Imperial College London

Department of Bioengineering

On the Structure of Natural Human Movement

Andreas Alexander Christian Thomik

October 31, 2015

Supervisor:

Dr. A. Aldo Faisal

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Bioengineering of Imperial College London and the Diploma of Imperial College London

Copyright Declaration

Declaration of Originality

I hereby declare that unless otherwise acknowledged this thesis and all the work therein are my own original research. Any concepts, ideas, methods or data which I did not develop myself are duly cited and a list references given in the bibliography.

Andreas Thomik

Abstract

Understanding of human motor control is central to neuroscience with strong implications in the fields of medicine, robotics and evolution. It is thus surprising that the vast majority of motor control studies have focussed on human movement in the laboratory while neglecting behaviour in natural environments. We developed an experimental paradigm to quantify human behaviour in high resolution over extended periods of time in ecologically relevant environments. This allows us to discover novel insights and contradictory evidence to well-established findings obtained in controlled laboratory conditions. Using our data, we map the statistics of natural human movement and their variability between people. The variability and complexity of the data recorded in these settings required us to develop new tools to extract meaningful information in an objective, data-driven fashion. Moving from descriptive statistics to structure, we identify stable structures of movement coordination, particularly within the arm-hand area. Combining our data with numerous published findings, we argue that current hypotheses that the brain simplifies motor control problems by dimensionality reduction are too reductionist. We propose an alternative hypothesis derived from sparse coding theory, a concept which has been successfully applied to the sensory system. To investigate this idea, we develop an algorithm for unsupervised identification of sparse structures in natural movement data. Our method outperforms state-of-the-art algorithms for accuracy and data-efficiency. Applying this method to hand data reveals a dictionary of *sparse eigenmotions* (SEMs) which are well preserved across multiple subjects. These are highly efficient and invariant representation of natural movement, and suggest a potential higher-order grammatical structure or "movement language". Our findings make a number of testable predictions about neural coding of movement in the cortex. This has direct consequences for advancing research on dextrous prosthetics and robotics, and has profound implications for our understanding of how the brain controls our body.

Acknowledments

A PhD may be an individual project, but it is not achieved alone. I am grateful to the many people with whom I had the chance to discuss my work, who have pointed out flaws, suggested new ideas, inspired me or "simply" been supportive throughout the time it took to achieve this work.

Many thanks to my supervisor, A. Aldo Faisal for kindling my interest in academic research and providing me with many opportunities to present my work all over the world. His trust in my work and my ability to perform independent research has been of great support.

My colleagues, many of which have become friends, have been instrumental to this work by allowing me to bounce ideas off them, helping me to solve seemingly impossible problems and patiently listening to me going on and on about minor details which were probably of interest to no one but myself. A special mention goes to the lab members, Alessandro, Ali, Anastasia, Ben, Chin-Hsuan, Feryal, Katya, Margarita, Marta and Scott. In particular Costa and Will have been absolutely key in developing the methodology for data collection and discussing ways of analysing it. Thank you to Ariadne for her patience and for providing a fresh look on the thesis. I am also thankful to Benoît, Claire, David and David, Gaëlle, Harris, JB, Jérémy, Laurent, Lucas, Mak, Marie, Michele and Michele, Niharika, Palpac, Pierre-Antoine, Simon, Sonja and Yesna for contributing to maintaining my sanity over the past four years.

Last, but certainly not least, thank you to my family for supporting me and believing in me even when things did not look too great.

This work would not have been possible without the generous financial support by the National Research Fund, Luxembourg (1229297).

Contents

List of Figures

List of Abbreviations

ADC	–	Analogue-to-Digital Converter	
BIC	–	Bayesian Information Criterion	
CMC	–	Carpometacarpal Joint	
DALI	–	Dynamic Activity Logging Interface	(Section 2.2.4)
DIP	–	Distal Interphalangeal Joint	
DOF	–	Degrees of Freedom	
CDF	–	Cumulative Distribution Function	
CNS	–	Central Nervous System	
CSR	–	Complete Spatial Randomness	
FWHM	–	Full-Width at Half Maximum	(Figure 3.5**A**)
GOF	–	Goodness Of Fit	(Section 4.3.4)
GOF_L	–	Local Goodness Of Fit	(Section 4.3.4)
GUI	–	Graphical User Interface	
HMM	–	Hidden Markov Model	
IMU	–	Inertial Measurement Unit	
IQR	–	Interquartile Range	
IT	–	Inferotemporal Cortex	
JS	–	Jensen-Shannon (divergence)	
KL	–	Kullback-Leibler (divergence)	
KS	–	Kolmogorov-Smirnov (test/statistic)	
M1	–	Primary Motor Cortex	
MCP	–	Metacarpophalangeal Joint	
MP	–	Matching Pursuit	
MTL	–	Medial Temporal Lobe	

OMP	–	Orthogonal Matching Pursuit	
PC	–	Principal Component	
PCA	–	Principal Component Analysis	
PIP	–	Proximal Interphalangeal Joint	
PMd	–	Dorsal Premotor Cortex	
PMv	–	Ventral Premotor Cortex	
PPCA	–	Probabilistic Principal Component Analysis	
ROI	–	Region of Interest	
SD	–	Standard deviation	
SEM	–	Sparse Eigenmotion	(Chapters 4 & 5)
SEMD	–	Sparse Eigenmotion Decomposition	(Chapter 4)
SS	–	Synchronous Synergies	(Chapter 3)
TVS	–	Time-Varying Synergies	(Chapter 3)
VAR	–	Variance	
VAF	–	Variance Accounted For	(Section 3.2)
VAF_L	–	Variance Accounted For locally	(Section 3.2)

Nomenclature

$atom$	–	Individual base used for reconstructing data	(see Chapter 4)
$dictionary$	–	Collection of $atoms$	(see Chapter 4)
$\mathbf{B}$	–	Dictionary of SEMs or PCs	
D	–	Dimensionality of the data	
Q	–	Dimensionality of the latent representation	
r	–	Pearson's product-moment correlation coefficient	
T	–	Length of the time-series	
T_w	–	Width of the ROI in the SEMD algorithm	(see Chapter 4)
$\mathbf{X}$, x	–	Observed data	
$\mathbf{Z}$, z	–	Latent representation	
θ	–	Joint position	

Notation and Other Conventions

We observe the following notation for mathematical symbols throughout this document: vectors are lower-case bold letters (e.g. $\mathbf{x}$), matrices the same but capitalised ($\mathbf{X}$). Italicised lower case letters (x) are either scalars or the realisation of a random variable, the meaning should be clear given the context. Italicised capital letters are either functions (in which case the parameters will be made explicit, as in $F(x|\theta)$) or constants. A variable dependent on time will be denoted $x(t)$ and its time-derivative $\dot{x}(t)$, though the explicit dependence on t may be dropped for clarity.

Statistical testing was performed rigorously and care was taken with underlying assumptions of statistical tests. Whenever a p-value is reported, the test performed is named immediately before or after the numerical value. To avoid too much repetition, this is done only when the test changes or at regular intervals (at least once per paragraph). If Students' t-test was used, an Anderson-Darling test was performed beforehand to check for normality. In convention with the literature, significance was defined at $p < 0.05$.

Related Publications and other Public Presentations

Publications

- Andreas AC Thomik, David Haber, A Aldo Faisal, et al. (2013). "Real-time movement prediction for improved control of neuroprosthetic devices." In: *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*. IEEE, pp. 625–628
- David Haber, Andreas AC Thomik, A Aldo Faisal, et al. (2014). "Unsupervised time series segmentation for high-dimensional body sensor network data streams." In: *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*. IEEE, pp. 121–126
- Michele Xiloyannis, Constantinos Gavriel, Andreas AC Thomik, and A Aldo Faisal (2015a). "Dynamic forward prediction for prosthetic hand control by integration of EMG, MMG and kinematic signals." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 611–614
- Michele Xiloyannis, Constantinos Gavriel, Andreas AC Thomik, and A Aldo Faisal (2015b). "Gaussian Process Regression for accurate prediction of prosthetic limb movements from the natural kinematics of intact limbs." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 659–662
- Andreas AC Thomik, Sonja Fenske, and A Aldo Faisal (2015). "Towards sparse coding of natural movements for neuroprosthetics and brain-machine interfaces." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 938–941
- Feryal MP Behbahani, Ruth Taunton, Andreas AC Thomik, and A Aldo Faisal (2015). "Haptic SLAM for context-aware robotic hand prosthetics-simultaneous inference of hand pose and object shape using particle filters." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 719–722
- Constantinos Gavriel, Andreas AC Thomik, Pedro Rente Lourenco, Sathiji Nageshwaran, Stavros Athanasopoulos, Anastasia Sylaidi, Richard Festenstein, and A Aldo Faisal (2015a). "Kinematic body sensor networks and behaviourmetrics for objective efficacy measurements in neurodegenerative disease drug trials." In: *IEEE Body Sensor Networks Conf.* IEEE

- Constantinos Gavriel, Andreas AC Thomik, Pedro Rente Lourenco, Sathiji Nageshwaran, Stavros Athanasopoulos, Anastasia Sylaidi, Richard Festenstein, and A Aldo Faisal (2015b). "Towards neurobehavioral biomarkers for longitudinal monitoring of neurodegeneration with wearable body sensor networks." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 348–351

Talks

- Andreas AC Thomik (2013), *Statistics and Structure of Human Movement*, UK Sensorimotor Meeting, Cambridge (UK), 02/12/2013
- Andreas AC Thomik (2014), *Symbolic Representation of Complex Action Sequences*, Hand, Brain & Technology Conference, Ascona (CH), 11/09/2014

Posters

- Andreas AC Thomik and A Aldo Faisal (2012), *Deriving Motion Primitives from Natural Hand Movements for Neuroprosthetic Control*, Bernstein Conference, Münich
- Andreas AC Thomik, Marie Tolkiehn, Ingrid Vella and A Aldo Faisal (2012), *Towards the Human Ethome*, Champalimaud Neuroscience Symposium, Lisbon (PT)
- Andreas AC Thomik and A Aldo Faisal (2013), *Symbolic Encoding of Complex Actions by Movement Primitives*, Neural Control of Movement Conference, San Juan, Puerto Rico (US)
- Andreas AC Thomik, Sonja Fenske and A Aldo Faisal (2014), *Similarity and Idiosyncrasy of Hand Kinematics in Daily Tasks*
- Andreas AC Thomik and A Aldo Faisal (2015), *Symbolic Representation of Complex Action Sequences*, COSYNE, Salt Lake City, Utah (US)
- Andreas AC Thomik and A Aldo Faisal (2015), *Sparse Encoding of "in-the-Wild" Hand Movements*, Society for Neuroscience, Chicago, Illinois (US)

1 Introduction

"We have now sunk to a depth at which the restatement of the obvious is the first duty of intelligent men."
— George Orwell

1.1 Human Movement: Research and Applications

The brain's primary function is arguably the generation and coordination of movement. Consequently, understanding the processes and computations underlying motor control has been a central question in neuroscience. More than 100 years after Sherrington's seminal work on reflex actions (Sherrington 1910) – and almost 50 years after Bernstein's description of coordination and regulation of movement (Bernstein 1967) – the question is still far from being resolved.

Our understanding of movement generation and control varies greatly depending on which level of hierarchy is being considered (Figure 1.1). The lower level is reasonably well understood: neurons from the (mainly) motor cortex innervate alpha motor neurons in the spinal cord, which transmit the signal to the neuromuscular junction of one or more muscles. There, the neurotransmitter acetylcholine is released which causes the muscle to contract which in turn generates torque around the joint and finally movement of the limb.

Coordination of multiple muscles or joints is less well understood. This is maybe

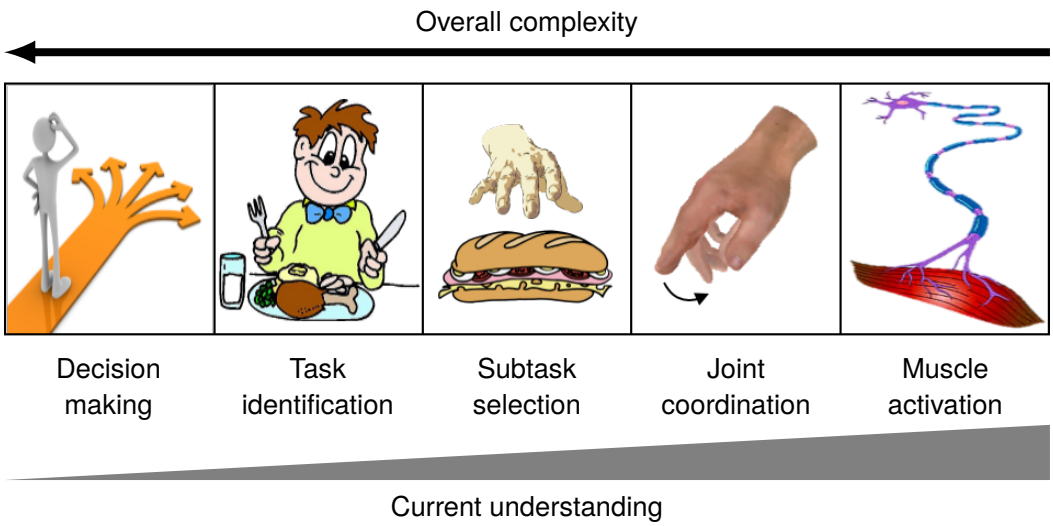

Figure 1.1: Natural movement can be schematised in decreasing stages of complexity. Decision making and task identification represent the two highest stages of movement planning. Subtask selection directly influences the decision of how to coordinate joint movement. This is executed via muscle contractions which are controlled by neurons in the spinal cord. While this last stage is well understood from a physiological point of view, there is still much debate regarding higher-level processing.

not surprising if we consider that the musculoskeletal system is highly non-linear and contains more than 600 muscles which are organised in a highly redundant fashion. How then does the brain achieve the feat of controlling such an immensely complex system so rapidly and seemingly effortlessly? Two theories emerge:

Optimal Feedback Control (OFC; Todorov and Jordan 2002) stipulates that to achieve a goal, the brain minimises a cost function (typically considered to be the squared target distance and energy expenditure). This theory implies that the brain must simulate possible muscle activation patterns in order to find the one which optimises the cost function. This makes a strong assumption on the existence of an accurate model of the body's dynamics in the brain, and supposes that the optimisation may be performed in real-time. Interestingly, this theory results in what is known as the "uncontrolled manifold" or "minimum intervention principle" where task irrelevant variables are

not controlled (e.g. when grasping a stick, the hand's position along the length of the stick does not matter particularly and is thus allowed to fluctuate). Conversely, any particular task will result in a stable muscle activation pattern which needs to be maintained to achieve the goal.

On the other end of the spectrum, the Muscle Synergy approach assumes that these activation patterns are not a consequence of the control method but rather a low-dimensional manifold on which the brain can act (d'Avella, Saltiel, and Bizzi 2003). Accordingly, the problem is not so much to optimise muscle activity globally but rather to select from one (or more) of many possible synergies which by themselves constrain the motor output. This approach has a number of shortcomings. However, its main critique is that the synergies observed from decomposition of movement data may simply be an artefact of the analysis method. In particular, synergistic muscle activity seems to occur even if the movement is not executed voluntarily but instead by directly manipulating the tendons (e.g. Valero-Cuevas, Venkadesan, and Todorov 2009).

Most research in motor control does not consider higher levels of processing. However, it seems obvious that the sequences in which tasks are being performed – though not necessarily critical to achieving one's goal – will influence the configuration of the body and consequently requires adaptation of the motor command to be executed.

A major shortcoming from most of the aforementioned research is that it almost exclusively considered simple movements (i.e. centre-out reaching tasks or similar) performed in laboratory settings. It thus artificially constrains the potential output of the motor system and consequently may lead us to see features which are not present or miss some. In particular, it prevents us from understanding how complex tasks are planned and executed. Are they monolithic or does the brain break them down into individual motor programs? In the latter case, how are these sequences organised? Yet, despite the limitations caused by laboratory experiments, the findings are widely applied in real-world settings:

In the clinical setting, the ability to perform certain movements is widely used for assessing the progress of disorders or rehabilitation, in particular when neurological damage is involved. These include the Fugl-Meyer Assessment of Sensorimotor Recovery After Stroke (FMA; Fugl-Meyer et al. 1974), the Unified Parkinson's Disease Rating Scale (UPDRS; Goetz et al. 2007) and the Spinocerebellar Ataxia Functional Index (SCAFI; Schmitz-Hübsch et al. 2008) to name but a few. These scales typically involve comparing the patient's ability to perform an "Activity of Daily Living" (ADL) with a baseline or with a control population. Yet, while hundreds of different ADLs have been proposed, the properties of natural human behaviour have yet to be mapped. This has two negative consequences: (1) a clinician rating the subject's ability to perform a task is highly subjective and (2) functional recovery might mask (possibly nefarious) compensatory mechanisms developed by the patient. Being able to compare natural movement with a baseline would thus be a major step towards creating an objective assessment of disease and recovery.

State of the art hand prostheses have the ability to perform a range of different grasp types, with the user required to choose the grasp to perform manually (Farina et al. 2014). The choice of available hand motions includes some of the standard grasps (e.g. power grasp, pinch, etc.) which have been identified through painstaking observation and manual classification (e.g. Cutkosky 1989). Due to the subjective nature of this procedure, the grasps it identifies may only represent a subset of possible hand configurations. Furthermore, it does not directly quantify the movements necessary to move from one pose to the next, but instead assumes that all movements start from an arbitrarily decided "neutral" pose.

A detailed analysis of natural hand movements would also be of particular interest because of their evolutionary significance. The shape of the modern human hand is thought to be intrinsically linked with our ancestor's development of stone tools (Napier 1962; Marzke and Marzke 2000). Relatively short fingers and a proportionally

longer thumb have been described as key features required to achieve a stable grip on a stone for shaping it (Napier 1962).

Hand shape, however, was probably not the most important factor in allowing early humans to design ever more complex tools. The crude Oldowan stone flakes (~2.5 million years ago) were produced both by our predecessor *Homo habilis* as well as the more primitive (and smaller brained) *Australopithecus*. Conversely, the much more sophisticated Acheulean stone axes (~1.5 million years ago) are the product of *Homo erectus* which is known to have a much larger brain volume than its ancestors (Ambrose 2001). Along the same line of reasoning, Faisal et al. (2010) demonstrated that the complexity of hand postures required for producing Acheulean tools is not significantly different from the ones required for the simpler Oldowan flakes. This leads the authors to conclude that the ability of performing the complex task sequences and force modulations required for the production of Acheulean stone axes is a consequence of a more evolved brain. The study of complex movements may thus give us direct insights into our evolution and the development of the brain.

Given the scientific and technological importance of understanding human movement, it seems surprising that research has mainly been focused on laboratory experiments. Human movement in ecologically valid settings is highly variable, which reflects our versatility in interacting with numerous different environments. This large variability, however, requires novel tools for analysing the data. We no longer have the luxury of averaging over repeated identical trials. Instead, new tools are required to identify interesting movements and comprehensively map the full extent of human movement.

1.2 Thesis Structure

Our work attempts to tackle some of these challenges. In Chapter 2 we introduce an experimental paradigm for collecting kinematic data from healthy or diseased human subjects in unconstrained settings. This overcomes the risk of overlooking important aspects of human motor control by experimentally constraining the subject's movement. We describe the statistics of natural movement in order to identify invariant features of human movement. These may be used as a baseline against which disease progression may be measured. Additionally, we test some of the typical assumptions in motor control for their validity in ecologically valid settings.

Chapter 3 goes further by looking into the structure of movement, both from a spatial and a temporal perspective. In particular, we highlight the problem of collecting data in constrained laboratory settings by contrasting results from a laboratory study with those from studies in natural environments. From an application perspective, we propose a new method for enhancing prosthetic control by exploiting information provided by the inherent structure of natural movement. Finally, we review current hypotheses of low-dimensional control strategies in the hand and propose a new approach inspired by work in the sensory system.

A key issue when dealing with large and unstructured datasets such as the ones we collected from freely behaving subjects is that most classical analysis techniques are not appropriate for investigating them. Consequently, Chapter 4 presents an algorithm we developed for this type of data, specifically to investigate our hypothesis laid out in the previous chapter. The method uses principled ways to identify meaningful local patterns in the stream of data. We validate our algorithm and compare it to state of the art applications.

This methodology is applied to hand movement data in Chapter 5. We show that our representation (termed *sparse eigenmotions*) is more efficient at representing the

data than current methodology, both for typical laboratory experiments and for unstructured data. We demonstrate the existence of higher-order structure in this representation, which points to a potential language of movement. Finally, we show that knowledge of these local structures vastly improves the ability to predict hand movements (and thus prosthetic control) by exploiting local structures.

Finally, Chapter 6 offers an outlook into the future potential of the methods presented here and summarises the questions raised by this study, as well as possible ways of answering them.

2

Human Ethomics

"Never confuse movement with action."

— Ernest Hemingway

Acknowledgements

The methods and data presented in this chapter are the fruit of collaboration between several people. As part of their Master's project, Marie Tolkiehn and Ingrid Vella designed the basic setting for data recording and collected motion capture data from over 40 subjects. William Abbott was of great help in extending the original paradigm to include collection of hand and eye-tracking data, as well as sharing the fun of collecting data from 8 additional subjects.

2.1 Introduction

The developing field of *Ethomics* combines classic ethology, i.e. the study of animal behaviour in natural settings, with the potential given by modern genomics in the hope of linking behavioural traits with gene expression. As such, the main subjects of the research have typically been small organisms whose genetic material is either well known or can be easily manipulated (Branson et al. 2009; Dankert et al. 2009).

On the other end, what we term *Human Ethomics* aims to quantify natural human behaviour, in particular by measuring *perception* (sensory inputs) and *action* (movement) in ecologically relevant settings. We move away from descriptive ethology towards a highly quantitative measure of human activity, opening up the potential for detailed studies of sensorimotor function, e.g. to produce novel biomarkers for ageing and disease.

Studying motor behaviour in a laboratory setting has given us many important insights into sensorimotor adaptation and integration in humans and primates (see Ingram and Wolpert (2011) for a review). However, laboratory tasks are, by their very nature, limited in their ecological validity. Thus, tasks which may appear impossible in the laboratory may well be achieved in natural settings. For example, Rhesus monkeys were seemingly incapable to use the experimenter's gaze to determine which prop to pick up (Anderson, Montant, and Schmitt 1996). Yet, a later study showed that similar animals would rather steal food from the experimenter who was looking away than the one watching the food (Flombaum and Santos 2005). The same ability was shown in chimpanzees when competing over various food resources with a more dominant conspecific: the non-dominant chimpanzee would go for food it knew the dominant one could not see (Hare et al. 2000; Hare, Call, and Tomasello 2001). Therefore, the monkey's ability to use a specific skill was related to the ecological validity of the context in which it had to use the skill. Similarly, researchers studying human cognition have recognised the necessity to move away from the rigid and featureless laboratory environment and study human behaviour in ecologically relevant contexts (Kingstone, Smilek, and Eastwood 2008). A prime example of this approach is the study by Land and Lee (1994), in which they studied eye movements of humans during a real-world driving task. Having studied the human behaviour in a natural environment, they then went on to copy some aspects of real driving in a laboratory setting to study the importance of the field of view (Land and Horwood 1995). These studies identified

that, contrary to prior belief, both near field (close to the car) as well as far-field vision are required for safe driving at medium to high speeds.

While it may be argued that cognitive tasks are more likely to be influenced by the context, studies show that this is equally true for motor tasks. Howard et al. (2012) showed that two conflicting motor tasks can be learnt simultaneously, as long as the context is suitably modified for each task, akin to our ability to lean the manipulation of various different tools. Bock and Beurskens (2010) studied age-related changes in walking both in the lab and in the park and showed there is an increased variability in step duration with age in the laboratory setting, but not in subjects freely walking in a park. Similarly, hand movements as part of a larger task were significantly different from isolated movements (Bock and Hagemann 2010; Bock and Züll 2013). These findings make a strong case for studying human motor behaviour in settings where subjects can move and behave freely.

Quantitative studies of human behaviour "in-the-wild" have often relied on indirect observation, such as diaries kept by the subjects themselves (Schlich and Axhausen 2003) or from observers shadowing the subject (Kilbreath and Heard 2005). Both studies, however, lack objective data and are sensitive to subjects correctly reporting results and the observer being capable of registering all the information rapidly enough. The development of cell phones, and in particular smartphones with a large range of sensors such as accelerometers, gyroscopes, magnetometers and GPS has made it easier to unobtrusively measure human behaviour in natural settings. Combined with body-worn sensors which can stream data directly to the phone or an independently carried recording unit, this can give us information about not only the subjects' location, but also their real-time activity[1].

[1] As example, (Győrbíró, Fábián, and Hományi 2009) use a phone and three sensors attached on ankle, wrist and hip to classify the subjects' data between six different activities (resting, typing, gesticulating, walking, running and cycling) with a neural network. While their system performs quite well, the activities do have very different kinematic signatures.

Although new technology has enabled the scientific community to collect kinematic data from humans in natural environments, most of the motion capture equipment which is commercially available – and has been available for a while now – is designed with a laboratory setting in mind. Though it is still possible to collect ecologically relevant data (e.g. the Carnegie Mellon University Motion Capture Database[2]), subjects are highly constrained in the space in which they can move. In addition to laboratory based tasks, at least two studies modified equipment to make it usable by subjects during their everyday life: Ingram et al. (2008) used a sensorised glove connected to a laptop and a battery in a backpack to record the movement of subjects' right hand in completely unconstrained environments over periods of around $3\,h$. From the same research group, Howard et al. (2009) adapted a magnetic positioning system to record the movement of the subjects' arms. Both studies confirmed findings from previous laboratory experiments, such a the symmetry bias for arm movements (Kelso 1984) or the dimensionality of hand movements (Santello, Flanders, and Soechting 1998), but fall short of identifying novel features specific to natural behaviour.

More recently, Body Sensor Networks (BSNs) have emerged as a way of measuring human movement in an unobtrusive way. These are typically small cases including a 3-axis accelerometer, 3-axis gyroscope and 3-axis magnetometer (thus called a 9-axis sensor) which are attached to the subject's body and can record its movement over a prolonged period of time, storing it onboard or streaming it to a base-station. Additional sensory modalities, such as skin galvanometry, surface electromyography (sEMG) or mechanomyography (MMG) (Fara et al. 2013; Gavriel and Faisal 2014) may be combined into these sensor blocks, termed nodes (as in nodes of a network). These sensor systems have been independently developed by a number of research groups, and some have recently become commercially available. Thanks to their small size and weight, these sensors are ideally suited for measuring human movement and

[2] Available at http://mocap.cs.cmu.edu/ & http://kitchen.cs.cmu.edu/

physiological parameters (respiratory and heart rate (Teichmann et al. 2014) and or metabolic rate (Bourke et al. 2014)) during everyday activity.

In the medical field, BSN nodes are being studied to augment the normal diagnostic protocol by providing objective data about the patient's movements during standard clinical test conditions, such as the $10\,m$ walk for Parkinson's disease (Giuberti et al. 2014). Despite promising advances, results obtained from BSN-augmented studies need to be taken with caution. Indeed, because of their recency, a significant amount of research is still going into refining sensor standards and data collection protocols. Currently, sensors are frequently attached at different locations on the limbs and consequently the data may be difficult to compare, especially if the location is not precisely recorded. Such limitations will undoubtedly disappear with the more wide-spread acceptance of such equipment in clinical environments and its commercial availability.

In the following, we present our work on the collection of human behavioural data in unconstrained settings. We collect the data using a variety of inertial or mechanical sensors, thus allowing the subject complete freedom over the space in which they move and avoiding typical problems of optical systems such as sensor occlusion. A special feature of this semi-unconstrained data collection paradigm is that it allows us to annotate the subject's behaviour online using custom software. Video recordings are also available for subsequent, detailed offline analysis. We first present the data collection environment along with the motion capture equipment and software used. We then illustrate the methodology with some representative data and analyse its potential as a baseline to assess natural and disease-related motor variability. We test some assumptions and findings of constrained motor tasks to verify their relevance to natural motor control. Limitations of the method are described in the discussion along with further analyses of our findings.

2.2 Materials and Methods

2.2.1 Subjects

A total of 62 healthy subjects (18 female, age 19–34) were recorded in various experiments. All subjects were naïve to the purpose of the experiment and all experimental procedures were approved by the Imperial College Ethics Committee.

2.2.2 Settings

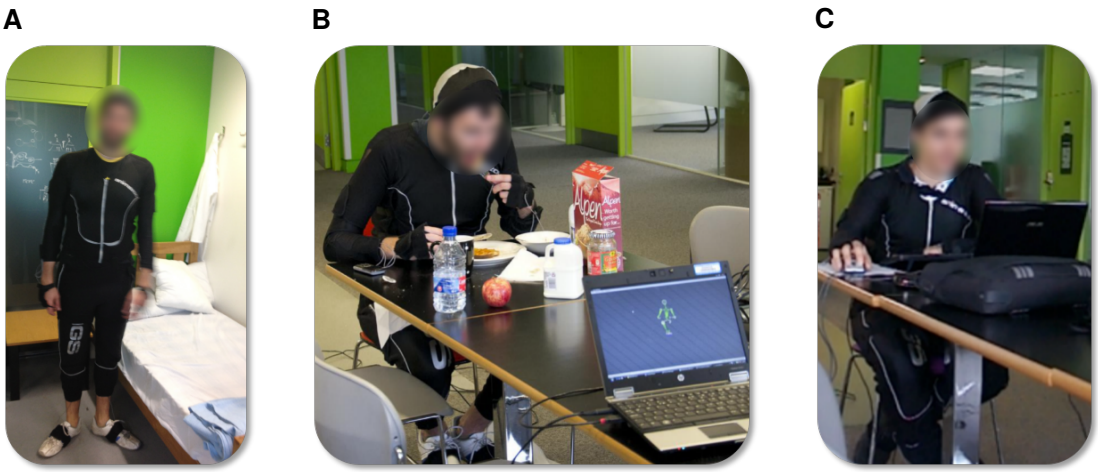

Figure 2.1: Settings constructed to emulate the subjects' natural environment. **A** A bedroom complete with bed, covers and clothes to put on over the motion capture suit. **B** A kitchen/breakfast setting for subjects to prepare and consume food. **C** An office environment with various computer and paper related tasks. The metal table and chair depicted here were later replaced by plastic furniture to avoid magnetic interference.

We recorded subjects in three different settings emulating their natural life (Figure 2.1): (1) a kitchen/breakfast setting, (2) a bedroom setting and (3) an office setting. These recording areas were set up inside the laboratory (bedroom) or in a communal area. In each setting, subjects were given a number of tasks from their daily life to perform (e.g. set the table, have breakfast) but were free to achieve these tasks in any way they liked, thus enabling us to collect naturalistic motion data. On average, we

collected 30–40 min of data for each subject and setting.

2.2.3 Movement Data Acquisition

The subjects' behaviour was recorded using a variety of equipment, although not all equipment was used simultaneously. We will refer to each different recording method as a *recording modality*. In addition to the modalities described below, some subjects also wore an eye-tracker to record their gaze. The data and recording method are extensively described in William Abbott's Ph.D. thesis and are not further discussed here.

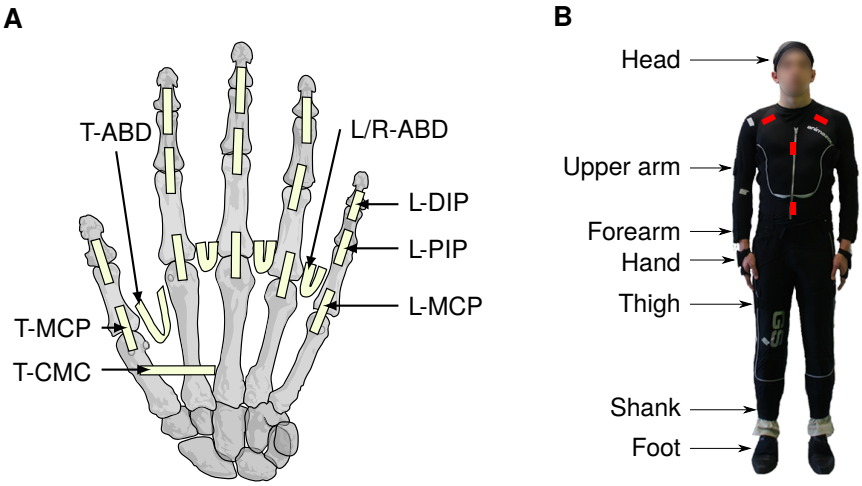

Figure 2.2: A Sensor placement for the right-hand Cyberglove. The left glove is similar but lacks sensors on the DIP joints. **B** Subject wearing the IGS-180 suit. Arrows indicate sensor placement, which is identical on the left hand side. Sensors on the hip, spine, and clavicles are on the back of the suit (red boxes). Figure **B** modified from (Gavriel and Faisal 2013).

Hand Data

The data from the subjects' left and right hand were recorded using a CyberGlove I and III respectively (CyberGlove LLC, San Diego, CA, USA). The CyberGloves are fabric gloves with stretch sensors placed over the joints of the fingers (see Figure 2.2**A** for

sensor placements). Note that the right hand glove also measures the movement of the distal interphalangeal (DIP) joints, while the left hand does not. Moving the fingers bends the sensors and thus modifies their electrical resistance, which can be measured. The hand movement is recorded at 140Hz and 90Hz for the left and right hand respectively, and converted using an 8-bit analogue-to-digital converter (ADC). This results in in an effective worst-case resolution of about 0.5°. Calibration is performed as a two stage-process: an initial calibration using software provided by the manufacturer and subsequent refinement by manually adjusting sensor gains and offsets while comparing the subject's hand with a visual rendering of the hand. The recorded data was streamed to a laptop where it was saved along with a time-stamp obtained from the system clock.

Body Data

The movement from the entire body (excluding fingers) was recorded using an IGS-150 or IGS-180 motion capture suit (Animazoo UK Ltd, Brighton, UK). The suits are mainly used in the movie and video-game industry to record the movement of characters which would be difficult or impossible to animate artificially. A suit consists of a pair of Lycra trousers and jacket with sensors embedded to measure the movement of various limbs (see Figure 2.2**B**). The sensors are classic 9-axis inertial measurement units (IMUs) and a complete data set (i.e. values from all sensors) are streamed to the laptop at 60Hz. The only noteworthy differences between the two models of the suit used are that the IGS-180 has two additional sensors placed behind the clavicles and is capable of streaming data wirelessly, potentially allowing for more natural movement by the subject. Calibration of the motion capture suit is done by a simple routine provided by the manufacturer as part of the control software.

The data provided by the suit comes in BVH format which saves each joint's position as Euler angles with rotation order ZXY (front, right, up from the subject's perspec-

tive). Because the rotations are not commutative and the rotation order is unlike the motion of our limbs, this encoding is difficult to interpret in biomechanical terms. To address this, we transform the rotation angles into biomechanically meaningful values by representing the Euler angles as quaternions and then mapping them back to Euler angles with different rotation orders as recommended by the Interational Society of Biomechanics (Wu et al. 2002; Wu et al. 2005).

Scene Recording

The working area of the subjects was monitored externally using a standard video recorder. The camera was placed on a tripod in a way to observe most of the subjects' movements without requiring frequent repositioning. The data thus recorded is not explicitly evaluated but serves as reference to compare data recorded from the suit and gloves, as well as assisting post-experimental data annotation.

Data Preprocessing

Prior to further analysis, the raw data from suit and gloves was synchronised to a common time vector at $100\,Hz$ by linearly interpolating between available samples using the spherical-linear interpolation method (Shoemake 1985). The position data thus obtained was filtered with three passes of a $110\,ms$ moving average filter to remove noise and artefacts due the ADC.

2.2.4 Data Annotation

The subjects' behaviour was annotated online by the experimenter using one of two custom-written tools implemented in Matlab (The Mathworks Inc., Nattick, MA). The first, termed *logING* has been extensively discussed in an M.Sc. thesis (Vella 2012) and will not be described in detail here.

