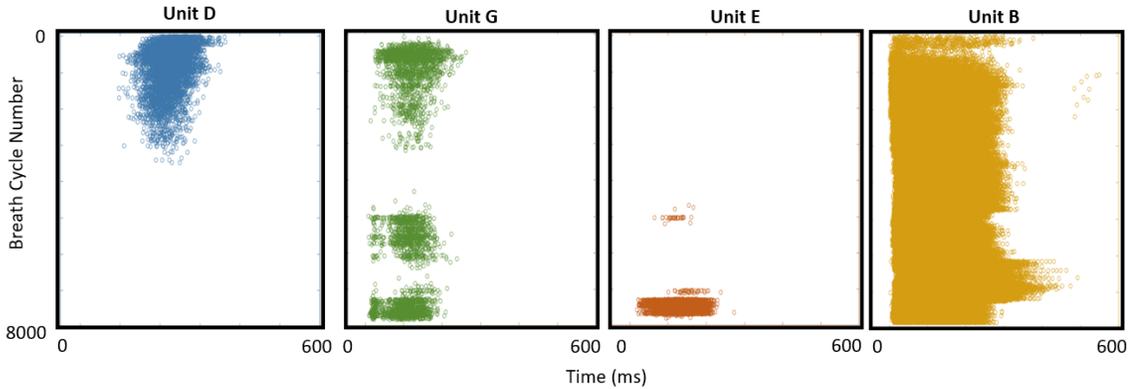


Figure 3.1: Four different motor units exhibit distinct spiking characteristics throughout the data collection time. Unit D (blue) exhibited robust spiking throughout the beginning of the data collection, and then stopped. Unit G (green) exhibited sparse spiking throughout the entire data collection. Unit E (orange) exhibited sparse spiking specifically towards the end of the data collection. Finally, Unit B (yellow) exhibited robust tonic spiking throughout the entire data collection.

### 3.2 Spike patterns of different motor units are coordinated in their spike counts.

We first asked whether the motor units we recorded exhibited coordination in their spike counts when we collectively considered the entire data collection. We define spike counts to be the total number of spikes produced by each motor unit within a breath cycle (Figure 2.1, Equation 2.8). The marginal distributions of spike counts for the sparse units (D, G, and E) are relatively similar to each other (Figure 3.2 a, blue, red, and green traces). All of the sparse units have zero spikes for most breath cycles. Unit E has the widest spread of different spike counts of the sparse units, yet this unit spikes for the fewest number of cycles. The tonic unit, Unit B, has the widest overall spread of different spike counts, ranging from zero to 44 spikes per cycle. Unit B also has fewest cycles with zero spikes. The example joint distributions illustrate three distinct cases of coordination at the timescale of the behavior that occurred cumulatively throughout the duration of the data collection (Figure 3.2 c). The joint distribution of Units D and G illustrates the coordination between the two sparse

units that were active throughout the most breath cycles. The joint distribution of Units D and E illustrates that units, which never fired simultaneously, exhibited coordination when their spike patterns throughout the whole data collection were considered. Finally, the joint distribution of Units B and G illustrate the coordination between the tonic unit (Unit B) and a sparse unit, which were both active throughout most of the data collection.

We estimated the mutual information between the spike counts for each pair of units as a measure of their coordination at the timescale of the behavior. The count-count MI estimates depend on estimates of the marginal and joint entropies, which depend on the marginal and joint distributions of spike counts (Methods: Overview of Mutual Information and the KSG Method, Figure 2.4). Consistent with what we would expect based on the marginal and joint distributions, we found that all pairs of motor units were coordinated in their spike counts - all the count-count MI values were positive and of larger magnitude than the standard deviations (Figure 3.2 b). Across each unit pair, the total MI ranged from about  $0.03 \pm 0.01$  –  $0.35 \pm 0.01$ . Of the three example unit pairs, the most active units (Units B and G) had the highest MI, and the two units which were never active simultaneously (Units D and E) had the lowest MI and the MI that was the closest to being within error of zero (Figure 3.1, 3.2 b and c).

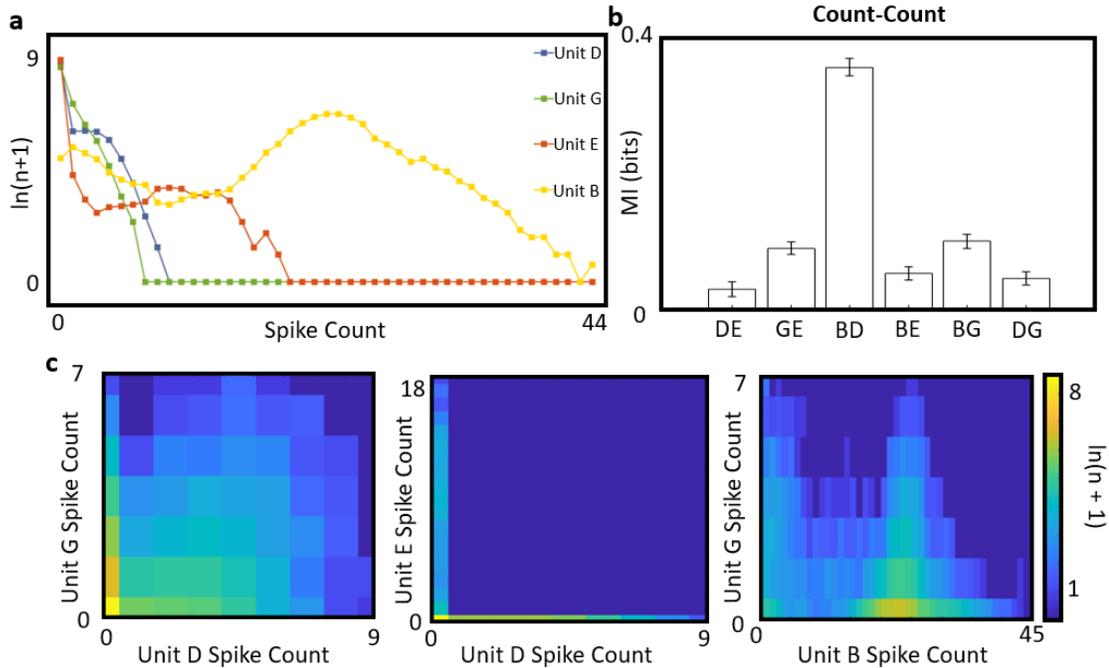


Figure 3.2: Motor units are coordinated at the timescale of the behavior. **a.** Histograms of spike count for each unit illustrate the spike count distributions for the entire data collection. Mutual information was found to be greater than zero for all pairs of units. **b.** Count-count MI values for all unit pairs are shown. Error bars show the estimated standard deviation the MI **c.** Joint histograms of spike counts for each example pair illustrate the joint distributions of spike count across the whole data collection.

### 3.3 The spike timings of some motor units may be coordinated with the spike counts of others, but no coordination between the spike timings of pairs of motor units was detected.

We next asked whether motor units exhibit coordination at mixed or sub-behavioral timescales. Since we already captured the count-count MI, we specifically examined the timing-count and timing-timing MI of all pairwise combinations of motor units (Figure 2.1). For the timing-count MI, we separated out the contribution of spike

count from the spike timing variable by finding a weighted sum of the conditional mutual information between spike timing and spike count for all given spike counts of the timing unit (Equation 2.9) [77]. Similarly, for the timing-timing MI, we separated out the contribution of spike count from both spike timing variables by finding a weighted sum of the conditional mutual information between the spike timings of each motor unit given each spike count value for each pair of units (Equation 2.10). We limited both the timing-count and the timing-timing analyses to consider only breath cycles where the spike timing unit(s) spiked at least one time. Therefore, we re-weighted the probability term to discard these breath cycles where the timing unit had a spike count of zero (see Methods: 2.3.3). In doing so, we found the coordination between motor units at mixed and sub-behavioral timescales given that the spike timing unit(s) spike(s) at least one time in a breath cycle and thereby estimated an upper bound on the error (see Appendix).

We examined a few example distributions to see if there was any qualitative coordination that could be observed at mixed and sub-behavioral timescales, which might be indicative of non-zero MI at these timescales. We selected example marginal and joint histograms to visualize the structure in the two example pairs of units from Figure 3.2 that spike simultaneously. The spike times of each unit vary for the example spike count values (Figure 3.3 a). The example joint distributions for the relationship between spike timing and spike count seem to exhibit coordinated structure (Figure 3.3 b). This structure is particularly of note in the joint distribution of Unit B spike timing and Unit G spike count. The final spike of Unit B becomes more and more spread out and later in time as the spike count of Unit G decreases. A similar trend seems to exist for the Unit D spike timing and Unit G spike count joint histogram (Figure 3.3 b). The trends for the timing-count joint distributions between units D and G are less obvious. Unit G seems to have an early spike time when Unit D has a lower spike count and when Unit B has a higher spike count. The example

timing-timing joint distributions vary in whether they exhibit qualitative correlation (Figure 3.3 c). There seems to be a positive correlation at the sub-behavioral scale between the spike timings of Unit D and the spike timings of Unit G when Unit D spikes twice and Unit G spikes once (Figure 3.3 c). When Unit B spikes 23 times and Unit G spikes once, Unit G spike times closer to average seem to be correlated with a later last spike in Unit B (Figure 3.3 c).

The total MI estimates for mixed and sub-behavioral timescales depend on the conditional marginal entropies (and therefore conditional marginal distributions) of spike patterns and the conditional joint entropies (and therefore conditional joint distributions). Qualitative correlation may be reflected in the respective MIs, although because many count-conditioned joint distributions made up each timing-count and timing-timing MI value, it may be harder to predict the MI results from visualizing the joint distributions. We found that some of the motor unit pairs exhibited coordination between spike timing and spike count. Eight out of the 12 timing-count MIs were larger in magnitude than the upper bound on the standard deviation of the MIs (Figure 3.3 d). However, all but two pairs (the timing of Unit B and the counts of Units D and G) were within two standard deviations of zero. The timing-count MIs for the spike timing of Unit B and the spike count of each sparse unit were not within error of zero - which supports mixed timescale coordination for the tonic unit. These timing-count MIs where Unit B served as the timing unit ranged from  $0.02 \pm 0.01 - 0.10 \pm 0.03$ . We note that significantly more breath cycles were able to be included in the analysis for the timing-count MIs where Unit B served as the timing unit. It is possible that these MI values were the least impacted by limited data, thereby providing a more accurate representation of mixed timescale coordination (see Discussion). All of the timing-timing MIs were within one standard deviation of zero with the exception of Units D and E and Units B and E for which there was not enough data in any count-conditioned subgroup to estimate the MI

(Figure 3.3 e). Although we see some qualitative correlation in the example joint histograms, based on our MIs, we did not detect a robust correlation across all unit pairs for timing-count or timing-timing.

### **3.4 The structure of motor unit coordination may change throughout the data collection.**

Because the spike patterns of each motor unit varied drastically throughout the the data collection (Figure 3.1), we were concerned that the data may violate a key assumption to MI and many other common analyses: stationarity in the marginal and joint distributions. In other words, we wanted to measure whether the distributions of the spike patterns of each motor unit or the joint distributions of their spike patterns changed across the duration of the data collection. If there were changes to these variables, it would impact the interpretation of our results. In order to assess the stationarity of the data, we decided to assess how the coordination between the units changed from the first half of the data collection to the second half. We split the data into two equally sized, consecutive halves and repeated the behavioral timescale coordination analyses for each half with the example unit pairs from Figure 3.2. We chose these three unit pairs for Figures 3.2 and 3.4 because each illustrates a different type of possible temporal structure to coordination from the first half to the second half of the data collection.