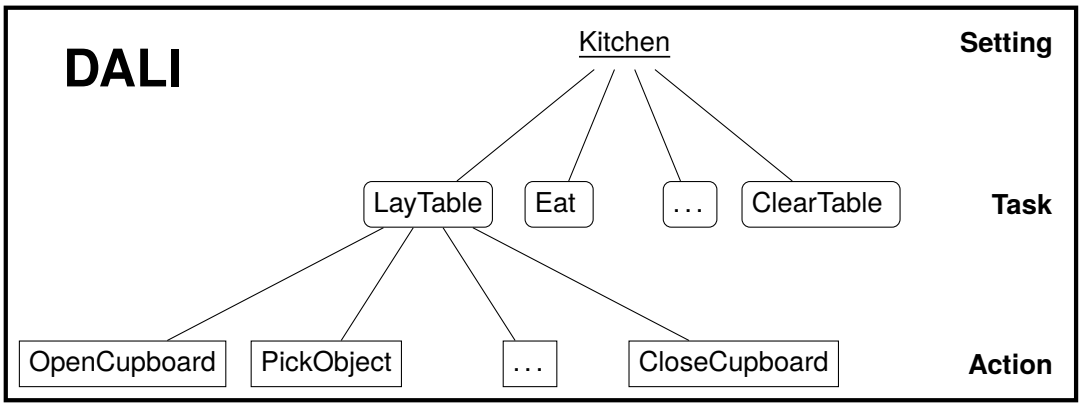

Figure 2.3: Example annotation using the Dynamic Activity Logging Interface. Each *setting* allows for a number of different *tasks* which may be performed using a sequence of predetermined *actions*.

Dynamic Activity Logging Interface

A major drawback of *logING* is the requirement for the experimenter to remember the mapping between keyboard short-cuts and actions or to look them up while they are annotating the data, thus running the risk of voiding the initial work of creating a large database of potential actions. Further, the number of possible actions is inherently limited by the number of keys on the keyboard. To address these shortcomings, we developed the **D**ynamic **A**ctivity **L**ogging **I**nterface. DALI simplifies the process of online annotation by introducing two features:

Graphical User Interface Instead of remembering countless actions and their corresponding keyboard short-cut, DALI displays all possible actions on a grid of clickable buttons. In the event an action is not available, e.g. because the experimenter had not thought about it before, a special "NEW" action allows to register the time-stamp of manual annotations.

Hierarchical Structure While the number of actions a subject can perform is virtually endless, they are intrinsically tied to the setting they are in. For instance, the action

"Sit down on bed" may occur multiple times when the subject is in the bedroom, but is totally irrelevant when the subject is in the kitchen. To take this into account, DALI introduces a hierarchical annotation structure where the annotator first identifies the setting the subject is in (e.g. "Kitchen"), followed by the task the subject is currently executing (e.g. "Clearing Table"). The list of actions available is then tailored to the specific task and setting (see Figure 2.3).

Unlike *logING*, DALI does not provide the ability to record start and end times of an action. This follows from the analysis of over 40 recordings made with *logING* which suggest that the annotator has rarely the time to record both start- and end-times of events.

2.2.5 Modelling

All analysis was performed using Matlab. Unless otherwise stated, optimisations were performed using an interior-point algorithm implemented by the `fmincon` function.

Position Data Fitting

To obtain a parametric distribution of the joint position data and enable outlier rejection, we fitted the data obtained for each joint and movement direction (i.e. flexion, pronation, etc.) with a variety of the following distributions (and mixtures thereof):

- the normal distribution: $\theta \sim \mathcal{N}(\mu, \sigma^2)$;
- the generalised extreme value distribution: $\theta \sim \mathcal{GEV}(\mu, \sigma, \xi)$;
- the "negative" generalised extreme value distribution: $-\theta \sim \mathcal{GEV}(\mu, \sigma, \xi)$; and
- the continuous uniform distribution $\theta \sim \mathcal{U}(-180, 180)$.

The $\mathcal{GEV}$ distribution was added to this selection as it enables us to capture skewed distributions and long tails. Alternatives (e.g. the skew-normal distribution) are

possible but lack a closed-form analytical expression. As the raw data contains a considerable number of erroneous data points (e.g. due to magnetic interference of large metallic objects with the suit sensors), we estimated the distribution parameters in multiple steps. (1) A mixture distribution was fitted which always contained a uniform distribution spanning all possible values of the data. For the best fitting distribution (as measured by the Bayesian Information Criterion (BIC)), the data points which were more likely to come from the uniform distributions than from the other mixture components were removed for the next step. (2) The uniform distribution was removed from the mixture and the parameters re–estimated on the cleaned data set. (3) We reconsider the entire data set, excluding only data points whose log-likelihood was less than −5.83 and estimate the maximum likelihood parameters on this final data set[3].

The choice of distributions to add to the mixture, as well as initial parameter guesses were chosen by visual inspection of the data. At each step of the fitting, the optimal parameter estimate was obtained by maximising the likelihood function.

Velocity Data Fitting

In agreement with general observation and similarly to previous authors (Morasso 1981; Abend, Bizzi, and Morasso 1982; Flash and Hogan 1985), we assumed that the velocity profile at the joints $\dot{\theta}(t)$ is given by a bell-shaped curve. More complex movement are simply considered to be a superposition of multiple such curves. Thus, any particular movement can be described by:

$$\dot{\theta}(t) = \sum_{i=1}^{N} \phi_i(t - \tau_i) \tag{2.1}$$

where N is the number of superimposed movements, τ_i the delay between movements

[3]This is the likelihood of a data point 5σ away from the mean in the normal distribution.

and ϕ_i is approximated by:

$$\phi_i = \alpha_i \exp\left(-\frac{t^2}{b_i^2}\right) \tag{2.2}$$

where α_i and b_i respectively modulate the amplitude and duration of the movement. Consequently, the probability of a single movement without superposition is given by:

$$P(\dot{\theta}_i|\alpha, b) = C^{-1} b \sqrt{\ln(\alpha) - \ln(\dot{\theta})} \qquad s.t. \qquad \dot{\theta} \in (0, \alpha] \tag{2.3}$$

where $C = \frac{1}{2}\alpha b\sqrt{\pi}$ is a normalisation constant. Note that after normalisation, b disappears from Equation 2.3, thus making the probability of any given velocity independent of the movement duration. Further, under the assumption that $\tau_i \gg b_i$, the probability distribution over the velocity is simply given by:

$$P(\dot{\theta}) = \sum_{\alpha_i \in A} \omega_i P(\dot{\theta}|\alpha_i) \qquad s.t. \qquad \sum_i \omega_i = 1 \tag{2.4}$$

Here, A is the space of all possible amplitudes and ω_i the probability of a movement with peak velocity α_i. The assumption that $\tau_i \gg b_i$ is obviously problematic, as it imposes that the movements are clearly separated from each other. The impact of this assumption will be analysed in the discussion. To estimate the values of ω for the observed velocities, $\dot{\theta}_o$, we used the following procedure: (1) we picked 50 values for α, logarithmically spaced between 0.06 deg/s and $\lceil \dot{\theta}_o \rceil$. (2) The corresponding 50 initial values of ω_i were set to 1/50, and (3) we optimised the values of ω_i by minimising the cost function given by:

$$E = \int_0^{\lceil \dot{\theta}_o \rceil} (F(\dot{\theta}_o) - F(\dot{\theta}))^2 d\dot{\theta} + \lambda \| \sum_{i=2}^{M} \omega_i - \omega_{i-1} \|_2^2 \tag{2.5}$$

where F is the cumulative distribution function (CDF) of P, λ is a regularisation

parameter (see below), and $\sum_{i=2}^{M} \omega_i - \omega_{i-1}$ imposes smoothness onto the weights.

To determine the best value of λ, we randomly picked a data set and optimised Equation 2.5 with 100 λ values logarithmically sampled between 10^{-3} and 10^3, as well as $\lambda = 0$. For each $\boldsymbol{\omega}$ thus obtained, we computed the Kolmogorov-Smirnov (KS) statistic between $F(\dot{\theta}_o)$ and $F(\dot{\theta})$. The quality of the fit remained largely unchanged for increasing values of λ until it started deteriorating for λ between 0.1 and 1. To determine the ideal value of λ, we adapted the method of Cheung et al. (2005) to select the point at which the KS statistic starts increasing. The corresponding value of λ, 0.7, was used for the optimisation of all other velocity fits as it provides a good balance between accuracy of representation and smoothness of the function.

2/3 Power Law

We examined how the relationship between movement velocity and curvature holds in natural behaviour by fitting a power law between the two variables. As shown previously (Schaal and Sternad 2001), the resulting fit parameters vary greatly depending on whether one fits a power law to the raw data using a non-linear least squares method or by fitting a straight line to the data plotted on a log–log axis. To maintain comparability with previous studies, we report values from both methodologies. An additional difficulty which arises from analysing the relationship in natural movement data is that the original formulation of the 2/3 power law (Lacquaniti, Terzuolo, and Viviani 1983) assumes that the scaling coefficient of the power law, k, is only piecewise constant. Thus, fitting a power law to the entire data set may give erroneous results by confounding the effect of k. To avoid this, we fitted a power law to non-overlapping 500 ms (50 samples) windows of data and evaluated the R^2 of the fit to decide whether the scaling coefficient was approximately constant within the window. Only windows with an $R^2 > 0.9$ were considered for further analysis.

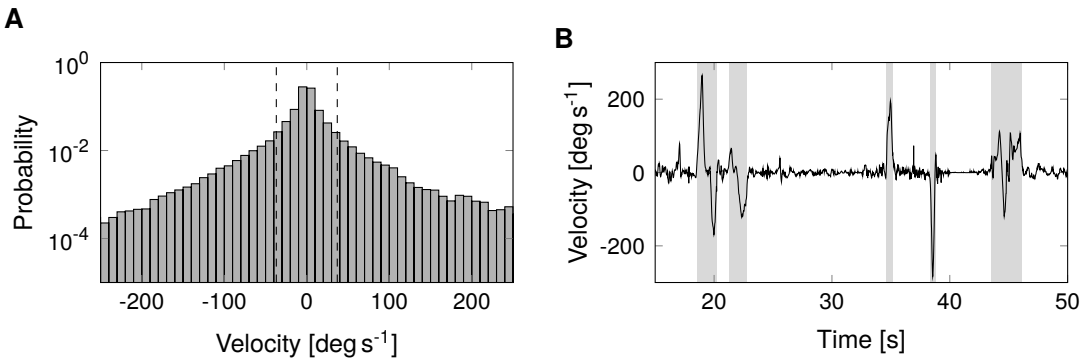

Figure 2.4: Movement detection example on flexion movements of the right elbow. **A** Empirical distribution of joint velocities and computed thresholds (dashed lines). Note the logarithmic scale on the y axis. **B** Result on the time series. Areas shaded in grey are those for which a movement is detected. Zero-crossings are bridged by the morphological closing operation.

2.2.6 Movement Decomposition

We identify movements by decomposing the time-series of each joint angle into periods of *movement* and *no movement*. A *movement* segment is presumed to be one continuous action. To identify the beginning and end of movements, we compute the empirical probability distribution of the velocity data from a histogram with 10^4 evenly spaced bins over the range of the data. The data points with an empirical p-value < 0.3 were considered to be movement. To mitigate the effect of noise, bridge zero-crossings and capture the entirety of the movement, we perform a series of morphological opening and closing operations with increasing window size (3, 11, 11 and 55 samples respectively) on the binary classification. An example of the ability of this method to decompose a time-series into areas of movement and stillness is given in Figure 2.4.

2.2.7 Movement Plane Orientation

To analyse the planarity of hand movements, we identified start- and endpoint of each movement using the aforementioned method and performed PCA on the 3-

dimensional trajectory data in between. If the first PC accounted for less than 90% of the variance (i.e. the movement was not deemed linear), we define **n**, the normal vector to the movement plane as the third PC. As the sign of the **n** does not matter, we ensured that they were all pointing into the northern hemisphere prior to further analysis. To identify preferred orientations, we first projected **n** onto the surface of a unit sphere before unwrapping it using the Lambert azimuthal equal–area projection. As the name suggests, this type of mapping preserves areas and thus point density. Preferred orientations were identified by comparing the point cloud obtained in this way with a model of complete spatial randomness (CSR). In particular, we compare the so-called G-function of inter-point distances (Bailey and Gatrell 1995, pp. 83–110):

$$G(w) = \frac{\#(w_i \leq w)}{n} \tag{2.6}$$

where $\#(w_i \leq w)$ indicates the number of inter-point distances less or equal to some value w and n is the total number of points in the area. To be able to tell whether the observed function G was significantly different from a purely random pattern, we performed 2000 simulation runs by computing G with uniformly randomly distributed data.

2.3 Statistics of Natural Movements

We analyse the natural movement data under a number of different aspects. First, we present the raw probability distributions for joint position and velocity. To assist with the evaluation of specific recordings, we compare individual recordings of healthy people with the grand average, thus qualifying the individuality of movement. Next, we show the relationship between joint position and instantaneous velocity, demonstrating modulation of the velocity bandwidth by position. Finally, we examine the existence of the 2/3 power law (Lacquaniti, Terzuolo, and Viviani 1983) in natural movement.

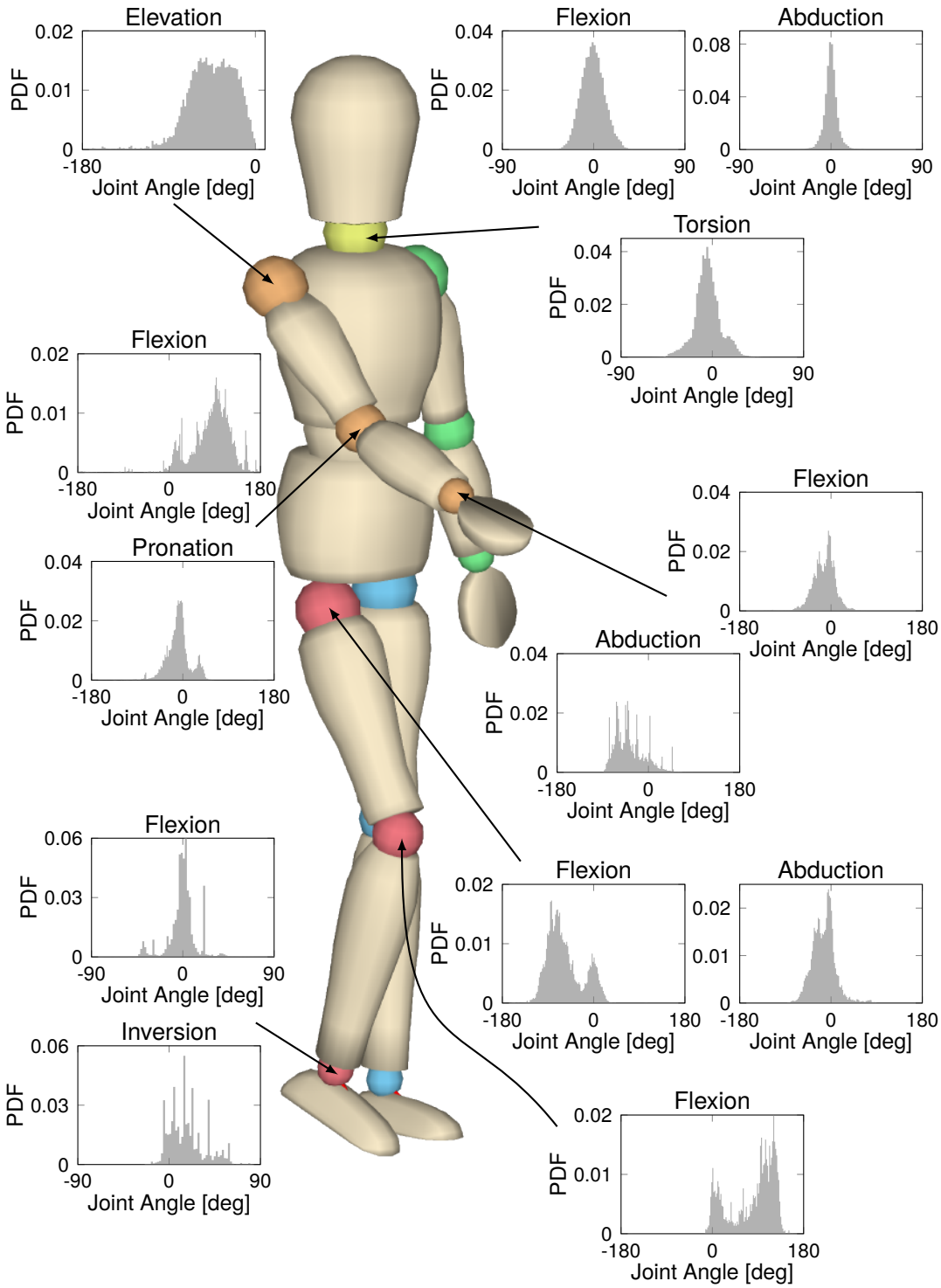

Figure 2.5: Probability distributions of joint positions.

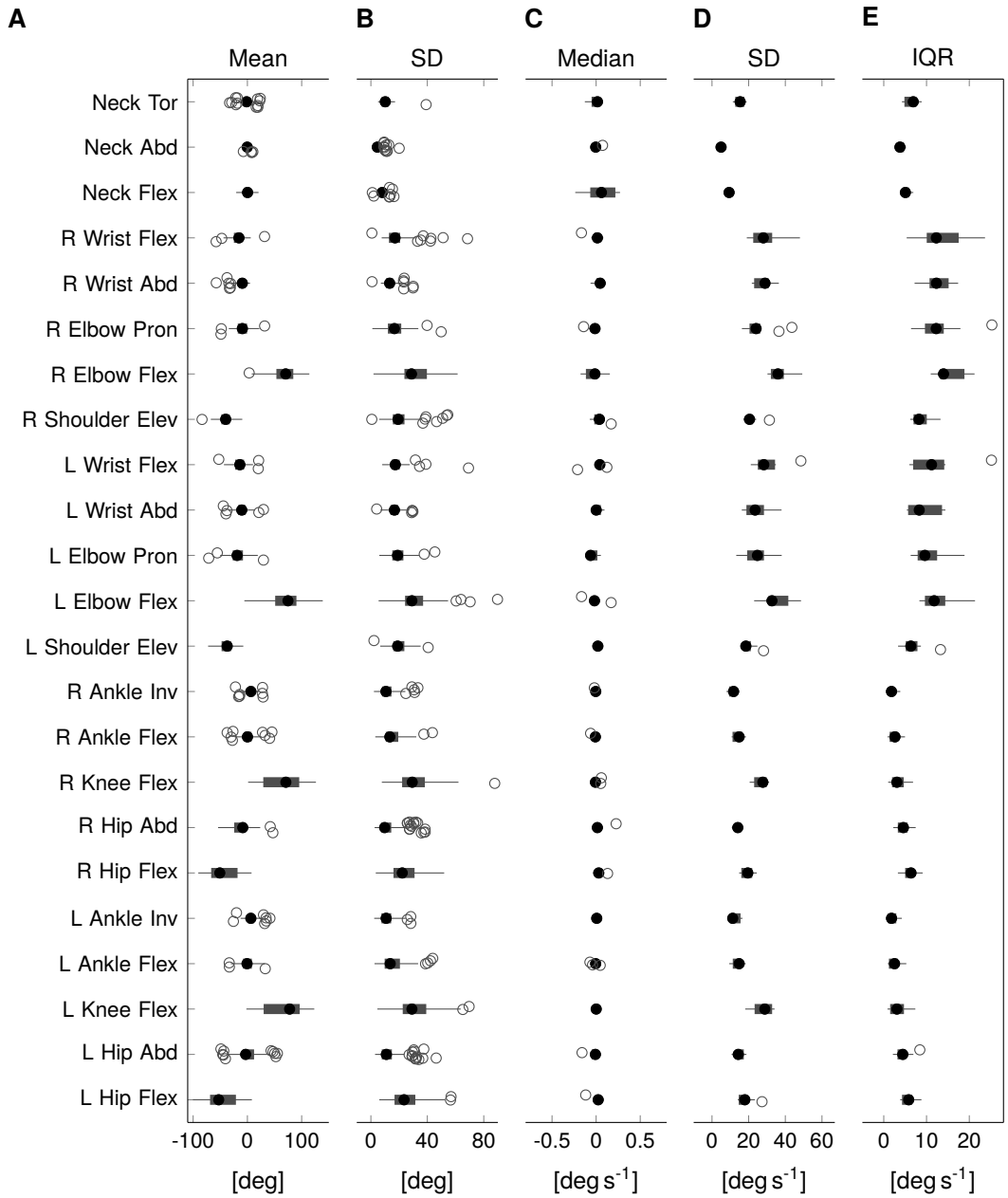

Figure 2.6: Summary statistics of natural movement data for all joints considered. **A** & **B** Mean and standard deviation of joint position respectively. **C** – **E** Median, standard deviation and interquartile range of velocity distributions. Note the large difference between SD and IQR which is characteristic of a heavy tailed distribution.

2.3.1 Position

The distributions of joint angles computed from the aggregate data of all recordings are shown in Figure 2.5 for the right side of the body (the statistics for the left are comparable). Outliers were removed according the the procedure outlined in Section 2.2.5. Unsurprisingly, knee flexion, hip flexion and elbow flexion show very marked bimodal distributions characteristic of sitting and standing. The other distributions all deviate from normality by either being skewed or displaying an substantial excess kurtosis[4] resulting in a distribution with higher probability density around the mean. The deviation from normality is also found for each individual mode of the bimodal distributions: the best fit for each mode of the distribution was found to be a generalised-extreme value distribution which is characterised by a heavy tail and a marked excess kurtosis.

While the average distribution of limb position may be interesting to understand the general control of the body, we wanted to know how much individual samples recordings varied from the mean. This baseline deviation may eventually help identifying abnormal deviations from normal behaviour caused by injury or disease. To do so, we compared the individual recordings using two metrics: (1) the Jensen-Shannon divergence (JS, a symmetrised version of the Kullback-Leibler (KL) divergence) and the average log-likelihood of individual data points. The results are reported in Figure 2.7 with different colour codes for each part of the body measured. Note that the two metrics do not measure exactly the same property of the movement. The Jensen-Shannon divergence measures the difference between the two distributions. Thus, someone standing still manipulating an object with his hands will have a very large JS divergence value for the lower limbs as their distribution will be almost singular and thus very different from the global average. Conversely, the average log-likelihood looks at every data point individually and computes how likely that observation is given the

[4]This is defined as the kurtosis of the distribution minus 3, the kurtosis of the normal distribution.

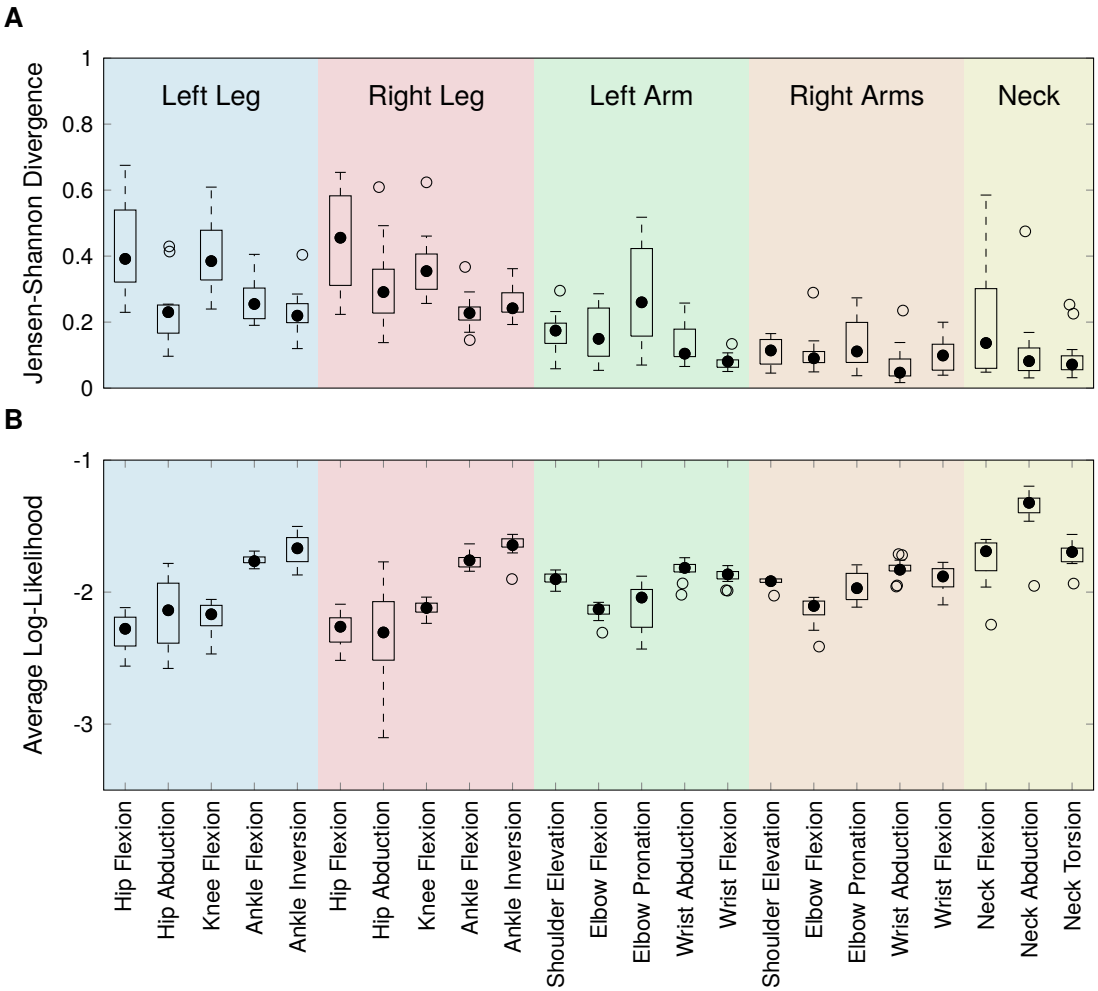

Figure 2.7: Variation of individual recordings from the grand average as calculated by the **A** Jenson-Shannon Divergence (lower is more similar) and **B** average log–likelihood (higher is more similar) of recorded data points. The different coloured areas of the plot indicate various body parts. As expected, right and left limbs show very similar patterns with respect to their similarity score. The dissimilarity between the results from the divergence metric and the average likelihood comes from the fact that they measure slightly different aspects of the movement distribution (see text for details). Full circles indicate medians, boxes the interquartile range and whiskers extend at most 1.5 IQR. Hollow circles indicate outliers.

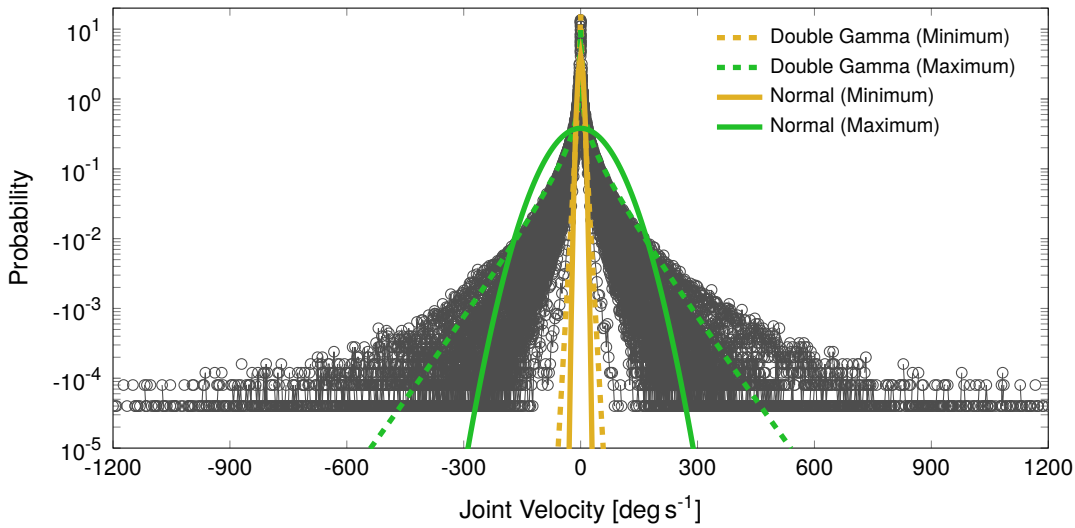

Figure 2.8: Probability distributions for the velocity of all joints of the body. The shape is largely preserved across joints, but the variance is very different. Note that neither the double gamma distribution nor the Normal distribution can fit the data well, as neither capture the heavy tails of the data.

grand average. This conceptual difference explains that the values of the JS divergence and the average log-likelihood are only weakly correlated ($r = -0.56$).

Interestingly, the values for the JS divergence (Figure 2.7**A**) suggest that the upper limb statistics of different subjects are more similar to each other than is the case for the lower limbs. While this may at first seem surprising, it simply suggests that the upper body is used in a similar fashion, independently of whether the subject is sitting or standing. Conversely, the lower limb position are largely constrained by the subject's pose (sitting, standing, etc.) which causes a narrow distribution of joint angles and thus a larger JS divergence.

2.3.2 Velocity

Unlike the position distributions which are very characteristic of the joints observed, the shape of the distributions of joint velocities are largely independent thereof (see Figure 2.8): the probability density falls off very sharply, thus indicating a majority of

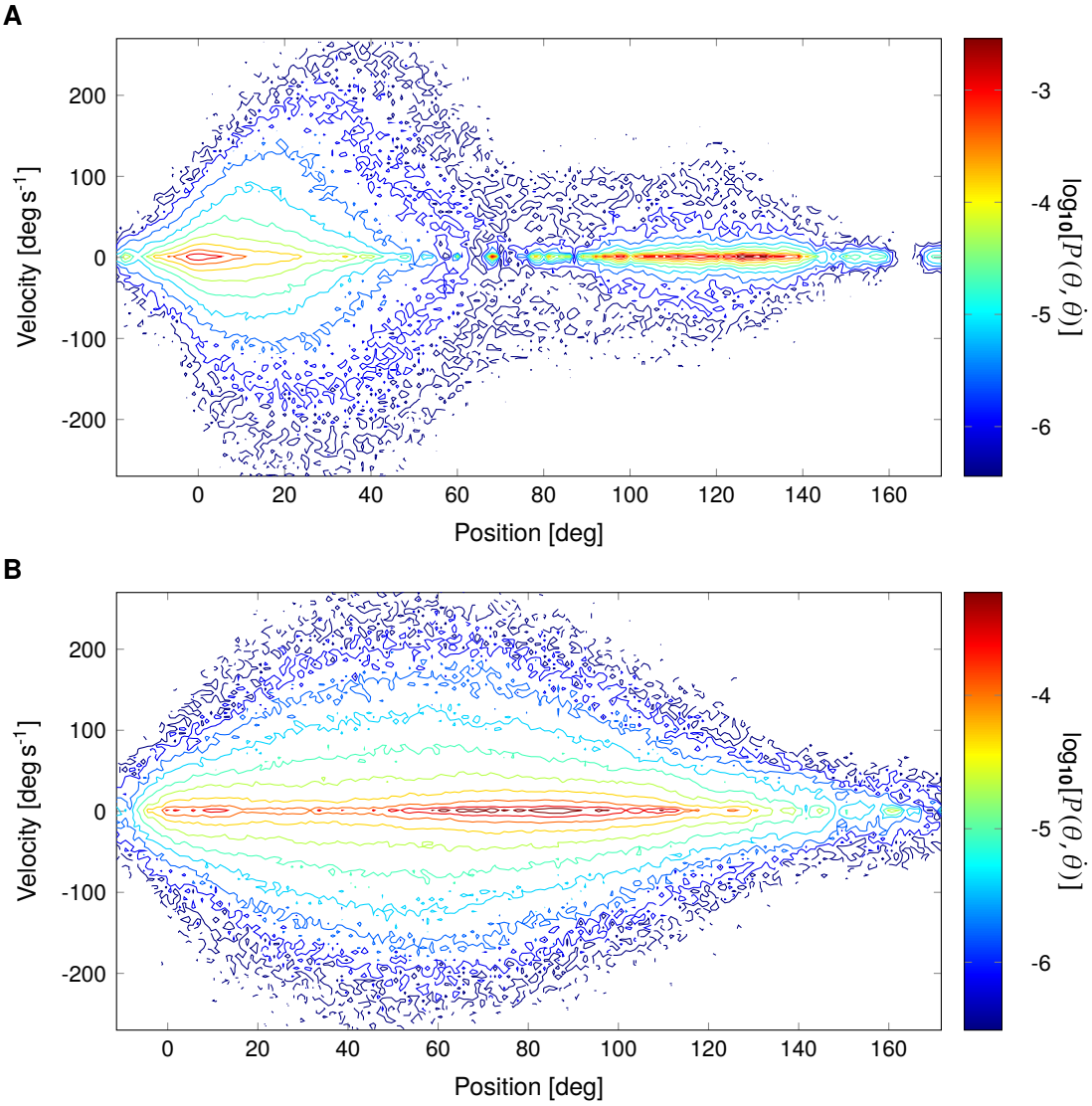

Figure 2.9: Contour plots of $\log_{10} P(\theta, \dot{\theta})$ for flexion of the **A** right knee and **B** right elbow obtained from a histogram with bins of $5° \, s^{-1}$ for velocity and $1°$ for position.

Identifier	Mean	95% C.I.		Identifier	Mean	95% C.I.	
Hip Flex. L	0.49	-62.32	63.31	Elbow Pron. L	-21.73	-60.19	16.74
Hip Abd. L	-35.83	-90.26	18.61	Wrist Abd. L	0.3	-0.24	0.83
Knee Flex. L	10.05	-8.79	28.88	Wrist Flex. L	-1.53	-21.67	18.61
Ankle Flex. L	2.05	1.75	2.35	Shoulder El. R	0.05	-1.97	2.08
Ankle Inv. L	3.29	1.28	5.3	Elbow Flex. R	-7.16	-31.29	16.96
Hip Flex. R	11.26	-16.97	39.5	Elbow Pron. R	-6.10	-33.64	21.44
Hip Abd. R	13.32	-11.1	37.74	**Wrist Abd. R**	0.37	0.18	0.55
Knee Flex. R	13.08	-11.65	37.81	Wrist Flex. R	-6.67	-21.89	8.55
Ankle Flex. R	-16.37	-65.59	32.85	Neck Flex.	-0.41	-7.46	6.64
Ankle Inv. R	-0.25	-0.64	0.14	Neck Abd.	7.67	-327.47	342.81
Shoulder El. L	-0.04	-0.64	0.55	Neck Tors.	7.69	-38.31	53.69
Elbow Flex. L	8.96	-11.29	29.21				

Table 2.1: Bootstrapped mean skewness and confidence intervals for the velocity distribution of various joints. Joints for which the estimate did not include 0 are indicated in bold.

movements at relatively low velocity. However, this does not mean that fast movements are non-existent. On the contrary, rapid movements maintain a much higher likelihood than could be expected if the distribution of velocities followed a Normal or "double" Gamma[5] distribution (Figure 2.8 solid and dashed lines respectively). The PDFs for individual joints differentiate themselves mainly in terms of the standard deviation of the distribution, neck abduction having the narrowest distribution (SD = 6°/s) and the right elbow and wrist the largest (SD ≃ 40°/s).

Next, we asked whether the movement speed in a given direction (e.g. flexion) was markedly different than the movement in the opposite direction (extension). We expected joints where the muscles acting in opposing direction have very different strengths (e.g. the elbow) to show a skew towards faster movements in the direction in which the muscles are stronger. As shown in Table 2.1, this is generally not the case, most skewness estimates having confidence intervals comprising 0. In the three cases

[5]This is the symmetrised Gamma distribution defined as $P(x|k,\theta) = \frac{1}{2\Gamma(k)\theta^k}|x|^{k-1}e^{-|x|/\theta}$ with support $x \in (-\infty, \infty)$.

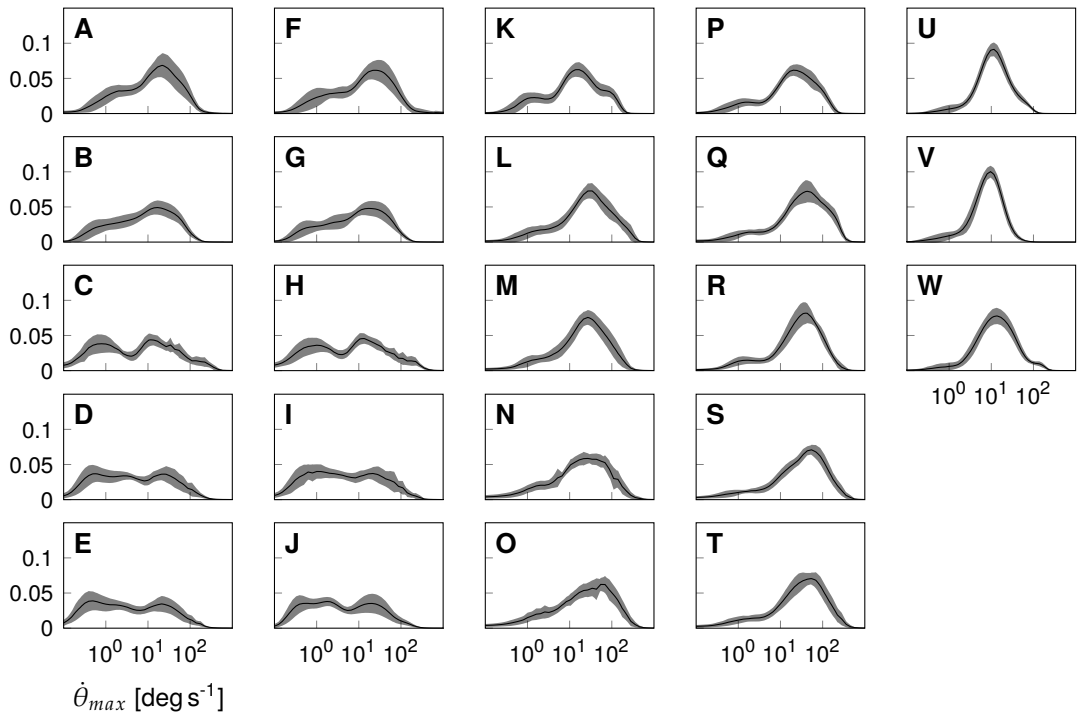

Figure 2.10: A–W Probability distributions for peak velocities of all joints of the body. **A–E** Left leg (Hip flexion, abduction; knee flexion; ankle flexion, inversion). **F–J** Equivalent for the right leg. **K–O** Left arm (shoulder elevation; elbow flexion, elbow pronation; wrist abduction, flexion). **P–T** Equivalent for right arm. **U–W** Neck flexion, abduction and torsion. Note the variability in the shape of distributions, as well as the consistency within equivalent joints on the left and right side of the body. Shaded areas indicate standard deviation across $N = 15$ recordings.

where the 95% confidence interval did not include 0, the estimated skewness is so close to zero that the distribution can still reasonably be approximated as symmetric. While we found no direction dependence we did observe that fast movements were much more likely to occur at positions with high likelihood (see Figure 2.9). This relationship is characterised by a strong positive correlation between the likelihood of a given position θ and the width of the contour lines of $P(\theta, \dot{\theta})$ (Pearson's $r = 0.86 \pm 0.06$). An exception to this rule is the knee (Figure 2.9**A**) which has an offset of 15–30°, consistent with the region of peak velocity in the swing phase of walking (Chao et al. 1983).

We attempted to model the observed velocity distributions by assuming that each

movement is composed of sub-movements (Morasso 1981; Milner 1992; Burdet and Milner 1998) whose velocity profile can be represented by a bell curve with variable amplitude (see Section 2.2.5). We chose 50 amplitude parameters, α_i, $i = 1 \ldots 50$, logarithmically spaced between the slowest and fastest observed movement of each joint. The resulting curves fitted the empirical velocity distribution very well (GOF > 0.99 for all fits). From this modelling approach, we can extract the probability distributions for peak velocities of individual joints (Figure 2.10). These distributions have three noteworthy characteristics: (1) The shape of the distributions vary greatly from joint to joint, but are similar for the same joint on either side of the body. (2) The distributions for the right arm (dominant) are shifted towards higher peak velocities relative to the left arm. This is not the case for the right leg. (3) They are extremely similar across individuals. Indeed, the JS divergence between individual probability distributions and the average is about an order of magnitude lower than it is the case for position distributions (see Figure 2.11).

2.3.3 Movement Linearity and the 2/3 Power Law

As a final consideration, we verified whether features of movement recorded during constrained experiments still hold in natural behaviour. We consider two long-standing observations: (1) movements are mainly straight (Flash and Hogan 1985; Atkeson and Hollerbach 1985; Uno, Kawato, and Suzuki 1989) and (2) the $2/3$ power law (Lacquaniti, Terzuolo, and Viviani 1983; Viviani and Cenzato 1985; Schaal and Sternad 2001).

Movement Duration and Linearity

We analysed the straightness of hand movements by identifying start and end times of motions and decomposing the time-series accordingly (see Section 2.2.6). We identified $N = 3523$ movements with an average time of $1.6s$. The duration of individual movements could accurately be described by a log-normal distribution with parame-

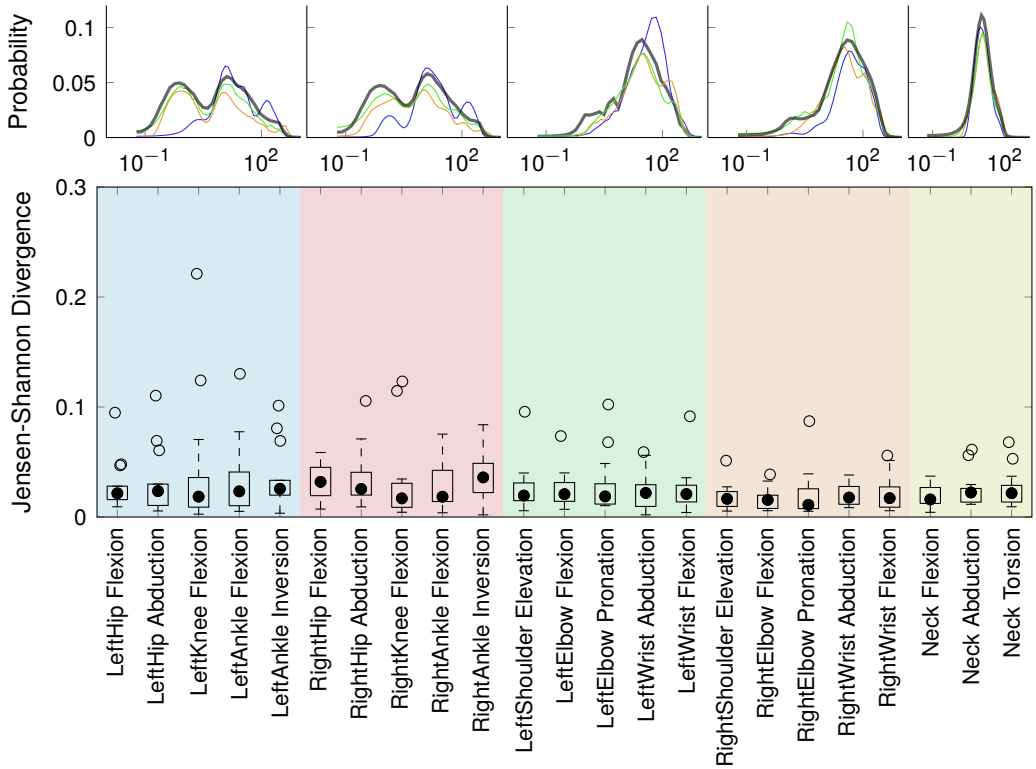

Figure 2.11: Top Representative examples of peak velocity distributions for three subjects (coloured lines) and the average (black). From left to right, they represent the left and right knee flexion, left and right elbow flexion and neck flexion. **Bottom** Jensen-Shannon divergence between individual peak velocity distributions and the mean. These values should be contrasted with the distance of position distributions (Figure 2.7**A**), which are about an order of magnitude larger.

ters $\mu = -0.15 \pm 0.035$ and $\sigma = 1.05 \pm 0.025$ (mean± 2SD; Figure 2.12**A**). Movement straightness and planarity were analysed using Principal Component Analysis (PCA). As shown in Figure 2.12**B**, it results that 69% of movements are straight, i.e. more than 90% of their variance is captured by the first principal component (PC). Two PC are sufficient to accurately represent 99.2% of movements, meaning that they are effectively all executed in a plane. Linearity and movement duration were negatively correlated, meaning that longer movements were less likely to be straight lines (Spearman's $\rho = -0.66$, $p < 10^{-5}$). The plane of movement is rarely parallel to the ground.

Instead, its orientation is extremely varied as shown by the projection of the plane's normal vector onto the unit sphere (Figure 2.12**C**). Further analysis did not reveal any particular structure or preferred orientation for planar movements: the observed G function, which gives a sense of spatial randomness was not significantly different from a simulated G-function with truly random data (Figure 2.12**D**).

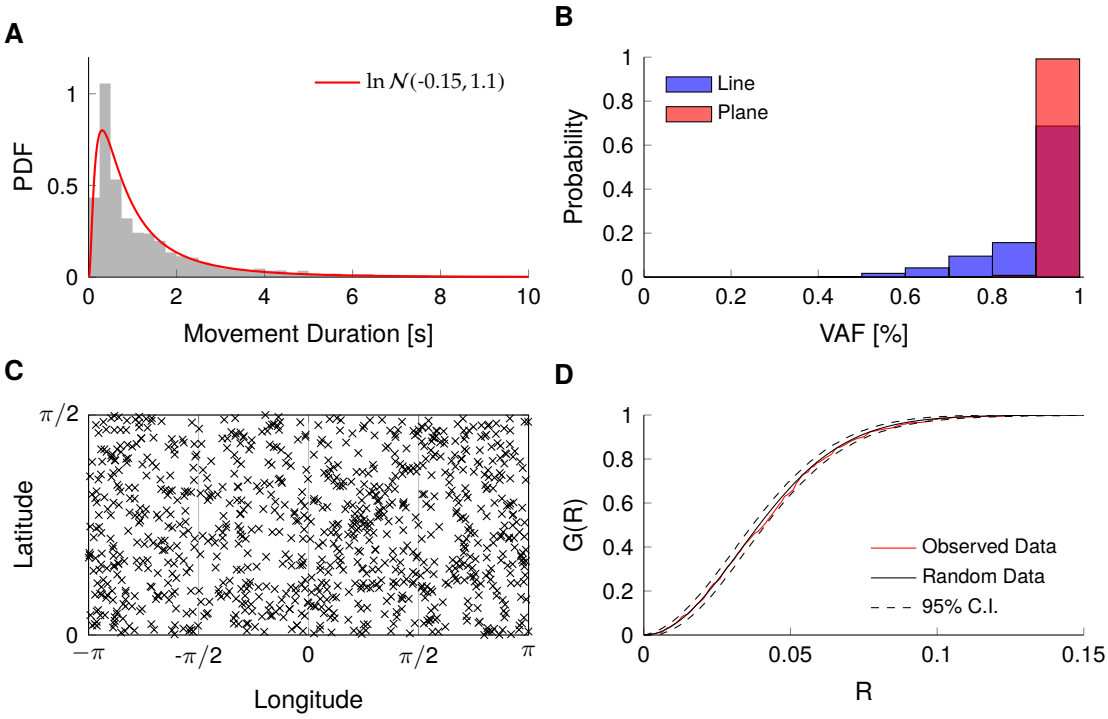

Figure 2.12: **A** Distribution of movement durations and exponential fit. **B** Our results indicate that almost all (> 99%) movements are executed in a plane, while only about 70% can be described as linear. **C** For the planar movements, the normal vector to the plane did not appear to point in a consistent direction, as shown in this equal area projection of the unit-sphere. **D** The randomness of plane orientation was confirmed by comparing the distribution of points in **C** with a model of complete spatial randomness.

The 2/3 Power Law

A number of studies have observed that the angular velocity, ω, of the hand along its path is related to the curvature of the path by a power law with exponent $\beta = 2/3$

(Lacquaniti, Terzuolo, and Viviani 1983; Schaal and Sternad 2001). Mathematically, this takes the form of:

$$\dot{\omega} = kc^{2/3} \tag{2.7}$$

or, alternatively,

$$\dot{x} = kr^{1/3} \tag{2.8}$$

where $\dot{x}$ is the tangential velocity, k is a constant, c is the curvature and $r = 1/c$ is the radius of the path and $\beta = 1/3$. For any given path, k is supposed to be a piecewise constant function (Viviani and Cenzato 1985). A major concern with this finding is that with one exception (Schaal and Sternad 2001), all the data supporting the power law comes from experiments in the plane. Further, the vast majority of experiments looked at specific, predetermined movement patterns, in particular ellipses and lemniscates (figures of eight). We wondered whether the power law would still hold in natural, 3D movements and if so, whether it was consistent with previous research.

To this end, we calculated the instantaneous curvature and velocity for each individual recording Figure 2.13**A** and fitted a power law to the relationship between the two variables. The result for a representative subject is shown in Figure 2.13**B**. The dashed blue and red line correspond to power law fits using non-linear least squares on the raw data ($\beta_{NL} = 0.49$) and linear fits on log-log plots ($\beta_{LL} = 0.64$) respectively. These coefficients are significantly different from each other and from the ideal value of $\beta = 0.33$ (Bootstrap test, all $p < 10^{-3}$). This approach, however, neglects the assumption that k be piecewise constant (Viviani and Cenzato 1985). To take this into account, we divided the data into non-overlapping windows of 500 ms and fitted the power law in each individual window (Figure 2.13**D**. We only considered windows in which the fit had an $R^2 > 0.9$. In case of non-linear fits, this accounted for 13.6% of windows for the

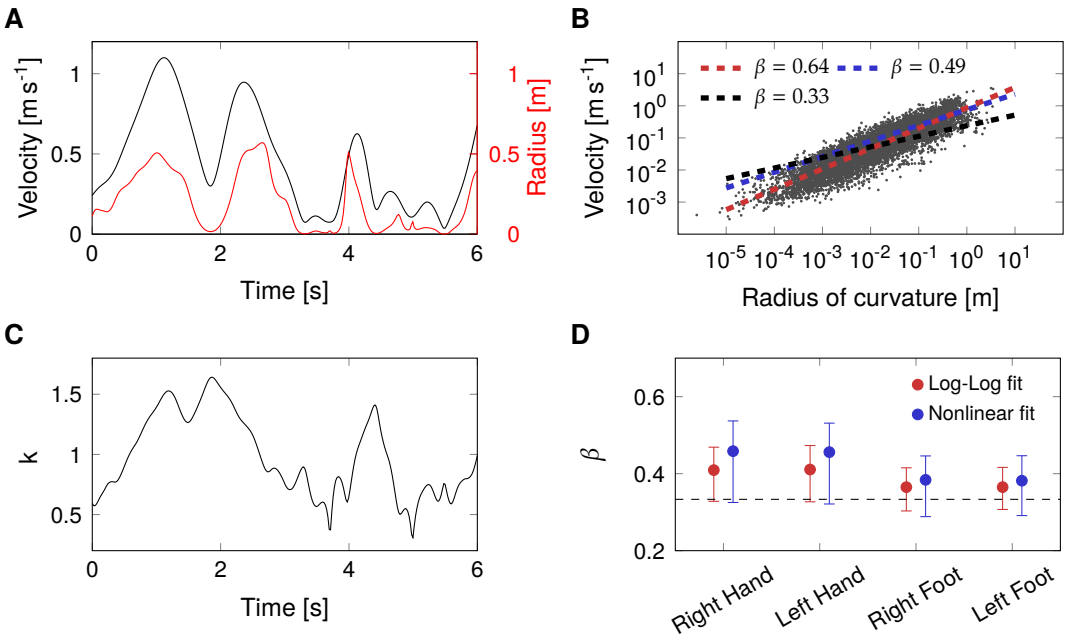

Figure 2.13: **A** Sample velocity (black) and radius of curvature (red) traces. **B** A global power-law relationship exists between the tangential velocity and the radius of curvature, but the exponent value is very dependent on the method used to compute it (and in any case not 0.33). **C** The value of k estimated from the data in **A** is smooth and shows no trace of discontinuities. **D** Distribution of power-law exponents computed on windowed data. Error bars indicate IQR.

right hand, 13.7% for the left hand, 6.3% for the right foot and 7% for the left foot. These values were higher for linear fits on log-log plots, being 19.4%, 19%, 10.7% and 11.5% respectively. This method did not reveal an exponent value of 0.33 either. Instead, the values obtained were very variable and broadly distributed, with medians $\beta = 0.46$ for hand movements (0.41 when fitting on a log-log plot) and $\beta = 0.38$ (0.36) for foot trajectories. Using these empirically determined exponents, we could then determine the scaling coefficient k. As shown in Figure 2.13**C**, this was not a piecewise constant function with sharp discontinuities but instead varied smoothly. Computing k using a different value than the fitted exponent of $\beta = 0.41$, e.g. by using the theoretical value of 0.33 did not produce the postulated result, either.

2.4 Discussion

2.4.1 Methodology

We collected behavioural data of freely behaving subjects in simulated natural environments. Unlike previous studies (Ingram et al. 2008; Howard et al. 2009; Faisal et al. 2010; Belic 2010), we captured the full perception-action loop by simultaneously measuring sensory inputs (in particular visual stimuli) and the brain's output in terms of movements (both of limbs and the eyes). We are thus, for the first time, mapping the entire statistics of natural human behaviour. While such studies have been done previously on a smaller scale (i.e. measuring only movement of certain limbs (Ingram et al. 2008; Howard et al. 2009) or in more constrained environments (Torre et al. 2009), it is important to review the findings presented here terms of the ecological validity of the proposed methodology. Two points in particular could argue against the fact that we are collecting "natural behaviour": (1) due to experimental constraints, recordings only lasted 2-3 hours for each subject during which they were placed in various naturalistic settings which emulate their normal environment. The area in which subjects could move was further constrained by connectivity of the equipment, some of which was tethered to a laptop by means of a USB cable. This restriction in space and time implies that subjects could not go by their daily routine completely unhindered. (2) the suit, gloves and eye-tracker had to be worn in addition of the subjects' normal clothing, potentially hampering their movements. The latter issue is an inherent limitation to current motion capture technology and is likely to affect recordings using either optical or inertial sensors. Markerless motion technology which could overcome these restrictions does exist (e.g. Corazza et al. 2010; Schmitz et al. 2015) but is still very inaccurate and suffers from the same limitations as marker-based optical systems (see Section 2.1). Despite these constraints, no subject reported any problems with executing movements when asked after the experiment. It should, however, be noted that

the use of gloves for measuring hand movements does limit tactile feedback, though previous studies did not report problems with this (Ingram et al. 2008; Faisal et al. 2010; Belić and Faisal 2015). Regarding the nature of the data collected, we acknowledge that due to the very active nature of the subjects' behaviour, the probability distributions of joint position and velocity are likely to be somewhat biased. An adult in working age spends on average 9.4 h sitting (Miller and Brown 2004) and 7.7 h sleeping (Basner et al. 2007). During that time, the joint position is likely to vary very little, which would cause sharper peaks in the probability distributions for position as well as skew the distribution of velocities more towards zero. Thus, the data presented here should be used with care when trying to draw conclusions about a specific subject, i.e. when trying to determine disease progression (Gavriel et al. 2015a).

2.4.2 Joint Position and Velocity Distributions

The data collected during our experiments enables us to quantify, for the first time, the probability distribution of the position of individual joints during natural tasks (Figure 2.5). We note that the distributions are extremely varied, in particular for the lower body (Figure 2.7). This is likely caused by the fact that our subjects spent most of their time sitting at a table, where they were allowed to sit any way they liked. This is reflected in very sharp distributions for leg positions which may explain the observed variability. In contrast, the probability distribution of joint velocity was similarly shaped for all joints considered and consistent across subjects. This invariance may go some way in explaining why humans are much better in identifying biological motion (i.e. dot patterns which appear to move like humans (Johansson 1973)) than biological configuration (Krakowski et al. 2011; Buzzell et al. 2013).

The consistency of velocity distributions was preserved when we analysed the amplitude of individual movements by fitting a specifically constructed probability distribution to the data (Figure 2.10). This method makes the simplifying assumption that

movements are far enough apart to avoid significant overlap between them. Since there is a considerable amount of literature pointing towards the existence of submovements, i.e. overlapping movements (e.g. Morasso 1981; Burdet and Milner 1998), it is worth considering the bias induced by our method: our estimate of peak velocities is likely to be slightly inflated since it does not take into account the contribution of surrounding submovements. However, based on evidence from detailed analyses of submovements (Milner 1992; Rohrer and Hogan 2006), this effect is likely to be small, in particular if assuming a Gaussian shape for submovements, as we do. It should in theory be possible to quantify exactly how much our estimate is off by using "exact" submovement decomposition algorithms (Morasso and Ivaldi 1982; Milner 1992; Rohrer and Hogan 2003; Rohrer and Hogan 2006), however, the computational demands of those methods make them unsuitable for the amount of data we were analysing (Rohrer and Hogan 2006).

2.4.3 Similarity Between Subjects

We investigated two measures of behavioural similarity between subjects: (1) the Jensen-Shannon divergence between the probability distribution of individual subjects' position and movement and the grand average (Figure 2.7**A**) and (2) the mean log-likelihood of subjects' data given the grand average (Figure 2.7**B**). Both measures are based on the same probability distribution and are thus related. However, differences exist in the quantity they measure: the JS divergence measures distances between distributions. It is thus unsuited for small quantities of data and – in particular for position distributions – highly dependent on the activity performed. It is thus best suited to observe consistent differences in position/velocity in longitudinal studies where large amounts of data are available and fine differences need to be detected. Conversely, the average log-likelihood only measures the probability of a given point (or sequence of points) given the overall data available. It could thus be used to rapidly identify

odd body configurations during specific experiments or when monitoring patients. In either case, we maintain our previous reservations regarding the generalisation of this data to other circumstances. In particular, we would recommend to use the (peak) velocity distributions (Figure 2.10) when comparing subjects as these are less prone to perturbations of the motion capture equipment, characteristic of the joint under consideration and highly preserved between subjects (Figure 2.11).

2.4.4 Planarity of Hand Movements

Unlike the standard centre-out reaching movements as performed in countless laboratory-based studies (e.g. Georgopoulos et al. 1982; Shadmehr and Mussa-Ivaldi 1994; Moran and Schwartz 1999, to name but a few), the shape of self-paced movements was straight only ~70% of the time. The choice of straight movements to study the effect of perturbations on motor control can thus be justified in that they are so common that learning of the basic task is not required and should be well represented in the brain. The fact that these studies normally involve movement in a horizontal plane can also be justified by considering that all movement planes appear to be equally probable: moving in a horizontal plane is thus unlikely to be a special case of general reaching. On the other hand, the apparent universality of planarity for reaching movements offers an intriguing new possibility for experimental setups: 3D movements. Although movements in 3D-space have been studied previously (e.g. Morasso 1983; Desmurget and Prablanc 1997; Schaal and Sternad 2001; Domkin et al. 2005; Biess, Liebermann, and Flash 2007), the trajectory of the endpoint is mostly planar in all cases. Indeed, Morasso (1983) found that even when subjects were free to "scribble" in 3-D space, their velocity profile could be decomposed into planar sections, very much in the same way as e.g. Viviani and Cenzato (1985) decompose planar movements into submovements based on a change in scaling factor in the 2/3 power-law. Forcing subjects to trace a clearly 3-dimensional path and generate perturbations on that space might thus lead

to valuable insights into human motor control.

2.4.5 The 2/3 Power-Law

The 2/3 power-law relationship between movement speed and curvature, as well as the discontinuous nature of the scaling coefficient, has often been pointed to as evidence for segmented motor control (Morasso 1981; Morasso and Ivaldi 1982; Morasso 1983; Viviani and Cenzato 1985; Flash and Hogan 1985; Ivanenko et al. 2007) and as a central constraint to movement generation by the CNS (Lacquaniti, Terzuolo, and Viviani 1983; Massey et al. 1992; Ivanenko et al. 2002). Its apparent universality has even been used as a validating feature for theories of motor control which do not directly depend on motion segmentation (e.g. Harris and Wolpert 1998; Sternad and Schaal 1999). Yet, our findings relating to the 2/3 power-law in natural movements do not reflect previous findings Figure 2.13. First of all, the power-law was rarely a good fit of the relationship between velocity and radius of curvature, applying to only 10–15% of all data windows analysed depending on the fitting method used. This low number could reflect the arbitrary nature of our data segmentation. However, the small data window (500 ms) combined with a median movement duration of 860 ms make the low percentages observed unlikely to be caused by bad luck. Secondly, the exponent values observed, in particular for hand movements, are not in accordance with the expected value of 0.33. Instead, we observe very variable values with medians between 0.36 and 0.46, depending again on the fitting method Figure 2.13**D**. We are not the first to observe violations of the 2/3 power-law (Wann, Nimmo-Smith, and Wing 1988; Schaal and Sternad 2001; Ivanenko et al. 2002), although previous studies mainly found deviations occurring in the exponent value. That is not to say that a (power-law) relationship between velocity and curvature does not exist. In fact, we find that without segmenting the data, a noisy power-law relationship exists between velocity and radius of curvature (Figure 2.13**B**). While the reasons for this being a power-law are unclear,

it is not surprising that such a relation exists: basic physics imply that changes in direction imply accelerations proportional to v^2/r or, alternatively, that $v \sim kr^{0.5}$ for a constant acceleration. This exponent value is much closer to the actual value observed in our study. Of course, we do not directly apply forces to the endpoint but have to move it by applying torque to multiple joints. The impact of endpoint acceleration on joint torques through inverse kinematics and the implications for motor control are left to future studies.

2.4.6 Conclusion

Analysing large data sets of naturalistic movement opens the potential for verifying assumptions commonly made in laboratory settings, as well as providing a ground on which to build new experiments targeting specific observations. Our work provides a framework as well as a tool-set for performing such large-scale experiments which will undoubtedly be easier to perform in the near future with less intrusive monitoring technology. Our results enable us to quantify key statistics of natural behaviour which would normally not be measurable and put into question previous findings of experiments performed in restricted laboratory settings. These outcomes will hopefully stimulate further studies and highlight the necessity of scientific work performed in the lab as much as in the real world.

3

Structure of Natural Movement

"The whole is greater than the sum of the parts."

— Aristoteles

3.1 Introduction

A central question in motor control is to understand how the brain controls the multiple degrees of freedom of our body (Bernstein 1967). Whether we consider the problem from a kinematic or a muscular perspective, the number of degrees of freedom (DOF) accessible to the brain to achieve a given goal are typically larger than strictly necessary, resulting in a underdetermined system. In this case, which combination of DOF should the brain choose? Attempts to answer this question fall broadly into two categories: (1) optimisation of a cost function (Uno, Kawato, and Suzuki 1989; Todorov and Jordan 2002; Li, Todorov, and Pan 2004; Todorov, Li, and Pan 2005) and (2) dimensionality reduction (Bizzi, Mussa-Ivaldi, and Giszter 1991; Giszter, Mussa-Ivaldi, and Bizzi 1993; Santello, Flanders, and Soechting 1998; d'Avella, Saltiel, and Bizzi 2003; Cheung et al. 2005; Gentner and Classen 2006; Ingram et al. 2008). We will mainly focus on dimensionality reduction in the kinematic space which suggests that motor control takes place in a low-dimensional subspace of all possible joint configurations. However, note that linking various kinematic variables, e.g. the power-law relationship between

velocity and path curvature (see Section 2.3.3 and work by Lacquaniti, Terzuolo, and Viviani (1983)) is also a form of dimensionality reduction as it reduces the number of possible combinations of variables.

Initial studies on hand movements suggested that the 21 DOF of the hand could be reduced to 1 *synergy* (i.e. all joints of the hand move proportionally to each other) during reach-to-grasp movements of conical objects (Mason, Gomez, and Ebner 2001) or 2–3 synergies during imagined grasping of everyday objects (Santello, Flanders, and Soechting 1998). These studies, however, fail to take into account the use of the hand during object manipulation. A study of synergy use during a haptic exploration task reported 7 synergies required to explain 90% of the data (Thakur, Bastian, and Hsiao 2008), but the methodology used differs markedly from previous studies, making a direct comparison difficult. To the best of our knowledge, only a single study has investigated synergistic finger use across a large range of behaviours (Ingram et al. 2008). This study outfitted the dominant hand of three subjects with a sensorised glove measuring joint position and allowed them to wear it for the duration of a day, yielding a data set of natural movement. This data could be well described by a set of 6 synergies, 2 of which were common to all subjects. This value is estimated from data which has not been normalised, suggesting that the real number may actually be higher. Finally, Jarrassé et al. (2014) looked into synergies during controlled bimanual manipulation of objects, finding that 3 synergies per hand explain almost all the variance in their data. It should be noted that all of these studies used principal component analysis (PCA) for their dimensionality reduction, effectively equating a synergy with a principal component (PC).

A further aspect of synergistic motor control is the timing, which can be divided into two categories: (1) synchronous synergies (SS) (e.g. (Santello, Flanders, and Soechting 1998; Ingram et al. 2008)) and (2) time-varying synergies (TVS) (d'Avella, Saltiel, and Bizzi 2003). The former assumes that the command is applied to all

actuators (muscles or joints) simultaneously, while the latter allows for time shift in the activation. This distinction has to date only been made for muscle synergies: discovering low-dimensional control manifolds using PCA intrinsically enforces a SS model. Allowing for TVS in kinematic data may reveal an even lower-dimensional control than seen to date.

In this chapter, we build upon previous work by analysing natural hand movement of both the dominant and non-dominant hand in natural settings, allowing us to examine not only low-dimensional control structures within each hand but to compare the synergies across hands and investigate the existence of bimanual synergies. To assess the assumption that the kinematics of hand movements underlie a SS model, we analyse the temporal structure of the data in search of time delays. Next, we transpose the analysis performed on the hand data onto all other joints of the body. Is synergistic control the general rule, or is it specific to the hand? Finally, we link both analyses by exploring the structure of arm and hand movement, with the particular goal of evaluating its potential to improve prosthetic control by exploiting natural movement statistics.

3.2 Methods

3.2.1 Data

We use data collected simultaneously from two CyberGloves (CyberGlove LLC, San Diego, CA, USA) and an IGS-180 motion capture suit (Animazoo UK Ltd, Brighton, UK). Data collection protocols and preprocessing were performed as described in the previous chapter. Any additional processing will be highlighted in the relevant sections.

3.2.2 Comparison with Previous Studies

We extracted previously published PCs and VAF curves from Ingram et al. (2008) and Jarrassé et al. (2014) using a free online tool (WebPlotDigitizer). Subsequently, PCs were normalised for them to have unit length and VAF curves modified to display the cumulative variance explained. The data from Ingram et al. (2008) contains values for the distal interphalangeal (DIP) joints, which were removed from the PC vectors before computing subspace distances. It was further necessary to apply an orthogonalisation procedure to the PCs extracted as the data reported in both cases is the average from multiple subjects which effectively destroys the orthogonality between the PCs. As linear subspaces are being compared, transforming PCs such that they are orthogonal to each other while still defining the same space does not affect the results. Since direct PC-to-PC comparison would be affected, the re-orthogonalisation procedure was not applied in those cases.

3.2.3 Comparison of PCA Subspaces

To compare low-dimensional subspaces generated by multiple PCs we use the method described by Krzanowski (1979). In brief, given any two sets of PCs $\mathbf{L} = \{\mathbf{l}_1, \ldots, \mathbf{l}_Q\}$ and $\mathbf{M} = \{\mathbf{m}_1, \ldots, \mathbf{m}_Q\}$ the method computes θ, the angle between $\mathbf{l}_i, i = 1 \ldots Q$ and $\mathbf{l}_i^{'}$, it's representation in the space of $\mathbf{M}$. Thus, the total distance between the subspaces defined by $\mathbf{L}$ and $\mathbf{M}$ can be described as $\sum_i \theta_i$. Since this value is dependent on the number of PCs considered, we report the value $d = \sum_i \theta_i / \pi$, which is bound between $[0, 1]$.

3.2.4 Prediction of Hand State

We define the hand state, $\mathbf{y}(t)$, to be the 30-dimensional vector describing the hand's configuration and instantaneous velocity at any point time. To predict the state of

the hand without any knowledge of its previous state (such as would be the case for prosthetic control), we performed multivariate linear regression (MVLR) between $\mathbf{y}(t)$ and the position and velocity of the contralateral (relative to the hand being predicted) hand as well as of both elbows (flexion and pronation), shoulder elevations and neck movement (termed $\mathbf{x}(t)$ hereafter). The linear model was fitted using the Matlab function `lmfit` without additional parameters. To avoid over-fitting, we performed 10-fold cross-validation by randomly dividing the data into 10 blocks of equal size, training the model on 9 such blocks and testing its performance on the remaining data. To assess the chance level of such a regression, we performed the same analysis but first randomly shuffled $\mathbf{x}(t)$ to destroy any temporal relation to $\mathbf{y}(t)$.

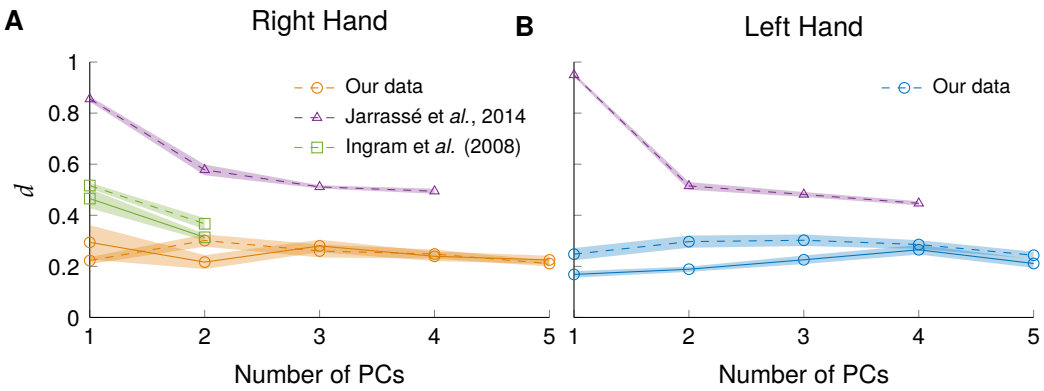

Figure 3.1: Distance between PCA subspaces. **A** Inter-subject distance ("our data") and distance to data published by Ingram et al. (2008) and Jarrassé et al. (2014). **B** Equivalent for the left hand. Dashed lines indicate position, solid lines velocity. Shaded areas represent standard error.

3.3 Spatio-temporal Structure of Movements

3.3.1 Hand Movements

Similarly to previous studies (Santello, Flanders, and Soechting 1998; Mason, Gomez, and Ebner 2001; Ingram et al. 2008; Faisal et al. 2010) we analyse the complexity of hand

movements by means of principal component analysis. Our results are in agreement with the data from Ingram et al. (2008), the only prior study on natural hand movements (see also Figure 3.2**F**), requiring ~33% of PCs to explain 80% of the variance. This value is somewhat misleading, as it combines both correlation structure in the movement as well as vastly different movement amplitudes of different joints (e.g. MCP flexion *vs.* abduction). If the latter contribution is removed by normalising the data, the fraction of PCs required to explain 80% of the variance increases to 40% (compare Figure 3.2**F** to Figure 3.3).

To test the similarity between the latent spaces defined by the principal components, we computed the distance between them (see Figure 3.1). As expected, the latent spaces of natural movement data were much closer to each other than those obtained from data in the laboratory (Jarrassé et al. 2014). At a more detailed level, the first two PCs from our study and from Ingram et al. (2008) are in good agreement both in terms of position (R^2 = 0.8 and 0.83) and velocity (R^2 = 0.79 and 0.91, see also Figure 3.2**A**–Figure 3.2**D**). In contrast, the data collected in the lab (Jarrassé et al. 2014) demonstrates a very different first PC as well as substantially less variability (Figure 3.2**E**&**F**).

In contrast to Ingram et al. (2008) which looked only at the dominant hand, we are in the position of comparing both hands and analysing bimanual structure in natural movements. In terms of complexity, we find no difference between the dominant (right) and non-dominant hand (Wilcoxon signed rank test; position $p = 0.7$; velocity $p = 0.07$) as shown by the almost perfect overlap of the VAF curves in Figure 3.3. Structure common to both hands is very limited but statistically significant in terms of position (AUC Right: 0.81 ± 0.02 (mean ± SD), Left: 0.81 ± 0.01; Combined: 0.82 ± 0.02; ($p < 10^{-3}$ in both cases). In the velocity domain, the difference is almost imperceptible (AUC Right: 0.77 ± 0.01; Left: 0.77 ± 0.01; Combined: 0.78 ± 0.02) and significant only between right hand and common structure ($p < 10^{-3}$).