Each of the four units and the three individual pairings that we examined exhibited qualitative changes in their distributions from the first half of the data collection to the second half - indicative of nonstationarity in the data. Each of the sparse units spiked during distinct epochs throughout the data collection (Figure 3.1). Largely due to this variability, the marginal distributions of Unit D and Unit E changed from the first half of the data collection to the second. In contrast, the marginal distribution

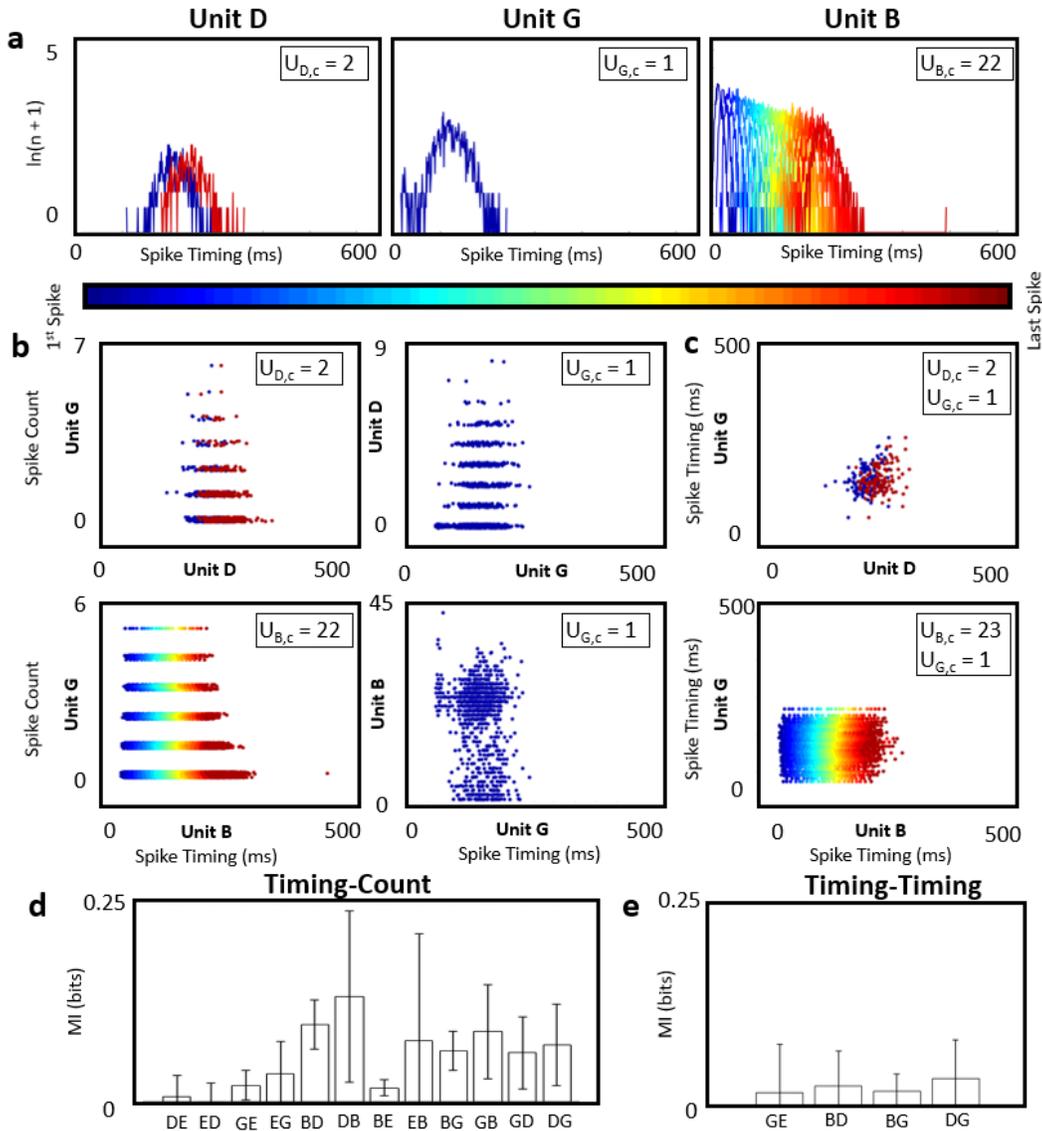


Figure 3.3: Motor units may be coordinated at the sub-behavioral scale. **a.** Example histograms illustrating the distribution of spike timing from three example units are shown, each for one spike count condition. The distributions of spike timings for each unit for all breath cycles where Unit D spikes two times, where Unit G spikes once, and where Unit B spikes 22 times are shown. **b.** Joint distributions for mixed timescale coordination: spike timing and spike count are illustrated for one spike count condition for two example unit pairs: Units D and G and Units B and G. **c.** Additionally, the joint distributions for sub-behavioral scale coordination are illustrated for Units D and G for all breath cycles in which Unit D spiked twice and Unit G spiked once. The sub-behavioral timescales joint distribution is shown for Units B and G for all breath cycles in which Unit B spiked 23 times and Unit G spiked once. **d,e.** The weighted sum of mutual information for mixed timescales and for sub-behavioral timescales for all spike count conditions between Units D and G and Units B and G are shown.

of the other sparse unit, Unit G was qualitatively similar between the first half and the second half (Figure 3.4). The tonic unit, Unit B, spiked consistently throughout the data collection (Figure 3.1), yet, this unit exhibited lower spike counts in the first half of the data collection compared with the second half (Figure 3.4 a and b).

We wanted to capture a range of possible variabilities in the joint distributions of a few pairs of units to see how various distribution changes impact the MIs across consecutive halves. We chose to study the coordination between the spike counts of the following unit pairs: Units D and G - the two sparse units which spiked for the highest number of breath cycles, units D and E - two sparse units which never spiked simultaneously, and units B and G - a tonic unit and a sparse unit, both of which spiked throughout the duration of data collection.

The coordination between units D and G differs from the full data collection to each half and from the first half to the second half. Unit D does not spike at all during the second half of the data collection (Figure 3.1). Based on the joint distributions of spike counts of units D and G for each half of the data collection, we would expect to find zero MI between the spike counts of Unit D and Unit G for the second half of the data collection because there was no variability in the spike counts of Unit D for the this epoch (Figure 3.4 e and f). As expected, there is zero MI between the spike counts of units D and G when the second half of the data collection is considered alone (Figure 3.4c). Similarly, due to the structure in the joint distribution of spike counts of units D and G for the first half of the data collection, we would expect to find non-zero MI between the spike counts of these units for this half, and we do. Interestingly, the count-count MI for units D and G for the first half of the data collection is roughly three times the magnitude of the count-count MI for these units when the full data set was considered (Figure 3.4). Although the sparse units exhibit variability in the breath cycles when they are active, the shift in MI from the first half of the data collection to the second half illustrates that the full data set MI was

likely impacted by nonstationarities in the distributions, but that during an epoch when both units were simultaneously active, they were highly coordinated.

The coordination between units D and E exhibits differences from the full data set to each consecutive half, but is consistent from the first half to the second half. As stated above, Unit D does not spike during the second half of the data collection. In contrast, Unit E does not spike during the first half of the data collection (Figure 3.1). Unit D and Unit E have non-zero MI for the full data set. However, due to the lack of variability for at least one of the units illustrated in the each of joint distributions of spike counts of units D and E for each half of the data collection, we would expect to find zero count-count MI for these units for each consecutive half (Figure 3.4 e and f). As expected, we find zero count-count MI for these units for each consecutive half of the data collection. The contrasting count-count MIs for this pair of units in the full data set versus each consecutive half further illustrates that our results were impacted by nonstationarities in the distributions and that the non-zero MI in the full data set likely reflected which units were active when rather than coordination between varying spike counts. Finally, the coordination between units B and G changes from the full data set to each consecutive half and from the first half to the second half. As stated above, both units B and G spike throughout the duration of the data collection. Qualitatively, the joint distribution of spike counts for Unit B and Unit G for the first half of the data collection is more similar to the joint distribution for the full data set (Figures 3.2 c and 3.4 e). For this reason, we might expect to find comparable MI values for these two units in the full data set and the first half of the data collection. Surprisingly, the count-count MI for units B and G roughly doubled when the first half of the data collection is considered alone. In contrast, the joint distribution of spike counts for units B and G is qualitatively different from that of the full data set. Despite this difference, the count-count MI for units B and G for the second half of the data collection is within error of the

count-count MI between these units for the full data set. The differences in the joint distributions and MI values for units B and G highlights how units with qualitatively similar joint distributions can exhibit differences in MIs and units with qualitatively different joint distributions can be coordinated by consistent amounts. Additionally, these results illustrate that the coordination we found between the spike counts of units B and G for the full data set were impacted by non-stationarities - examining the data collection in consecutive halves reveals interesting detail about how this coordination changes across the data collection.

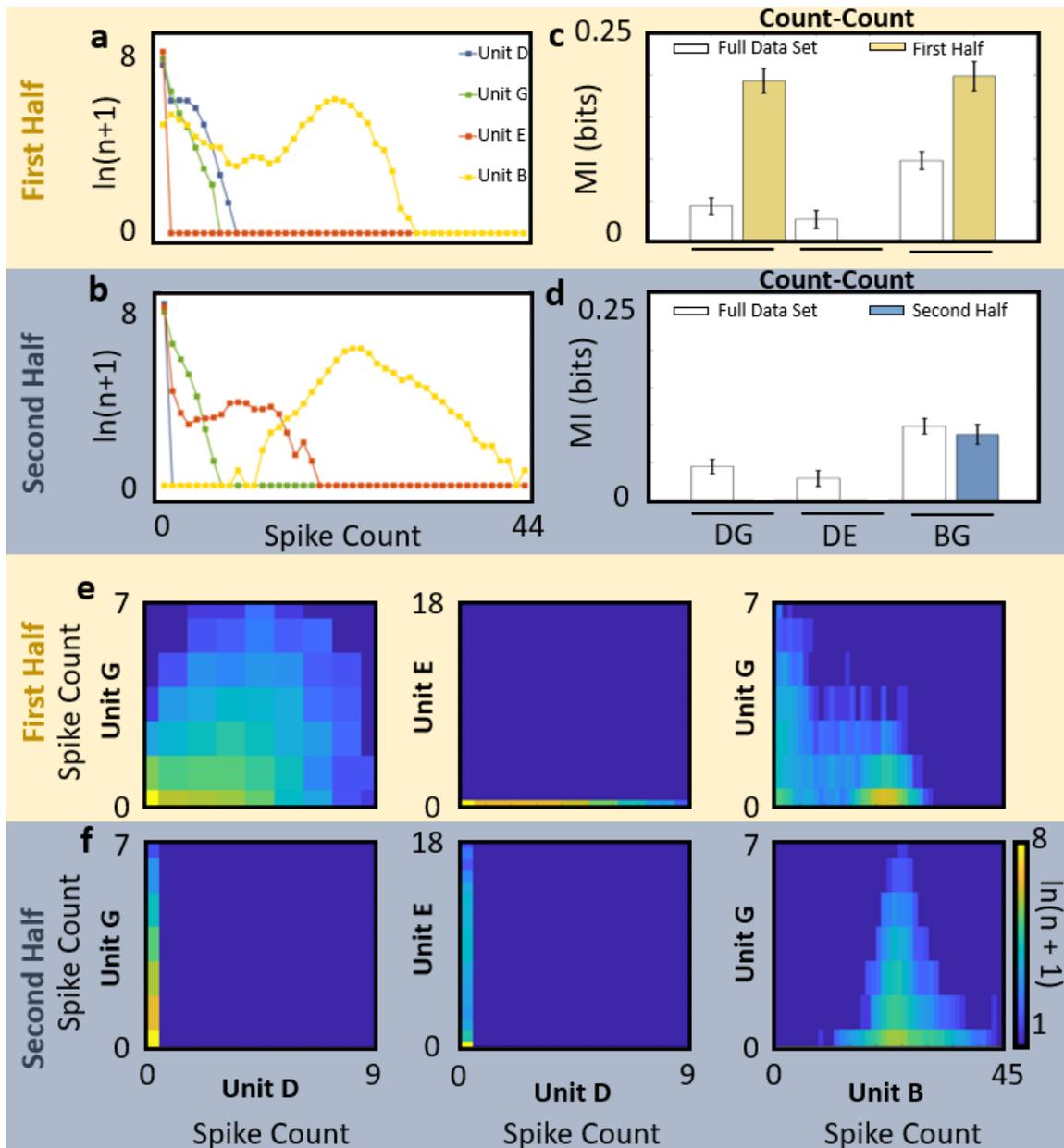


Figure 3.4: The coordination between motor units changes between consecutive halves of the data collection. The data were split into two consecutive halves and the behavioral scale coordination analyses were repeated for each half. **a** and **b**. The marginal histograms for each unit are shown for the first half (**a**) and the second half (**b**) of the data set. In particular, the marginal distributions of Unit D, Unit E, and Unit B change from the first half to the second half. **c** and **d**. The mutual information estimates for the same example pairs of units as in Figure 2 are shown for each data half (**c** and **d**, respectively). Each example unit pair illustrates a different changes in coordination when the data set is split. **e** and **f**. Joint distributions of spike count for each example pair are shown for each half (**e** and **f**, respectively).

# Chapter 4

## Discussion

Behavioral timescale changes to spike patterns (rate coding) have been well studied and accepted as mechanisms of motor control [40, 27, 71, 43, 19, 12, 61, 18, 50]. A growing body of evidence supports that sub-behavioral timescale spike patterns, in motor areas ranging from neurons in motor cortex to single motor units in muscle can affect motor control [82, 77, 64, 74, 73]. Overall, we found that the motor units we recorded exhibited coordination at the timescale of the behavior (Figure 3.2), and that the structure of this coordination changed throughout the experiment (Figure 3.4). We found that some pairs of units exhibit evidence of coordination at sub-behavioral timescales (timing-count), but the unit pairs were too data limited to draw robust conclusions for timing-timing MI (Figure 3.3).