The independence between dominant and non-dominant can examined further by

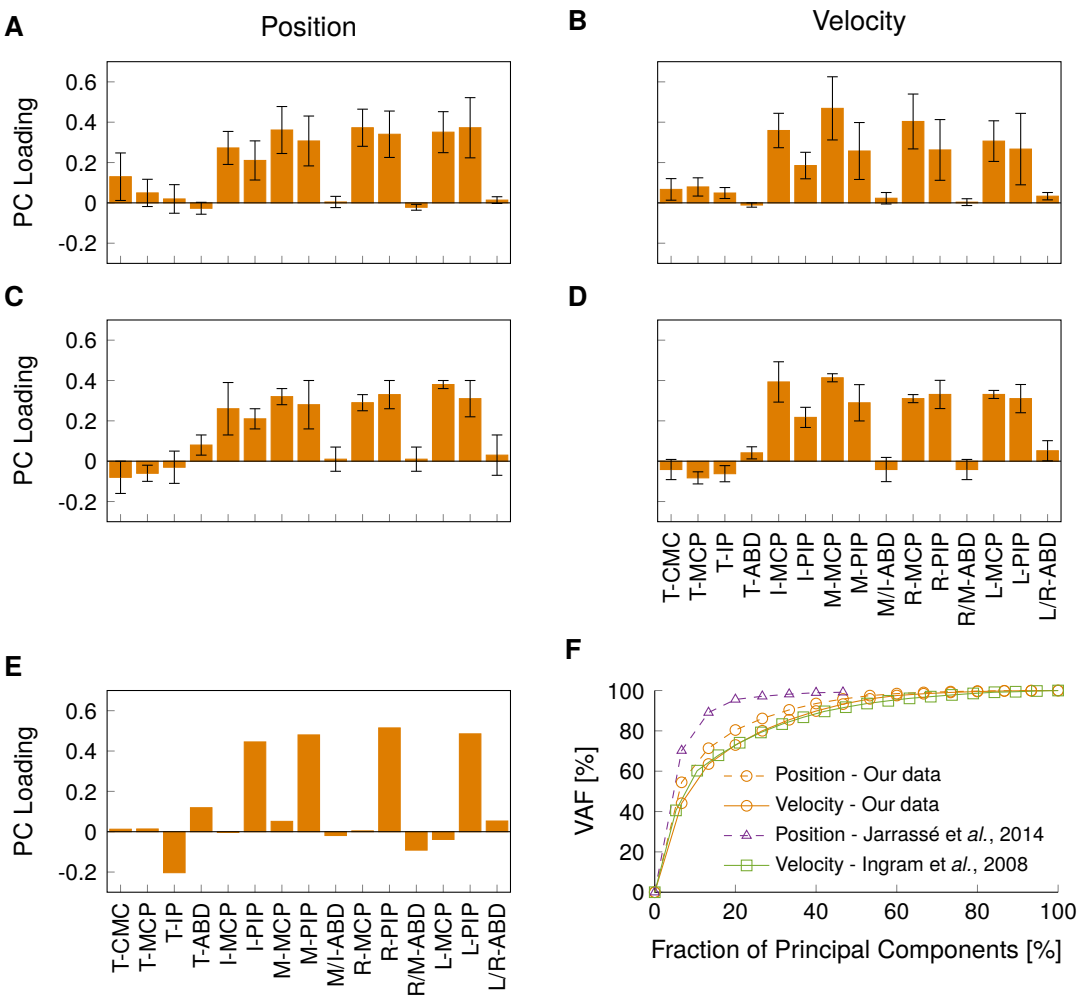

Figure 3.2: Comparison of the first PC extracted from **A**, **B** our data, **C**, **D** Ingram et al. (2008) and **E** Jarrassé et al. (2014). Note the similarity between the natural movement data (**A**–**D**). **F** Variance explained for all of the datasets. Note how the natural statistics are very similar, while the lab data (Jarrassé et al. 2014) is clearly less complex.

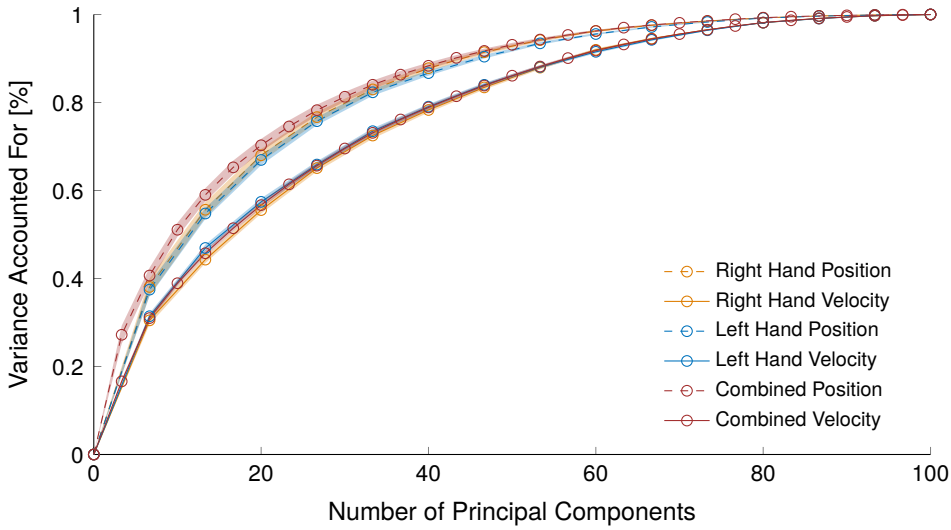

Figure 3.3: Total variance captured by an increasing number of principal components for the right hand, left hand and combined hands. Shading indicates standard error. The velocity lines overlap almost perfectly.

looking at the time during which their fingers move simultaneously (Figure 3.6). While the fingers of the same hand are in motion simultaneously an average 43.4 ± 6.1% (mean ± SD) of the time, this value drops to 24.2 ± 5.6% (paired Welsh's t-test; $p < 10^{-4}$) across hands. The difference between movement times of the dominant and non-dominant hand is not significant (Welsh's t-test; $p = 0.15$). Even if only considering the period of simultaneous movement, the VAF curve for both hands combined did not deviate from the globally computed one.

Differences between the dominant and the non-dominant hand become apparent when considering the number of joints moving simultaneously at any moment in time (Figure 3.4**A**). The probability of one to three fingers being in motion simultaneously is significantly higher in the dominant hand, whereas the probability of full hand movements (e.g. grasping) is more likely for the non-dominant hand. Joints most likely to be moved individually belong to the thumb, irrespective of the hand under consideration, whereas the middle and ring finger fail to move by themselves (Figure 3.4**B**). A more

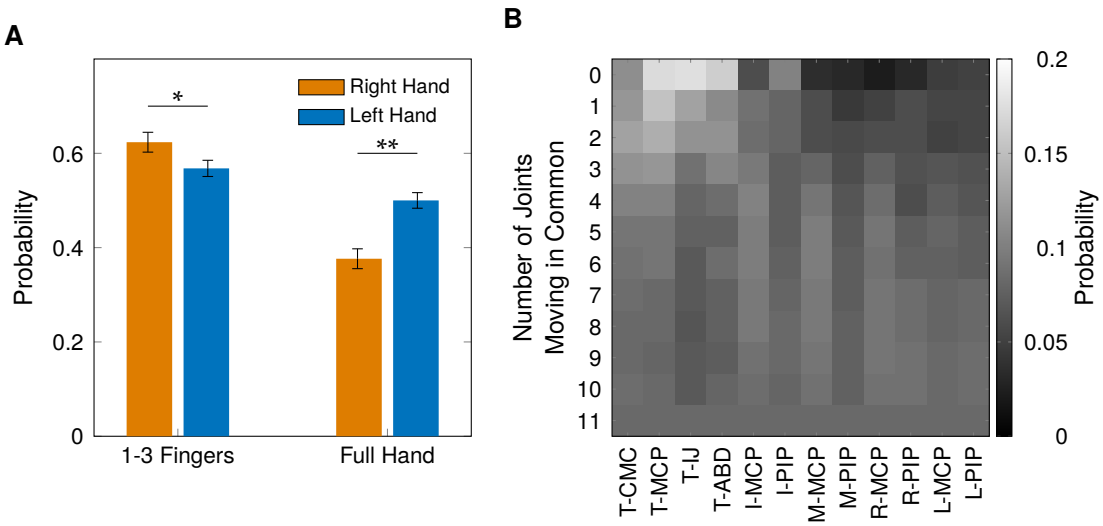

Figure 3.4: A Probability of 1–3 fingers moving simultaneously *vs.* full hand movments. The dominant (right) hand is more likely to use a small number of fingers while the non-dominant has a preference for full hand movements (e.g. grasps). One sided Wilcoxon signed rank test; *: $p < 0.05$, **: $p < 0.01$. **B** Probability of a given joint to move simultaneously with others. The autonomy of the thumb is highlighted by the light patch on the first row.

detailed analysis of the temporal structure of finger movement reveals no further differences between dominant and non-dominant hands. The full width at half-maximum (FWHM, see Figure 3.5**A**), which is an indicator for how rapidly movement changes is almost identical between the right and left hand (see Figure 3.5**B**. Mean right: $222ms$; left: $223ms$; t-test; n.s.). Cross-correlation analysis between joints of the same displayed only a few non-zero values (Figure 3.5**C**), though even then the lag is limited to 1-2 samples and not always consistent between subjects or across multiple recordings of the same subject. The exception to this rule is the precedence of movement of the R-MCP on the L-MCP which occurs in more than 60% of recordings. Cross correlation between hands displays no consistency and can probably be neglected given the low peak values (see Figure 3.5**D**).

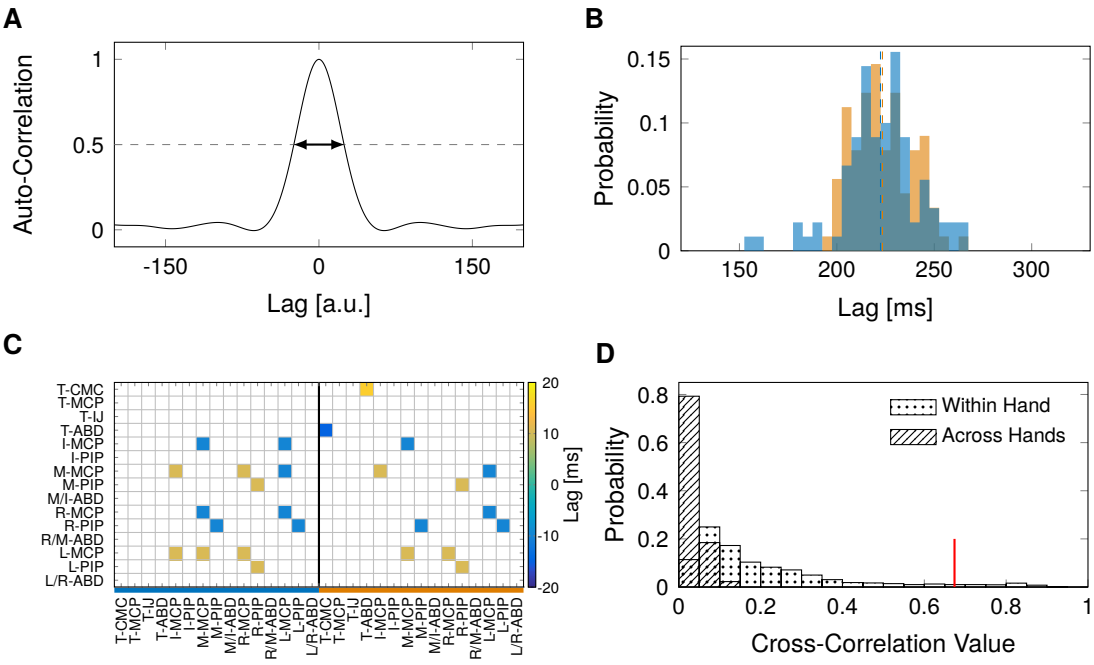

Figure 3.5: **A** Illustration of the auto-correlation function. The double arrow indicates the full-width at half-maximum (FWHM) of the function. **B** Distribution of the FWHM for all joints and subjects of the right (orange) and left (blue) hand. Dashed lines indicate the mean (223ms), which is not significantly different between hands. **C** Cross-correlation lag between the joints of the right and left hand. Negative (blue) values mean that the joint on the y axis leads on the joint on the x axis. Only lags where the peak value of the cross-correlation function exceeded 0.67 are displayed. **D** This value is the 95-percentile of the distribution of cross-correlation values within joints of the same hand (red line). Cross-correlation between hands is almost nil as shown by the very low peak values.

3.3.2 Spatio-Temporal Structure of Body Movements

While hand movements can be well described by a lower dimensional representation, this is not the case for the whole body (Figure 3.7**A**). On average, 16 (range: 15–17) out of 23 possible PCs were required to explain 80% of the variance in terms of velocity. Comparing the PCs of the body data gives a very different picture from the hands: the distance between PCA subspaces starts off very high and drops off almost linearly. This makes the comparison of individual PCs difficult, as they lack the consistency found

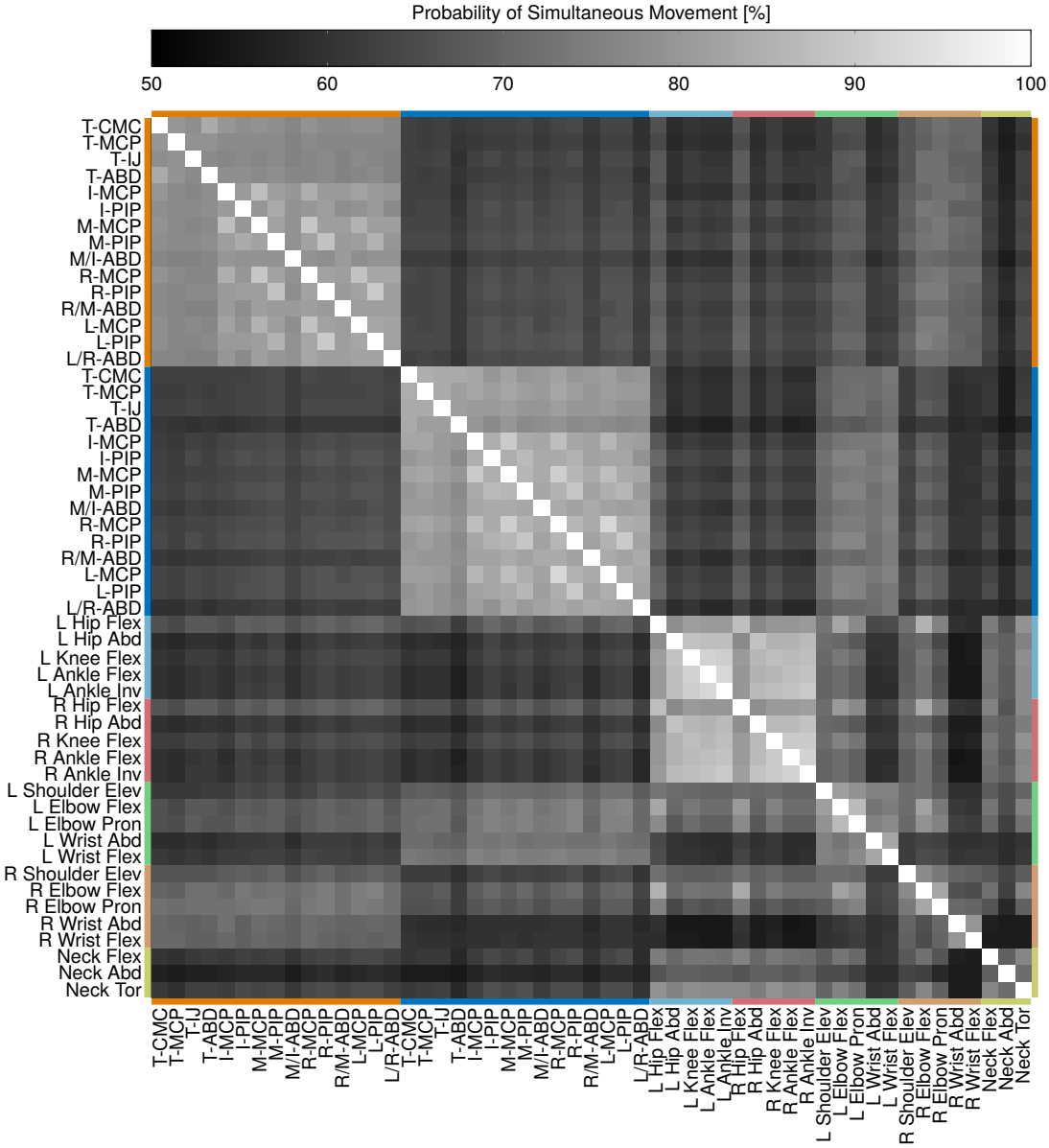

Figure 3.6: Probability of various joints moving simultaneously. The block-diagnoal structure suggests that individual limbs are more likely to move by themselves, with exception of the legs. See text for details. R: Right; L: Left; Flex: Flexion/Extension; Abd: Abduction/Adduction; Inv: Inversion/Eversion; Pron: Pronation/Supination; Tor: Torsion.

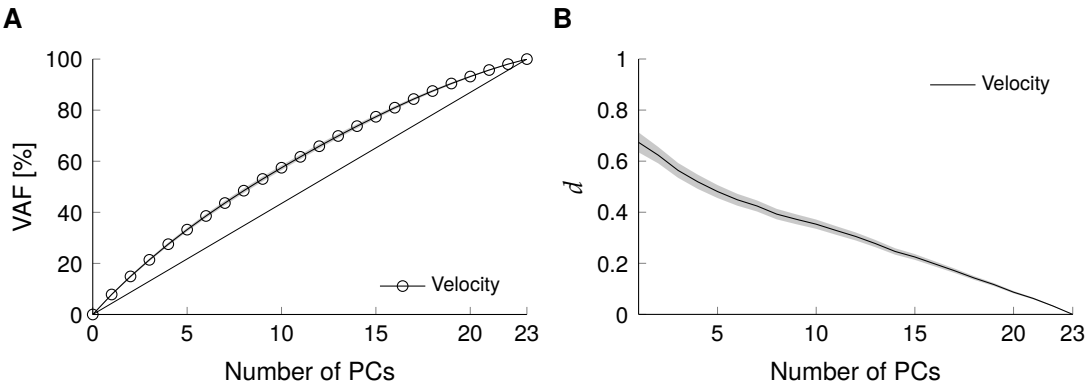

Figure 3.7: A Total variance captured by an increasing number of principal components for whole body, excluding hands. At least 16 PCs are required to exceed 80% VAF of the velocity. The solid line indicates the VAF curve expected for completely decorrelated data. **B** Inter-subject distance of PCA subspaces. Shading indicates standard error.

in the hand data. As a general observation, the first PC captures mainly movement of the legs, while the second PC is more representative of arm motion, in particular coordinated shoulder and elbow movement.

A potential explanation for this lack of low-dimensional structure may be given by Figure 3.6. Similarly to the hands, movement of one limb reduces the likelihood of a different limb moving, resulting in a block-diagonal structure, each block being one limb. An exception to this are the legs, which are likely to move simultaneously. Nevertheless, some additional structure can be seen in Figure 3.6: (1) arms and head occasionally move with the legs but wrists do not; (2) movement of the dominant hand's wrist suppresses head movements. This phenomenon is not as marked for the non-dominant hand; (3) finger movement is more likely while the respective arm is moving than when the opposite arm is moving (71% *vs.* 63%). This is particularly marked for wrist movements (71% *vs.* 61%).

In terms of temporal structure, body movements happen on very different time-scales depending on which limb is being considered, ranging from a median value of $\sim 400ms$ for movements of the ankle and wrist to more than $700ms$ for movements of

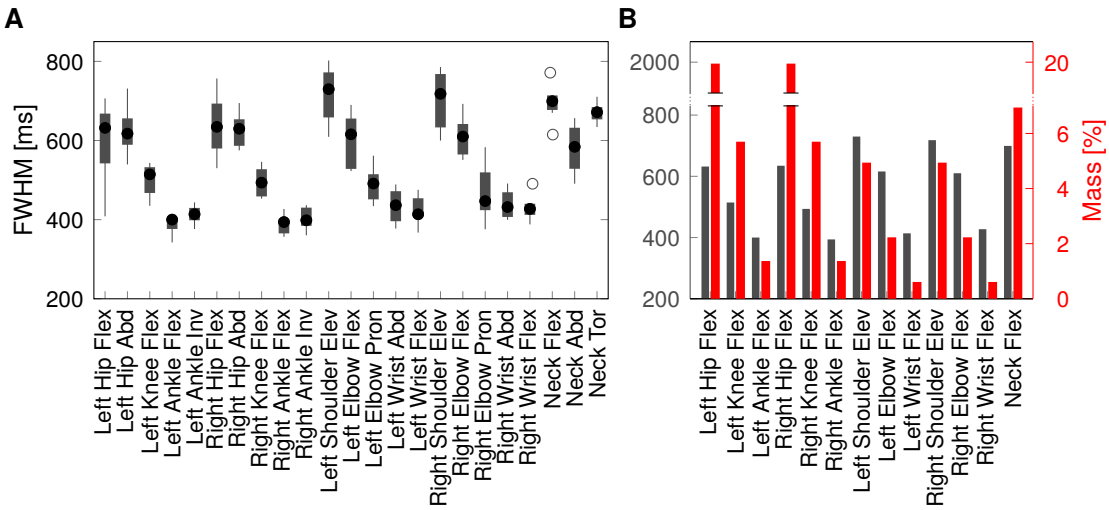

Figure 3.8: A FWHM for various parts of the body. **B** Comparison between the median FWHM and the percentage of body mass moved by the joint. A good relationship exists for most joints, with exception of the hip flexion which has a uncharacteristically low FWHM given the mass being moved.

the shoulders (Figure 3.8**A**). As for finger movements, there is no significant difference between the left and right side of the body. We hypothesised that the difference in time-scale is dependent on the mass attached to the joint moving, as a larger mass implies a larger moment of inertia which may translate into longer movement times[1]. This assumption holds to some extent (Figure 3.8**B**): within limbs, the FWHM of flexion movements decreases in the distal direction, consistent with the decreasing mass (e.g. the shoulder needs to move the mass of the entire arm, whereas the elbow only moves the forearm). The intercept of the function relating the FWHM to the mass is similar across limbs (legs: $397 \pm 21ms$; arms: $395 \pm 42ms$; paired t-test; $p = 0.85$) but the slope is very different (legs: $12 \pm 3\,[ms/\%]$; arms: $42 \pm 15\,[ms/\%]$; $p < 10^{-5}$). In all cases the GOF for the fit exceeded 0.8. Cross-correlation analysis of body movements did not reveal any consistent or strong patterns, almost all peaks (for non-zero lag) having values well below 0.2.

[1]The fraction of body mass attached to each joint was obtained from De Leva (1996).

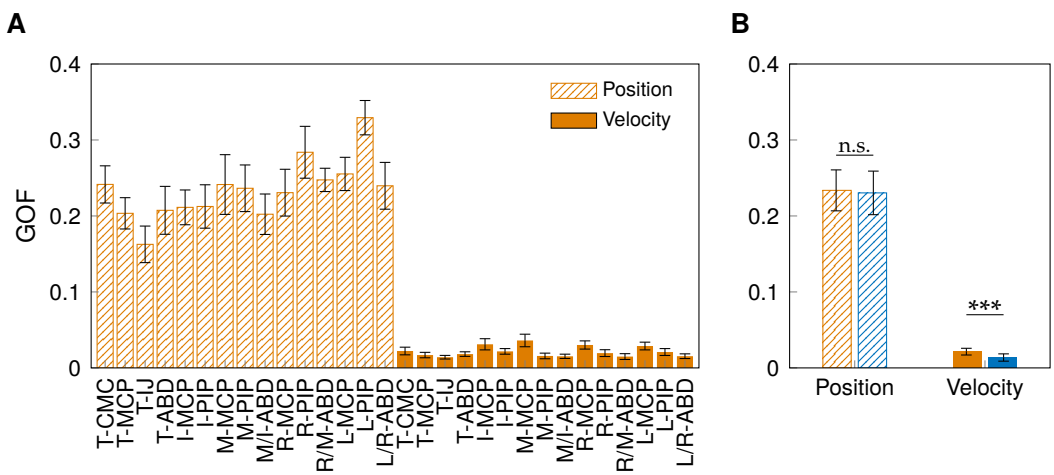

Figure 3.9: A Goodness of fit for the prediction of joint position and movement of the right hand as a linear combination of other body movements. **B** Comparison between the prediction quality of the right and left hand. Student's paired t-test; ***: $p < 10^{-3}$.

3.3.3 Arm-Hand Structure

With a view on improving user friendliness and facilitating control of multi-joint hand prostheses, we wanted to evaluate the possibility to control hand movements based on the motion and configuration of the rest of the body. In particular, we wanted to know whether it would be possible to predict finger movement from the ipsilateral elbow and shoulder, as well as the contralateral arm (including hand) and the neck. We chose a linear prediction model as previous studies reported no significant improvement by using more complex non-linear methods such as Gaussian Processes (Krasoulis, Vijayakumar, and Nazarpour 2015; Xiloyannis et al. 2015a). The results are shown in Figure 3.9. Prediction of both hand configuration and movement exceeded chance level (GOF: $\sim 10^{-4}$) but generally was not very good. The GOF for position is 0.23 ± 0.03, which is significantly better than for velocity (0.017 ± 0.005, see Figure 3.9**A**). There was no statistical difference for the prediction quality of hand configuration between the dominant and non-dominant hand (Figure 3.9**B**; Student's paired t-test, $p = 0.7$). Although significant difference exists for quality of the velocity prediction (GOF right:

0.021 ± 0.004; left: 0.014 ± 0.005; $p < 10^{-3}$), the quality remains very low.

Although the predictive quality of the model is relatively low, we can draw some interesting insight into the factors contributing to this prediction, as visualised by the relative contribution of the predictor variables' coefficient after correcting for different variances (Figure 3.10): (1) Position predicts position and velocity predicts velocity; (2) the ipsilateral arm contributes most to the prediction of hand state independently of the hand under consideration but (3) ipsilateral arm movement contributes significantly more when predicting motion of the right hand than for the left (Wilcoxon ranked sum test; $p < 10^{-6}$). These observations are consistent across all data sets examined.

3.4 Discussion

We compared kinematic synergies of both hand movements and body movements between different subjects and with synergies reported in the literature. We find that, for the hand, the synergies are extremely consistent across subjects and in good accord with the synergies described in a previous study of naturalistic hand movement (Ingram et al. 2008). Similarity with synergies from laboratory based tasks bore less resemblance with our data. Taken together, these two findings suggest that the synergistic structure of hand movements revealed by the analysis is not simply a representation of the biomechanical structure of the hand but most likely reveals something about the control of the hands by the brain. Note, however, that while individual synergies are different, the latent space they describe becomes increasingly similar with additional numbers of synergies (Figure 3.1). While none of the cited studies gave a detailed description of more than the first four synergies, we may expect that the similarity of laboratory based and naturalistic subspaces eventually drop down to levels of inter-subject variability. This would suggest that the low-dimensional structure observed is only partially caused by a simplification of the control structure by the brain, the

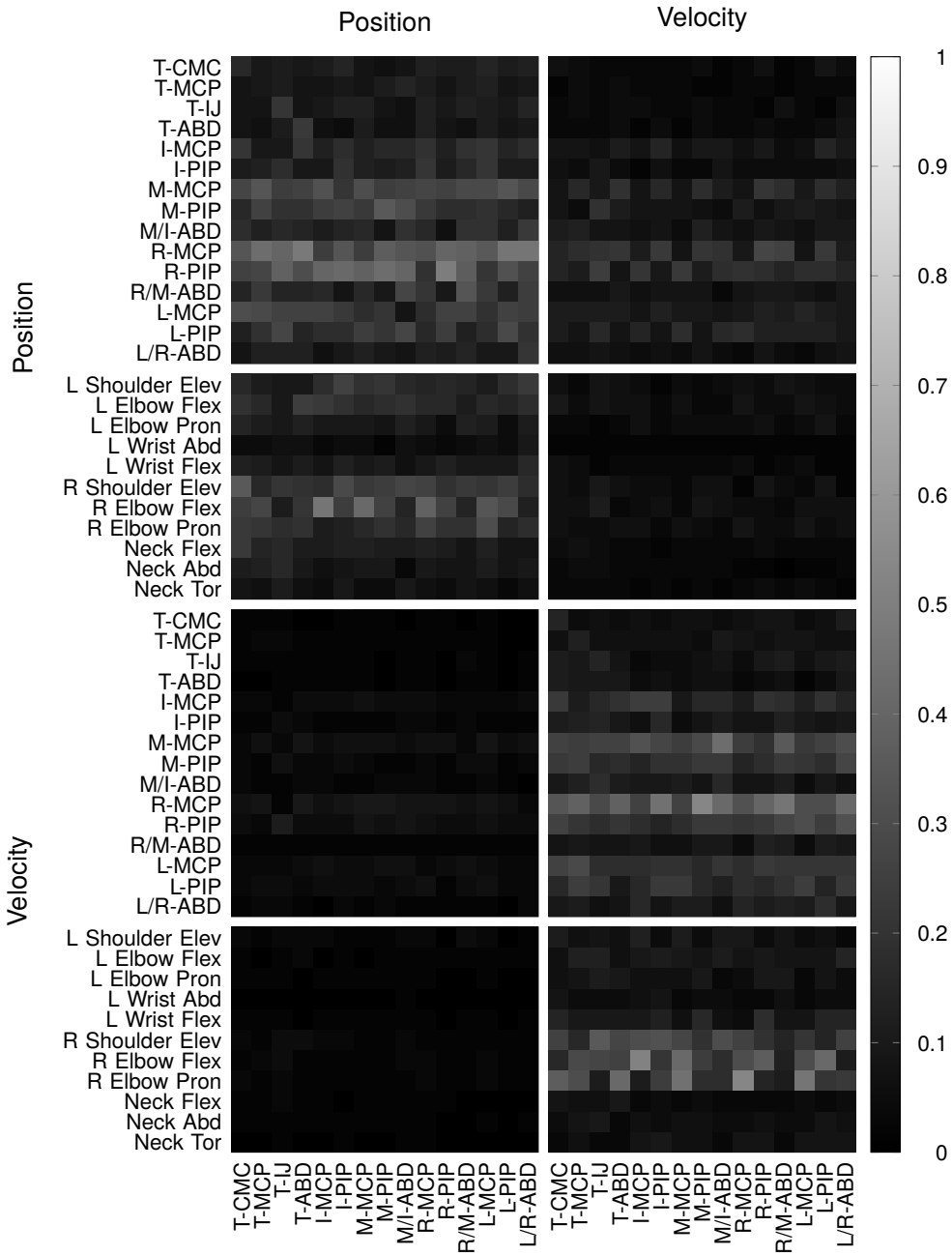

Figure 3.10: Normalised contribution of other joints on the prediction of the right hand's state, averaged across subjects. A clear separation can be seen when predicting the hand's position where the velocity component of the predictor variables have almost no influence at all. When predicting the hand's velocity, the separation is less clear but the strong influence of the ipsilateral shoulder and elbow are evident. The values were scaled by the predictor variable's SD for fair comparison.

rest being caused by the biomechanical structure of the hand which is well known to have muscles and tendons acting on multiple digits at once. This may also account for findings by (Todorov and Ghahramani 2004) which described a low-dimensional structure in hand movements even when subjects were required to move their fingers individually or at random.

The different number of synergies required to explain a given amount of variance in the laboratory data *vs.* natural tasks further emphasises the need to study behaviour in an unconstrained setting. This is particularly important when estimating the dimensionality of the space in which control occurs: it can be easily shown that PCA on a small amount of data with no spatial structure but significant auto-correlation at non-zero lags (such as is the case for movement data, see Figures 3.5**B** & 3.8**A**) will falsely reveal some amount of spatial structure. Thus, the short time-scales of movements in laboratory experiments combined with the repetitive and constrained nature of the tasks performed are likely to severely under-estimate the effective dimensionality of the latent subspace. A similar argument is made relative to muscle synergies by Valero-Cuevas, Venkadesan, and Todorov (2009).

Is this synergistic control a key feature of human motor control or is it specific to the hands? Both for movement of arms and legs, muscle synergies have been described (d'Avella et al. 2006; Ivanenko et al. 2007). Similarly, activity of muscles in the legs and the back has been shown to be synergistic in the context of gait and posture (Olree and Vaughan 1995; Torres-Oviedo, Macpherson, and Ting 2006; Torres-Oviedo and Ting 2007), but to the best of our knowledge, multi-limb synergies during natural behaviour have not been studied. Even if they had been, our results suggest that few muscle synergies bridging multiple limbs would have been found. From a purely kinematic perspective, very little structure (and even less consistency) is found when analysing movement of major body limbs (see Figure 3.6). The exception to this rule is movement of the legs, consistently with aforementioned studies.

The "hand exception" may be explained by a combination of kinematic and neurological factors: as revealed in our analysis of temporal structure of natural movement, the digits of the hand move together much more frequently than other limbs do (except maybe the legs). It thus makes more sense to have a synergistic control structure for fingers than for limbs which will move independently from each other anyway. This is added to (or maybe a consequence of) the fact that the joints of the hand are tightly linked mechanically. This linkage also requires muscle activation to prevent fingers from moving (e.g. Contraction of the index compartment of the finger extensor muscle when the middle finger is flexed). This antagonistic co-activation of muscles is reflected in the hand area of the motor cortex by two factors: (1) reduced somatotopy with neurons acting on the same muscle spread over a large area and (2) branching of descending connections where a single neuron may innervate muscles required for the movement of multiple, non-adjacent fingers (Schieber and Hibbard 1993; Weiss and Flanders 2004). These features of the neuronal substrate of hand movements have been suggested to be essential for the implementation of a synergistic controller at a cortical level (Schieber and Santello 2004; Santello, Baud-Bovy, and Jörntell 2013). Yet, while models for neural control of synergies have been proposed (Hart and Giszter 2010; Santello, Baud-Bovy, and Jörntell 2013), electrophysiological evidence is still lacking. Although TMS activation of the hand area in M1 causes hand movements resembling synergies (Gentner and Classen 2006), a recent study of multi-electrode array recordings in the primary motor (M1) and ventral premotor (PMv) cortex of non-human primates failed to find evidence for neuronal activation which was more correlated with synergies than with individual joint movement (Mollazadeh et al. 2014). Similarly, postural and muscle synergies in the hand do not display a clear one–to–one relationship. Instead, muscle synergies of the hand are active in an unpredictable fashion over a range of postural synergies (Weiss and Flanders 2004).

3.4.1 An Alternative Hypothesis

Given the inconsistency of kinematic synergies across subjects and experimental settings, as well as their lack of demonstrable neuronal representation one might argue that the apparent synergistic control of the hand is an artefact of the analysis method rather than an actual control strategy of the central nervous system (CNS). A similar argument has been made repeatedly against work on muscle synergies, arguing that what appears to be a low-dimensional control space is really only the consequence of the *minimum intervention* hypothesis (Kutch et al. 2008; Valero-Cuevas, Venkadesan, and Todorov 2009). This theory, which directly flows out of the concept of optimal feedback control (Todorov and Jordan 2002) suggests that the CNS does not try to control task-irrelevant variables, leaving them to fluctuate freely. Without getting drawn into that argument, it should nevertheless be noted that using PCA to find meaningful synergies in kinematic data has two major limitations: (1) PCs are required to be orthogonal to each other. While this has a lot of useful mathematical properties, it is unclear why it should apply to kinematic or muscular synergies. (2) The number of synergies will always be strictly smaller or equal the number of DOF being controlled. This is a necessary consequence of (1), but may result in sub-optimal encoding of data in some cases, as illustrated in Figure 3.11. We thus propose a slightly modified version of the PCA approach to find synergies in kinematic data which overcomes the aforementioned problems. As a reminder, the classic PCA methodology assumes that

$$\mathbf{X} = \mathbf{Z}\mathbf{B} \tag{3.1}$$

where $\mathbf{X} \in \mathbb{R}^{T \times D}$ is a matrix of T samples of kinematic data, $\mathbf{B} \in \mathbb{R}^{D \times D}$ a matrix of principal components and $\mathbf{Z} \in \mathbb{R}^{T \times D}$ the latent representation. Typically, $\mathbf{B}$ and $\mathbf{Z}$ are truncated to retain only the first $Q < D$ PCs. We modify this model in a small but critical way by (1) removing the requirement for all elements in $\mathbf{B}$ to be orthogonal

to each other and (2) allowing $Q \gg D$. This may seem counter-intuitive as we are now increasing the dimensionality of the control space, thus worsening the problem by adding complexity. We avoid this by stipulating that at any moment in time only $K \ll Q$ of the weights may be non-zero. This approach is similar to seminal work by Olshausen et al. (1996) relating to sparse coding in the primary visual cortex (V1) which elegantly demonstrated that a sparse code accurately describes the receptive field of V1 neurons, which PCA fails to do. In analogy, we suggest that the CNS may have access to a dictionary of *sparse eigenmotions* (SEMs, the new $\mathbf{b}_i$) from which it can pick a few elements and combine them to achieve its desired goal (Rizzolatti and Gentilucci 1988).

The hypothesis of sparse encoding of motor output has been mentioned before as interesting possibility (e.g. Tresch and Jarc 2009; Ingram et al. 2008) but has not been investigated further to the best of our knowledge. However, there has been some recent interest in using sparse encoding of kinematics (termed "efficient code" by the authors) to classify movements (Johnson and Ballard 2014). The authors find that sparse coding outperforms PCA and random projections in most cases but do not further investigate its application to motor control. We hypothesise that sparse encoding of movement may account for observed grasp-type specificity in neurons of the monkey F5 (equivalent to human PMv) area (Rizzolatti and Gentilucci 1988; Raos et al. 2006) and may help understand the neural mechanisms of learning by observation (Di Pellegrino et al. 1992). Testing of these hypotheses will require more work from a

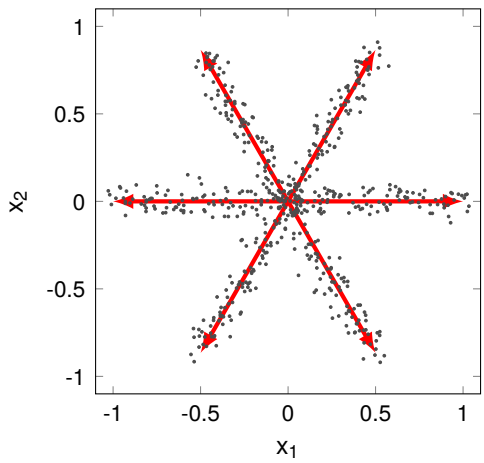

Figure 3.11: Example of six noisy 1-D structures embedded in a 2-D plane. PCA would fail to identify the low-dimensional structure whereas a sparse code should extract the red lines.

methodological (how do we identify sparse eigenmotions?) and electrophysiological perspective to verify whether neurons are responsive to this type of encoding.

4

Sparse Eigenmotion Decomposition

"The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them."

— Sir William Bragg

4.1 Introduction

The spatial structure of movements – and evidence for low-dimensional control manifolds – has typically been investigated using principal component analysis (PCA, Santello, Flanders, and Soechting 1998; Mason, Gomez, and Ebner 2001; Gentner and Classen 2006; Ingram et al. 2008; Faisal et al. 2010; Belić and Faisal 2015). Identifying a low-dimensional structure in hand movements has multiple desirable properties: (1) it simplifies the control problem, (2) provides insights into the dynamics governing the data and (3) allows for comparison of control strategies across subjects and settings. As argued previously (Section 3.4), Principal Component Analysis is severely limited by imposing orthogonality between the principal components. This imposes the additional constraint that the number of principal components cannot exceed the dimensionality of the input data. From a mathematical perspective, this low-dimensional encoding can be written as:

$$\mathbf{X} = \mathbf{B}\mathbf{Z} + \boldsymbol{\Xi} \tag{4.1}$$

where $\mathbf{X} \in \mathbb{R}^{DxT}$ is the actual movement , $\mathbf{B} \in \mathbb{R}^{DxQ}$ a matrix where each column is a principal component (PC) and $\mathbf{Z} \in \mathbb{R}^{QxT}$ refers to the individual activation of each PC over time, also referred to as latent representation. The matrix $\mathbf{\Xi} \in \mathbb{R}^{DxT}$ represents the noise and residuals which the linear model cannot capture. In the case of PCA, $Q \leq D$ and preferably $Q \ll D$ so as to capture as much of the variance in the data with as few PCs as possible.

To achieve a high dimensionality reduction while avoiding the limitations of PCA, Olshausen and Field (1997) implemented a method known as *sparse coding*. In essence, this modifies the terms in Equation 4.1 by allowing **B**, the *dictionary*, to be over-complete (i.e. $Q \geq D$ and potentially $Q \gg D$), which also implies that the elements of **B** no longer need to be orthogonal to each other. The sparsity resides in **Z** where for every time point $t \in T$ almost all elements are 0. However, an over-complete basis does not directly imply a sparse latent representation. Indeed, identifying such a sparse encoding from known data **X** and dictionary **B** is an active area of research.

Using over-complete dictionaries to represent the data has several advantages over the use of complete bases (such as e.g. PCA or the Fourier basis). As the atoms are more specialised, they are capable of capturing higher-order correlations in the data (Mairal et al. 2009b), and provide greater interpretability. This typically implies that the latent representation of the data can be more sparse than in the case of complete bases, thus allowing for higher compression of the data (e.g. the JPEG2000 image compression standard is based on (sparse) wavelet decomposition). Further, it has been suggested that coding based on an over-complete dictionary has benefits in terms of robustness of the latent representation in the presence of additive noise on the data being encoded (Simoncelli et al. 1992; Donoho, Elad, and Temlyakov 2006). Finally, sparse codes have been claimed to be useful for feature and novelty detection (Barlow 1989).

Although the advantages of using an over-complete representation are clear, defining an appropriate dictionary is a difficult task. Traditionally, these have been hand-

crafted to have specific desirable properties (such as being invertible), and being well suited to the data being compressed. As they originate from the image and signal processing field, many such dictionaries have been designed to capture time-frequency characteristics of the underlying signals, and are based on families of functions such as wavelets, curvelets (Candes and Donoho 2000), bandlets (Le Pennec and Mallat 2005) or controurlets (Do and Vetterli 2005). A drawback of using pre-defined dictionaries is that the atoms they contain may not be suited for the data considered, thus yielding sub-optimal compression and hindering interpretation. Deriving a dictionary directly from the data has the additional benefit of providing insights into the underlying statistics of the data being analysed. Indeed, learning the dictionary directly from the data has led to several significant improvements in terms of image processing, such as image restoration (Mairal, Elad, and Sapiro 2008; Mairal et al. 2009a; Mairal, Bach, and Ponce 2012), inpainting (Mairal, Elad, and Sapiro 2008; Peyré 2009), denoising (Simoncelli et al. 1992) and synthesis (Peyré 2009). Sound processing is another typical application of sparse coding (Grosse et al. 2007; Févotte, Bertin, and Durrieu 2009), where it is closely related to the problem of blind source separation. In particular, learnt dictionaries have been found to be particularly useful for learning classifiers by automatically identifying interesting features. This is helpful both in the supervised (Mairal et al. 2008; Yang et al. 2009; Mairal, Bach, and Ponce 2012) as well as the semi-supervised case (Mairal, Bach, and Ponce 2012). Remarkably, a learnt dictionary is also the basis for *self-learning* classifiers (Raina et al. 2007), a particular type of classifiers which learn features not only from a (small) sample of labelled data, but mainly from a large volume of unlabelled data, without assuming that the unlabelled data belong to the same classes as the labelled data.

Of particular interest to our research is the work by Bruno Olshausen and David Field (Olshausen et al. 1996; Olshausen and Field 1996; Olshausen and Field 1997) on sparse coding in the visual system. By learning an over-complete, sparse representation

of natural images, they show that the resulting atoms have high similarity with the receptive fields of neurons in the macaque V1 cortex (Olshausen and Field 1997). Given the brain's highly parallel structure and the many advantages provided by sparse coding, it seems plausible that this strategy evolved to encode sensory stimuli (Field 1994) and has been mentioned as an alternative strategy to muscle synergies in motor control (Tresch and Jarc 2009).

4.2 Dictionary Learning

Training of over-complete dictionaries directly from the data is an active area of research. While a number of different approaches have been developed, they commonly consider Equation 4.1 with an additional penalty term as an optimisation problem:

$$\arg\min_{\mathbf{B},\mathbf{Z}} \|\mathbf{X} - \mathbf{B}\mathbf{Z}\|_F^2 + \lambda \|\mathbf{z}_t\|_p \tag{4.2}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, p indicates the degree of the L^p–norm and λ is a Lagrange multiplier. A truly sparse solution can theoretically be obtained by penalising the activation of atoms by the L^0 pseudo-norm, which simply counts the number of non-zero elements in $\mathbf{z}$. This approach has major drawbacks as it makes Equation 4.2 non-continuous and is unable to cope with noise in the data (Wohlberg 2003). As a compromise, the L^1–norm is often used, and it has been shown that under certain conditions it yields the best sparse approximation to the data (Mairal et al. 2010).

Data-driven dictionary learning is first mentioned in work by Olshausen et al. (1996) which used patches from natural images as inputs and optimised Equation 4.2[1] using gradient descent[2]. Numerous modifications and enhancements were made to the

[1]Olshausen et al. (1996) use different cost functions to obtain sparsity: (1) $-e^{-z^2}$, (2) $\log(1 + z^2)$ and (3) $|z|$, which is equivalent to $\|\cdot\|_1$.

[2]It should be noted here that the problem is not jointly convex in $\mathbf{B}$ and $\mathbf{Z}$ and that it has a degenerate

original framework by Olshausen and Field to give it a probabilistic interpretation. These methods promote sparsity by implying a Cauchy (Olshausen and Field 1997) or Laplace[3] (Lewicki and Sejnowski 2000; Kreutz-Delgado et al. 2003; Bradley and Bagnell 2009; Mairal et al. 2009b) prior on the latent representation and finding $\mathbf{B}$ and $\mathbf{Z}$ based the Maximum Likelihood (ML) or Maximum-A-Posteriori (MAP) estimate of the data model. This has lead to a number of new algorithms for data-driven dictionary learning (Aharon, Elad, and Bruckstein 2006; Mairal et al. 2008; Mairal, Bach, and Ponce 2012) as well as variants capable of handling non-negativity constraints (Hoyer 2004) or dealing with data online (Mairal et al. 2010). A significant drawback of current methods is that the size of the dictionary needs to be specified in advance, which means that estimating the true dimensionality of the data is close to impossible since it requires sweeping over the number of elements in the dictionary as well as having a reasonable metric of model efficiency. Choosing too few elements is likely to generate a non-sparse representation, whereas too many will cause some elements to be either very similar or not used. Similarly, there is no guarantee that the bases chosen to represent a particular data point do not conflict with each other (i.e. cancel each other out along one dimension). Although this is not a problem in most applications, it would be preferable if two simultaneously acting eigenmotions did not try to pull a joint into different directions (but see Lesage et al. (2005) for a similar concept).

The following sections introduce a new method called *Sparse Eigenmotion Decomposition* (SEMD). It is a two stage algorithm which extracts the over-complete representation of the data by obtaining the principal components from small data windows strategically chosen within the time-series, thus guaranteeing local orthogonality. Once a

solution for $\mathbf{Z} \to \mathbf{0}$ and $\mathbf{B} \to \infty$. This problem is generally addressed by renormalising the elements in $\mathbf{B}$ at each iteration of the gradient descent, but it is unclear how this affects the objective function being optimised (Lewicki and Sejnowski 2000).

[3] This is equivalent to penalising $\mathbf{z}$ with an L^1–norm. Using the L^2–norm would be assuming a normal distribution. While mathematically a lot simpler to deal with, a normal prior does not typically provide a sparse solution (Tibshirani 1996).

satisfactory representation of the data is obtained, it then further refines this representation by reducing conflict among eigenmotions. This allows us to automatically estimate the number of elements within the dictionary, making SEMD an unsupervised learning method with the added benefit of low computational cost.

4.3 Methods

4.3.1 Synthetic Eigenmotion Generation

Synthetic sparse eigenmotions (SEMs) were generated following the method described in Algorithm 1. This results in each synthetic SEM having a random number of large non-zero values with the remaining entries being very small. Alternatively, the addition of random noise ϵ may be skipped during the generation process and added again each time the SEM is invoked in the data generation process. This has the advantage of adding extra variability to the eigenmotion, as would be expected in a biological setting.

Algorithm 1 Synthetic eigenmotion generation

```
function CREATESEMS(D, Q)
    for i = 1 ... D do
        b_i ← 0_{D×1}
        n ~ U{0, D}
        for j = 1 ... n do
            idx ~ U{0, D} s.t. b_{i,idx} = 0
            b_{i,idx} ← N(0, 1)
        end for
        ε ← N(0, 0.05)
        b_i ← b_i + ε
        b_i ← b_i / ||b_i||_2
    end for
return B
end function
```

4.3.2 Synthetic Data Generation

Synthetic data was generated for validation purposes following the methodology described in Algorithm 2. The velocity profile f may be chosen as any square-integrable function with finite support. For convenience, and because of its similarity to human movement (Morasso 1981; Abend, Bizzi, and Morasso 1982; Flash and Hogan 1985), we chose a truncated and randomly scaled bell curve with a width proportional to the width of the support l. This in turn was uniformly sampled in the range $[100, 200]$. The function's starting point was randomly shifted by a small value Δt in each dimension, such that the peaks of the bell curve were not perfectly aligned, in agreement with empirical observations. When deciding whether to start a new movement or not, we took into consideration whether (1) the previous movement started more than 20 samples

Algorithm 2 Synthetic data generation

```
function CREATEDATA(T, D, Q)
    t ← 0
    B ← createSEMS(D, Q)
    X ← 0_{T×D}
    while t < T do
        if can start new movement then
            ϵ ← N(0, 0.05)
            s ← random compatible SEM ∈ B
            s ← s + ϵ
            f ← arbitrary velocity profile of length l
            for all d ∈ D do
                if s_d ≠ 0 then
                    Δt ← N(0, 0.01l)
                    X_{i,d} ← s_d X_{i,d} + f, i = {t + Δt, ..., t + l + Δt}
                end if
            end for
            t ← t + 1
        else
            t ← t + 1
        end if
    end while
end function
```

ago; (2) the current number of active SEMs is lower than 10, (3) at least one dimension is not currently moving and (4) there are compatible SEMs available in the dictionary. SEM compatibility requires that all currently active dimensions have 0 weight in the new SEM to be selected.

4.3.3 Eigenmotion Comparison

When extracting SEMs from one or multiple time-series, it is crucial to be able to compare them. This avoids the extraction of duplicate SEMs from the same time-series and allows to compare the SEMs extracted from data belonging to different subjects. The task of comparing two or more SEMs is equivalent to comparing D-dimensional vectors. This is usually achieved by calculating the Pearson correlation coefficient (Santello, Flanders, and Soechting 2002; Chvatal et al. 2011; Delis et al. 2013) or occasionally the scalar product (d'Avella, Saltiel, and Bizzi 2003) between SEMs. However, visual inspection of SEMs deemed similar by these metrics revealed widely varying results and inconsistencies as some SEMs where considered similar when considering the scalar product but not with respect to their correlation coefficient, and vice-versa. To circumvent this problem, we compare SEMs using not one but six different metrics:

- Pearson correlation coefficient;
- Euclidean distance;
- Scalar product;
- Synergy intersection;
- Functional correlation; and
- Functional GOF.

The last three metrics will be explained in more detail in the following sections. For all of these metrics, significance values were estimated empirically from the distribution obtained by computing the corresponding metric for each possible pair of 10'000

principal components computed from randomly sampled data windows.

Eigenmotion Intersection

The idea of SEM intersection is derived from the histogram intersection metric proposed by Swain and Ballard (1991) for comparing colour histograms in images. In brief, the original histogram intersection metric compares two histograms A and B with n corresponding bins by calculating:

$$\frac{\sum_{i=1}^{n} \min(A_i, B_i)}{\sum_{i=1}^{N} B_i} \tag{4.3}$$

where A_i is the value in the i-th bin of the histogram.

We apply this method to SEMs by considering each dimension of the SEM as a bin. It should be noted that unlike the SEMs considered in this work, colour histograms cannot have negative values in their bins. Therefore, the metric was modified to ensure compatibility with bin values of opposing sign. Given two Q-dimensional SEMs **a** and **b** which have been normalised such that $\sum_{n=1}^{Q} |a_i| = 1$ (where $|\cdot|$ indicates the absolute value):

$$H_i = \begin{cases} \min(A^i, B^i) & if sgn(A_) = sgn(B_i) \\ -|A_i - B_i| & if sgn(A_i) \neq sgn(B^i) \\ 0 & if sgn(A_i) = 0 \text{ or } sgn(B_i) = 0 \end{cases} \tag{4.4}$$

where $sgn(\cdot)$ is the sign function and

$$\rho_{inter} = \sum_{i=1}^{Q} H_i \tag{4.5}$$

This gives a similarity measure which is bounded between −2 (if **a** = −**b**) and 1 (if **a** = **b**). As for all purposes of this work the sign of the SEM is irrelevant, any comparison

between **a** and **b** is also computed for **a** and –**b** and the maximum value taken, thus leading to a distribution in the interval $[-1, 1]$.

Functional Similarity

Instead of considering the SEMs themselves, the concept of functional similarity compares the relevance of the SEMs in the data. This is achieved by computing the responsibility, z, of each SEM in reconstructing a random set of data. For n SEMs and a $T \times Q$ time series, this results in n, $T \times 1$ time series of responsibility.

Subsequently, the *functional correlation* is defined as the Pearson correlation coefficient between the time series of responsibility of SEM **a** and **b**:

$$\rho_{FC} = \rho(\mathbf{z}_{\mathbf{a}}^2, \mathbf{z}_{\mathbf{b}}^2) \tag{4.6}$$

and the *functional* R^2 as the R^2 between the time-series:

$$\rho_{FR^2} = \mathrm{R}^2(\mathbf{z}_{\mathbf{a}}^2, \mathbf{z}_{\mathbf{b}}^2) \tag{4.7}$$

Voting Procedure

As mentioned in the introduction to this section, individual similarity metrics were found to regularly disagree on whether two SEMs are the same. Using the six metrics described previously, we introduce a majority voting system where two synergies are deemed similar with confidence α if 3 or more metrics find them similar at a significance level of $1 - \alpha$. Note that both functional similarity metrics get a single, combined vote as their results are very closely related.

4.3.4 Goodness of Fit (GOF)

Also known as *coefficient of determination* or R^2, the goodness of fit is generally used as a way of measuring how well a model fits the data. Typically[4], this is computed as

$$\mathrm{GOF} = 1 - \frac{\mathrm{VAR}[\boldsymbol{x} - \hat{\boldsymbol{x}}]}{\mathrm{VAR}[\boldsymbol{x}]} \tag{4.8}$$

where $\boldsymbol{x}$ represents the data and $\hat{\boldsymbol{x}}$ the data as estimated from the model. While this measure is perfectly acceptable for short and univariate data sets or simple linear fits, it has several drawbacks when considering multivariate and/or very long data. When considering multivariate data, a trivial extension of Equation 4.8 can be obtained by considering how well each dimension of the data is fitted and take the average GOF over all dimensions D:

$$\mathrm{GOF} = \frac{1}{D} \sum_{i=1}^{D} \mathrm{GOF}_i \tag{4.9}$$

This measure does not, however, take into account that some of the dimensions may be a lot more variable (and thus convey more information[5]) than others. This can be remedied by using a variance-weighted GOF – which is typically known as Variance Accounted For (VAF) in the dimensionality reduction literature (Cheung et al. 2005) – by modifying Equation 4.9 to account for the different variability of dimensions:

$$\mathrm{VAF} = \frac{1}{D} \sum_{i=1}^{D} w_i \mathrm{GOF}_i \tag{4.10}$$

where $w_i = \mathrm{VAR}[\boldsymbol{x}_i] / \sum_{i=1}^{D} \mathrm{VAR}[\boldsymbol{x}_i]$.

A further disadvantage of the classic GOF is the global nature of this measure as it

[4]Some authors define the R^2 as the square of the correlation coefficient between the model and the fit. This measure is misleading because it does not consider whether the mean value of the fit matches that of the data, nor whether the scaling is accurate.

[5]Under the assumption that all dimensions are affected by the same amount of noise.

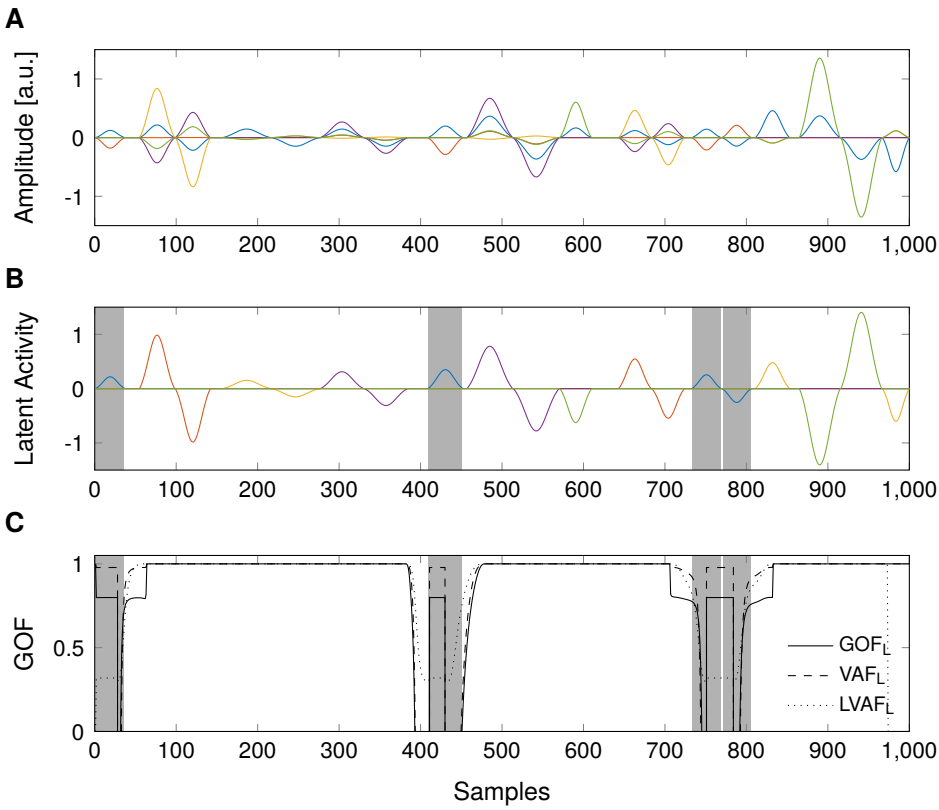

Figure 4.1: Comparison of goodness of fit measures. **A** Original data generated by Algorithm 2. **B** Latent activity of the atoms. The first atom (grey area) was removed from the dictionary during reconstruction. **C** GOF measures for the reconstruction. All three show a value below 1 in the shaded areas, but most significantly for the VAL$_L$ metric.

only captures an average goodness of fit over the entire time-series. In general, it would be interesting to know at what moment in time the model fits the data poorly, as this may reflect an event of particular interest. For the specific case of this work, this local knowledge is crucial as it allows to detect the presence of previously unseen kinematic synergies. We therefore introduce the concept of local GOF (GOF$_L$):

$$\mathrm{GOF}_L(t) = 1 - \frac{\mathrm{VAR}[\boldsymbol{x}_W(t) - \hat{\boldsymbol{x}}_W(t)]}{\mathrm{VAR}[\boldsymbol{x}_W(t)]} \tag{4.11}$$

where $\boldsymbol{x}_W(t)$ and $\hat{\boldsymbol{x}}_W(t)$ are the data and its estimate taken from a window of size W

around the time-step t. An example of its application is shown in Figure 4.1. In regions where the model fits the data well (e.g. between sample 100 and 350), the $\mathrm{GOF_L}$ is high. In regions where the fit is less good (e.g. because the atom is not present in the dictionary), the $\mathrm{GOF_L}$ is low, indicating that that particular region may require a different model. It should be noted that taking a windowed approach to computing variance causes unwanted effects at the boundaries (i.e. beginning and end) of the time-series as the local variance can no longer be accurately approximated. Thus, the $\mathrm{GOF_L}$ for time-points $t < \left\lceil \frac{W}{2} \right\rceil$ should be considered with care. This measure can be trivially extended to multivariate time-series by using different weights for each dimension: equal ($\mathrm{GOF_L}$ – Figure 4.1**C**, solid line), total variance ($\mathrm{VAF_L}$, dashed line) or local variance ($\mathrm{LVAF_L}$, dotted line). Note that the latter metric is particularly useful for identifying poor model fit as it does not assume the model is stationary, but allows for variable variance.

4.3.5 Saturating Norm

The aim of this metric is to provide a quasi-binary measure of similarity between objects. In effect, it gives a distance of 1 for any two points which are further apart than a certain radius. This translates mathematically into the form:

$$\|\mathbf{x}\|_s = 1 - \exp(-\|x\|_2^2/\sigma) \tag{4.12}$$

where $\sigma \in \mathbb{R}^+$ is a parameter which controls how rapidly the metric saturates to 1.

4.4 The SEMD Algorithm

The algorithm's function is outlined in Figure 4.2 and can be broadly divided into two stages which are computationally similar but follow different objectives. The first

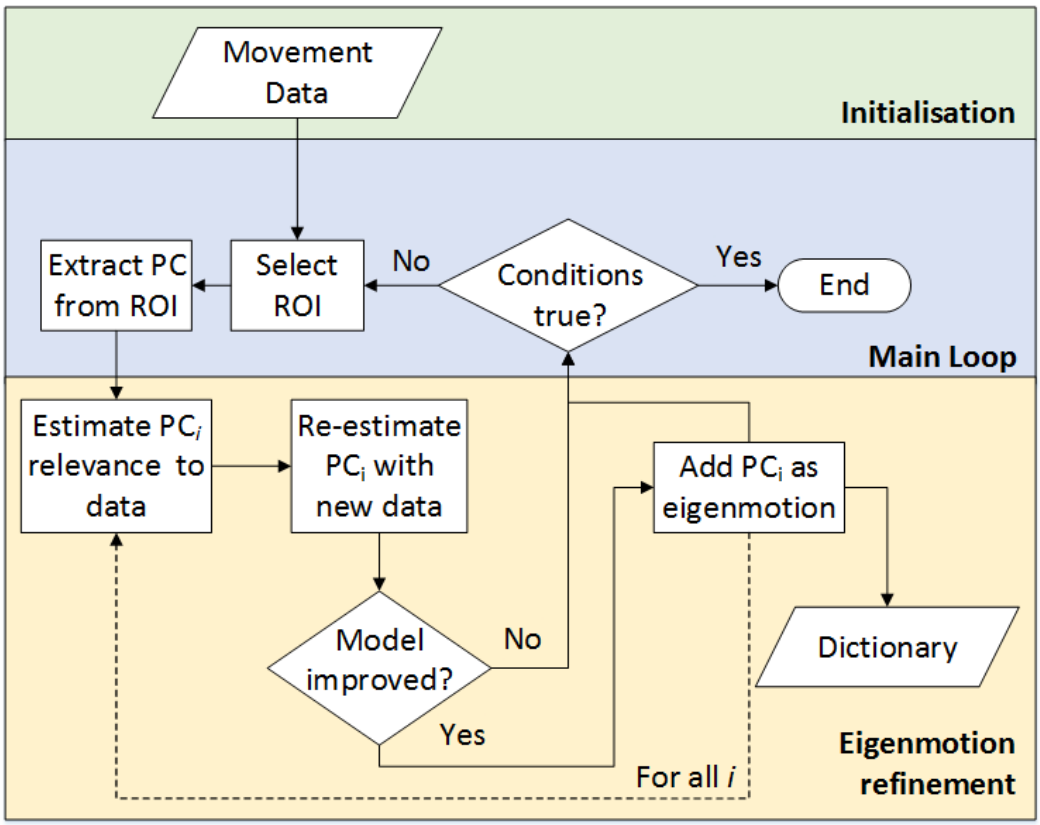

Figure 4.2: Sparse Eigenmotion Decomposition algorithm flowchart. After pre-processing of the data, a small chunk of data is selected and eigenmotions are extracted from it. These are refined and selected if they fulfil some usefulness criteria. After the first termination condition is reached, the algorithm proceeds to a second stage to reduce the interaction between atoms. Reaching of a further termination condition ends the algorithm. See text for details.

run aims at identifying a sufficient number of eigenmotions to accurately reconstruct the data; we will refer to this step as *reconstruction stage*. In the second iteration, we further reduce the local dimensionality of the latent variable structure by reducing conflict between bases and identifying additional eigenmotions which are better suited to capture the data locally; this step will be referred to as *interference reduction stage*. The following sections detail each processing stage of the algorithm and illustrate them as appropriate.

4.4.1 Initialisation

In the initialisation step, the time-series is separated into episodes of movement and immobility. The aim of this is to concentrate the analysis on active movements and avoid fitting the noise in the data. This is achieved by computing the empirical distribution for each dimension of the data. Any data point for which at least one dimension has an amplitude with corresponding $p - \text{value} < 0.2$ is classified as movement, the others being classified as being static. To avoid zero-crossings to be counted as static or noise as movement, a morphological closing operation followed by morphological opening is applied to the binary classification time-series using a linear structuring element. This has the effect of removing small islands of movement and immobility (see Figure 4.3). A disadvantage of this motion detection method is that very slow movements are potentially missed.

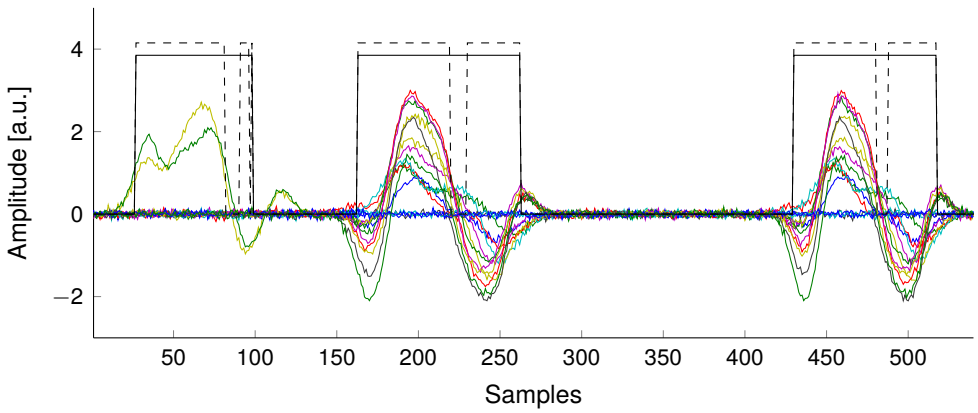

Figure 4.3: The data preprocessing stage discards all regions in which there is no significant change in the data to avoid the EMD algorithm trying to fit noise. Small islands of activity are removed by morphological opening, while excessive fragmentation of the output and inclusion of zero-crossings is obtained by morphological opening of the time-series. Compare the original output (dashed line) with the filtered output (solid line).

4.4.2 Region-of-Interest Selection

Identifying regions of interest (ROI) in the time-series which can provide us with novel information is a key step in the SEMD algorithm. When selecting the size of the ROI (we will refer to it as T_w), it is essential to take into account two potentially conflicting needs: (1) time-scale of the underlying dynamics and (2) numerical stability of the estimated statistics[6]. Selection criteria for the ROI depend on the stage of the algorithm, and can be divided into three different categories:

Initialisation

When the algorithm is confronted with a new time-series for the first time, the ROI is randomly selected from the data.

Reconstruction

During the reconstruction stage of the algorithm, the new ROI is selected as the data around the time-point with the lowest reconstruction score (see Section 4.4.4 for reconstruction of data with non-orthogonal basis functions and Section 4.3.4 regarding the reconstruction score).

Interference Reduction

In the interference reduction stage, the ROI is selected around the time-point with the highest basis conflict score C_{B}. This value is computed on a dimension-by-dimension basis by subtracting the contribution from the most important eigenmotion from the sum of all contributions. If only one eigenmotion influences a given dimension, the score is 0. Variations in data amplitude are compensated through non-linear scaling via a saturating 'norm' (see Section 4.3.5):

$$C_{\mathrm{B}} = \| \sum_{i=1}^{Q} |z_i| - \max(|\mathbf{z}|) \|_s \tag{4.13}$$

[6]For instance, the covariance matrix of a D-dimensional data set has $2D - 1$ DoF and thus requires at least as many data points to avoid over-fitting.

It should be noted that in all cases, the ROI selected is then discarded to ensure that it does not get selected on multiple occasions.

4.4.3 SEM Extraction and Refinement

The SEMs we seek can be considered as local principal components (which are equivalent to atoms of our dictionary). They can thus be estimated from the data within the ROI selected using classic PCA. The number of significant PCs extracted for each ROI is automatically determined using the approach described by Cheung et al. (2005) with a threshold set at 10^{-4}. This method iteratively fits a linear function through the VAF curve and terminates when the root mean squared error (RMSE) falls below a given value. Because of the limited amount of data present (see Section 4.4.2 for an explanation), this estimate may be inaccurate. This inaccuracy is mitigated in two different ways: (1) the sample covariance matrix used as input to the PCA algorithm is bootstrapped for a more accurate estimate and (2) the extracted PCs go through a refinement process:

Significance Estimate

For each atom extracted in the ROI, its relevance to the entire data set needs to be estimated before it can be taken up into the dictionary. This mitigates the risk of extracting "noise SEM" or linear combinations of SEMs. Relevance for each sample in the data is estimated by computing the empirical p-value of the reconstruction quality (specifically the $\mathrm{LVAF_L}$ metric, see Section 4.3.4). The *empirical* aspect emerges from the comparison of the reconstruction of the data by the extracted PC with the reconstruction from a large number (typically 1000) of randomly generated SEMs (see Section 4.3.1 for details). The only parameter to be set is the threshold for the p-value for which a SEM is considered significant. If chosen too low, much of the data may be missed, if chosen too high, we run the risk of including data for which that particular SEM is not truly

relevant. Figure 4.4 illustrates this aspect: the areas shaded in green denote the regions in the synthetic data in which a particular atom is truly active. The areas shaded in red indicate the regions in which the atom is considered relevant at $\alpha = 0.99$ levels (Figure 4.4**A**). Setting the threshold so high virtually eliminates the false positive rate. Lowering the threshold even slightly has almost no effect on the true detection rate but increases false positives. Given the sensitivity of PCA to outliers, the high threshold of 0.99 was chosen.

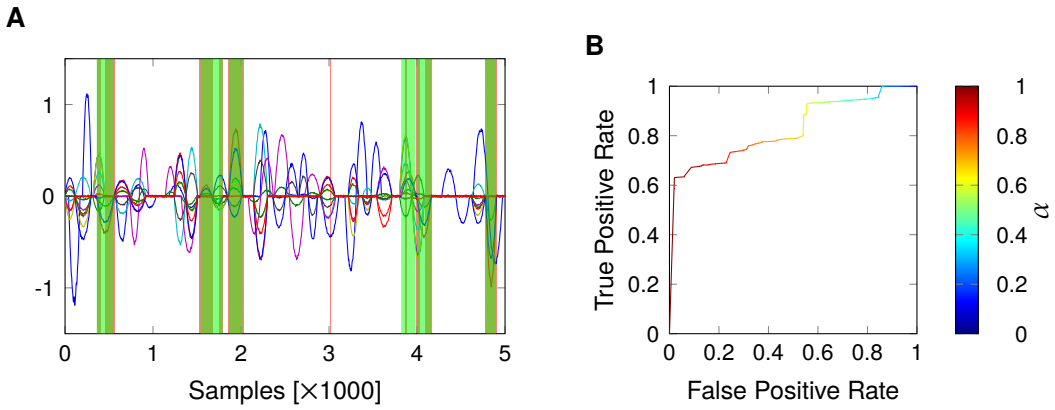

Figure 4.4: Empirical atom relevance determination. We compare the regions in which an atom is determined to be relevant (red) with those in which it actually generated the data (green). **A** Setting $\alpha = 0.99$ produces almost no false positives while still correctly identifying a large number of relevant data points. **B** Receiver operating characteristic. Only very high values for α produce low false positive rates.

When selecting the right threshold for use in real data, one should consider that the atom extracted in the first PCA step is likely to be noisy, i.e. the loadings may not be exact. This will cause the reconstruction by this extracted atom to be qualitatively less good than if the "true" atom was used, making it more similar to the random baseline. Thus, choosing a high threshold may cause the majority of the data to be discarded. To avoid this, it may be tempting to reduce the significance threshold but this has the averse effect of including irrelevant data. Nevertheless, we keep the value of α at 0.99 throughout our analysis since it minimises the risk of including false positives.

Eigenmotion Re-Estimation and Validation

We consider an atom to be relevant to the dataset if it is significant in at least T_w points outside the original ROI. In this case, all the data in which the atom was found to be significant are pooled and a single atom is extracted using the normal PCA algorithm. Finally, we check whether this atom adds anything to our knowledge of the data. This criterion is dependent on the stage of the algorithm: in the *reconstruction stage*, we check whether the reconstruction with the additional atom is significantly better than without. In the interference reduction stage, we only accept an additional element in the dictionary if it reduces the interference between atoms within the data.