### **4.1 Motor unit coordination at sub-behavioral and mixed timescales**

We found robust evidence of coordination at the timescale of the behavior (count-count MI) for all pairs of motor units (Figure 3.2). Although a growing body of evidence supports that motor units control behavior using precisely timed spikes

[74, 82, 77, 64], we found that only some pairs of motor units exhibited non-zero MI in timing-count, our mixed timescales coordination metric. For these MI estimates, 2/3 of the unit pairs exhibited MI values that were higher in amplitude than the standard deviation. However, most of these values were so low that they were within two standard deviations of zero. For the sub-behavioral timescales coordination metric, or the timing-timing MI, we found that all of the MIs were within one standard deviation of zero, except for two unit pairs for which there was insufficient data to complete the analysis. There are two potential reasons that we found such variability in our results for coordination between motor units at mixed and sub-behavioral timescales. (1) These motor units are coordinated at mixed and/or sub-behavioral timescales, but we were unable to detect this coordination. (2) These motor units are not coordinated at mixed and/or sub-behavioral timescales, and our methods successfully identified a substantial proportion of motor unit pairs with MIs within error of zero. We are confident that our analyses were too data limited to draw conclusions, and more work will be needed to distinguish between these two cases for coordination at mixed and sub-behavioral timescales. We present arguments for and against each case below.

#### **4.1.1 Case 1: These motor units are coordinated at mixed and/or sub-behavioral timescales, but we were unable to detect enough coordination.**

We believe that our variable results for mixed and sub-behavioral coordination at least partially reflects a data size insufficiency. Largely due the sparse firing of most of the motor units we recorded, our mixed and sub-behavioral timescale analyses became severely data-limited. The mixed and sub-behavioral timescales MIs consisted of weighted sums of MI values across spike count conditions (Equation 2.9, 2.10). For these MI estimates, we included only data where each timing unit spiked at least one

time. Therefore, we re-weighted the probability term to discard these breath cycles where the timing unit(s) had a spike count of zero (see Methods: 2.3.3). Large percentages (up to 94% of data) for almost every mixed and sub-behavioral timescales MI estimate were omitted from the analyses because of zero spike counts alone. Additionally, we artificially set MI to zero for certain spike count conditioned subgroups of data (terms corresponding to specific  $i$  and/or  $j$  values in equations 2.9 and 2.10) for which we could not estimate MI because of insufficient samples or unstable dependence on the parameter  $k$ , which further reduced the number of samples that contributed to our MI values. The two mixed timescale MIs found to be of a reasonable, nonzero magnitude (over two standard deviations away from zero) quantified the coordination between the counts of the the two sparse units that spiked the most frequently throughout the experiment and the timing of the tonic unit (Unit B, the unit for which the least amount of data were omitted). For each of these MIs, less than 2% of the data was omitted because of zero spike counts, and MI was artificially set to zero for only up to 4% of the remaining data (Figure 4.1). In contrast, the sparse units had spike counts of zero for 71% – 94% of breath cycles. So we omitted these percentages of data from the mixed timescales MIs where each of the sparse units served as the spike timing variable. Additionally, we artificially set MI to zero for 0.4% – 22% of the subset of data with at least one spike in the timing variable (setting MI to zero) for the mixed timescales MIs for which a sparse unit served as the timing variable (reasons described above). Figure 4.1 illustrates the dependence of the MI and the error on the percentage of breath cycles that were included in the analysis. Although there does not seem to be a relationship for the MI values, the standard deviation of the MI estimate does seem to be heavily dependent on the amount of data included. The mixed timescales MIs with the lowest error were the same MIs where the least (by a large margin) amount of data had to be omitted or artificially set to zero. The data size for the sub-behavioral timescales MIs was much worse -

we omitted 72% – 95% of the data because at least one of the units had zero spikes, and we artificially set MI to zero for 26% – 93% of the subset of data with at least one spike in both variables. In fact, for two pairs of units sub-behavioral timescales MI could not be estimated at all because of insufficient data. It is therefore possible that the extent to which our methods were able to detect mixed and small timescales coordination between motor units was artificially decreased by the small amount of the data that was analyzable for most of the mixed and sub-behavioral MIs. In fact, whether or not the units are coordinated at mixed and/or sub-behavioral timescales, we are confident that the sub-behavioral timescales MIs were data-limited. The mixed timescales analysis yielded MI values that were comparable in magnitude to the MI between a single unit and the breathing behavior in our previously published work [77]. However, our error bars were much larger. It is therefore possible, that having more analyzable data for the mixed and sub-behavioral timescales analyses would give us more certainty in the magnitude of the MI between the spike patterns of each unit pair. Our work here focused on just four motor units recorded from a single muscle in a single songbird. Many more data sets may help to resolve this limitation. However, in order for the future data sets to resolve the data limitation, we will need recordings from neurons that exhibit spiking activity during many more breath cycles (e.g. more tonic units) or much longer recordings, both of which will be technically challenging to acquire.

As stated above, our previously published work by Srivastava et al. found that single motor units share approximately  $0.05 \pm 0.01$  bits of information with the behavior (estimated from Figure 2A) [77]. This study was done in the same species, muscle, and behavior as our study. Our MI values for mixed timescales coordination ranged from  $0.02 \pm 0.01 - 0.1 \pm 0.03$  for the analyses where Unit B served as the timing unit (the MIs that were the least data limited). The mixed timescales MIs for all other unit pairs were of similar orders of magnitude. At face value, the

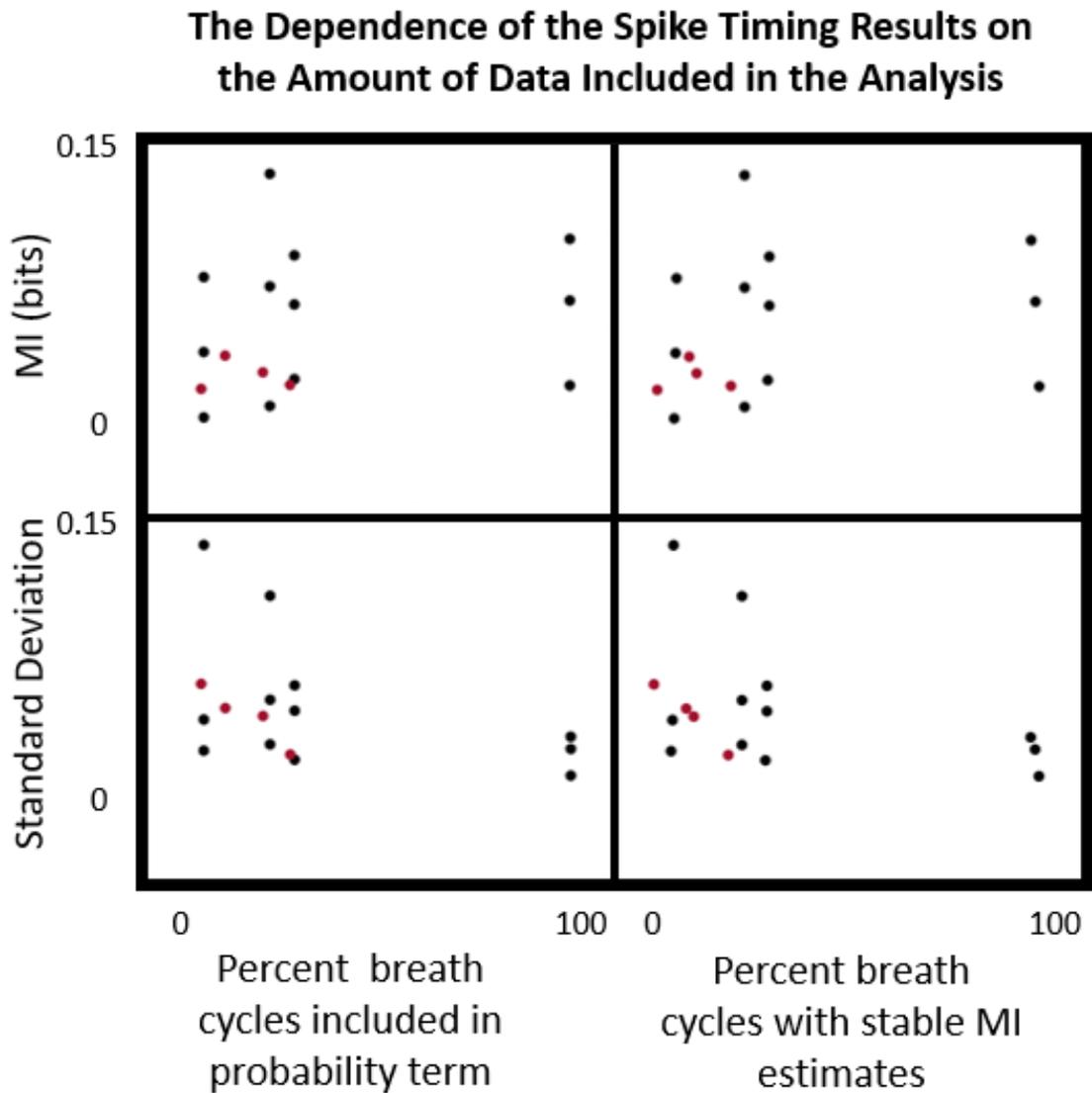


Figure 4.1: Our timing-count and timing-timing MI results appear to depend on the amount of data that was able to be included in the analysis. The top panel illustrates the dependence of the each timing-count and timing-timing MI on the percentage of data included in the probability term (all breath cycles with spike counts of one or higher for all timing units) and the percentage of breath cycles that yielded stable MI estimates (i.e. the percentage of breath cycles that was not omitted for any reason including zero spikes, dimensionality, data insufficiency, or instability across parameter values). The bottom panel illustrates the dependence of the error of each timing-count and timing-timing MI on the percentage of breath cycles included in the probability term and the percentage of breath cycles that yielded stable MI estimates. Timing-count unit pairs are shown in black, and timing-timing unit pairs are shown in red.

magnitude of our mixed timescales MIs seems low. However, Srivastava et al. found that individual single motor units share about the same amount of information with the air sac pressure, and we know that these motor units can play a causal role in modulating the air sac pressure during breathing [77].

In this context, new questions arise about the role of mixed timescale coordination between the motor units: Are the motor units independent predictors of behavior, coordinated only via shared noise? Do the spikes share the same information with each other as they share with the behavior? More analysis would be necessary to answer these questions. Estimating a quantity called interaction information for the spike patterns of two motor units (the spike timing of one and the spike count of another) and the behavior would determine whether the units are net redundant in how they predict the motor behavior [84, 64]. The interaction information can be calculated as follows [84]:

$$II = I(U_1, U_2; B) - [I(U_1; B) + I(U_2; B)] \quad (4.1)$$

Where  $II$  represents the interaction information,  $B$  represents the behavior.  $I(U_1, U_2; B)$  is the joint mutual information between the spike patterns of Unit 1 and Unit 2 and the behavior.  $I(U_1; B)$  is the mutual information between the spike patterns of Unit 1 and the behavior, and  $I(U_2; B)$  is the mutual information between the spike patterns of Unit 2 and the behavior. Similarly to our methods, one could consider only the spike patterns of each unit at the timescale of the behavior by only considering spike counts in these equations. Additionally, one could find interaction information for mixed timescales or sub-behavioral timescales by inputting equations 2.9 and 2.10 for the two mutual information terms as appropriate and by similarly expanding the joint mutual information terms into a weighted sum. If the interaction information is negative, then the units are net redundant, meaning that the two units

together share less information with the behavior than the sum of the individual MIs. This would indicate that net redundant units would share at least some of the same information with each other as they share with the behavior — that at least some of the mutual information between the units is behaviorally relevant.

An example of an interaction information analysis can be found in work from Putney et al [64]. In this study, tethered, flying hawk moths tracked an oscillating artificial flower while the ten major flight muscles were recorded via EMGs and the yaw torque produced by the moth was recorded. The researchers conducted an interaction information analysis for all pairwise combinations of 10 hawk moth flight muscles [64]. They examined interaction information for behavioral timescale spike patterns (spike count) and for sub-behavioral timescale spike patterns (spike timings) separately. They did not examine interaction information for mixed timescale spike patterns. Pairwise combinations of hawk moth muscle spike patterns were found to be net redundant primarily at the sub-behavioral timescale. This indicates that pairwise combinations of spike timing patterns of moth muscles are correlated with each other to some degree - the spike timings of these muscles would exhibit non-zero mutual information, and the information that they share with each other is shared with the behavior as well [64]. This analysis in combination with the analysis that we have done would help to determine whether our motor units share behaviorally relevant information.