4.4.4 Data Reconstruction

Given a dataset $\mathbf{X}$ and an over-complete and/or non-orthogonal dictionary $\mathbf{B}$, determining the optimal selection of atoms to represent a given data point is a non-trivial task, and a number of methods have been developed over the years. The possibly simplest solution to the problem, termed Matching Pursuit (MP Mallat and Zhang 1993), projects each $\mathbf{x}_i$ onto $\mathbf{B}$ and selects greedily the best matching atom. This process is then repeated with the residuals until a pre-determined number of atoms has been selected or the residuals are smaller than a given threshold. Orthogonal Matching Pursuit (OMP Pati, Rezaiifar, and Krishnaprasad 1993) takes the same approach as MP but restricts the selection of additional bases to those which have not yet been chosen. These methods are, however, very sensitive to noise in the data and yield unstable representations which may change significantly depending on the signal-to-noise ratio. Moving away from these greedy methods, Basis Pursuit (BP) (Chen, Donoho, and Saunders 1998) and the Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani 1996) both consider the task of finding the best matching bases as a classical optimisation problem to be solved by any valid optimisation algorithm, where the cost

function to be minimised is:

$$C = \|\mathbf{X} - \mathbf{BZ}\|_2^2 + \lambda \sum_i \|z_i\|_1 \tag{4.14}$$

which is the same formulation as Equation 4.2 for a known basis **B**. Finally, the Least Angle Regression (LARS) algorithm by Efron et al. (2004) provides a highly efficient method for combining all of the above[7] and giving an intuitive geometrical solution to the problem. For the SEMD algorithm, we reconstruct the data using the LASSO-LARS algorithm implemented in the SPAMS Toolbox by Julian Mairal (Mairal et al. 2010).

4.5 Algorithm Validation

4.5.1 Validation Data

We validate our method on data of various length and dimensionality, as well as by varying model parameters. To this end, we generate synthetic data sets with known dictionary and latent structure (see Section 4.3 for details on data generation) on which we tested the algorithm by varying number of data points (T), dimensionality of the data (D) and latent dimensionality (Q) as detailed in Table 4.1. For all cases, the window size T_w was set to 51 samples. In particular, we are interested in measuring

Table 4.1: List of parameters for the data with which the SEMD algorithm was validated. In total, 300 trials were run. Note that the value for Q is only indicative and may slightly fluctuate given to flexibility in the algorithm generating the data.

	T	D	Q	Runs
1	1000	16	3	50
2	4000	16	10	100
3	15000	16	32	50
4	4000	10	10	50
5	4000	54	10	50

[7] By changing some parameters of LARS, all other methods can be reproduced.

the robustness of our method, the impact of the interference reduction stage on the algorithm's ability to extract the eigenmotions from the data and its ability to determine the underlying dimensionality of the dataset. To do so, we evaluate the performance of our method on the basis of the following metrics:

1. The fraction of the number of extracted atoms, N_E, over the dimensionality of the latent representation, Q;
2. The fraction of extracted atoms which match the input atoms, N_M, over Q, where similarity between input atoms and extracted atoms is measured with the method described in Section 4.3.3;
3. The GOF of our model to the data; and
4. The amount of interference between the SEMs normalised by the number of data points.

4.5.2 Example Output

An example of the algorithm's application is given in Figure 4.5. The top panel indicates the set of 10 16-dimensional atoms which were used to generate the data in the middle panel. The bottom panel indicates the algorithm's output. Note that, except for some noise, the extracted atoms almost perfectly match the input. The shaded areas in the data indicate the windows from which the atoms can be extracted. The small amount of data required makes the algorithm extremely data efficient.

4.5.3 Overall Performance

The performance across all trials, as well as the comparison between the reconstruction stage of the algorithm (orange) and its overall performance (green), are summarised in Figure 4.6. If we first focus on the reconstruction stage, we note that it very accurately determines the dimensionality of the data (N_E/Q median(iqr): 1.02 (0.9–1.2)), and

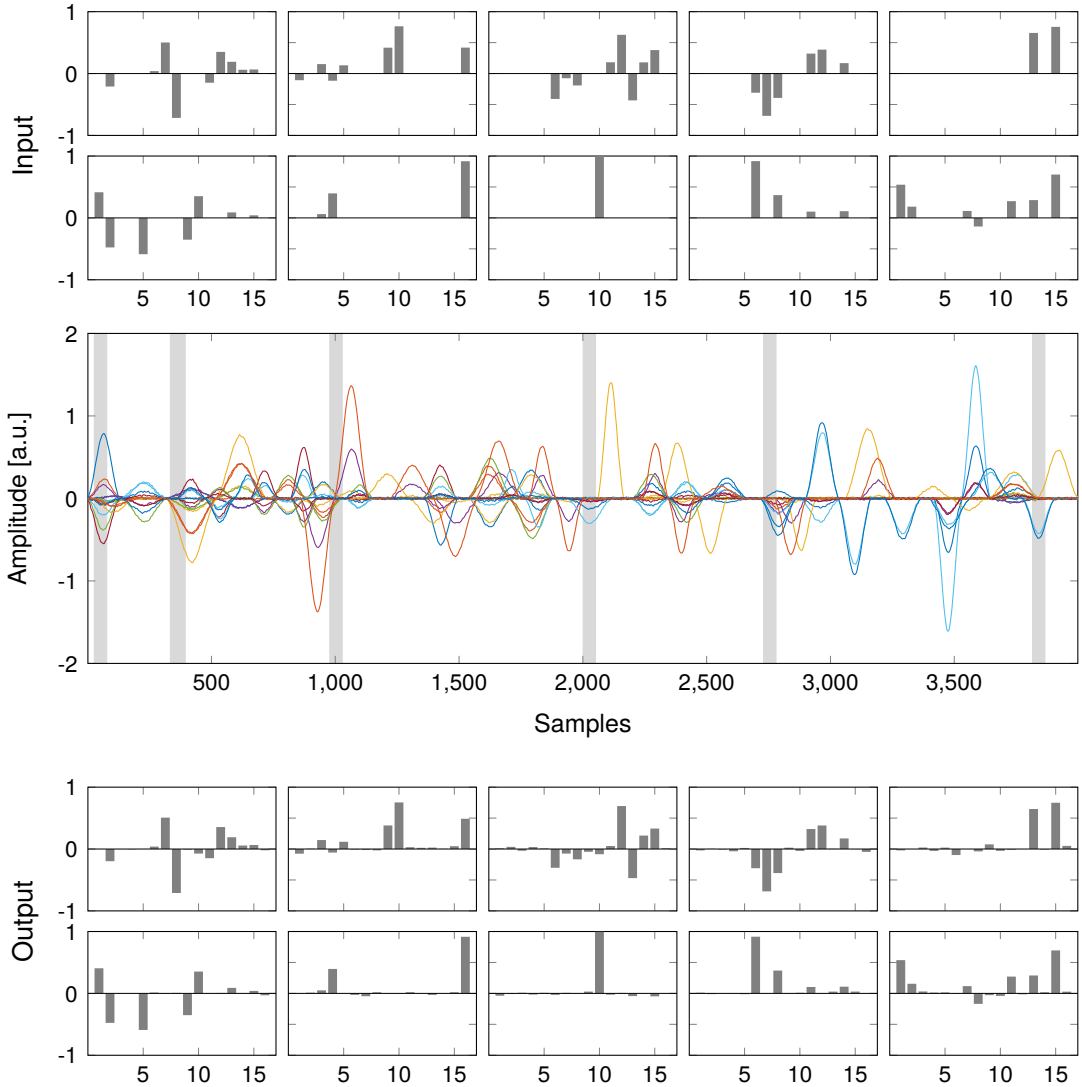

Figure 4.5: Example data and output. **Top** Input atoms which are used to generate the data (**Middle**). The algorithm identifies a dictionary of the correxct size, and with atoms which are almost identical to the input (**Bottom**). The shaded areas in middle plot indicate the windows from which the atoms were identified.

correctly identifies 77.6 (60–89.5)% of the input atoms. In contrast (Figure 4.6**A**), the complete algorithm compares unfavourably when estimating the latent dimensionality, significantly overestimating it 1.13 (1–1.33) (median(iqr), Wilcoxon signed rank test: $p < 0.001$), though it performs significantly better at identifying the correct input atoms 81.25 (70–90)% ($p < 0.001$). In terms of the model's GOF, the interference reduction step only slightly improves the overall result, but pulls up the models with the poorer fit: median(iqr) 0.73 (0.66–0.76) *vs.* 0.74 (0.69–0.77) (Figure 4.6**B**). It does however

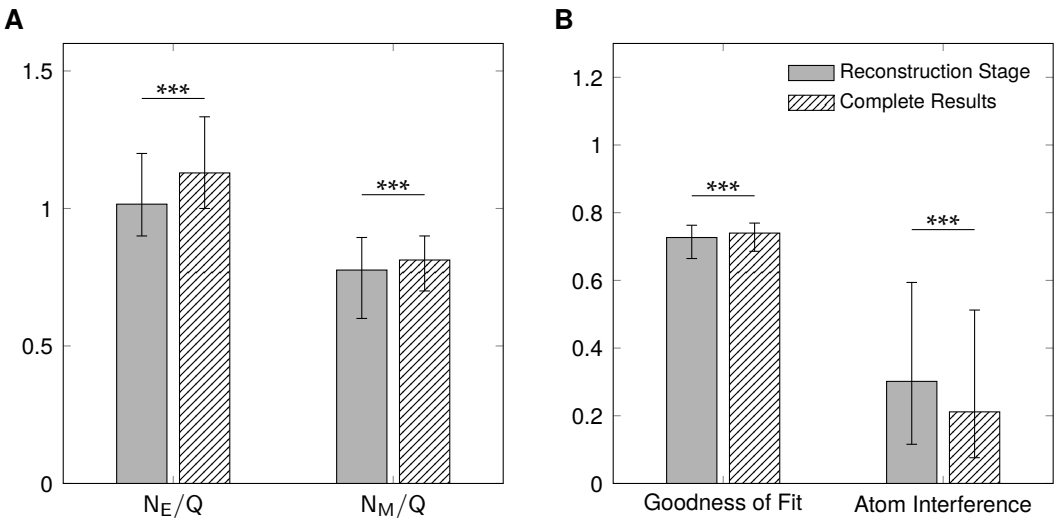

Figure 4.6: Performance of the SEMD algorithm and comparison between the reconstruction stage and the full method. **A** (Left) The *reconstruction stage* accurately identifies the dimensionality of the latent representation, with the ratio N_E/Q being very close to 1. The full method, however, significantly overestimates it. (Right) In contrast, the full algorithm performs significantly better when it comes to identifying the correct input atoms, achieving >80% matches on average. **B** (Left) The *dimensionality reduction* stage of the EMD algorithm slightly, though significantly improves the model's goodness of fit, and (Right) significantly reduces the interference between atoms within the data. Whiskers indicate IQR. Wilcoxon signed rank test; ***: $p < 0.001$.

achieve its main purpose by greatly reducing the interference between atoms, despite a substantial increase in dictionary size: 0.30 (0.12–0.59) *vs.* 0.21 (0.08–0.51) ($p < 0.001$). In the following sections we detail SEMD performance when confronted with data sets

of different lengths and dimensions.

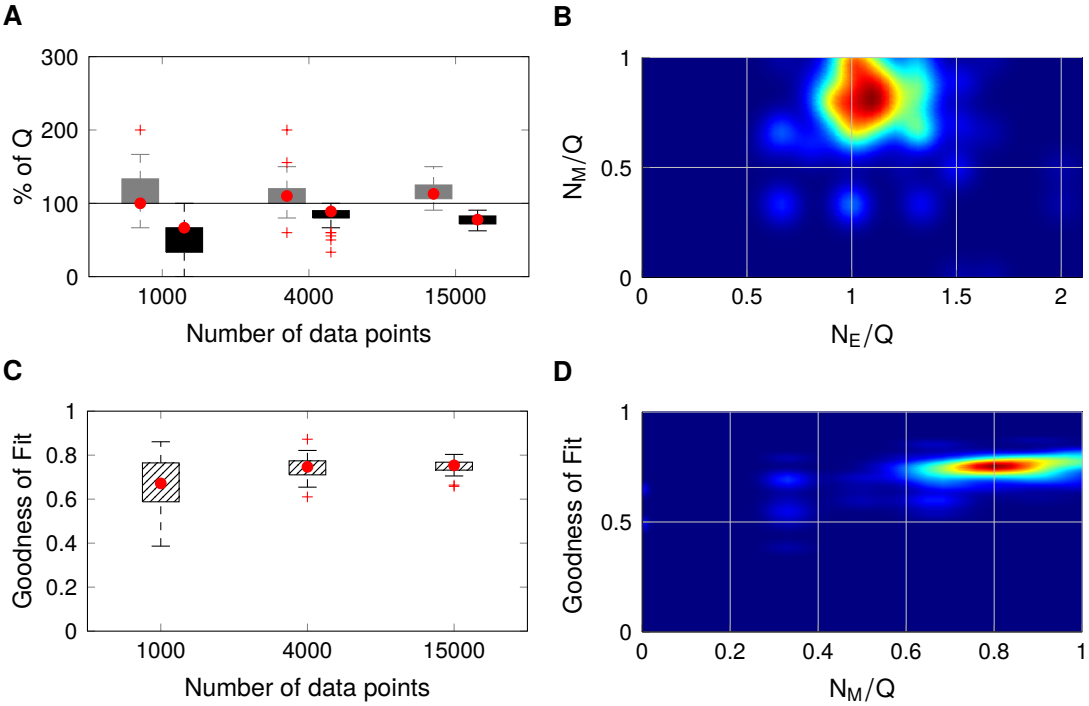

Figure 4.7: Influence of number of data points on reconstruction stage performance. **A** Distribution of the number of extracted (grey) and matching SEMs (black). Small data sets or data with high latent dimensionality slightly degrade performance. **B** Joint distribution of the fraction of extracted atoms and of the fraction of matching atoms. The extraction of "correct" atoms is not influenced by the total number of atoms extracted. Ideal performance would be at (1,1). **C** The goodness of fit improves with the amount of available data and **D** is related to the algorithm's ability to identify the correct atoms in the data. See text for details.

4.5.4 Impact of Data Set Size

The algorithm's performance when confronted with data sets of varying length and latent dimensionality is shown in Figure 4.7 and Figure 4.8 for the reconstruction stage and the complete algorithm respectively. As for the overall results (see Section 4.5.3), we note that the reconstruction stage by itself provides a good estimate of Q (Figure 4.7**A**, grey boxes), although slightly more upwards biased (median: 110%, iqr: 100%–127%),

and is particularly variable for small data sets with low values of Q. This variability is reflected in the numbers of input atoms identified (black boxes). We further observe that the algorithm's ability to correctly identify the input atoms appears to be highly dependent on both data size and latent dimensionality, as it identifies significantly less input atoms for both very small and very large data sets, which implies a high Q. The difference between small and large data sets is not significant ($T = 1000$ *vs.* $T = 4000$, median: 0.66 *vs.* 0.89, $p < 0.001$; $T = 4000$ *vs.* $T = 15000$, median: 0.89 *vs.* 0.78, $p < 0.01$; Kruskal-Wallis one-way ANOVA, Bonferroni corrected). Interestingly, we find no correlation between N_E and N_M (Figure 4.7**B**[8]).

In terms of goodness of fit, the results are globally in line with the overall statistics (Figure 4.7**C**). However, the algorithm has trouble with small data sets, significantly under-performing situations where more data is available ($T = 1000$ *vs.* $T = 4000$ *vs.* $T = 15000$, median: 0.67 *vs.* 0.75 *vs.* 0.75, $p < 0.001$ for $T = 1000$ *vs.* rest, not significant between $T = 4000$ and $T = 15000$; Kruskal-Wallis one-way ANOVA, Bonferroni corrected). Maybe unsurprisingly, we find a correlation between the GOF and N_M, although the values remain generally high (Figure 4.7**D**, Spearman's $\rho = 0.46$, $p < 0.001$).

Considering the algorithm's performance after adding the interference reduction stage (Figure 4.8), we first note that both the goodness of fit and the ratio N_M/Q improve significantly across all conditions ($p < 0.001$, Wilcoxon signed rank test with Bonferroni correction), thus demonstrating the usefulness of the second stage. In particular, the interference reduction stage has the benefit of levelling the field between data sets of various lengths. Thus, the ratio N_M/Q becomes the same for both $T = 1000$ and $T = 4000$ time steps, although it still remains a bit lower for the long data set with high value of Q (Figure 4.8**A**, black boxes).

[8]This figure, as well as Figure 4.7**D** and similar figures in this section were obtained by estimating the distributions using a kernel density approximation with a Gaussian kernel with a bandwidth of 0.1.

This homogenisation across conditions is further exemplified in Figure 4.8**B**, where the distribution $P(N_E/Q, N_M/Q)$ is much better defined, and in Figure 4.8**C** for the model's goodness of fit where there is no more significant difference between conditions (Kruskal-Wallis one way ANOVA, Bonferroni corrected). The goodness of fit is also distributed more narrowly, though the dependency on N_M persists, albeit to a lesser degree (Figure 4.8**D**; Spearman's $\rho = 0.31$, $p < 0.001$).

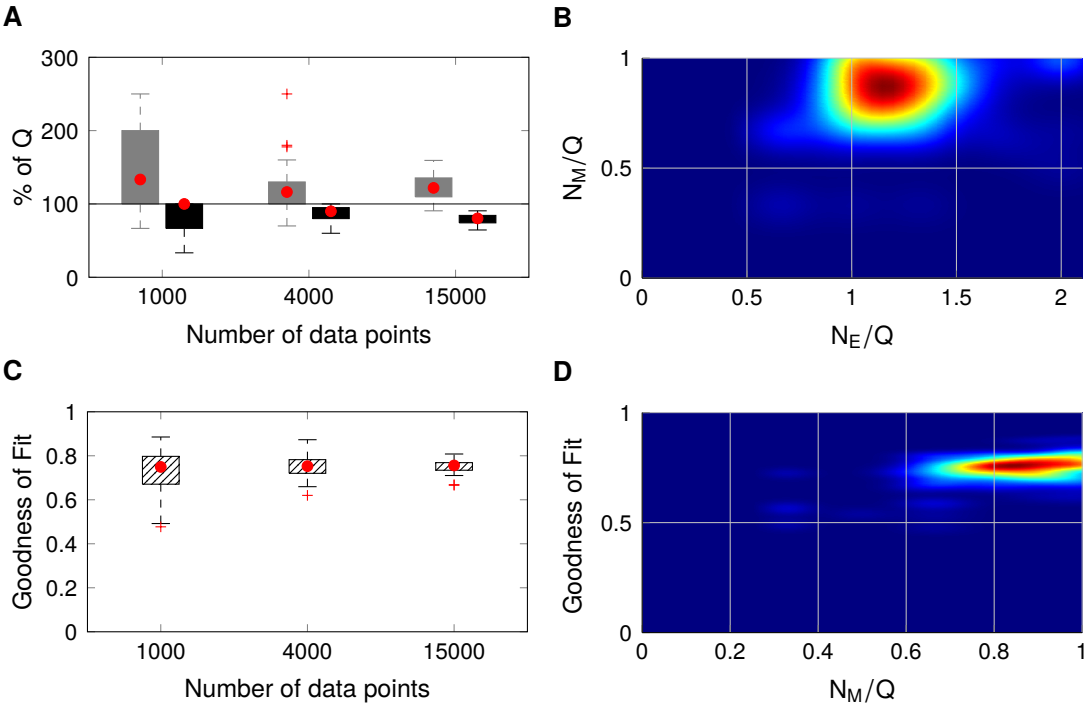

Figure 4.8: Influence of number of data points on the performance of the complete SEMD algorithm. **A**–**D** The second stage significantly improves both N_M/Q and the goodness of fit with respect to the reconstruction stage alone, as well as levelling the results across various conditions. **A**, grey boxes: The downside is that the estimate for Q is further biased, in particular for very short time-series.

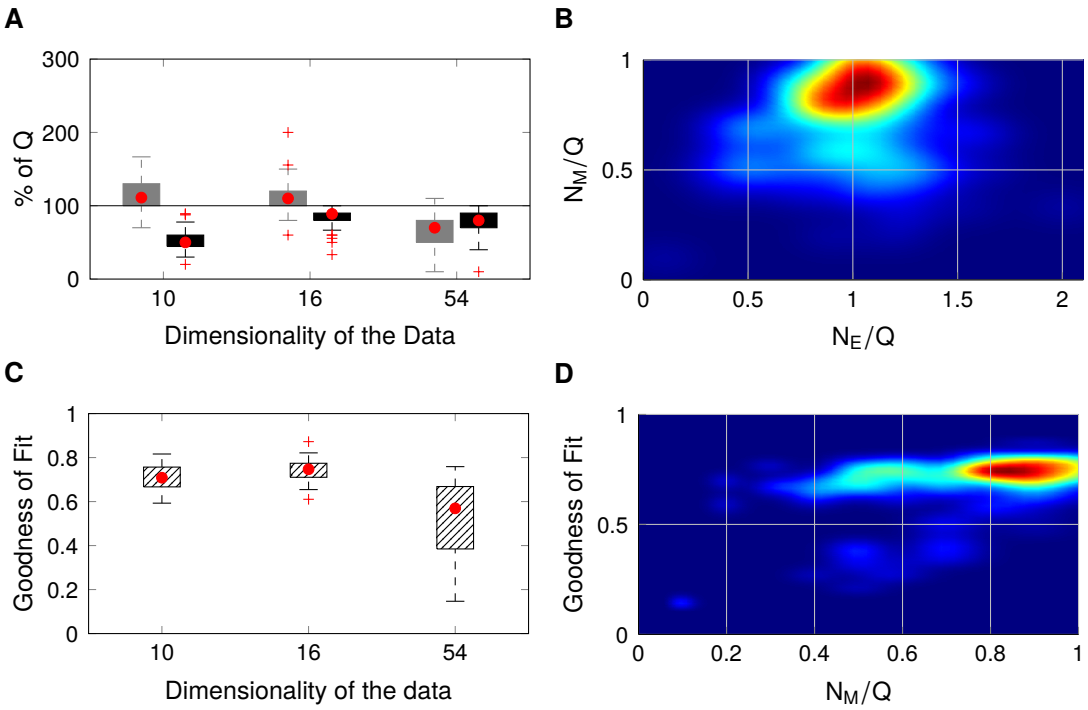

Figure 4.9: Influence of dimensionality the performance of the reconstruction stage. **A** Distribution of the number of extracted (grey) and matching SEMs (black). **B** Joint distribution of the fraction of extracted atoms and of the fraction of matching atoms. **C** Goodness of fit of the learnt model. Contrast the poor performance in extracting the correct atoms for 10-dimensional data in **A** with the comparatively high GOF. This is further detailed in **D** which displays a variable dependency of GOF on the ratio N_M/Q. See text for details.

4.5.5 Impact of Data Dimensionality

Next, we illustrate the algorithm's ability to deal with data of varying dimensionality, as well as some of its limitations. As with the previous analysis, we first present the output of the reconstruction stage (Figure 4.9), followed by the results for the complete algorithm (Figure 4.10). Again, we note a slight bias in the estimate of Q, though it appears that the value is significantly dependent on the overall dimensionality of that data (Figure 4.9**A**): upwardly biased for low dimensional (10 & 16) and downwardly for high dimensional (54) data. The number of input atoms identified also varies with

the dimensionality of the data: the results for $D = 10$ is significantly lower than for $D = 16$ or $D = 54$, which do not differ from each other ($p < 0.001$, Kruskal-Wallis one way ANOVA, Bonferroni corrected). This result may seem surprising as it should be a lot easier to identify a low-dimensional atom than a high-dimensional one as there are less numbers to be estimated. This paradoxical finding is an artefact caused by the method used for computing the similarity, and is discussed further below as well as in Section 4.3.3. As previously, we find no correlation between N_E and N_M although the distribution is quite a bit wider, mainly due to the aforementioned artefact (Figure 4.9**B**).

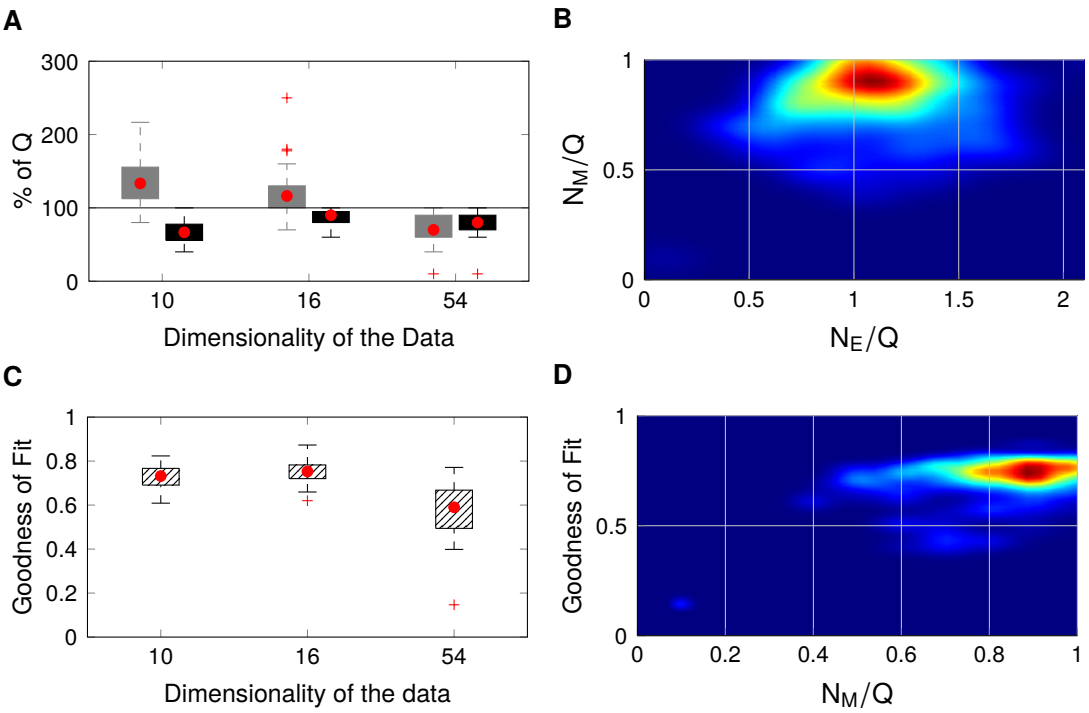

Figure 4.10: Influence of dimensionality on the performance of the complete SEMD algorithm. **A**–**D** The second stage significantly improves both N_M/Q and the goodness of fit with respect to the reconstruction stage alone. The estimate for Q worsens for 10 and 16 dimensional data but improves for 54. In general, the algorithm performs best with 16 dimensional data and worst with the high-dimensional one.

Akin to the results for N_M, the goodness of fit is significantly worse for $D = 10$ than

for $D = 16$ (median: 0.71 *vs.* 0.75). However, it is even lower for $D = 54$ (median: 0.57) (Figure 4.9**C**). This result can be explained by considering that the dependency of the GOF on N_M is much higher for the 54-dimensional data than the 10-dimensional data (Figure 4.9**D**, Spearman's $\rho = 0.83$ *vs.* 0.46). This is not surprising: because of the curse of dimensionality, reconstructing data with 'wrong' bases will lead to larger errors in high-dimensional space than in low-dimensional space, thus pulling down the goodness of fit.

The second stage of the algorithm further improves both N_M/Q as well as the goodness of fit significantly ($p < 0.001$, Wilcoxon signed rank test, Bonferroni correction), as found previously. Its ability to equalise results across conditions is, however, less clear for varying dimensions: the number of correctly identified atoms in the 10-D data is still significantly below the values for the other cases (Figure 4.10**A**; $p < 0.001$ Kruskal-Wallis one way ANOVA, Bonferroni correction), although the spread of the distribution is again reduced (Figure 4.10**B**). Thanks to the second stage, the goodness of fit for $D = 10$ becomes similar to the one for $D = 16$. In the case of the high-dimensional data, the gof improves, although it remains significantly worse (Figure 4.10**C**, $p < 0.001$, Kruskal-Wallis one way ANOVA, Bonferroni correction). The relationship between the fraction N_M/Q and the gof remains present, but its impact on high-dimensional data is reduced (Figure 4.10**D**; Spearman's $\rho = 0.67$, $p < 0.001$).

4.6 Comparison with the State of the Art

To compare our algorithm's performance with competitor algorithms, we analysed the same dataset which we validated our algorithm with using two alternative dictionary learning algorithms: K–SVD (Aharon, Elad, and Bruckstein 2006), a widely used algorithm and standard in numerous image processing applications and the method proposed by Mairal et al. (2010), which implements recent developments in the field

and is claimed to work online. We compare the results in terms of the same metrics we used for the validation process. A summary of the results is provided in Figure 4.11. Note that both these methods require an initial dictionary or at least with a guess as to the size of the dictionary. This was set to be the true Q as to avoid bias.

We compared the output dictionaries with the input dictionaries to determine the number of input atoms correctly identified by the algorithm (Figure 4.11**A**). The SEMD algorithm shows clear superiority to both other methods, correctly identifying 81.25(70–90)% (median(iqr)) of the input atoms. This is significantly higher than the performance of Mairal et al. (2010) (70(60–80)%, $p < 0.001$, Wilcoxon signed rank test) and the K–SVD algorithm (54.09(40–70)%, $p < 0.001$). Similarly, the interference between atoms is significantly lower for our algorithm than for the other cases (Figure 4.11**B**). None of the methods, however, achieve the ideal value of 0. This is not particularly surprising considering that to obtain such a low value would require extracting the input atoms with absolute precision, a task made very difficult by the presence of noise in the data. The goodness of fit does not vary very much between the methods (median; SEMD: 0.740; Mairal (2010): 0.737; K–SVD: 0.717), though SEMD outperforms the other algorithms significantly. These values should be contrasted with the goodness of fit obtained using the exact input atoms (Figure 4.11**C**, dashed and dotted lines (median and iqr respectively)). Even using the exact input atoms, the GOF remains well below its ideal value of 1. This can be explained by three factors: (1) noise in the data, (2) the linearity of the model which makes it unable to deal with time-shifts in the carrier functions (Δt in Algorithm 2) and more importantly (3) the sparsity constraints of the reconstruction algorithm. Depending on the value used for λ (see Equation 4.2), the reconstruction will sacrifice some of the goodness of fit for sparsity, thus potentially suppressing low-amplitude signals.

The major drawback of our method is its computational cost which increases linearly with data size (Figure 4.11**D**). The main contribution to this cost is calculating the

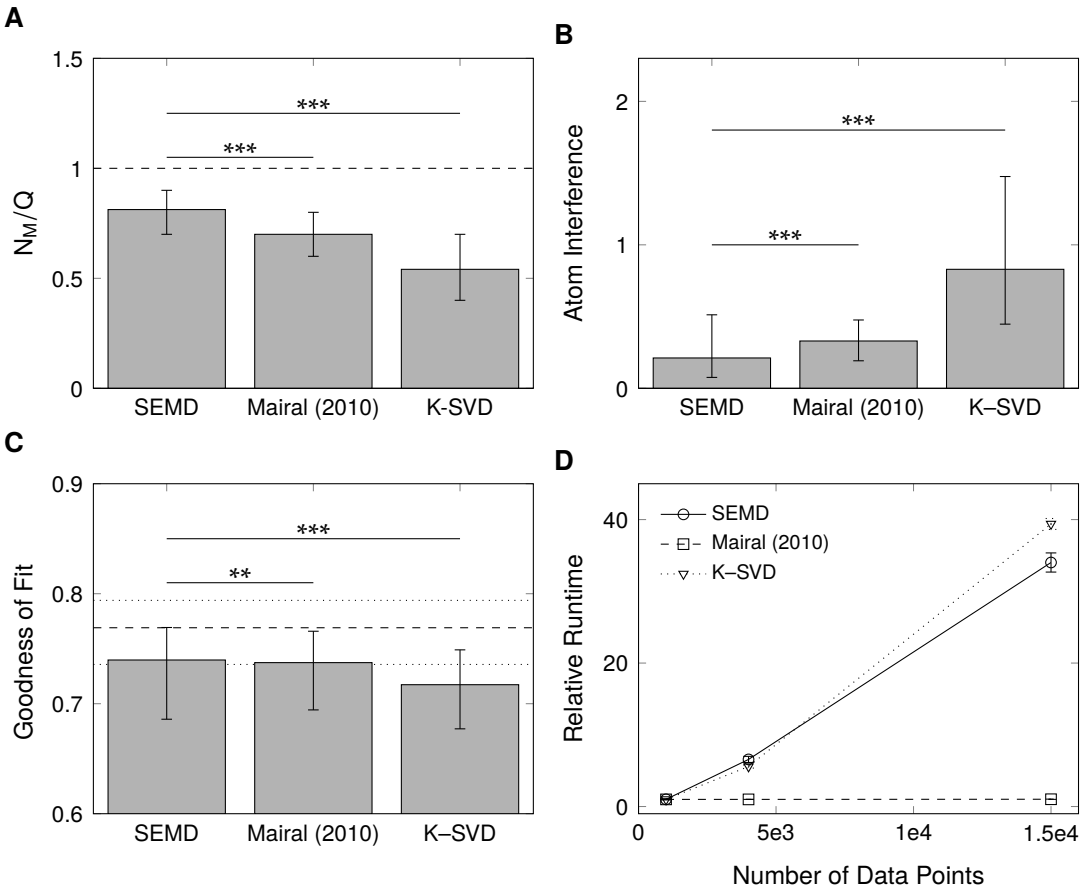

Figure 4.11: Comparison of the SEMD algorithm with the state of the art. **A** Fraction of correctly identified input atoms. SEMD significantly outperforms its competitors by correctly identifying more than 80%. **B** Similarly, the interference between atoms is significantly lower for the results obtained by the SEMD algorithm that the competitors. **C** The goodness of fit for all models is sensibly the same, but SEMD slightly outperforms the other algorithms. **D** In terms of run-speed, the online algorithm by Mairal et al. (2010) effectively remains approximately constant with additional data. SEMD slightly outperforms K–SVD for longer data sets. All values are medians and inter-quartile range. Dashed and dotted lines indicate ideal performance, except for **B** where it is indistinguishable from the x–axis. See text for details. **: $p < 0.01$; ***: $p < 0.001$, Wilcoxon signed rank test.

distributions for estimating the significance of atoms within the data (see Section 4.4.3), as well as calculating the GOF at each iteration. While it is unlikely that the overall complexity of the algorithm can be reduced, the implementation can certainly be sped up significantly, though this is not the aim of this work.

4.7 Discussion

We presented a new algorithm for dictionary learning from time-series of arbitrary dimensionality. Our method can be understood as both an extension and an improvement on the technique used by Olshausen and Field (1997) and refined by numerous others (e.g. Aharon, Elad, and Bruckstein 2006; Mairal et al. 2010). Three aspects set our approach apart from previous work: (1) data efficiency, (2) consideration of spatiotemporal instead of only spatial structure and (3) automatic estimate of the latent dimensionality.

The first aspect is achieved by a highly selective sampling approach which attempts to maximise the information gathered by each iteration. This selective, windowed sampling approach – combined with the morphological filtering operations – also implies that despite using PCA (a method for spatial structure) locally, we effectively recover bases which have a spatiotemporal function in the data. This is in stark contrast with the previously mentioned global methods which are invariant to random shuffling of the input data. Our algorithm's ability to automatic estimate the latent dimensionality of the data is also a consequence of the selective sampling method: when no more novel information is available, or the data is accurately represented by a sufficiently sparse code, it can be assumed that the number of atoms obtained are equal to the latent dimensionality. Using a window, of course assumes that the frequency spectrum of the data is not white, i.e. that there is some temporal structure. It also adds an additional parameter to be set. Fortunately, the time-scale of the dynamics governing the

data being analysed can be obtained relatively easily by either considering theoretical aspects, or by studying the auto-correlation function (e.g. Figure 3.5**A**&**B**).

We validated the algorithm's behaviour for data of varying size and dimensionality, demonstrating its ability to recover both under-complete and over-complete bases from the data. In contrast to our method, the only other algorithm capable of determining the (over-complete) dimensionality of the data, Generalised Principal Component Analysis (Vidal, Ma, and Sastry 2005), has the disadvantage of being computationally extremely costly and thus unusable for high-dimensional data.

A disadvantage of our approach is that unlike the method described by Mairal et al. (2010), it is not possible to parallelise. However, it is conceivable to chunk a very large dataset into smaller blocks and process these in parallel, although there is a real risk of increasing the number of dictionary elements thus identified. This adds a further complication as it requires a method for merging dictionary elements which are very similar. This is the approach taken by (Mukherjee and Seelamantula 2014) which learns a sparse representation of the union of individual dictionaries. However, a simple clustering approach may provide equally good results.