Consistent with Case 1, it is also possible that the mixed and sub-behavioral timescale coordination exhibited by motor units is more complex than our methods could detect. For example, the joint distributions in Figure 3.3 b and c exhibit qualitative correlation. Of particular note is the joint distribution that illustrates the timing of Unit B versus the spike count of Unit G for all breath cycles where Unit B has a spike count of 22 (Figure 3.3 b). The spike timings of Unit B seem to spread out in time as the spike count of Unit G decreases. It is possible that mixed and/or sub-

behavioral coordination between motor units occurs for certain spike count conditions (terms corresponding to specific  $i$  and/or  $j$  values in equations 2.9 and 2.10) and not others. If this were the case, large MI values corresponding to individual count-conditioned subgroups might have become muted in the weighted sum (equations 2.9 and 2.10). For this reason, we examined what proportion of count-conditioned subgroups had MI estimates over twice the magnitude of the error on the estimate. Across all of the mixed timescale pairwise estimates, these proportions ranged from 0% to 80%, and there did not seem to be a relationship between the proportion of MI subgroups with MI over two standard deviations above zero and whether the overall MI was found to be larger than error. Unfortunately these proportions seem to leave us with more questions than answers with regard to whether individual count-conditioned subgroups exhibit mixed timescales coordination. It would be very surprising and interesting if mixed and/or sub-behavioral timescales coordination between two motor units only occurs for specific spike count conditions depending on the motor units. Future work may examine subgroups with non-zero MI more in detail in order to investigate case-specific mixed and sub-cycle coordination.

#### **4.1.2 Case 2: Lack of coordination at mixed and subcycle timescales**

We believe that our variable results for mixed and sub-behavioral coordination reflects a data insufficiency, yet it is possible that our results at least partially reflect genuine absence of coordination at these timescales for these motor units. Although we had to discard and set MI to zero for large percentages of data from the mixed and sub-behavioral timescales analyses, in particular for the mixed timescales estimates, most of the data that was analyzable yielded stable MI estimates. As mentioned above, for the mixed timescales estimates, only 0.4% – 22% of data was artificially set to zero because of insufficient samples or instability whereas for the sub-behavioral

timescales estimates 26% – 93% of data was artificially set to zero. The fact that over 77% of the analyzable data produced stable mixed timescales MI estimates tells us that for most of the data where the timing unit spikes at least one time, a data limitation did not prevent us from accurately estimating the mutual information. Even still, the magnitude of the MI for all but two pairs of units was within two standard deviations of zero, and all of the mixed timescales MI estimates were within three standard deviations of zero. Therefore, it would be reasonably likely for us to estimate those MI values by chance even if there truly is zero MI. We are confident that our MI estimates for sub-behavioral timescales coordination were impacted by insufficient data. For this reason, our sub-behavioral timescale results are inconclusive - regardless of whether the units we recorded are truly coordinated at sub-behavioral timescales, we were unable to detect coordination. That being said, because the mixed timescales MI estimates were less data-limited, it is possible that our low magnitude MIs for mixed timescales analyses reflects a genuine lack of coordination at mixed timescales.

Motor units from EXP have previously been shown to control behavior via sub-behavioral changes to spike patterns. Therefore, a lack of coordination at the sub-behavioral and mixed timescales could indicate that the nervous system may employ separate information channels for motor control: one by which actuators control behavior (spike count and spike timing) and another by which multiple actuators coordinate together (spike count alone). This would contrast with the interaction information results from Putney et al. described above. The magnitude of the behavioral timescale interaction information for the moth was close to zero, which would be expected if the behavioral timescale spike patterns act as independent channels [64]. However, in contrast with our methods, the interaction information does not capture all possible coordination between units - it focuses solely on the behaviorally relevant coordination. For a fair comparison of results, we would need to find the

mutual information between pairwise combinations of moth muscles at behavioral, mixed, and sub-behavioral timescales. Our method of measuring coordination captures both the behaviorally relevant component of coordination in addition to any coordinated noise or other variability that impacts both motor units [26]. It is possible that Putney et al. would have found more coordination at the timescale of the behavior if they had considered both the behaviorally relevant and noise components of coordination between muscles [64]. It also is possible that different organisms employ different strategies for coordination - songbirds may employ primarily spike patterns at the timescale of the behavior (spike counts) for coordination whereas hawk moths may employ primarily spike patterns at sub-behavioral timescales, even down to the millisecond. Alternatively, this discrepancy may reflect different mechanisms of coordination between motor units in different muscles versus within the same muscle.

## 4.2 Motor unit coordination across timescales

Our findings that motor unit pairs are coordinated in their spike counts may be consistent with existing hypotheses about motor control and coordination. Muscle force is thought to be modulated via changes in motor unit recruitment and via changes to spike rates of individual motor units [18, 50]. The size principle is a well known hypothesis governing motor unit recruitment. According to this principle, motor neurons are recruited in order of cell body size, which tends to correlate with the force production capabilities of the motor units. When the maximum force producible by smaller units is achieved, the larger units are recruited to produce additional force [41, 32, 33, 35]. The size principle states that smaller units begin spiking before and cease spiking after larger motor units for a given synaptic drive [32]. Therefore, if we assume that the spike rate of the smaller unit is greater than or equal to the spike rate of the larger motor unit during a movement when two units are simultaneously active,

according to the size principle, we would expect to find non-zero MI between the spike counts of motor units. Therefore, our behavioral timescale MI results that considered the full data collection (Figure 3.2) may support these hypotheses of motor control. Moreover, it is possible that the sparse units that we recorded (Figure 3.1) corresponded to larger motor neurons and the tonic unit corresponded to a smaller motor neuron. If that is the case, the sparse spiking activity that we observed throughout the data collection may further support the size principle - larger motor neurons activating as needed for increased muscle force. However, it would be possible for other patterns of motor unit recruitment that contradict the size principle to result in non-zero MI between spike counts. More motor unit recordings combined with precise histology would be necessary to differentiate between motor unit types in these experiments. Additionally, more detailed examination of the onset and offset times of motor unit bursts within each breath cycle could examine whether there is variability in the recruitment patterns of the motor units [48].

Because of the relationship between inter-spike-intervals and spike rate, coordination at mixed timescales may support the hypothesis that muscle force is controlled via changes to spike rates [18, 50]. The joint histogram of spike counts from Unit G and spike timings from Unit B in Figure 3.3 b provides an example of qualitative results that may support this hypothesis. It appears that the average inter-spike-interval for Unit B decreases (spike rate increases) as the total spike count of Unit G increases. This inter-spike-interval trend implies that the spike rate of Unit B may be correlated with the spike count of Unit G. It is possible that the spike rate (with constant spike count) of Unit B increased during the same breath cycles when spike counts of Unit G were increased in order to increase the overall force production of the EXP group.

To our knowledge, previous work has not examined multiple single motor units within a muscle at the resolution necessary to examine coordination via spike tim-

ings. Although muscles have previously been shown to coordinate at behavioral or intermediate timescales (sub-behavioral, but on the order of tens to hundreds of milliseconds) [32, 33, 35, 86, 87, 88], there is neither sufficient evidence to support nor reject the hypothesis that motor units within a muscle coordinate together at mixed and sub-behavioral timescales. Important future work will distinguish between the Case (1) and Case (2) explanations of our mixed and sub-behavioral timescales results.

Alternative analysis methods may help to resolve between the Case (1) and Case (2) explanations of our results. Discrete estimators of mutual information, such as the NSB method, developed by Nemenman et al [53, 54], may help to resolve our data limitation. As mentioned previously, two reasons that our analyses were impacted by a data limitation were that we only included breath cycles where the spike timing unit spiked at least one time (see Methods) and that we separated our data into spike count-conditioned subgroups prior to estimating the MI (Equations 2.9 and 2.12). The count-conditioning step served to separate the behavioral timescale contribution(s) to coordination from the sub-behavioral timescale contribution(s), and it was necessary for the KSG continuous estimator of MI [44, 36, 77, 64]. The NSB method of MI estimation was used in previous work from our lab in a songbird motor cortex analog to illustrate that spike patterns at sub-behavioral timescales, down to the millisecond, are better predictors of motor behavior than behavioral timescale spike patterns. To apply this method to our data, one would consider all of the breath cycles for each pairwise combination of neurons. The spike patterns for each breath cycle would be represented by a matrix of spike counts within each time bin ranging in size from full length of a breath cycle down to  $1ms$ , and these matrices would serve as inputs to the discrete MI estimator. This type of method would help to address our data limitation because the full data set (including all cycles with zero spike counts), rather than count-conditioned subgroups of the full data set would serve as the input

to the MI estimator for the sub-behavioral and mixed timescales estimates, [82, 65]. A major difference between this method and our method is that the binning method would capture MI from both small and large timescales. We would have to subtract the MI between motor units with large time bins from the MI between motor units with small time bins to isolate the small timescale component of coordination.

### 4.3 A Challenge to the idea of a static code for motor control

Information theory has been applied to neuroscience as a means of reading neural codes by considering neurons and their spike patterns from the lens of communication systems. Pioneered in sensory systems, under this school of thought, scientists apply tools from information theory, such as mutual information, to understand the coding properties of sensory neurons. Information theoretic approaches have found great success in seeking to understand how peripheral sensory neurons encode features of a stimulus and how the central nervous system decodes the stimulus from the spike patterns of the peripheral neurons [6, 15, 80, 66]. These types of approaches have been applied to studies of the nervous system more broadly, investigating correlations between neurons [70, 63] and the coding properties of individual neurons or populations in motor systems [77, 82, 64, 65]. Because information theoretic approaches examine the marginal and joint probability distributions of the stimulus and response, there is an inherent assumption, particularly for analyses that involve mutual information, that these distributions are stationary- that they do not change over the time of the data collection [14]. Therefore, these approaches assume that the neural code governing the system under interrogation is static. A strength of applying mutual information analyses to sensory systems is that identical stimuli can be repeated for thousands of trials while recording from the same neuron [6]. In these sorts of stud-

ies, even if the distribution of the sensory neurons change over time, an experimenter manipulates the stimulus, and can therefore guarantee stationarity for at least one of the probability distributions of interest. In contrast, a challenge to applying mutual information analyses to motor control is that no animal makes the same movement twice. The experimenter can not control the marginal distribution of the behavior. The marginal distributions of the behavior and the neural activity, in addition to the joint distribution may change over time. Although researchers can place constraints on experimental conditions to improve the validity of stationarity assumptions, no living system ever perfectly occupies a stationary probability distribution. Changes in neuromodulator concentration [79], metabolic conditions [90], circadian rhythms [1], attention [37], environmental conditions [90], and many other factors are likely to impact the probability distributions of sensory stimuli and neural activity and/or the distributions of neural activity and behavior. Even still, few studies have investigated violations to the stationarity assumption or applied analysis techniques that do not assume stationarity. Here, we have identified nonstationarity in the probability distributions of motor signals governing anesthetized breathing, a cyclic, consistent motor behavior (see Nonstationarity in the coordination between motor units, below). Our results emphasize the importance of investigating the stationarity of data prior to interpreting results of analyses techniques that assume stationarity.

### **4.3.1 Nonstationarities in the coordination between motor units**

We identified qualitative and quantitative changes in the coordination between motor units from the beginning to the end of the data collection. When we split our data into two consecutive halves, we found qualitative differences in the marginal and joint distributions of spike counts per breath cycle as well as changes to the MI values between spike counts from the first half to the second half of the data collection (Figure

3.4) and from each half to the full data set (Figures 3.4 and 3.2) - all indicative of violation to the stationarity assumption of mutual information. We do not believe that this violation detracts from the credibility of our results in Figures 3.2 and 3.3. However, we must carefully consider how this violation affects the interpretation of our results. The results in Figure 3.2 simply represent a time-averaged picture of the coordination that exists between motor units. For example, Unit D and Unit E never exhibit simultaneous spiking activity. Unit D spikes only during the first half of the data collection whereas Unit E spikes only during the second half of the data collection (Figures 3.1 and 3.4 a, b, e, and f). Units D and E share non-zero MI across the full data set (when the distributions are considered collectively across the full time of the data collection) but zero MI in each consecutive half of the data collection (Figure 3.4 c and d). The non-zero MI found between these units in the full data set may not reflect coordination between the units on a single-trial timescale, but it may reflect the holistic structure of which units are active at which times throughout data collection. This particular example highlights the importance of applying careful reasoning to the interpretation of MI analyses in studies of systems in which the stationarity assumption may be violated. Applying MI analyses or other analyses that assume stationarity to a data set that exhibits a lot of temporal dependence may output misleading results - units may appear to be coordinated, when in reality they are just active at different times. In contrast to units D and E, our other example pairs of units for which we examined the coordination separately for the two consecutive halves of the data collection provide results with more variability in the coordination between the halves and full data set. Unit D and Unit G exhibited non-zero MI for the full data set (Figure 3.2 b), and Unit D only spiked during the first half of the data collection (Figure 3.1 and 3.4 a and b). However, the MI between units D and G roughly tripled when the first half of the data was considered alone (Figure 3.4 c). Therefore, during the epoch when the units spiked simultaneously, they were found to be coordinated in

their spike counts. Additionally, Unit B and Unit G exhibited simultaneous spiking throughout the data collection (Figures 3.1 and 3.4 a, b, e, and f). Similarly to units D and G, units B and G exhibited non-zero MI for the full data set. The value of the MI estimate roughly doubled when the first half of the data collection was considered alone, whereas the MI estimate for the second half of the data collection was comparable to the MI estimate for the full data set (Figure 3.4 c and d). That the MI estimate changed between the first and second half of the data collection, even for units which exhibited simultaneous spiking throughout the data collection further supports that the stationarity assumption of MI is not met by this data - the extent to which these units are coordinated changes from the first half of the data collection to the second half. Although the subject performed a cyclic, consistent, breathing behavior throughout the duration of the experiment, the coordination underlying that breathing changed for all three example pairs of units that we examined across each data half. These results emphasize the importance of considering stationarity in the interpretation of mutual information approaches to studying the motor system and challenges the idea of a static motor code, suggesting that the nervous system may employ dynamic coding strategies for motor behaviors. Future work with these data may segment the data collection into smaller time epochs for which the stationarity assumption is more reasonable (the less time that has passed, the less probability distributions have been able to change). The same analyses could be repeated for these smaller epochs to examine how the behavioral timescale coordination changes throughout the data collection.

### **4.3.2 Addressing non-stationarity concerns for motor control**

Common analysis methods including mutual information [77, 82, 44, 36] and Bayesian statistics [34, 25, 22] require stationarity in the probability distribution across trials. Particularly our results from units D and E illustrate how erroneous interpretation

may result from blindly applying these types of analyses to non-stationary data. The fact that many prevalent analysis techniques have an assumption that biological data cannot perfectly satisfy emphasizes the importance of developing new, more biologically relevant computational techniques. An example of analyses applied to non-stationary experimental data can be found in work by Fairhall et al. The dynamics of the fly visual neural code was studied in the context of sensory adaptation [20]. Flies were presented with a white noise velocity visual stimulus, while recordings of action potentials from a visual neuron were taken. The variance of the visual stimulus was periodically shifted between two values. Therefore, this task was designed such that the distribution of the sensory stimulus was non-stationary in that it shifted between two constant distributions, but the distributions and the shift time were completely known by the experimenter. The change in spike rate and the change in the relationship between the stimulus and response that resulted from the shift in the stimulus distribution were investigated. Overall, Fairhall et al. found that the fly visual neurons respond to a change in the distribution of the sensory stimulus within tens of milliseconds [20]. This type of approach offers promise in the sensory system, where non-stationary stimuli can be carefully designed to meet constraints of computational methods. However, it would be challenging to adapt this approach to studies of motor systems for the same reason stated above- an experimenter can not control the distribution of motor behavior or neural activity.

A simple way to combat stationarity concerns for studies of motor control, similarly to our methods, is to assume quasi-stationarity for subsets of the data and to repeat analyses with a sliding window that defines each quasi-stationary subset of data [29]. The question remains- how to define and identify quasi-stationarity regimes. Although many of the variables that may affect the probability distributions of neural activity and their impact on behavior remain elusive, a wealth of knowledge from the fields of neuromechanics [56] and muscle physiology [57] can help to set crite-

ria for quasi-stationary regimes by contributing to a more thorough understanding of whether, how, and why the probability distributions of motor commands and/or their impact on motor behaviors change over time. Many complex interactions between the brain, the body, and the environment affect the dynamics of motor behaviors [73]. The force of a muscle depends on the activating signal from the motor neuron(s) and the length and velocity of the muscle at the time the signal arrives in addition to the history of the muscle before the signal arrives and which fibers are active simultaneously [85, 67]. The behavioral implications of the same muscle force may vary depending on the states of other muscles that contribute to the behavior and the mechanical properties of the body and environment. Therefore the same neural signal may result in drastically different motor outputs - the joint probability distribution of neural signal and behavior is liable to change depending on each of these factors [85, 67]. A thorough approach to defining a quasi-stationary regime would be to examine as many of these variables as possible and identify time epochs where these variables are held constant. Practically, it would be extraordinarily challenging to measure all of these variables, much less expect all of them to be simultaneously held constant for multiple time epochs. A more reasonable approach would be to choose one or a few neuromechanical variables to set the criteria for quasi-stationary epochs. After conducting analyses for these quasi-stationary epochs, it would then be important to incorporate variables that were not included in the stationarity criteria into the discussion of limitations and interpretation.

To apply this quasi-stationary epoch approach to our data, more computational and experimental work may be required. We have shown that the probability distributions and the joint probability distributions of the spike counts of most of the motor units we recorded changed from the first half to the second half of the experiment (Figure 3.4). We measured action potentials from the EXP group in addition to air sac pressure. Future computational work on this data set may examine the how dis-

tribution of the air sac pressure waveforms changed. Segmenting the data into epochs where the air sac pressure is relatively consistent could serve as a reasonable quasi-stationary criteria. It would be interesting to study how the coordination motor units changes across epochs of quasi-stationary air sac pressure. However, other neuromechanical variables that may impact the neuromuscular transform include muscle fiber length and velocity, location of the active fibers, fiber type, activation history, activation patterns of synergistic or antagonistic muscles, etc. These variables would need to be considered in the interpretation of this new quasi-stationary epoch analysis for our data set. The primary muscle we targeted from the EXP group was the obliquus externus abdominis, which is a large muscle that stretches from the pubic bone to the sternum. It attaches and inserts on aponeuroses, which are flat, sheet-like tissues that function as tendons with larger attachment sites [21]. The location of the fibers within the muscle may be particularly important for this muscle because of the sheet-like attachment and insertion sites. Muscles which attach to tendons converge to exert force on a single point, whereas the force of different fibers at different locations on a sheet-like attachment cite may have result in very different behavioral outputs. Additionally, since many different muscles comprise the EXP group, synergistic interactions between the different EXP muscles or even just different fibers within the obliquus externus abdominis may impact the transformation from the spike patterns of motor units to respiratory behavior. For example, different activation patterns of other muscles in the EXP group may affect the length of the obliquus externus abdominis at its time of activation, and thereby change the relationship between the motor unit spikes and their behavioral output. Additionally, changes in activation patterns of different muscle fibers within the obliquus externus abdominis may affect the transformation from spikes to behavior. Future work experimental work could track some of these variables by measuring the cycles of length change [7] of the obliquus externus abdominis during breathing while simultaneously recording spike patterns from single

motor units [52]. Ideally, in these same experiments, histology could be used to identify what motor unit types were recorded during the experiment. Once the muscle length change cycles have been measured, workloop experiments [17, 7, 52] could be conducted to measure the force that the muscle produces in response to stimulation patterns that mimic the measured biological conditions. Any subset of these variables could be examined in conjunction with the distribution of spike patterns and/or the distribution of air sac pressure to identify epochs of quasi-stationarity to analyze with our methods.

## 4.4 Conclusions and future directions

Overall, we have determined that the pairwise combinations of the motor units we recorded in the EXP muscle group are coordinated at the timescale of the behavior when the whole data collection is considered together. When we considered the two consecutive halves of the data collection separately, the coordination between motor units at this timescale changed from the first half to the second, indicating that the coordination between motor units is non-stationary - the neural code for motor control is not static, but dynamic! Finally, our investigation of mixed and sub-behavioral timescale coordination yielded inconclusive results, likely due to insufficient data, but possibly reflecting a genuine lack of coordination at those timescales. Given the data set size limitations, we did not consider whether the mixed or sub-behavioral timescale coordination changed from the first half of the data collection to the second half. Because we examined the mixed and sub-behavioral timescale spike patterns by conditioning them on spike counts, the stationarity assumption may be more reasonable for these analyses. In future work, we could examine how the distribution of sub-behavioral timescale spike patterns for each spike count subgroup changes from one quasi-stationary epoch to the next, but much more data would be required to

overcome the same data limitations that we faced in this work.

The results from this work point to future work in two major directions: (1) further investigation of the how spike patterns across different timescales code for and cause motor control and coordination and (2) investigation of the nonstationarities in the neural code for motor control.

Increasing evidence illustrates that spike patterns from the timescale of the behavior down to the millisecond correlate with and cause motor behaviors in songbirds [82, 77]. Anesthetized breathing is thought to occur largely without input from the songbird motor cortex analog, the robust nucleus of the archipallium. However, it is thought that during singing, there is a great deal of communication between song and respiratory brain structures [49, 69, 95, 94]. Although our lab has previously shown that during anesthetized breathing, millisecond-scale time shifts in spike patterns cause changes to the airsac pressure [77] and that during singing, the activity of neurons in the song motor cortex analog are most correlated with song acoustics at millisecond timescales [82], future work in the brain and periphery may help identify differences in the timescales at which motor units are coordinated for breathing at rest, a relatively slow behavior versus singing, a dynamic behavior. Perhaps the inputs from the cortical song circuitry that innervate respiratory areas increase the sub-behavioral timescale coordination between motor units in respiratory muscles [95, 94]. Even still, as stated above, further investigation of motor control and coordination particularly at mixed and sub-behavioral timescales will require the exploration and/or development of analysis techniques to overcome the data limitation we faced or the development of better experimental methods to collect larger data sets, which will be particularly challenging to do during singing.

Analyses with assumptions of stationarity in the data are pervasive in neuroscience. We show here that the neural code for motor control is dynamic, even for a cyclic, constant behavior- anesthetized breathing. Throughout their life times

animals are subject to physical changes that can result in major changes to biomechanical variables. Natural processes such as juvenile growth and development [59], injury, etc. in addition to changes in the surrounding environment may result in an altered transformation from spikes to motor behavior and require an animal to alter its motor control strategy. After initially learning their song, songbirds maintain a remarkably stable song over time [51], but in many species, deafening causes the song to deteriorate over time [51, 58]. Songbirds can alter the variability in the acoustics of their song depending on context [42], and can learn to adjust the pitch of certain syllables [89]. Since the song behavior is so consistent across renditions, it becomes an interesting question whether the neural code governing song is more or less dynamic than the neural code governing anesthetized breathing. Because song is highly dynamic and stereotyped, the neural code governing song may exhibit remarkable stationarity within days or weeks, but may shift dramatically over long periods of time or after a event that causes dramatic shifts in the biomechanics of the vocal and respiratory muscles. It may be particularly advantageous for the motor code underlying breathing and other processes that are associated with song to be dynamic in other non-singing contexts in order to promote the flexibility necessary to maintain a stable song despite natural changes to the variables that affect the impact of neural signals on behavioral output. In of the brain, it is well-known that the anterior forebrain pathway (AFP) is necessary for adult songbirds to alter the acoustics of their song [9, 93, 51]. Although it is well known that the AFP plays a role in altering the distribution of motor behavior by outputting signals to the song specific analog of motor cortex [42, 93, 51], our results illustrate that even in the periphery, at the level of the muscles, there are dynamics in the neural code for motor behavior. Future work will investigate whether and how these mechanisms of variability interact to result in learning, maintenance, and successful motor control throughout animals' life times. Overall, our work highlights these two major initiatives as important next

steps in motor control research. Future investigations in the proposed directions have the potential to result in significant contributions to our understanding of how the brain controls the body.

# Appendix A

## Proof of Upperbound on Error for Spike Timing MI Estimates

### A.1 Mixed Timescales: Timing and Count

We begin with Equations 2.9, 2.11, and 2.14. As mentioned in Methods because of the limitations of the KSG MI estimator [44, 36], we omitted some of the data from the mixed timescales MI analysis due to three different concerns: (1) We omitted data for which  $U_{1,c} = 0$  because there are no spike timing values for breath cycles when the motor unit did not spike. (2) We omitted data for which  $n_i < U_{1,c}$  because the number of samples must exceed the dimensionality of the data for the KSG estimator. (3) We omitted data for which  $\frac{n_i}{4} < 5$  because the estimator required at least five samples to run and we needed at least four data fractions in order to properly estimate the variance of the mutual information estimate (see Holmes and Nemenman [36]).