5

Sparse Representation of Natural Movement

> "Nature uses only the longest threads to weave her patterns, so each small piece of her fabric reveals the organization of the entire tapestry."
>
> — Richard Feynman
>
> *The Character of Physical Law*

5.1 Introduction

When we formulated our alternative hypothesis to hand control in Chapter 3, we drew inspiration from a previous study which showed that sparse encoding of natural images reflects the receptive fields of neurons in the primary visual cortex (Olshausen and Field 1997). Setting aside the argument that "it worked elsewhere", why is a sparse code a good idea? And what evidence is there that such a concept is being implemented in the brain? The first question is typically answered using three distinct arguments: (1) energy consumption, (2) stimulus discrimination and (3) memory capacity and retrieval.

Because the generation of action potentials in the brain consumes such a large amount of energy (ca. 80%; Attwell and Laughlin 2001), metabolic constraints place an upper limit on the number of neurons which may be active at any given time. This imbalance between energy requirements to maintain a neuron and to transmit a signal imply that

an efficient neural code should try to maximise the information transmitted per spike. In the most extreme case, this would result in "grandmother" cells – neurons which represent one and only one concept (Konorski 1967). Thus, if that neuron did spike, one would be certain that a particular stimulus has been presented (e.g. the face of the grandmother). Without going to such lengths, a sparse code which represents high-level features has been shown to be a promising candidate for encoding information under energy constraints (Levy and Baxter 1996; Laughlin 2001).

By encoding these higher-order statistics of the stimulus, sparse codes make it also much easier to discriminate between different inputs (Barlow 1972; Rolls and Treves 1990). This is particularly relevant when the input statistics are not completely random but contain a lot of structure, such as natural images (Olshausen and Field 1997) and sounds (Lewicki 2002).

Finally, sparse coding has been demonstrated theoretically to increase the memory capacity and information retrieval ability of associative memory networks (Rolls and Treves 1990; Okada 1996; Palm 2013). This is a memory model in which memories are stored as the interaction of multiple neurons.

Evidence for sparse coding in the nervous system is abundant in the literature. However, "sparsity" is often used interchangeably for two different concepts: *Lifetime sparsity*, which implies that a neuron will only rarely fire throughout its life, i.e. codes for a specific feature, and *population sparsity* which suggests that only a small part of the population is active at any given time. Given data, separating both concepts is not a trivial task as it requires monitoring a large population of neurons over prolonged periods of time. However, it can be safely said that lifetime sparsity implies population sparsity, although the opposite is not necessarily true.

In terms of sparse representation of high-level features, the most prominent example is found in the visual system of monkeys (Baddeley et al. 1997; Vinje and Gallant 2000; Vinje and Gallant 2002) and humans (Quiroga et al. 2005; Quiroga et al. 2008).

Building on the results of Olshausen and Field (1997), Vinje and Gallant (2000) and Vinje and Gallant (2002) demonstrate sparse firing of neurons in V1 of awake and behaving monkeys, and link it to interactions between the cells' classical and non-classical receptive field. Baddeley et al. (1997) follow the visual processing stream up to the inferotemporal (IT) cortex and show an increase in neural firing sparsity, although this was somewhat dependent on the metric. In the human medial temporal lobe (MTL) which is one processing stage higher than IT, cells were evidenced which were responsive to the representation of a single person, independently of the perspective (Quiroga et al. 2005). Although reminiscent of the grandmother cells discussed before, the authors themselves note that considering the number of stimuli used in the study and the total number of neurons in MTL, it is more likely that these cells respond to many more stimuli, thus making them a sparse code (Quiroga et al. 2008).

Similar evidence is available for the olfactory system of insects: in locusts, after a population code type encoding of smells in the first processing stage, these are projected to a secondary processing stage, termed mushroom body, whose neurons show extremely high selectivity for individual smells (Perez-Orive et al. 2002). Interestingly, each neuron in the second processing stage is innervated by 50% of the primary neurons. This lead the authors to conclude that albeit the neurons investigated only responded to a single smell, the cells are likely to respond to a much greater number of stimuli (Jortner, Farivar, and Laurent 2007). Concurrent results were demonstrated in the mushroom body of drosophila (Honegger, Campbell, and Turner 2011) and even in mice (Rinberg, Koulakov, and Gelperin 2006; Poo and Isaacson 2009).

Evidence for sparse coding in other sensory modalities is somewhat more spurious, but present. Recordings in the rat's auditory cortex, revealed that at most 5% of the neural population respond to any given sound (Hromádka, DeWeese, Zador, et al. 2008), and that the neuron's tuning curves for sound frequency are very narrow (DeWeese, Wehr, and Zador 2003). Furthermore, Lewicki (2002) showed that sparse coding of

speech produces filters not dissimilar to those of the cochlea. In the somatosensory cortex, whisker movement and contact with objects result in only a small part of the neural population firing (Brecht and Sakmann 2002; De Kock et al. 2007; O'Connor et al. 2010; Crochet et al. 2011). The neurons' response was not correlated with whisker position or velocity, suggesting that higher-order features may be encoded (Kock and Sakmann 2009). The most compelling evidence for sparse coding, however, originates from a study by Houweling and Brecht (2007) which demonstrated that rats were capable of responding to stimulation of a single neuron in the sensory cortex. Further evidence for sparse firing (and sometimes coding) in the brain can be found in place cells in the hippocampus (Thompson and Best 1989), the prefrontal cortex of monkeys (Abeles, Vaadia, and Bergman 1990), their gustatory system (Rolls, Yaxley, and Sienkiewicz 1990) and the electrosensory system of weakly electric fish (Vonderschen and Chacron 2011).

What about motor cortex? At least two studies point conclusively towards a sparse representation of complex motor patterns by a small group of cells: in rats, stimulation of a single pyramidal neuron in the vibrissae motor cortex is capable of eliciting complex whisker movements which far outlast the duration of the stimulation (Brecht et al. 2004). The movement amplitudes generated in this fashion were much smaller than those observed behaviourally, suggesting that a small number of similar neurons firing together may be the source of whisking in normal situations. In the zebra finch (a songbird), neurons in the high vocal centre – a premotor area – were shown to represent a temporal code of song generation. Spiking of one of these neurons generates a unique pattern in the motor area it projects to, and is precisely time-locked to the bird's song (Hahnloser, Kozhevnikov, and Fee 2002).

In light of this evidence, and given the limited success of other methods to describe motor output (see Section 3.4), we investigate the existence of a *lifetime* sparse representation of motor outputs during natural human behaviour. If a meaningful sparse

representation exists, we expect it to be task dependent but with relatively large overlap between tasks. To examine this question, we analyse two datasets: (1) in a controlled setting (Belić and Faisal 2015) and (2) from freely behaving subjects. We demonstrate that such task dependence exists, that dictionaries are likely to be finite in size and that higher-order structure exists in the sparse representation of the data.

5.2 Materials and Methods

5.2.1 Data

We learnt dictionaries of sparse eigenmotions (SEMs) on two different data sets: (1) a previously published database containing data of 7 subjects performing 10 repetitions of 16 different tasks from everyday life (Belić and Faisal 2015). This data was recorded using a Cyberglove (Cyberglove Systems LLC, San Diego, CA) from right-handed subjects performing the tasks with their left hand and contains approximately 1 million samples equivalent to 3.5 hours of data. (2) A collection of recordings in natural settings described in Chapter 2. This contains data from the right and left hand of 7 subjects recorded in two different settings (kitchen and bedroom) and totals more than 4.5 million data points or 13 hours of recording.

5.2.2 Dictionary Learning

Dictionaries of SEMs were learnt for each individual recording using the algorithm described in Chapter 4. The width of the window in which local principal components are extracted was informed by the auto-correlation function of natural finger movements (see Figure 3.5**B**) and set to 510 ms. Because of the large number of data points contained in our recordings of natural movement data, we separated the data into blocks of ca. 30000 samples, equivalent to 5 minutes of data. To obtain a unified dictionary (and simultaneously estimate the number of SEMs used over time), we performed an

iterative merger of the individual dictionaries as described below.

5.2.3 Dictionary Merger

We merge dictionaries using an iterative approach: given any two sets of SEMs $\mathbf{B}_1$ and $\mathbf{B}_2$, we pool the elements and merge those deemed similar by averaging their values. This procedure is done starting with very high thresholds for similarity and then repeated with gradually lower thresholds. This guarantees that only similar SEMs are actually merged and that "chains" of elements are not lumped together into one large cluster. For our purposes, the similarity measures are the same as used in Chapter 4 and detailed in Section 4.3.3. The threshold started at 0.999 and was decreased in steps of 0.001 until reaching 0.98.

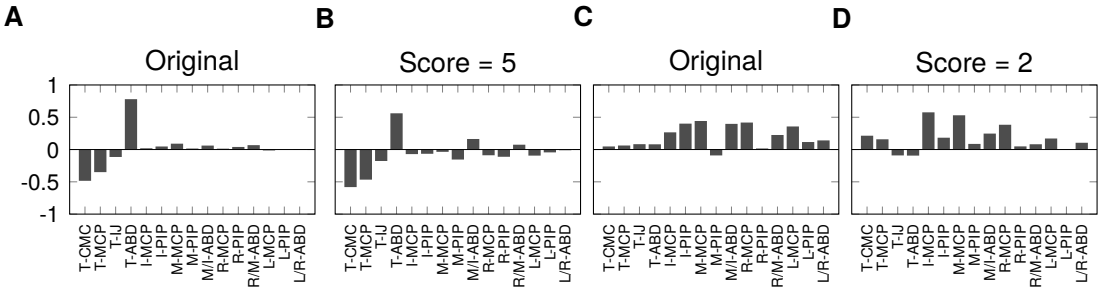

Figure 5.1: Comparison of sparse eigenmotions. **A** and **B** agree across in all five measures and are considered a match. **C** and **D** only agree on two measures (intersection and functional similarity) and are thus not considered to be equal.

5.2.4 Comparison with PCA

We use two different techniques to contrast the ability of SEMs and PCs to capture natural movement data: (1) Variance explained with a latent dimensionality of 1 and (2) the area under the VAF curve. In both cases, the latent SEM representation is determined using the OMP algorithm (Pati, Rezaiifar, and Krishnaprasad 1993) which enables us to precisely set the latent dimensionality. For fairness, the same method is

used for finding the latent representation of PCA. That is, we do not restrict ourselves to using a single PC throughout the data set but allow the algorithm to change PC at every time-step. In addition – but for the repetitive movement data only – we restrict size of the SEM dictionary and the number of PCs to be the same. When truncating a dictionary to size Q, we pick the Q most prominent (in terms of squared latent activity) elements. This guarantees that the number of parameters being compared in both situations is the same.

5.2.5 Manipulative Complexity

To determine the complexity of a given task, we use the metric of *manipulative complexity* as defined in Belić and Faisal (2015), which takes the mathematical form of:

$$C_m = 1 - \frac{2}{D-1} \sum_{j=1}^{D} \sum_{i=1}^{j} (VAF_i - 1/N) \tag{5.1}$$

where VAF_i is the variance captured by the i-th principal component. Values close to 0 indicate low complexity, i.e. one PC explains most of the variance in the data. A value of 1 would mean that all the PCs contribute equally to the variance in the data. Intuitively, this represents the area between between the diagonal and the VAF curve, and is equivalent to the Gini coefficient (Gini 1912).

5.2.6 Hidden Markov Model Training

We investigate the temporal structure of the sparse movement representation using a hidden Markov model (HMM). The data is compressed to a latent dimensionality of 1 by using the OMP algorithm and the 40 first elements of the SEM dictionary of subject 4. We chose this dictionary because it generalised best across all subjects, but most others probably could have been used. Data points during which the hand was static were identified using the previously described method (Section 2.2), and

removed. Non-zero values in the latent representation were replaced with the ID of the corresponding SEM. Finally, self transitions were removed, leaving only a sequence of different SEMs.

Using these sequences from all recordings, we learn a HMM using the Baum-Welch algorithm. We trained HMMs with increasing numbers of hidden states (from 1 to 25). For each number of hidden states, we trained the HMM 10 times using a random initialisation for the transition and emission matrices. The best HMM (i.e. the one with the highest log-likelihood) was chosen for further analysis. To compare the trained HMMs, we compute the Bayesian Information Criterion (BIC) for the HMM learnt. The BIC is a principled way of comparing the quality of models when they have different numbers of parameters. Lower BIC values indicate a better model.

5.2.7 Prediction of Hand Movements

We attempt to predict movements of the hand ($\mathbf{y}(t)$) from the position and motion of other body parts ($\mathbf{x}(t)$), similar to the methodology used in Chapter 3. Instead of directly predicting hand movements as done previously, we first find a one dimensional, latent representation of the data using the OMP algorithm and a single dictionary (see Section 5.2.6 for details). For each SEM, we take the data assigned to it and learn a linear regression model from $\mathbf{x}$ to $\mathbf{y}$ with 10-fold cross-validation. We then take the entire data and learn a classification tree to predict the current SEM from $\mathbf{y}$, also with 10-fold cross validation. This results in a two-stage prediction method: (1) predict SEM from $\mathbf{y}$ and (2) use a SEM specific regression model to predict hand movements.

5.3 Results

We analyse two data sets of hand movements using the dictionary learning algorithm presented in Chapter 4. The first data set contains multiple repetitions of the same task

by various subjects (Belić and Faisal 2015). It is used as validation for the algorithm's performance when confronted with real data. The second data set contains much longer time-series obtained from freely behaving subjects. On this data, we investigate the potential of sparse coding for natural movement statistics, consistency of sparse eigenmotions (SEMs) between subjects and higher-order structure in the time-series.

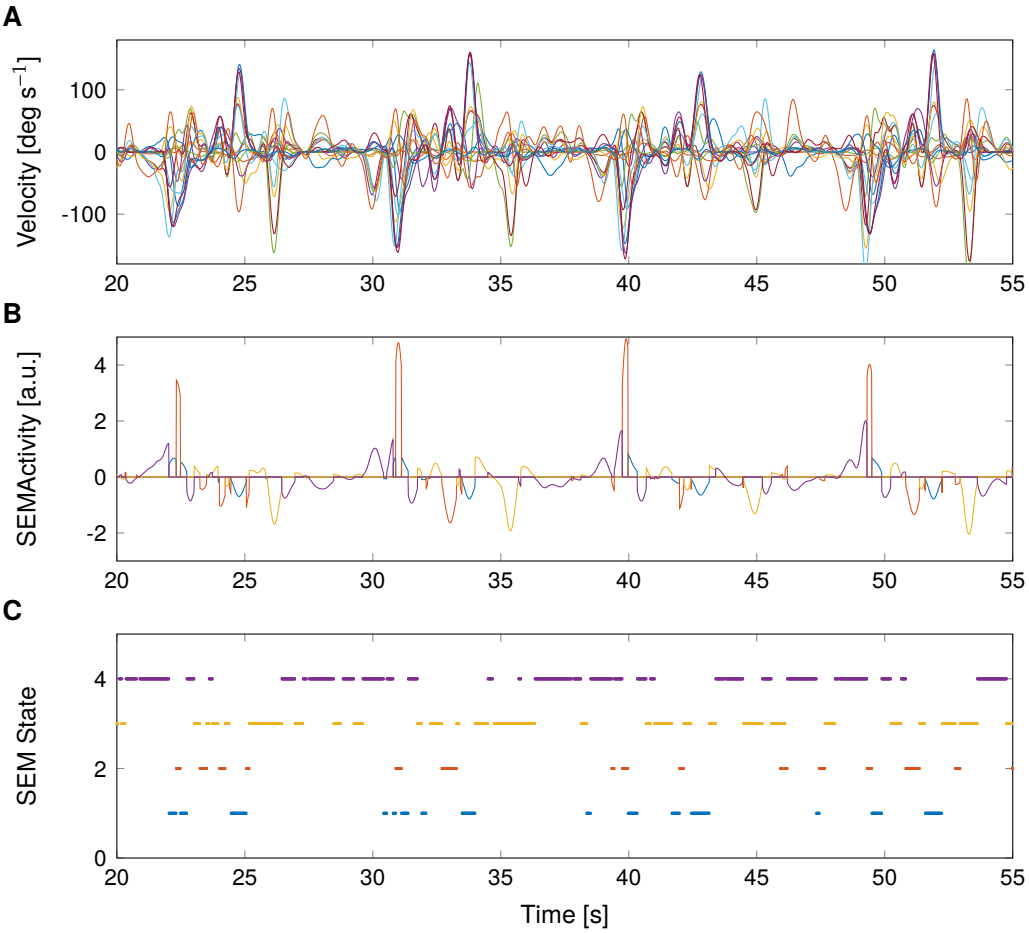

Figure 5.2: Example of encoding data using a SEM dictionary. **A** Raw data captured from the Cyberglove. Each colour represents a different channel. **B** Latent representation of the data. This example uses OMP with a single non-zero latent dimension as reconstruction method. **C** Binary encoding of **B** where the latent dimension is either on or off.

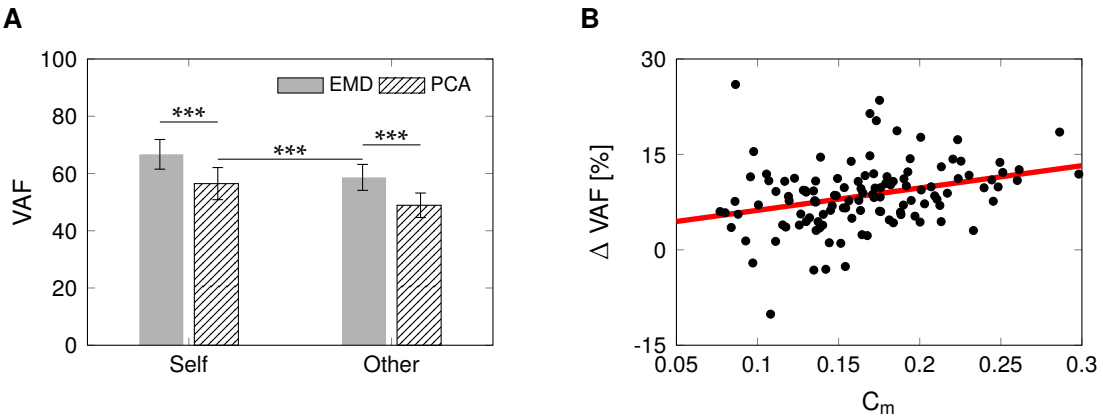

Figure 5.3: A Comparison of the variance captured by 1 SEM *vs.* 1 PC. The proposed method captures significantly more variability and generalises better across subjects. Error bars indicate SD. Wilcoxon signed rank test. ***: $p < 0.001$. **B** The additional variance captured by a single SEM positively correlates with the manipulative complexity of the task.

5.3.1 Repetitive Actions

We learnt dictionaries of sparse eigenmotions on 112 recordings of 7 subjects performing 16 different actions. Each recording contains 10 repetitions of the same action. The SEMD algorithm identified an average of 23 ± 6.5 (mean±SD) SEMs per recording. These enable us to compress the raw movement data (Figure 5.2**A**) into a sparse representation (Figure 5.2**B**) or even a discrete code (Figure 5.2**C**). Figure 5.3**A** demonstrates the ability of extracted SEMs to represent the data. The variance captured by a single SEM was significantly higher than the variance explained by a single PC (SEM: 66.8 ± 5.3%, PC: 56.4 ± 5.6%; Wilcoxon signed rank test, $p < 10^{-3}$). This was equally true for the ability of SEMs to explain the data of other subjects' performing the same task (SEM: 58.1 ± 4.4%, PC: 48.9 ± 4.3%; $p < 10^{-3}$). At the same time, the SEM dictionaries learnt were task-specific as they generalise significantly better to other subjects performing the same task than the same subject performing different tasks (same task, different subject: 58.1 ± 4.5%, same subject, different task: 56.6 ± 4.3%; $p < 10^{-3}$). Interestingly, a single SEM still captured more variance when generalising to other subjects

than a single PC could explain for the data it was trained on ($p < 10^{-3}$), pointing to the generality of SEMs to represent movement data as well as the inadequacy of forcedly orthogonal PCs for this task. The difference in variance captured by SEMs over PCs also correlates positively (Pearson's $r = 0.31$, $p < 10^{-4}$) with the manipulative complexity (Belić and Faisal 2015) of the task (Figure 5.3**B**). This supports the underlying assumption that natural behaviour can be described as a – possibly overlapping – sequence of simpler actions.

5.3.2 Natural Movement Data

Next, we analyse hand movement data collected during natural behaviour using the same methodology as described above. The algorithm identified 88 ± 3.5 SEMs in data of the right hand and 91 ± 4 for left hand data (Wilcoxon ranked sum test, n.s.). A representative selection of SEMs from a single subject and their influence on hand shape are displayed in Figure 5.4. Visual inspection of the loadings (i.e. weights of joints) on individual SEMs showed that most act on multiple fingers simultaneously. The only finger consistently being individuated in multiple SEMs was the thumb, which fits with its special role in natural hand movement (see Ingram et al. 2008, and Chapter 3). The SEMs explained significantly more variance for the same latent dimensionality than PCA (Figure 5.5**A**; Wilcoxon signed rank test, $p < 10^{-3}$ for both hands). On average, encoding the data using a single SEM explained ~140% of the variance captured by a single PC. This value decreases smoothly before attaining 100% (i.e. equality) at a latent dimensionality of 10 (Figure 5.5**B**, top). In terms of absolute difference, this amounted to ~16% for a latent dimensionality of 1 or 2 and then dropped off smoothly to 0 (Figure 5.5**B**, bottom), suggesting that a local latent dimensionality of 2 may be ideal for taking advantage of the properties of SEMs.

As for the previous data set, we tested how well the SEMD or PCA dictionary generalise across subjects (Figure 5.5**C**). Again, SEMs outperformed PCs even when

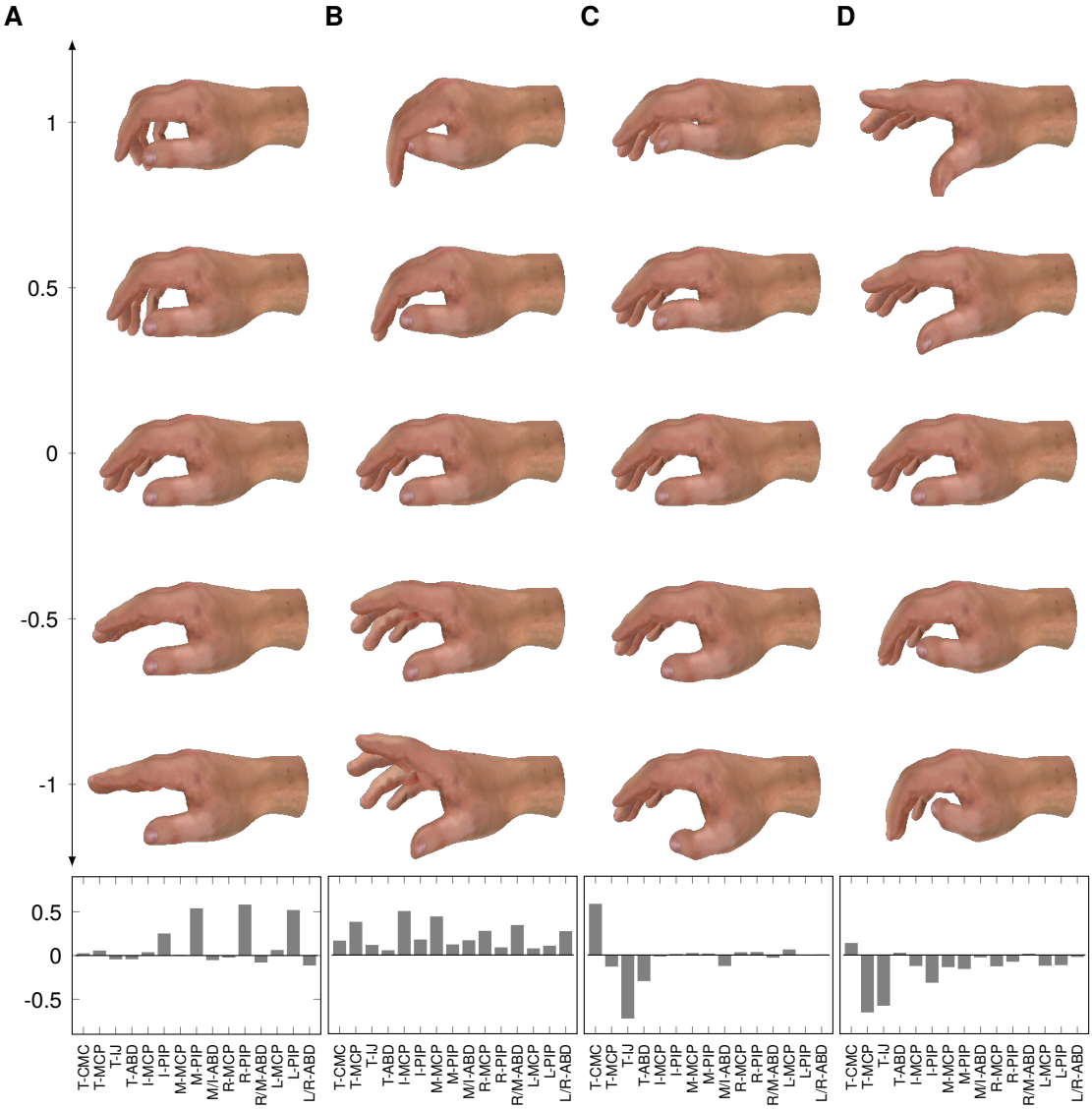

Figure 5.4: A–D Example activity of the 4 major SEMs in the dictionary of subject 4. **Top** Impact of SEM at different weightings (left axis) on hand pose. **Bottom** Mathematical representation of the SEMs illustrated above. Bar height indicates relative movement with respect to the other joints. Figures generated using LibHand (Šarić 2011).

generalising to data they were not trained on, although the loss in variance captured was not significantly different between SEMs and PCs (Wilcoxon signed rank test, right: $p = 0.79$, left: $p = 0.08$). We did not restrict SEMD and PCA dictionaries to have the same size as one of the key advantages of SEMs is to be over-complete. However, this means that we are not quite comparing models of equal complexity. Interestingly, while there was no difference in SEMD or PCA performance between hands when encoding the data they were trained on (SEMD: $p = 0.24$, PCA: $p = 0.90$), the left (non-dominant) hand did generalise significantly better in both cases (SEMD: $p = 0.004$, PCA: $p < 10^{-3}$), indicating a less individuated control structure. As a different measure of generalisation, we counted the number of matching SEMs in dictionaries learnt on different recordings. This overlap was $51.4 \pm 7.9\%$ for dictionaries learnt on right hand data and $39.1 \pm 7.3\%$ for left hand data. Within subject, between hands generalisation was $43.6 \pm 8.9\%$. This discrepancy between the dictionaries' ability to explain data they were not trained on and the lack of direct overlap of dictionary elements is intriguing. It is not related to a globally less complex structure of left hand movements (see Section 3.3.1), nor is it due to a smaller dictionary size (Figure 5.5**D**). We investigated whether this observation could be explained by a skewed distribution of SEM activity (i.e. a few SEMs represented in all dictionaries monopolising most of the activity), but this was not the case either. A possible explanation lies in the inter-subject similarity of PCA subspaces (Figure 3.1). As the similarity is significantly higher for subspaces of the left hand than the right hand, SEMs of the left hand are more likely to reside in a similar subspace, which allows them to better represent the data even though they are not identical. This is consistent with the fact that although the left hand dictionaries generalise better to other subjects, this advantage only holds when considering the VAF curve as a whole. For latent dimensionalities up to 3, right hand SEMs explain significantly more variance with respect to PCA than this is the case for the left (1 SEM right: $126.5 \pm 6\%$ of PCA; left: $123.7 \pm 6\%$. Wilcoxon ranked sum test, $p < 10^{-3}$).

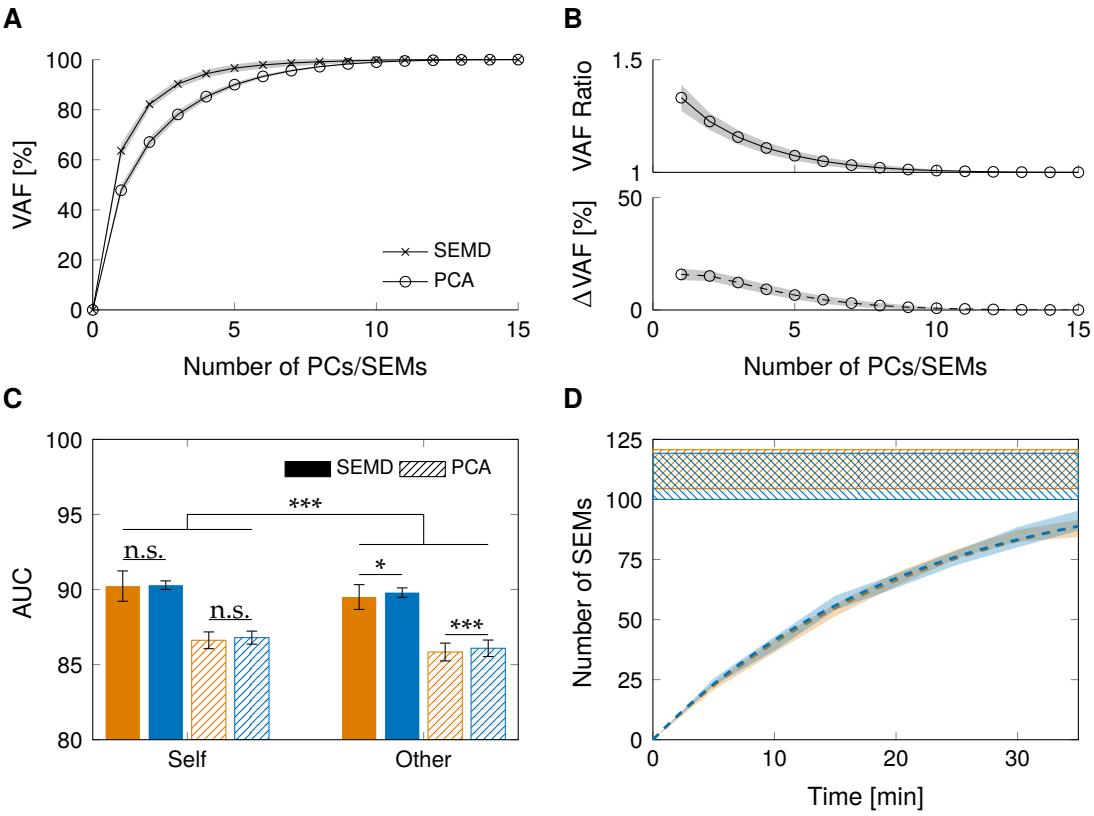

Figure 5.5: **A** Comparison of the variance captured by an increasing number of PCs and SEMs for the right hand. Results for the left hand are similar. **B** Ratio and absolute difference of the VAF by each algorithm. Black lines indicate means and shaded areas standard deviation. **C** The area under the VAF curve (AUC) is significantly higher for encoding via SEMs than PCs both for the data they were trained on (self) as well as the generalisation to other subjects (other). Significant differences appear between dominant and non-dominant hand when generalising the encoding to other subjects. Wilcoxon ranked sum test; *: $p < 0.05$, ***: $p < 0.001$. **D** Evolution of dictionary size over time. Shaded curves indicate 1 standard error around the mean across all recordings (mean not shown). Dashed lines are model fits. The hashed areas at the top are the 95% confidence intervals for total dictionary size. See text for details. Colours represent right (orange) and left (blue) hand as elsewhere in this document.

Since we analysed the data by separating it into blocks of $5\,min$ duration, we can ask the question of how many unique SEMs are being used over time, i.e. how does the dictionary size grow with additional data. As shown in Figure 5.5**D**, the number of individual SEMs grows at an almost identical rate between the right and the left hand. To model this growth, we fitted various functions to this curve. The best fit was consistently achieved by a function of the shape $y = A(1-\exp(-t/\tau))$ (see Figure 5.5**D**). This form gave better results than the two alternatives considered (a power-law and a logarithmic function) although the logarithmic function was a close second. An interesting feature of the exponential function is that it gives us an indication of the total dictionary size by examining the parameter A. This was $A_R = 112.6$ (104.5–120.7) (mean (95%C.I.)) for the right hand and $A_L = 109.6$ (100–119.2) for the left. Of course, given a logarithmic model of dictionary growth there is no upper limit to the number of SEMs.

5.3.3 Structure of the Sparse Representation

To further investigate the sparse encoding of the data, we look at a one-dimensional representation, i.e. a vector indicating which SEM is most likely to have generated each specific data point in the time-series (see Methods). Across all recordings, the probability distribution of a point being represented by any given SEM was quite flat (Figure 5.6**A**). The most represented SEM made up 5.3 ± 0.8% (mean ± s.d.) of the data, while the least represented SEM still accounted for 0.9 ± 0.2% of the data. This is characteristic of a sparse code where a single unit represents a specific high-level feature of the data. Conversely, a dense code would concentrate the bulk of the activity in a small number of SEMs.

Next, we investigated the transition matrix between SEMs. The matrix was predominantly diagonal, indicating that most transitions were self-transitions ($p_{\text{self}} = 0.85 \pm 0.016$). This is consistent with the fact that movements are not instantaneous but

have a certain duration. The time between SEM transitions followed an exponential curve with decay constant $\tau = 75\,ms$ (Figure 5.6**B**), a value which is substantially lower than the full-width at half-maximum for hand movements ($\sim 220\,ms$, see Figure 3.5**B**). This is a consequence of finding the sparse representation using the OMP algorithm. Given its greedy nature and disregard for surrounding time points and noise, it often causes discontinuities in the latent representation. The LASSO (Tibshirani 1996) would avoid this problem but is not suitable for our purposes as it does not allow control over the latent dimensionality.

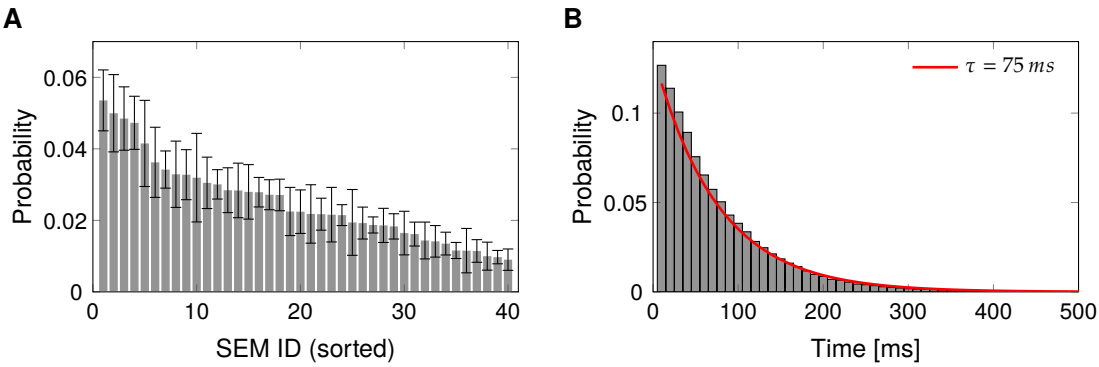

Figure 5.6: A Mean probability distribution over SEM assignment for all data points. The values are ordered from most to least common SEM. Error bars indicate standard deviation across all recordings. **B** Probability distribution of SEM duration and exponential fit.

To see whether higher-order structure was present in the latent representation, we trained hidden Markov models (HMM) with variable numbers of hidden states on the SEM assignments (Figure 5.7**A**). The quality of the fit was assessed with the Bayesian Information Criterion (BIC, Figure 5.7**B**). The optimal HMM was determined to be the one with the lowest BIC overall, and had 16 hidden states. This suggests that there may be some temporal structure in the data.

The HMM can be further examined by investigating the transition and emission matrices (Figure 5.8). The transition matrix displayed a strongly block-diagonal structure.

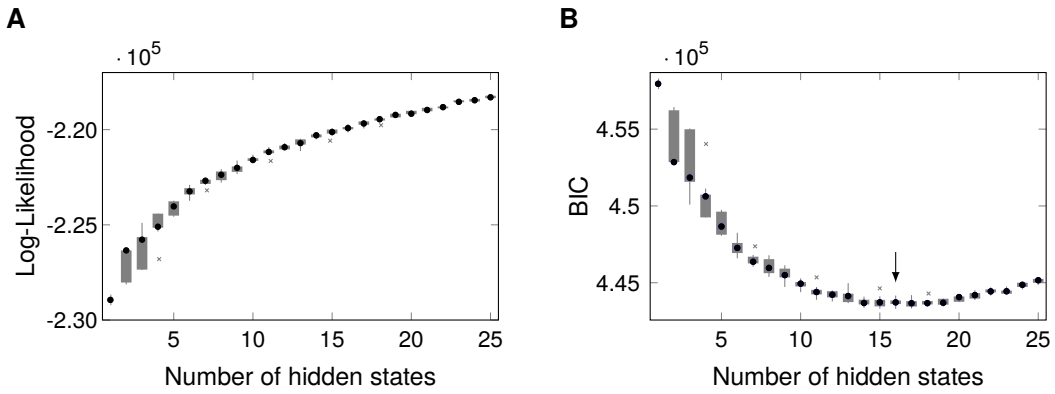

Figure 5.7: A Log-likelihood of the HMM with increasing number of hidden states. **B** Distribution of BIC values for HMMs with variable numbers of hidden states. The lowest value (16) is indicated with an arrow. Black dots indicate median values.