We rescaled the probability term in equation 2.9 so that the total MI estimate reflects the MI of the all the breath cycles for which the spike timing variable spiked at least one time. In other words, we considered only the conditional mutual information given that the spike timing neuron spiked at least once during the breath cycle. In

order to find this MI, we rescaled the probability term to sum to one when all of the breath cycles with zero spikes in the spike timing variable were excluded from the dataset. We propagated this rescaling to equation 2.14 so that we estimated an upperbound on the error of the MI estimate. As written in the methods, the equations accounting for the re-weighted probabilities are as follows:

$$I(U_{1,t}; U_{2,c}) = \sum_{i=1}^{C_{max}} I(U_{1,t}; U_{2,c} | U_{1,c} = i) p'(U_{1,c} = i) \quad (\text{A.1})$$

$$\sigma_{t,c}^2 = \sum_{i=1}^{C_{max}} \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{\sigma_{p'_i}^2}{\langle p'_i \rangle^2} \right) \langle I_i \rangle^2 \langle p'_i \rangle^2 \quad (\text{A.2})$$

$$\sigma_{p'_i}^2 = \frac{(1 - p'_i)}{n_i p'_i} p_i'^2 \quad (\text{A.3})$$

Where  $p'_i = \frac{n_i}{N'}$  and  $N' = \sum_{i=1}^{C_{max}} 1$ .

The equations for MI and error without re-weighting the probability term would be as follows:

$$I_{full}(U_{1,t}; U_{2,c}) = \sum_{i=1}^{C_{max}} I(U_{1,t}; U_{2,c} | U_{1,c} = i) p(U_{1,c} = i) \quad (\text{A.4})$$

$$\sigma_{full,t,c}^2 = \sum_{i=1}^{C_{max}} \left( \frac{\sigma_{I_{full,i}}^2}{\langle I_{full,i} \rangle^2} + \frac{\sigma_{p_i}^2}{\langle p_i \rangle^2} \right) \langle I_{full,i} \rangle^2 \langle p_i \rangle^2 \quad (\text{A.5})$$

$$\sigma_{p_i}^2 = \frac{(1 - p_i)}{n_i p_i} p_i^2 \quad (\text{A.6})$$

Where  $p_i = \frac{n_i}{N}$  and  $N = \sum_{i=0}^{C_{max}} 1$ .

To show that the error estimated is an upper bound, we need to show that:

$$\sigma_{full,t,c}^2 \leq \sigma_{t,c}^2 \quad (\text{A.7})$$

or by substitution:

$$\begin{aligned} & \sum_{i=1}^{C_{1,max}} \left( \frac{\sigma_{I_{full,i}}^2}{\langle I_{full,i} \rangle^2} + \frac{\sigma_{p_i}^2}{\langle p_i \rangle^2} \right) \langle I_{full,i} \rangle^2 \langle p_i \rangle^2 \\ & \leq \sum_{i=1}^{C_{max}} \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{\sigma_{p'_i}^2}{\langle p'_i \rangle^2} \right) \langle I_i \rangle^2 \langle p'_i \rangle^2 \end{aligned} \quad (\text{A.8})$$

Because we are summing across all  $i > 0$  on both sides, we can cancel the summation.

Therefore we are left to show that:

$$\begin{aligned} & \left( \frac{\sigma_{I_{full,i}}^2}{\langle I_{full,i} \rangle^2} + \frac{\sigma_{p_i}^2}{\langle p_i \rangle^2} \right) \langle I_{full,i} \rangle^2 \langle p_i \rangle^2 \\ & \leq \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{\sigma_{p'_i}^2}{\langle p'_i \rangle^2} \right) \langle I_i \rangle^2 \langle p'_i \rangle^2 \end{aligned} \quad (\text{A.9})$$

Each  $\sigma_{I_{full,i}}$  and  $\sigma_{I_i}$  refer to the variance of individual MI estimates. Since these terms have no dependence on the probability distribution of spike counts for the spike timing unit, but are estimated by the KSG estimator for identical, count-conditioned data, these terms are equivalent:

$$\sigma_{I_{full,i}}^2 = \sigma_{I_i}^2 \quad (\text{A.10})$$

Additionally, re-weighting the probabilities has no impact on the individual MI estimates for each spike count condition, therefore:

$$\langle I_{full,i} \rangle^2 = \langle I_i \rangle^2 \quad (\text{A.11})$$

So we are left with

$$\begin{aligned}
& \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{\sigma_{p_i}^2}{\langle p_i \rangle^2} \right) \langle I_i \rangle^2 \langle p_i \rangle^2 \\
& \leq \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{\sigma_{p'_i}^2}{\langle p'_i \rangle^2} \right) \langle I_i \rangle^2 \langle p'_i \rangle^2
\end{aligned} \tag{A.12}$$

We know that:

$$\sigma_{p_i}^2 = \frac{(1-p_i)}{n_i p_i} p_i^2 \tag{A.13}$$

and that

$$\sigma_{p'_i}^2 = \frac{(1-p'_i)}{n_i p'_i} p_i'^2 \tag{A.14}$$

We can fill in these values, and we are left with

$$\begin{aligned}
& \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{\frac{(1-p_i)}{n_i p_i} p_i^2}{p_i^2} \right) \langle I_i \rangle^2 \langle p_i \rangle^2 \\
& \leq \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{\frac{(1-p'_i)}{n_i p'_i} p_i'^2}{p_i'^2} \right) \langle I_i \rangle^2 \langle p'_i \rangle^2
\end{aligned} \tag{A.15}$$

We can then simplify.

$$\begin{aligned}
& \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{(1-p_i)}{n_i p_i} \right) \langle I_i \rangle^2 \langle p_i \rangle^2 \\
& \leq \left( \frac{\sigma_{I_i}^2}{\langle I_i \rangle^2} + \frac{(1-p'_i)}{n_i p'_i} \right) \langle I_i \rangle^2 \langle p'_i \rangle^2
\end{aligned} \tag{A.16}$$

We can do algebra to get

$$\begin{aligned}
& \left( \frac{\sigma_{I_i}^2 n_i p_i + (1 - p_i) \langle I_i \rangle^2}{n_i} \right) \langle p_i \rangle \\
& \leq \left( \frac{\sigma_{I_i}^2 n_i p'_i + (1 - p'_i) \langle I_i \rangle^2}{n_i} \right) \langle p'_i \rangle
\end{aligned} \tag{A.17}$$

Recall that  $p_i = \frac{n_i}{N}$  and  $p'_i = \frac{n_i}{N'}$ .  $N$  refers to the number of breath cycles in the entire data set, whereas  $N'$  refers to the number of breath cycles where the spike count of the timing motor unit is greater than zero ( $i > 0$ ). Therefore,

$$p_i \leq p'_i \tag{A.18}$$

The  $p_i$  and  $p'_i$  factors taken by themselves satisfy the inequality. Therefore, we can omit them from the remainder of the proof. We are now left to show that

$$\begin{aligned}
& \sigma_{I_i}^2 n_i p_i + \langle I_i \rangle^2 - p_i \langle I_i \rangle^2 \\
& \leq \sigma_{I_i}^2 n_i p'_i + \langle I_i \rangle^2 - p'_i \langle I_i \rangle^2
\end{aligned} \tag{A.19}$$

We can simplify this inequality to get

$$p_i (\sigma_{I_i} n_i + \langle I_i \rangle^2) \leq p'_i (\sigma_{I_i} n_i + \langle I_i \rangle^2) \tag{A.20}$$

By canceling the identical terms, we are left with

$$p_i \leq p'_i \tag{A.21}$$

which is true by definition as illustrated above. This proof confirms that by re-weighting the probability term in the mixed timescales MI estimates, we estimate an upperbound on the error of our MI estimates. Therefore, our results are conservative.

## A.2 Subcycle Timescales: Timing and Timing

We begin with Equations 2.10, 2.12, and 2.16. Because of the limitations of the KSG MI estimator [44, 36], we omitted some of the data from the mixed timescales MI analysis due to three different concerns: (1) We omitted data for which  $U_{1,c} = 0$  or  $U_{2,c} = 0$  because there are no spike timing values for breath cycles when the motor unit did not spike. (2) We omitted data for which  $n_i < U_{1,c}$  or  $n_j < U_{2,c}$  because the number of samples must exceed the dimensionality of the data for the KSG estimator. (3) We omitted data for which  $\frac{n_i}{4} < 5$  or for which  $\frac{n_j}{4} < 5$  because the estimator required at least five samples to run and we needed at least four data fractions in order to properly estimate the variance of the mutual information estimate (see Holmes and Nemenman [36]).

We rescaled the joint probability term in equation 2.10 so that the total MI estimate reflects the MI of all breath cycles for which the each spike timing variable spiked at least one time. In otherwords, we considered only the conditional mutual information given that both neurons spiked during the breath cycle. In order to find this MI, we rescaled the probability term to sum to one when all of the breath cycles with zero spikes in either variable were excluded from the dataset. We propagated this rescaling to equation 2.16 so that we estimated an upperbound on the error of the MI estimate. By estimating an upperbound on the error, we can be confident that any MI values that fall outside of error of zero are indeed significant. As written in the methods, the equations accounting for the re-weighted probabilities are as follows:

$$I(U_{1,t}; U_{2,t}) = \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} I(U_{1,t}; U_{2,t} | U_{1,c} = i, U_{2,c} = j) p'(U_{1,c} = i, U_{2,c} = j) \quad (\text{A.22})$$

$$\sigma_{t,t}^2 = \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{\sigma_{p'_{i,j}}^2}{\langle p'_{i,j} \rangle^2} \right) \langle I_{i,j} \rangle^2 \langle p'_{i,j} \rangle^2 \quad (\text{A.23})$$

$$\sigma_{p'_{i,j}}^2 = \frac{(1 - p'_{i,j})}{n_{i,j} p'_{i,j}} p'_{i,j} \quad (\text{A.24})$$

Where  $p'_{i,j} = \frac{n_{i,j}}{N'}$  and  $N' = \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} 1$ . The equations for MI and error without re-weighting the probability terms would be as follows:

$$I_{full}(U_{1,t}; U_{2,t}) = \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} I(U_{1,t}; U_{2,t} | U_{1,c} = i, U_{2,c} = j) p(U_{1,c} = i, U_{2,c} = j) \quad (\text{A.25})$$

$$\sigma_{full,t,t}^2 = \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} \left( \frac{\sigma_{I_{full,i,j}}^2}{\langle I_{full,i,j} \rangle^2} + \frac{\sigma_{p_{i,j}}^2}{\langle p_{i,j} \rangle^2} \right) \langle I_{full,i,j} \rangle^2 \langle p_{i,j} \rangle^2 \quad (\text{A.26})$$

$$\sigma_{p_{i,j}}^2 = \frac{(1 - p_{i,j})}{n_{i,j} p_{i,j}} p_{i,j} \quad (\text{A.27})$$

Where  $p_{i,j} = \frac{n_{i,j}}{N}$  and  $N = \sum_{i=0}^{C_{1,max}} \sum_{j=0}^{C_{2,max}} 1$ .

To show that the error estimated is an upper bound, we need to show that:

$$\sigma_{full,t,t}^2 \leq \sigma_{t,t}^2 \quad (\text{A.28})$$

or by substitution:

$$\begin{aligned} & \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} \left( \frac{\sigma_{I_{full,i,j}}^2}{\langle I_{full,i,j} \rangle^2} + \frac{\sigma_{p_{i,j}}^2}{\langle p_{i,j} \rangle^2} \right) \langle I_{full,i,j} \rangle^2 \langle p_{i,j} \rangle^2 \\ & \leq \sum_{i=1}^{C_{1,max}} \sum_{j=1}^{C_{2,max}} \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{\sigma_{p'_{i,j}}^2}{\langle p'_{i,j} \rangle^2} \right) \langle I_{i,j} \rangle^2 \langle p'_{i,j} \rangle^2 \end{aligned} \quad (\text{A.29})$$

Because we are summing across all  $i > 0$  and all  $j > 0$  on both sides, we can cancel the summation. Therefore, we are left to show that

$$\begin{aligned}
& \left( \frac{\sigma_{I_{full,i,j}}^2}{\langle I_{full,i,j} \rangle^2} + \frac{\sigma_{p_{i,j}}^2}{\langle p_{i,j} \rangle^2} \right) \langle I_{full,i,j} \rangle^2 \langle p_{i,j} \rangle^2 \\
& \leq \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{\sigma_{p'_{i,j}}^2}{\langle p'_{i,j} \rangle^2} \right) \langle I_{i,j} \rangle^2 \langle p'_{i,j} \rangle^2
\end{aligned} \tag{A.30}$$

Each  $\sigma_{I_{full,i,j}}^2$  and  $\sigma_{I_{i,j}}$  refer to the variance of individual MI estimates. Since these terms have no dependence on the probability distribution of spike counts of either unit, but they are estimated by the KSG estimator for identical count-conditioned data, these terms are equivalent.

$$\sigma_{I_{i,j}}^2 = \sigma_{I_{i,j}}'^2 \tag{A.31}$$

Additionally, re-weighting the probabilities has no impact on the individual MI estimates for each spike count condition, therefore:

$$\langle I_{full,i,j} \rangle^2 = \langle I_{i,j} \rangle^2 \tag{A.32}$$

So we are left with

$$\begin{aligned}
& \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{\sigma_{p_{i,j}}^2}{\langle p_{i,j} \rangle^2} \right) \langle I_{i,j} \rangle^2 \langle p_{i,j} \rangle^2 \\
& \leq \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{\sigma_{p'_{i,j}}^2}{\langle p'_{i,j} \rangle^2} \right) \langle I_{i,j} \rangle^2 \langle p'_{i,j} \rangle^2
\end{aligned} \tag{A.33}$$

We know that

$$\sigma_{p_{i,j}}^2 = \frac{(1 - p_{i,j})}{n_{i,j} p_{i,j}} p_{i,j}^2 \tag{A.34}$$

and that

$$\sigma_{p'_{i,j}}^2 = \frac{(1 - p'_{i,j})}{n_{i,j} p'_{i,j}} p'_{i,j} \quad (\text{A.35})$$

We can fill in these values, and we are left with

$$\begin{aligned} & \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{\frac{(1-p_{i,j})}{n_{i,j} p_{i,j}} p_{i,j}^2}{p_{i,j}^2} \right) \langle I_{i,j} \rangle^2 \langle p_{i,j} \rangle^2 \\ & \leq \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{\frac{(1-p'_{i,j})}{n_{i,j} p'_{i,j}} p'_{i,j}{}^2}{p'_{i,j}{}^2} \right) \langle I_{i,j} \rangle^2 \langle p'_{i,j} \rangle^2 \end{aligned} \quad (\text{A.36})$$

We can then simplify.

$$\begin{aligned} & \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{(1 - p_{i,j})}{n_{i,j} p_{i,j}} \right) \langle I_{i,j} \rangle^2 \langle p_{i,j} \rangle^2 \\ & \leq \left( \frac{\sigma_{I_{i,j}}^2}{\langle I_{i,j} \rangle^2} + \frac{(1 - p'_{i,j})}{n_{i,j} p'_{i,j}} \right) \langle I_{i,j} \rangle^2 \langle p'_{i,j} \rangle^2 \end{aligned} \quad (\text{A.37})$$

We can do algebra to get

$$\begin{aligned} & \left( \frac{\sigma_{I_{i,j}}^2 n_{i,j} p_{i,j} + (1 - p_{i,j}) \langle I_{i,j} \rangle^2}{n_{i,j}} \right) \langle p_{i,j} \rangle \\ & \leq \left( \frac{\sigma_{I_{i,j}}^2 n_{i,j} p'_{i,j} + (1 - p'_{i,j}) \langle I_{i,j} \rangle^2}{n_{i,j}} \right) \langle p'_{i,j} \rangle \end{aligned} \quad (\text{A.38})$$

Recall that  $p_{i,j} = \frac{n_{i,j}}{N}$  and  $p'_{i,j} = \frac{n_{i,j}}{N'}$ .  $N$  refers to the number of breath cycles in the entire data set, whereas  $N'$  refers to the number of breath cycles where the spike counts of each neuron are greater than zero ( $i > 0$  and  $j > 0$ ). Therefore,

$$p_{i,j} \leq p'_{i,j} \quad (\text{A.39})$$

The  $p_i$  and  $p'_i$  factors taken by themselves satisfy the inequality. Therefore, we

can omit them from the remainder of the proof. We are now left to show that

$$\begin{aligned} & \sigma_{I_{i,j}}^2 n_{i,j} p_{i,j} + \langle I_{i,j} \rangle^2 - p_{i,j} \langle I_{i,j} \rangle^2 \\ & \leq \sigma_{I_{i,j}}^2 n_{i,j} p'_{i,j} + \langle I_{i,j} \rangle^2 - p'_{i,j} \langle I_{i,j} \rangle^2 \end{aligned} \tag{A.40}$$

We can then simplify to get

$$p_{i,j} (\sigma_{I_{i,j}} n_{i,j} + \langle I_{i,j} \rangle^2) \leq p'_{i,j} (\sigma_{I_{i,j}} n_{i,j} + \langle I_{i,j} \rangle^2) \tag{A.41}$$

By canceling the identical terms, we are left with

$$p_{i,j} \leq p'_{i,j} \tag{A.42}$$

which is true by definition as illustrated above. This proof confirms that by re-weighting the joint probability term in the sub-cycle timescales MI estimates, we estimate an upperbound on the error of our MI estimates. Therefore our results are conservative.

# Bibliography

- [1] Carolina Abarca, Urs Albrecht, and Rainer Spanagel. Cocaine sensitization and reward are under the influence of circadian genes and rhythm. *Proceedings of the National Academy of Sciences*, 99(13):9026–9030, 2002.
- [2] Tyson N Aflalo and Michael SA Graziano. Relationship between unconstrained arm movements and single-neuron firing in the macaque motor cortex. *Journal of Neuroscience*, 27(11):2760–2780, 2007.
- [3] Ehsan Arabzadeh, Stefano Panzeri, and Mathew E Diamond. Deciphering the spike train of a sensory neuron: counts and temporal patterns in the rat whisker pathway. *Journal of Neuroscience*, 26(36):9216–9226, 2006.
- [4] Robin C Ashmore, J Martin Wild, and Marc F Schmidt. Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *Journal of Neuroscience*, 25(37):8543–8554, 2005.
- [5] Michael J Berry, David K Warland, and Markus Meister. The structure and precision of retinal spike trains. *Proceedings of the National Academy of Sciences*, 94(10):5411–5416, 1997.
- [6] William Bialek, Fred Rieke, RR De Ruyter Van Steveninck, and David Warland. Reading a neural code. *Science*, 252(5014):1854–1857, 1991.
- [7] ANDREW A Biewener, WILLIAM R Corning, and BRET W Tobalske. In vivo

- pectoralis muscle force-length behavior during level flight in pigeons (*Columba livia*). *Journal of Experimental Biology*, 201(24):3293–3307, 1998.
- [8] Stuart Binder-Macleod and Trisha Kesar. Catchlike property of skeletal muscle: recent findings and clinical implications. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, 31(6):681–693, 2005.
- [9] Michael S Brainard and Allison J Doupe. Interruption of a basal ganglia–forebrain circuit prevents plasticity of learned vocalizations. *Nature*, 404(6779):762–766, 2000.
- [10] Stefanie Brändle, Syn Schmitt, and Matthias A Müller. A systems-theoretic analysis of low-level human motor control: application to a single-joint arm model. *Journal of mathematical biology*, 80(4):1139–1158, 2020.
- [11] Naama Brenner, Steven P Strong, Roland Koberle, William Bialek, and Rob R de Ruyter van Steveninck. Synergy in a neural code. *Neural computation*, 12(7):1531–1552, 2000.
- [12] Mark M Churchland, John P Cunningham, Matthew T Kaufman, Stephen I Ryu, and Krishna V Shenoy. Cortical preparatory activity: representation of movement or first cog in a dynamical machine? *Neuron*, 68(3):387–400, 2010.
- [13] Mark M Churchland, John P Cunningham, Matthew T Kaufman, Justin D Foster, Paul Nuyujukian, Stephen I Ryu, and Krishna V Shenoy. Neural population dynamics during reaching. *Nature*, 487(7405):51–56, 2012.
- [14] T Cover and J Thomas. Elements of information theory, 2ndwiley. *Hobo-ken, NJ*, 2006.
- [15] R. De Ruyter Van Steveninck and William Bialek. Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information

- transfer in short spike sequences. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, pages 379–414, 1988.
- [16] Roger Eckert, David J Randall, and George Augustine. Muscle and movement. In *Animal physiology mechanisms and adaptations*, chapter 10, pages 329–367. W.H. Freeman, 1988.
- [17] Coen PH Elemans, Andrew F Mead, Lawrence C Rome, and Franz Goller. Superfast vocal muscles control song production in songbirds. *PloS one*, 3(7):e2581, 2008.
- [18] Roger M Enoka and Jacques Duchateau. Rate coding and the control of muscle force. *Cold Spring Harbor perspectives in medicine*, 7(10):a029702, 2017.
- [19] Edward V Evarts. Relation of pyramidal tract activity to force exerted during voluntary movement. *Journal of neurophysiology*, 31(1):14–27, 1968.
- [20] Adrienne L Fairhall, Geoffrey D Lewen, William Bialek, and Robert R de Ruyter van Steveninck. Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(6849):787–792, 2001.
- [21] MR Fedde. Respiratory muscles. *Bird respiration*, 1:3–37, 1987.
- [22] Charles K Fisher and Pankaj Mehta. Bayesian feature selection for high-dimensional linear regression via the ising approximation with applications to genomics. *Bioinformatics*, 31(11):1754–1761, 2015.
- [23] Felix Franke, Rodrigo Quian Quiroga, Andreas Hierlemann, and Klaus Obermayer. Bayes optimal template matching for spike sorting—combining fisher discriminant analysis with optimal filtering. *Journal of computational neuroscience*, 38(3):439–459, 2015.

- [24] WILLIAM A Friedman, GEORGE W Sypert, JOHN B Munson, and JAMES W Fleshman. Recurrent inhibition in type-identified motoneurons. *Journal of neurophysiology*, 46(6):1349–1359, 1981.
- [25] Yun Gao, Michael J Black, Elie Bienenstock, Shy Shoham, and John P Donoghue. Probabilistic inference of hand motion from neural activity in motor cortex. In *Advances in neural information processing systems*, pages 213–220, 2002.
- [26] Christophe Gardella, Olivier Marre, and Thierry Mora. Modeling the correlated activity of neural populations: A review. *Neural computation*, 31(2):233–269, 2019.
- [27] Apostolos P Georgopoulos, Ronald E Kettner, and Andrew B Schwartz. Primate motor cortex and free arm movements to visual targets in three-dimensional space. ii. coding of the direction of movement by a neuronal population. *Journal of Neuroscience*, 8(8):2928–2937, 1988.
- [28] Franz Goller and Ole N Larsen. A new mechanism of sound generation in songbirds. *Proceedings of the National Academy of Sciences*, 94(26):14787–14791, 1997.
- [29] Sonja Grün, Markus Diesmann, and Ad Aertsen. Unitary events in multiple single-neuron spiking activity: Ii. nonstationary data. *Neural computation*, 14(1):81–119, 2002.
- [30] William J Heitler. Dataview: A tutorial tool for data analysis. template-based spike sorting and frequency analysis. *Journal of Undergraduate Neuroscience Education*, 6(1):A1, 2007.
- [31] WJ Heitler. Practical tools for analysing rhythmic neural activity. *Journal of neuroscience methods*, 185(1):151–164, 2009.

- [32] Elwood Henneman. Relation between size of neurons and their susceptibility to discharge. *Science*, 126(3287):1345–1347, 1957.
- [33] Elwood Henneman, George Somjen, and David O Carpenter. Excitability and inhibibility of motoneurons of different sizes. *Journal of neurophysiology*, 28(3):599–620, 1965.
- [34] Damián G Hernández, Samuel J Sober, and Ilya Nemenman. Unsupervised bayesian ising approximation for revealing the neural dictionary in songbirds. *arXiv preprint arXiv:1911.08509*, 2019.
- [35] Emma F. Hodson-Tole and James M. Wakeling. Motor unit recruitment for dynamic tasks: Current understanding and future directions. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 179(1):57–66, 2009. ISSN 01741578. doi: 10.1007/s00360-008-0289-1.
- [36] Caroline M Holmes and Ilya Nemenman. Estimation of mutual information for real-valued data with error bars and controlled bias. *Physical Review E*, 100(2):022404, 2019.
- [37] Sheng-Hsiou Hsu and Tzyy-Ping Jung. Modeling and tracking brain nonstationarity in a sustained attention task. In *International Conference on Augmented Cognition*, pages 209–217. Springer, 2016.
- [38] Robert K Josephson. Dissecting muscle power output. *Journal of Experimental Biology*, 202(23):3369–3375, 1999.
- [39] Shinji Kakei, Donna S Hoffman, and Peter L Strick. Muscle and movement representations in the primary motor cortex. *Science*, 285(5436):2136–2139, 1999.
- [40] John F Kalaska. From intention to action: motor cortex and the control of reaching movements. *Progress in Motor Control*, pages 139–178, 2009.

- [41] Eric R Kandel, James H Schwartz, Thomas M Jessell, Steven Siegelbaum, and A James Hudspeth. *Principles of neural science*, volume 5. McGraw-hill New York, 2013.
- [42] Mimi H Kao, Allison J Doupe, and Michael S Brainard. Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature*, 433(7026):638–643, 2005.
- [43] Ronald E Kettner, Andrew B Schwartz, and Apostolos P Georgopoulos. Primate motor cortex and free arm movements to visual targets in three-dimensional space. iii. positional gradients and population coding of movement direction from various movement origins. *Journal of Neuroscience*, 8(8):2938–2947, 1988.
- [44] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical review E*, 69(6):066138, 2004.
- [45] Daeyeol Lee, Nicholas L Port, Wolfgang Kruse, and Apostolos P Georgopoulos. Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *Journal of Neuroscience*, 18(3):1161–1170, 1998.
- [46] Robert C Liu, Svilen Tzonev, Sergei Rebrik, and Kenneth D Miller. Variability and information in a neural code of the cat lateral geniculate nucleus. *Journal of neurophysiology*, 86(6):2789–2806, 2001.
- [47] Emily L Mackevicius, Matthew D Best, Hannes P Saal, and Sliman J Bensmaïa. Millisecond precision spike timing shapes tactile perception. *Journal of Neuroscience*, 32(44):15309–15317, 2012.
- [48] Najja J Marshall, Joshua I Glaser, Eric M Trautmann, Elom A Amematsro, Sean M Perkins, Michael N Shadlen, LF Abbott, John P Cunningham, and Mark M Churchland. Flexible neural control of motor units. *bioRxiv*, 2021.

- [49] Jorge M Méndez, Gabriel B Mindlin, and Franz Goller. Interaction between telencephalic signals and respiratory dynamics in songbirds. *Journal of neurophysiology*, 107(11):2971–2983, 2012.
- [50] HS Milner-Brown, Richard B Stein, and R Yemm. Changes in firing rate of human motor units during linearly changing voluntary contractions. *The Journal of physiology*, 230(2):371–390, 1973.
- [51] Richard Mooney. Neural mechanisms for learned birdsong. *Learning & memory*, 16(11):655–669, 2009.
- [52] Charlotte R Morris and Graham N Askew. The mechanical power output of the pectoralis muscle of cockatiel (*nymphicus hollandicus*): the in vivo muscle length trajectory and activity patterns and their implications for power modulation. *Journal of Experimental Biology*, 213(16):2770–2780, 2010.
- [53] I Nemenman, F. Schafee, and W Bialek. Entropy and inference, revisited. In T.G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advanced Neural Information Processing Systems*, chapter 14. MIT Press, Cambridge, MA, 2002.
- [54] Ilya Nemenman. Coincidences and estimation of entropies of random variables with large cardinalities. *Entropy*, 13(12):2013–2023, 2011.
- [55] Ilya Nemenman, Geoffrey D Lewen, William Bialek, and Rob R de Ruyter van Steveninck. Neural coding of natural stimuli: information at sub-millisecond resolution. *PLoS computational biology*, 4(3):e1000025, 2008.
- [56] Kiisa Nishikawa, Andrew A Biewener, Peter Aerts, Anna N Ahn, Hillel J Chiel, Monica A Daley, Thomas L Daniel, Robert J Full, Melina E Hale, Tyson L Hedrick, et al. Neuromechanics: an integrative approach for understanding motor control. *Integrative and Comparative Biology*, 47(1):16–54, 2007.

- [57] Kiisa C Nishikawa, Jenna A Monroy, and Uzma Tahir. Muscle function from organisms to molecules. *Integrative and comparative biology*, 58(2):194–206, 2018.
- [58] Kathy W Nordeen and Ernest J Nordeen. Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and neural biology*, 57(1):58–66, 1992.
- [59] Bence P Ölveczky, Timothy M Otchy, Jesse H Goldberg, Dmitriy Aronov, and Michale S Fee. Changes in the neural control of a complex motor sequence during learning. *Journal of neurophysiology*, 106(1):386–397, 2011.
- [60] Chethan Pandarinath, Paul Nuyujukian, Christine H Blabe, Brittany L Sorice, Jad Saab, Francis R Willett, Leigh R Hochberg, Krishna V Shenoy, and Jaimie M Henderson. High performance communication by people with paralysis using an intracortical brain-computer interface. *Elife*, 6:e18554, 2017.
- [61] Chethan Pandarinath, Daniel J O’Shea, Jasmine Collins, Rafal Jozefowicz, Sergey D Stavisky, Jonathan C Kao, Eric M Trautmann, Matthew T Kaufman, Stephen I Ryu, Leigh R Hochberg, et al. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature methods*, 15(10):805–815, 2018.
- [62] Eric J Perreault, Charles J Heckman, and Thomas G Sandercock. Hill muscle model errors during movement are greatest within the physiologically relevant range of motor unit firing rates. *Journal of biomechanics*, 36(2):211–218, 2003.
- [63] Jason L Puchalla, Elad Schneidman, Robert A Harris, and Michael J Berry. Redundancy in the population code of the retina. *Neuron*, 46(3):493–504, 2005.
- [64] Joy Putney, Rachel Conn, and Simon Sponberg. Precise timing is ubiquitous, consistent, and coordinated across a comprehensive, spike-resolved flight motor program. *Proceedings of the National Academy of Sciences*, 116(52):26951–26960, 2019.

- [65] Joy Putney, Tobias Niebur, Rachel Barker, and Simon Sponberg. An information theoretic method to resolve millisecond-scale spike timing precision in a comprehensive motor program. *bioRxiv*, 2021.
- [66] Fred Rieke, David Warland, Rob De Ruyter Van Steveninck, William S Bialek, et al. *Spikes: exploring the neural code*, volume 7. MIT press Cambridge, 1999.
- [67] JM Ritchie and DR Wilkie. The effect of previous stimulation on the active state of muscle. *The Journal of physiology*, 130(2):488, 1955.
- [68] Marc F Schmidt and Franz Goller. Breathtaking songs: coordinating the neural circuits for breathing and singing. *Physiology*, 31(6):442–451, 2016.
- [69] Marc F Schmidt and J Martin Wild. The respiratory-vocal system of songbirds: anatomy, physiology, and neural control. In Gert Holstege, Caroline M. Beers, and Hari H. Subramanian, editors, *Breathing, Emotion, and Evolution*, volume 212, chapter 15, pages 297–335. Elsevier, 2014.
- [70] Elad Schneidman, William Bialek, and Michael J Berry. Synergy, redundancy, and independence in population codes. *Journal of Neuroscience*, 23(37):11539–11553, 2003.
- [71] Andrew B Schwartz, Ronald E Kettner, and Apostolos P Georgopoulos. Primate motor cortex and free arm movements to visual targets in three-dimensional space. i. relations between single cell discharge and direction of movement. *Journal of Neuroscience*, 8(8):2913–2927, 1988.
- [72] Krishna V Shenoy, Maneesh Sahani, and Mark M Churchland. Cortical control of arm movements: a dynamical systems perspective. *Annual review of neuroscience*, 36:337–359, 2013.

- [73] Samuel J Sober, Simon Sponberg, Ilya Nemenman, and Lena H Ting. Millisecond spike timing codes for motor control. *Trends in neurosciences*, 41(10):644–648, 2018.
- [74] S Sponberg and TL Daniel. Abdicating power for control: a precision timing strategy to modulate function of flight power muscles. *Proceedings of the Royal Society B: Biological Sciences*, 279(1744):3958–3966, 2012.
- [75] Simon Sponberg, Andrew J Spence, Chris H Mullens, and Robert J Full. A single muscle’s multifunctional control potential of body dynamics for postural control and running. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1570):1592–1605, 2011.
- [76] Simon Sponberg, Thomas L Daniel, and Adrienne L Fairhall. Dual dimensionality reduction reveals independent encoding of motor features in a muscle synergy for insect flight control. *PLoS Comput Biol*, 11(4):e1004168, 2015.
- [77] Kyle H Srivastava, Caroline M Holmes, Michiel Vellema, Andrea R Pack, Coen PH Elemans, Ilya Nemenman, and Samuel J Sober. Motor control by precisely timed spike patterns. *Proceedings of the National Academy of Sciences*, 114(5):1171–1176, 2017.
- [78] Nicholas A Steinmetz, Cagatay Aydin, Anna Lebedeva, Michael Okun, Marius Pachitariu, Marius Bauza, Maxime Beau, Jai Bhagat, Claudia Böhm, Martijn Broux, et al. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science*, 372(6539), 2021.
- [79] Klaus M Stiefel, Boris S Gutkin, and Terrence J Sejnowski. The effects of cholinergic neuromodulation on neuronal phase-response curves of modeled cortical neurons. *Journal of computational neuroscience*, 26(2):289–301, 2009.

- [80] Steven P Strong, Roland Koberle, Rob R De Ruyter Van Steveninck, and William Bialek. Entropy and information in neural spike trains. *Physical review letters*, 80(1):197, 1998.
- [81] Roderick A Suthers, Franz Goller, and J Martin Wild. Somatosensory feedback modulates the respiratory motor program of crystallized birdsong. *Proceedings of the National Academy of Sciences*, 99(8):5680–5685, 2002.
- [82] Claire Tang, Diala Chehayeb, Kyle Srivastava, Ilya Nemenman, and Samuel J Sober. Millisecond-scale motor encoding in a cortical vocal area. *PLoS Biol*, 12(12):e1002018, 2014.
- [83] Frédéric Theunissen and John P Miller. Temporal encoding in nervous systems: a rigorous definition. *Journal of computational neuroscience*, 2(2):149–162, 1995.
- [84] Nicholas Timme, Wesley Alford, Benjamin Flecker, and John M Beggs. Synergy, redundancy, and multivariate information measures: an experimentalist’s perspective. *Journal of computational neuroscience*, 36(2):119–140, 2014.
- [85] Lena H Ting and Hillel J Chiel. Muscle , Biomechanics , and Implications for Neural Control. In Buschges Ansgar Hooper, Scott L., editor, *Neurobiology of Motor Control: Fundamental Concepts and New Directions*, chapter 12, pages 365–416. Wiley and Sons, Inc., 2017.
- [86] Lena H Ting and Jane M Macpherson. A limited set of muscle synergies for force control during a postural task. *Journal of neurophysiology*, 93(1):609–613, 2005.
- [87] Lena H Ting and J Lucas McKay. Neuromechanics of muscle synergies for posture and movement. *Current opinion in neurobiology*, 17(6):622–628, 2007.
- [88] Lena H Ting, Hillel J Chiel, Randy D Trumbower, Jessica L Allen, J Lucas McKay, Madeleine E Hackney, and Trisha M Kesar. Neuromechanical princi-

- ples underlying movement modularity and their implications for rehabilitation. *Neuron*, 86(1):38–54, 2015.
- [89] Evren C Tumer and Michael S Brainard. Performance variability enables adaptive plasticity of ‘crystallized’ adult birdsong. *Nature*, 450(7173):1240–1244, 2007.
- [90] Cortney A Turner, Mark C Yang, and Mark H Lewis. Environmental enrichment: effects on stereotyped behavior and regional neuronal metabolic activity. *Brain Research*, 938(1-2):15–21, 2002.
- [91] Catherine R Von Reyn, Patrick Breads, Martin Y Peek, Grace Zhiyu Zheng, W Ryan Williamson, Alyson L Yee, Anthony Leonardo, and Gwyneth M Card. A spike-timing mechanism for action selection. *Nature neuroscience*, 17(7):962, 2014.
- [92] Le Wang, Rajiv Narayan, Gilberto Graña, Maoz Shamir, and Kamal Sen. Cortical discrimination of complex natural stimuli: can single neurons match behavior? *Journal of Neuroscience*, 27(3):582–589, 2007.
- [93] Timothy L Warren, Evren C Tumer, Jonathan D Charlesworth, and Michael S Brainard. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *Journal of neurophysiology*, 106(4):1806–1821, 2011.
- [94] J Martin Wild. The avian nucleus retroambigualis: a nucleus for breathing, singing and calling. *Brain research*, 606(2):319–324, 1993.
- [95] J Martin Wild. Descending projections of the songbird nucleus robustus archistriatalis. *Journal of Comparative Neurology*, 338(2):225–241, 1993.
- [96] JM Wild, Franz Goller, and RA Suthers. Inspiratory muscle activity during bird song. *Journal of neurobiology*, 36(3):441–453, 1998.

- [97] Felix E Zajac. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Critical reviews in biomedical engineering*, 17(4):359–411, 1989.
- [98] M. Zia, B. Chung, S. J. Sober, and M. S. Bakir. Fabrication and characterization of 3d multi-electrode array on flexible substrate for in vivo emg recording from expiratory muscle of songbird. In *2018 IEEE International Electron Devices Meeting (IEDM)*, pages 29.4.1–29.4.4, Dec 2018. doi: 10.1109/IEDM.2018.8614503.
- [99] Muneeb Zia, Bryce Chung, Samuel Sober, and Muhannad S Bakir. Flexible multi-electrode arrays with 2d and 3d contacts for in vivo electromyography recording. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2020.