(Figure 5.8**A**). This was still the case when the number of hidden states was varied up or down by 1 or 2, although the number of blocks varied slightly. This structure results in relatively separate sub-networks of hidden states within the model (Figure 5.9) between which switching is rare. The probability of self-transition was generally low (median: 9%, min: 0%, max: 71%), which may suggest that hidden states may encode a single SEM. A brief look at the emission matrix (Figure 5.8**B**) or the HMM map (Figure 5.9) reveals that this is not the case. Instead, most hidden states had a broad emission spectrum, emitting on average 19 SEM with a probability exceeding 1% (range: 13–29). While the number of non-zero emission probabilities was high, it was non-uniformly distributed with on average 4 (1–9) SEM accounting for more than 50% of the emission probabilities for each state. This analysis also highlights the importance of SEMs 1, 2, 3, 6 and 20, which dominate the emission probabilities in states 3, 2, 7, 5 and 11 respectively. SEM 1 and 2 are depicted in Figure 5.4**A** and **B** respectively (Figure 5.8**B**). The others (not shown) generate flexion along all MCP joints, thumb abduction and thumb flexion respectively.

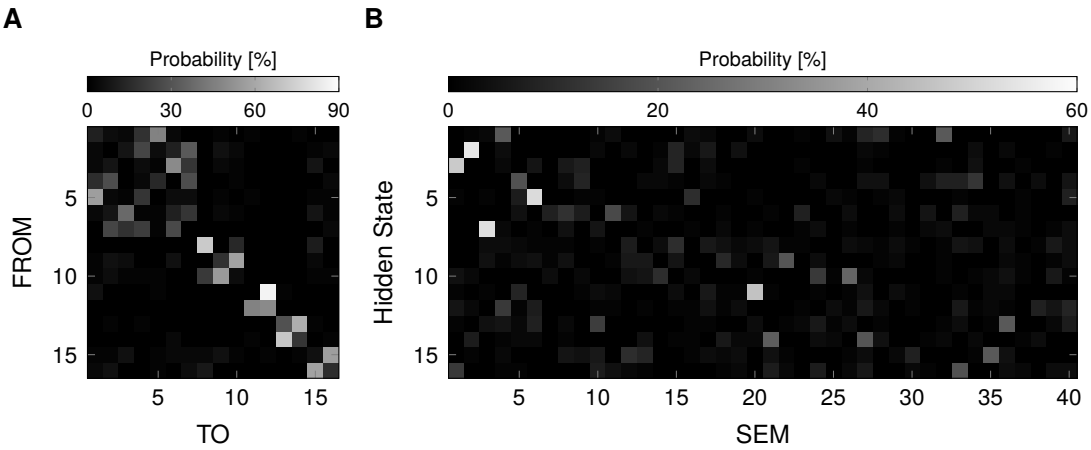

Figure 5.8: A Transition and **B** emission matrices of the best fitting HMM.

5.3.4 Prediction of Hand Movements

Similar to our work in Section 3.3.3, we attempt to predict the velocity of hand movements based on the movement and position of the rest of the body. In this case, however, we take a two stage approach. In the first phase, we use body data to estimate the SEM having generated the data using a classification tree. In the second step, individually trained regressors are used to predict hand movement data. The results are shown in Figure 5.10. SEM prediction from body data was well above chance at 46.5 ± 5% (mean ± s.d.) accuracy for right hand data (left: 49.5 ± 5%; Student's t-test, $p = 0.007$). No obvious pattern was visible in the misclassification rate of SEMs (Figure 5.10**A**). In terms of hand movement prediction, the regression model with known SEM attribution achieves 13.5 ± 1.3% accuracy for right hand movements (left: 10.3 ± 1.4%). If we add the uncertainty given by the SEM classification stage, these values drops to 6.0 ± 1% (left: 4.2 ± 1.5%). This remains significantly higher than for direct prediction without knowledge of SEM state (Wilcoxon ranked sum test, $p < 0.001$). As previously, prediction of right hand movements was significantly better than for the left hand.

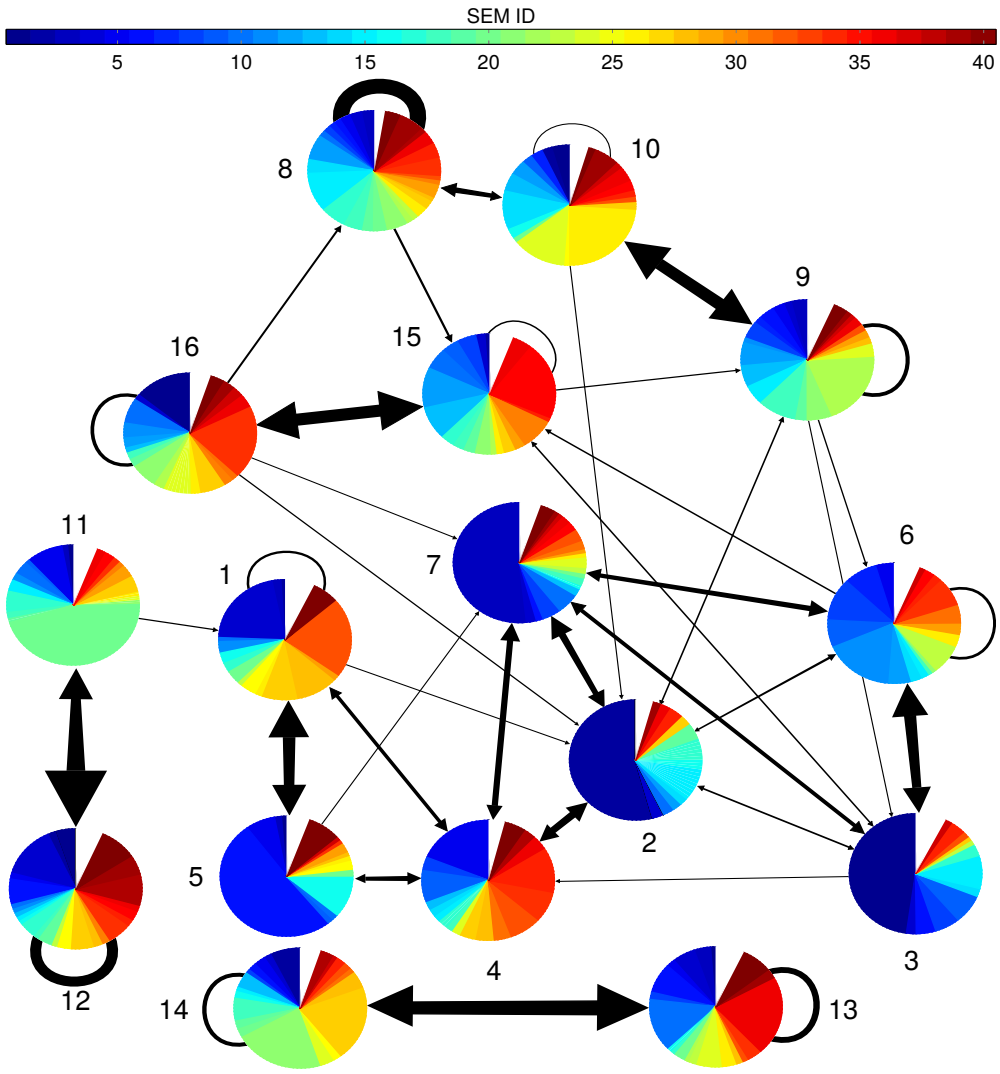

Figure 5.9: Network diagram of the hidden Markov model. Transition probabilities < 5% and emission probabilities < 1% are not displayed for clarity. Note the sub-networks (e.g. 11 & 12) which result from the block-diagonal structure of the transition matrix. Hidden states 2 and 3 are dominated by hand openings (the dominant SEM for hidden state 2 is depicted in Figure 5.4**B**), 5 by a SEM resembling a precision grasp and 11 by thumb flexion. Other hidden states show a much broader distribution of emission probabilities.

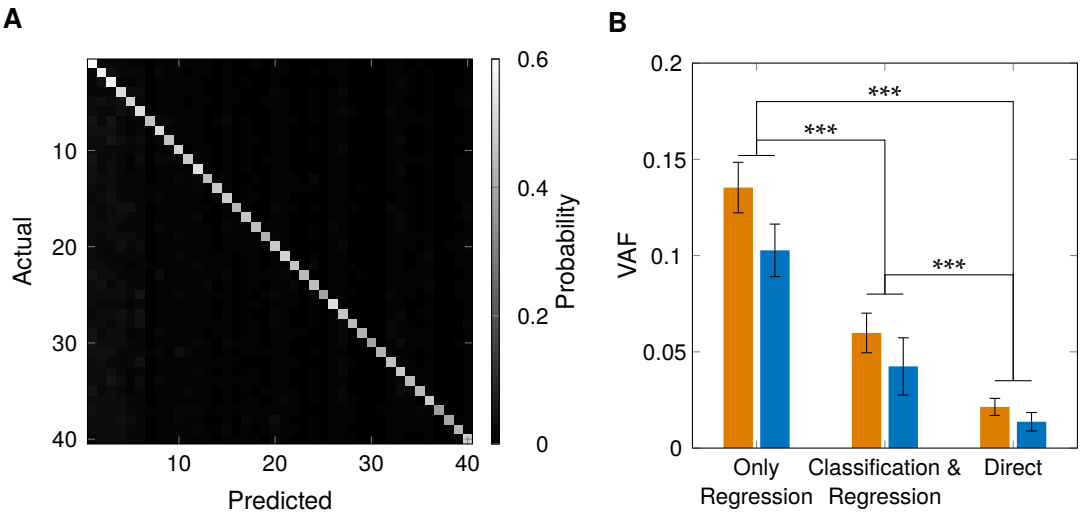

Figure 5.10: A Average confusion matrix for the prediction of SEM across all recordings. **B** Quality of hand movement prediction using known SEM assignment ("only regression"), regression after prior classification and direct regression. Wilcoxon ranked sum test, ***: $p < 10^{-3}$.

5.4 Discussion

We used a previously introduced algorithm (Chapter 4) to analyse natural movement data. Our goal was to identify whether there is the possibility that akin to sensory data (Olshausen and Field 1997; Lewicki 2002), the statistics of natural motor output may be represented by a sparse code. We demonstrate that such an encoding may be learnt, and that it is much more efficient at describing the data than PCA, which has been previously used to argue for a low-dimensional representation of hand movements (Santello, Flanders, and Soechting 1998; Ingram et al. 2008; Faisal et al. 2010; Belić and Faisal 2015). This advantage over PCA is directly dependent on task complexity (Figure 5.3**B**), agrees with the idea of sparse eigenmotions (SEMs) representing stereotyped individual movements. Crucially, these SEMs are well preserved across subjects (> 50% between dominant hands), suggesting a certain amount of invariance in hand representation in humans. Intriguingly, the sparse components could well be

predicted from movement of other body parts (arms, contralateral hand and neck), which suggests that the structure may extend beyond finger movements to include the wrist, elbow and shoulder.

5.4.1 Differences Between Dominant and non-Dominant Hand

Our analysis reveals some intriguing differences between dominant (right) and non-dominant hand: (1) although the SEM dictionaries are comparable in size (Figure 5.5**D**), the overlap between dictionary elements is significantly higher in the dominant hand. This contrasts with (2) the ability of non-dominant hand dictionaries to generalise to other subjects, which is significantly higher than for the dominant hand. However, (3) predicting dominant hand movements from body state was significantly better than for the non-dominant hand.

How do we reconcile these findings? Neurophysiological data suggests that the area of primary motor cortex representing the dominant hand is larger (Amunts et al. 1996; Volkmann et al. 1998) and more excitable (Tarkka and Hallett 1990; Boggio et al. 2006) than that of the non-dominant hand. However, similar studies for premotor areas – where we assume the sparse representation to reside – are lacking. If the results from primary motor cortex extrapolate to higher cortical areas, this would explain why the dictionaries of the non-dominant hands are able to generalise better. Yet, it does not explain the similarity in dictionary size.

The difference in hand predictability is likely to have its origin elsewhere. In Section 3.3, we showed that the correlation structure between fingers of both hands was not significantly different. This is in line with results from Ejaz, Hamada, and Diedrichsen (2015) which showed no difference between the functional representations of digits in either hand. An increased correlation between fingers of the dominant hand is thus unlikely to be the cause of the observed differences in prediction quality. We suggest that this discrepancy comes from the reliance of the linear regression model

on movements of the ispilateral arm (Figure 3.10): multiple studies have shown that neurons in M1 code both for arm as well as hand movements (Vargas-Irwin et al. 2010; Hendrix, Mason, and Ebner 2009). Although to the best of our knowledge no study has examined how laterality affects these representations, we predict that in the dominant hemisphere the number of neurons encoding both arm and hand areas is greater or, alternatively, that corticocortical connections between the arm and hand areas of the motor cortex are stronger.

5.4.2 Sparse Coding of Hand Movements in the Brain

The sparse representation of hand movements is supported by numerous electrophysiological studies of the motor and premotor cortex of monkeys, as well as the anterior intraparietal lobe (Rizzolatti and Gentilucci 1988; Jeannerod et al. 1995; Murata et al. 1996; Murata et al. 2000; Raos et al. 2004; Raos et al. 2006). These studies reveal a complex network responsible for visuo-motor transformation of object type shape and orientation for grasping. Studies in all these areas, but predominantly macaque area F5 (equivalent to the human PMv; Raos et al. 2006) exhibit neurons which code for specific grasp types (e.g. precision grip or hook grasp), equivalent to our concept of sparse eigenmotions. In a separate study, Schaffelhofer, Agudelo-Toro, and Scherberger (2015) found 20 different hand configurations from monkeys grasping 50 objects of different shapes and sizes. They demonstrate that during the planning phase the grasp type can be best decoded from neurons in area F5 and the AIP, but less well from M1. This trend is reversed during the movement execution phase, with optimal decoding performance coming from M1 and F5, and AIP being considerably worse.

These findings support the view of a hierarchical motor control system in which the premotor cortex determines the overall hand movement shape and the execution is performed at a more detailed level in M1. In particular, it should be noted that these grasp types do not reflect an orthogonal basis of hand configurations as would

be retrieved through PCA (Santello, Flanders, and Soechting 1998; Ingram et al. 2008), and for which neural correlates could not be demonstrated (Mollazadeh et al. 2014).

The aforementioned studies are subject to two limitations: (1) all hand movements performed were grasps and typically started from a fixed configuration. Consequently, we do not have information about the neural response to changes in hand shape which occur naturally. Furthermore, it neglects the fact that only ~13% of monkeys hand movements involve reach-to-grasp (Graziano et al. 2002). (2) Electrophysiological studies based on single unit recordings or on data obtained from multi-electrode arrays with subsequent spike sorting (this includes all those cited in this section), are prone to sampling bias by identifying neurons which are particularly active (Barth and Poulet 2012). More studies are thus required to determine the total number of neurons in these areas and those responsive to specific hand movements.

5.4.3 A Language of Movements?

With the original discovery of grasp specific neurons in area F5, Rizzolatti and Gentilucci (1988) suggested that this representation may correspond to a *"vocabulary of elementary motor acts"*, in which individual neurons code for specific words. The original formulation of the hypothesis thus reflects a "grandmother cell" code of motor actions, but more recent findings suggest this is more likely to be a sparse code where multiple neurons respond to a given grasp type (Raos et al. 2006; Schaffelhofer, Agudelo-Toro, and Scherberger 2015). While this hypothesis is currently cited in later work, no follow-up study has addressed this particular point. Indeed, the very paradigm at the heart of current research (i.e. reach-to-grasp) prevents a detailed analysis of this point, as the sequence of words is pre-imposed.

On the other end, Hidden Markov models have long been used for part-of-speech tagging in natural language processing (DeRose 1988). This operation attempts to assign a tag (noun, verb, adjective, etc.) to words in a body of text. Although most of

the time the tags are predefined by an expert, it should in theory be possible to learn the optimal number of tags (and subsequently assign a meaning to them) by learning a HMM on natural language. Key features of this approach are that (1) the number of tags is much smaller than the number of possible words and (2) words may belong to more than one tag category (e.g. *run* may be both a noun and a verb).

Our analysis of the discrete sequences constructed by encoding the natural movement data with a HMM reveals exactly that. The optimal number of hidden states (or tags) as determined by the BIC is 16. Given the flatness of the BIC curve for 14–17 hidden states (Figure 5.7**B**), the precise optimum is somewhat unclear and additional data and/or simulations may clarify this aspect. It is, however, indisputable that the number of hidden states is significantly smaller than the number of observed symbols. Moreover, individual SEMs were mostly assigned to a small number of hidden states: the median number of hidden states accounting for more than 50% of the emission probability of a SEM was 2 (range: 1–4). These results make a strong argument for the existence of a hierarchical, potentially language-like structure of human motor control which is intuitively consistent with the decomposition of a complex task into a sequence of simple subtasks.

6 Conclusion

"I may not have gone where I intended to go, but I think I have ended up where I needed to be."

— Douglas Adams

The Long Dark Tea-Time of the Soul

The previous four chapters introduced various experimental methods, data sets and analyses. In each case, we discussed the results and analysed weaknesses of our approach, as well as the implications of our findings in the context of related research. To finish, we summarise our findings, and put our work into the broader context of neuroscientific research. Finally, we discuss potential avenues for further research.

6.1 Summary of Results

Following a general introduction in Chapter 1, Chapter 2 describes an experimental paradigm measuring human in unconstrained, natural settings. We describe the statistics of joint positions and velocity and develop metrics for comparing behaviour between humans. We find that while position is extremely variable, the distribution of peak velocities in natural behaviour is highly stereotyped. We suggest to use this feature in future studies for comparing movement of healthy subjects to patients with neurological disorder and as target for rehabilitation. In addition the the characterisation of human behaviour, we examined two hypotheses in light of the newly collected

data: (1) movement linearity and (2) the 2/3 power law (Lacquaniti, Terzuolo, and Viviani 1983). We find that only 70% of movements are straight, but almost all (> 99%) are planar. Interestingly, movement planes had no preferred orientation. This supports current experimental procedures where subjects perform movements in a horizontal plane. In contrast, the 2/3 power law is not supported by our analysis. Instead of the expected exponent value of $\beta = 0.33$, we obtain very variable exponents, with a median of 0.5. We argue that this new value can be related directly to the physics of movement. The prevalence of the 2/3 power law in previous studies is unclear, and may be related to the experimental protocol. These results make a strong argument for validating results from laboratory experiments in ecologically relevant situations.

Chapter 3 goes beyond descriptive statistics by analysing spatial and temporal structure of the movement. In support of work in the previous chapter, we find that the correlation structure of hand movements contrasts significantly whether they were collected in the laboratory (Jarrassé et al. 2014) or from natural behaviour (Ingram et al. 2008, and our data). Interestingly, we find no difference between the dominant and the non-dominant hand either in terms of correlation structure or movement frequency (as characterised by the auto-correlation function). The only consistent difference between both hands is that the dominant hand tends to be used more for precision movements (i.e. only a few joints move simultaneously) while the non-dominant hand use more frequently used for grasping (i.e. all joints move simultaneously). Similar to movements of the hand, we characterise the correlation structure of the major joints of the body. In contrast to the hand, we find very few correlations between limbs: the only ones which appear consistently between multiple recordings are correlations between the lower limbs (probably induced by walking) and between shoulder and elbow (a typical occurrence during grasping). Principal components of body movements are consequently very variable between subjects. However, the little structure which is present enables us to predict hand configuration and movement above chance levels.

The findings of this chapter lead us to argue that PCA may not be the optimal way of analysing natural hand movements and that, instead, a sparse code may be more appropriate.

To identify a potential sparse code, Chapter 4 introduces an algorithm for data driven dictionary learning. In contrast to most available methods, the method proposed identifies not only spatial structure, but forces it to have temporal continuity. In addition, we the algorithm is very data efficient and capable of automatically estimating the latent dimensionality of the data. We show that these features are robust to changes in sample size, data dimensionality and latent dimensionality. Although it is not quite as computationally efficient as other algorithms, our approach consistently outperforms them when it comes to identifying the underlying generating dictionary.

Chapter 5 sees the application of the previously developed algorithm on natural movement data. We show that our *sparse eigenmotions* (SEMs) are significantly more efficient at encoding the data than principal components. In particular, the increased efficiency scales with the complexity of the task. Applied to hand movement data of freely behaving subjects, we identify dictionaries which are not only similar in size, but also largely overlap between subjects. This is the strongest argument for the existence of sparse coding of natural movement statistics. Compressing the data to a one-dimensional latent representation and analysing the resulting time-series points to the existence of a higher-level structure. We argue that this may be interpreted as a language for movement. Finally, we take advantage of the newly discovered structure to propose a new different approach to predicting hand movements from other body parts. This variation on the method first proposed in Chapter 3 provides a 10-fold increase in prediction quality. This improvement is an additional argument for the existence of a sparse structure of hand movements.

6.2 A Window Into the Brain

What does measuring natural movement tell us about motor control? If we make the simplifying assumption that the motor plant can be seen as a linear system with quadratic control cost and Gaussian noise (i.e. a LQG system) then a direct duality exists between an optimal state estimator (a Kalman filter) and the optimal controller (Kalman 1960). This implies that – given knowledge of the state space – we can estimate the brain's control method. This obviously raises three issues: (1) the assumption that the motor plant is a LQG system, (2) the use of optimal control by the brain and (3) knowledge of the state space.

The first question is somewhat beyond the scope of our work, although it is obvious that non-linearities are present at every level of the motor control system, from the integration of neural inputs to muscle activity. However, recent research suggests that this duality might extend beyond the LQG setting to non-linear (but control-affine) systems (Todorov 2008).

Whether the brain uses optimal control for planning and execution of motor tasks is a much debated question. Multiple experiments demonstrated that motor control is largely in agreement with optimal control which minimises energy expenditure (e.g. Todorov and Jordan (2002), for a review see Scott (2004)). However, whether this is due to active optimal control or is simply habitual as the consequence of evolution is unclear (De Rugy, Loeb, and Carroll 2012).

From our point of view, the most interesting question regards the state space on which the brain acts when controlling the motor plant. The output is obvious: movement clearly happens in kinematic space and forces in muscle space. What does the brain care about? Studies in the primary motor cortex of monkeys, mostly focused on M1, have found neurons responsive to instantaneous velocity (Paninski et al. 2004), peak velocity (Churchland, Afshar, and Shenoy 2006) and direction in extrinsic (Geor-

gopoulos, Schwartz, and Kettner 1986) and intrinsic space (Kakei, Hoffman, and Strick 1999), force (in extrinsic space (Hepp-Reymond et al. 1994)), muscle tension (Kakei, Hoffman, and Strick 1999) and, more recently, even muscle synergies (Overduin et al. 2015). Given the high amount of redundancy between these representations, it seems unlikely that all of them are controlled simultaneously. In particular, Barlow's efficient coding hypothesis (Barlow 1961) suggests that in light of the physical constraints on information processing by the brain (energy consumption, space limitation, processing speed and thermodynamic noise), the central nervous system should utilise a representation which is adapted to the statistical structure of the information being processed. This approach was mainly developed for the sensory system where it has been particularly successful in explaining the properties of neurons in the visual and the auditory system (Olshausen and Field 1997; Lewicki 2002). These systems have the advantage that it is relatively straightforward to characterise the statistical properties of the input stimuli (e.g. picture of natural environments (Olshausen and Field 1997) or natural sounds (Lewicki 2002)) and to derive efficient codes for them. Characterising the natural output of the central nervous system is much more technically challenging. Our work in Chapter 2 and Chapter 3 provides a step in that direction. For the first time, we have measured and described the statistical properties of the kinematics of natural movement from various perspectives. While these results do not in themselves tell us how movements are encoded in the brain, they provide a powerful tool for identifying the correct encoding in electrophysiological studies as it should be expected that the "true" representation is provided by a neural population with tuning curves optimally adapted to the statistics of the natural output.

The problem might, however, be more complicated. We showed in Chapter 3 that our limbs do not move completely independently of each other, but that a correlation structure exists between them, in particular for the hand. It may thus be that instead of encoding the statistical properties of individual joints or limbs, the brain encodes linear

(or even non-linear) combinations thereof. In Chapter 5 we attempt to identify such linear structure in hand movements by using the purpose-built algorithm described in Chapter 4. We demonstrated that hand movements may be described by a sparse code of *eigenmotions* which is task specific, generalises well across subjects and allows for a hierarchical representation and control of movement. Whether this representation is related to the way the brain controls movement remains to be studied in electrophysiological experiments. The fact that sparse eigenmotions (SEMs) are predictable from other body parts suggests that there may even be structure spanning multiple limbs.

These findings have not only neuroscientific interest, but lay the foundations for new approaches for context-aware robotic and prosthetic control. Recently, Abramova, Faisal, and Kuhn (2011) demonstrated that intelligent switching between randomly placed (in state space) linear controllers provides significant advantages over traditional methods for controlling non-linear robotic devices. Our sparse encoding of natural movement, although not directly concerned with the control problem, suggests a more intelligent way of placing the linearisation points and hints towards a hierarchical approach to control switching. In terms of prosthetic control, we previously investigated the possibility that hand movements can be described as a sequence of movement primitives (Thomik, Haber, Faisal, et al. 2013; Haber, Thomik, Faisal, et al. 2014). Although that representation does not generalise as well as SEM encoding, it already demonstrated that the approach of segmenting the possible output has significant benefits in terms of intention prediction for prosthetic control. In addition to the benefits of segmented control, we demonstrated that there is significant structure between movements of the affected arm and the hand (Chapter 3) which is even more obvious when the movement is segmented according to the generating SEM (Chapter 5). These features are worth exploring in terms of a multi-modal control structure for myoelectic hand prostheses. Movement of the arm and knowledge of the action executed previously may act as a prior for decoding the user's intention and could easily

be combined with recent developments in prosthesis control (see Farina et al. 2014, for a review). Together, these results should provide users with a prosthetic device which is more versatile and intuitive to control.

6.3 Future Perspectives

In order to fully exploit the statistics of natural movement and provide a deeper insight into the neuroscience of motor control, we need to pursue two avenues of research.

First of all, new analytical and computational methods need to be developed for analysing large, unstructured data sets such as natural behaviour. Controlled experiments in which subjects repeatedly perform very specific actions or tasks simply lack the statistical richness required to unravel the complexity of motor control. Analysing the data from freely moving subjects is, however, extremely complex. Traditional methods such as (cross-) correlation analysis, PCA, etc. are ill-suited to the task because they consider global effects and thus average out local structure. Performing the analysis on windowed data is itself problematic. How does one choose the window size and placement? Consequently, it is necessary to develop novel methods which are capable of identifying relevant local structure in a principled way while accounting for artefacts which such an analysis may create. The SEMD algorithm is one such method designed with a very specific application in mind. Other approaches might be needed to explore higher order statistics or spatio-temporal structure of the data.

The bigger task is to relate the statistics of natural movement to its neural representation. Initially, this will require to develop experimental paradigms which enable us to collect both behavioural and electrophysiological data in a rich, ecologically relevant environment while minimising the typical constraints of experimental setups. This will likely require creation of entirely new recording equipment and experimental protocols. Such studies have the advantage that they may be carried in a large range of

organisms, be it worms (e.g. *C. elegans*) rodents or more complex organisms such as primates. The analysis of neural recordings in light of the natural statistics of motor output offer great potential in identifying the relevant variables and will help us better understand the core purpose of our brain: movement.

Bibliography

Abeles, M, E Vaadia, and H Bergman (1990). "Firing patterns of single units in the prefrontal cortex and neural network models." In: *Network: Computation in Neural Systems* 1.1, pp. 13–25.

Abend, W., E. Bizzi, and P. Morasso (1982). "Human arm trajectory formation." In: *Brain* 105.2, p. 331.

Abramova, Ekaterina, Aldo Faisal, and Daniel Kuhn (2011). "Combining Markov Decision Processes with Linear Optimal Controllers." In: *ICCSW*. Citeseer, pp. 3–9.

Aharon, Michal, Michael Elad, and Alfred Bruckstein (2006). "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation." In: *Signal Processing, IEEE Transactions on* 54.11, pp. 4311–4322.

Ambrose, Stanley H (2001). "Paleolithic technology and human evolution." In: *Science* 291.5509, pp. 1748–1753.

Amunts, Katrin et al. (1996). "Asymmetry in the human motor cortex and handedness." In: *Neuroimage* 4.3, pp. 216–222.

Anderson, James R, Marie Montant, and Didier Schmitt (1996). "Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object-choice task." In: *Behavioural Processes* 37.1, pp. 47–55.

Atkeson, Christopher G and John M Hollerbach (1985). "Kinematic features of unrestrained vertical arm movements." In: *The Journal of Neuroscience* 5.9, pp. 2318–2330.

Attwell, David and Simon B Laughlin (2001). "An energy budget for signaling in the grey matter of the brain." In: *Journal of Cerebral Blood Flow & Metabolism* 21.10, pp. 1133–1145.

Baddeley, Roland et al. (1997). "Responses of neurons in primary and inferior temporal visual cortices to natural scenes." In: *Proceedings of the Royal Society of London B: Biological Sciences* 264.1389, pp. 1775–1783.

Bailey, Trevor C and Anthony C Gatrell (1995). *Interactive spatial data analysis*. Vol. 413. Longman Scientific & Technical Essex.

Barlow, H.B. (1972). "Single units and sensation: A neuron doctrine for perceptual psychology?" In: *Perception* 1, pp. 371–394.

Barlow, Horace B (1961). "Possible principles underlying the transformations of sensory messages." In: *Sensory Communication*. MIT press.

Barlow, Horace B. (1989). "Unsupervised learning." In: *Neural Computation* 1.3, pp. 295–311.

Barth, Alison L and James FA Poulet (2012). "Experimental evidence for sparse firing in the neocortex." In: *Trends in Neurosciences* 35.6, pp. 345–355.

Basner, Mathias et al. (2007). "American time use survey: sleep time and its relationship to waking activities." In: *Sleep* 30.9, p. 1085.

Behbahani, Feryal MP et al. (2015). "Haptic SLAM for context-aware robotic hand prosthetics-simultaneous inference of hand pose and object shape using particle filters." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 719–722.

Belic, J. (2010). "Prediction of object manipulation actions for hand neuroprosthetic control from the structured variability of finger movements." MA thesis. Imperial College London.

Belić, Jovana J and A Aldo Faisal (2015). "Decoding of human hand actions to handle missing limbs in Neuroprosthetics." In: *Frontiers in Computational Neuroscience* 9.

Bernstein, N.A. (1967). *The co-ordination and regulation of movements*. Pergamon Press Ltd.

Biess, Armin, Dario G Liebermann, and Tamar Flash (2007). "A computational model for redundant human three-dimensional pointing movements: integration of independent spatial and temporal motor plans simplifies movement dynamics." In: *The Journal of Neuroscience* 27.48, pp. 13045–13064.

Bizzi, E., F.A. Mussa-Ivaldi, and S. Giszter (1991). "Computations underlying the execution of movement: a biological perspective." In: *Science* 253.5017, pp. 287–291.

Bock, Otmar and Rainer Beurskens (2010). "Changes of locomotion in old age depend on task setting." In: *Gait & Posture* 32.4, pp. 645–649.

Bock, Otmar and Anne Hagemann (2010). "An experimental paradigm to compare motor performance under laboratory and under everyday-like conditions." In: *Journal of Neuroscience Methods* 193.1, pp. 24–28.

Bock, Otmar and Anne Züll (2013). "Characteristics of grasping movements in a laboratory and in an everyday-like context." In: *Human Movement Science* 32.1, pp. 249–256.

Boggio, Paulo S et al. (2006). "Enhancement of non-dominant hand motor function by anodal transcranial direct current stimulation." In: *Neuroscience Letters* 404.1, pp. 232–236.

Bourke, Alan Kevin et al. (2014). "Energy expenditure estimation using accelerometry and heart rate for multiple sclerosis and healthy older adults." In: *Wearable and Implantable Body Sensor Networks Workshops (BSN Workshops), 2014 11th International Conference on*. IEEE, pp. 1–5.

Bradley, David M and J Andrew Bagnell (2009). "Differential sparse coding." In: *Advances in Neural Information Processing Systems*, pp. 113–120.

Branson, Kristin et al. (2009). "High-throughput ethomics in large groups of Drosophila." In: *Nature Methods* 6.6, pp. 451–457.

Brecht, Michael and Bert Sakmann (2002). "-Dynamic representation of whisker deflection by synaptic potentials in spiny stellate and pyramidal cells in the barrels

and septa of layer 4 rat somatosensory cortex." In: *The Journal of Physiology* 543.1, pp. 49–70.

Brecht, Michael et al. (2004). "Whisker movements evoked by stimulation of single pyramidal cells in rat motor cortex." In: *Nature* 427.6976, pp. 704–710.

Burdet, Etienne and Theodore E Milner (1998). "Quantization of human motions and learning of accurate movements." In: *Biological Cybernetics* 78.4, pp. 307–318.

Buzzell, George et al. (2013). "Speed of human biological form and motion processing." In: *PLoS One* 8.7, e69396.

Candes, Emmanuel J and David L Donoho (2000). *Curvelets: A surprisingly effective nonadaptive representation for objects with edges*. Tech. rep. DTIC Document.

Chao, EY et al. (1983). "Normative data of knee joint motion and ground reaction forces in adult level walking." In: *Journal of Biomechanics* 16.3, pp. 219–233.

Chen, Scott Shaobing, David L Donoho, and Michael A Saunders (1998). "Atomic decomposition by basis pursuit." In: *SIAM Journal on Scientific Computing* 20.1, pp. 33–61.

Cheung, V.C.K. et al. (2005). "Central and sensory contributions to the activation and organization of muscle synergies during natural motor behaviors." In: *The Journal of Neuroscience* 25.27, pp. 6419–6434.

Churchland, Mark M, Afsheen Afshar, and Krishna V Shenoy (2006). "A central source of movement variability." In: *Neuron* 52.6, pp. 1085–1096.

Chvatal, Stacie A et al. (2011). "Common muscle synergies for control of center of mass and force in nonstepping and stepping postural behaviors." In: *Journal of neurophysiology* 106.2, pp. 999–1015.

Corazza, Stefano et al. (2010). "Markerless motion capture through visual hull, articulated icp and subject specific model generation." In: *International Journal of Computer Vision* 87.1-2, pp. 156–169.

Crochet, Sylvain et al. (2011). "Synaptic mechanisms underlying sparse coding of active touch." In: *Neuron* 69.6, pp. 1160–1175.

Cutkosky, Mark R (1989). "On grasp choice, grasp models, and the design of hands for manufacturing tasks." In: *Robotics and Automation, IEEE Transactions on* 5.3, pp. 269–279.

Dankert, Heiko et al. (2009). "Automated monitoring and analysis of social behavior in Drosophila." In: *Nature Methods* 6.4, pp. 297–303.

d'Avella, A., P. Saltiel, and E. Bizzi (2003). "Combinations of muscle synergies in the construction of a natural motor behavior." In: *Nature Neuroscience* 6.3, pp. 300–308.

d'Avella, A. et al. (2006). "Control of fast-reaching movements by muscle synergy combinations." In: *The Journal of Neuroscience* 26.30, pp. 7791–7810.

De Kock, CPJ et al. (2007). "Layer-and cell-type-specific suprathreshold stimulus representation in rat primary somatosensory cortex." In: *The Journal of Physiology* 581.1, pp. 139–154.

De Leva, Paolo (1996). "Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters." In: *Journal of Biomechanics* 29.9, pp. 1223–1230.

De Rugy, Aymar, Gerald E Loeb, and Timothy J Carroll (2012). "Muscle coordination is habitual rather than optimal." In: *The Journal of Neuroscience* 32.21, pp. 7384–7391.

Delis, Ioannis et al. (2013). "Quantitative evaluation of muscle synergy models: a single-trial task decoding approach." In: *Frontiers in computational neuroscience* 7.

DeRose, Steven J (1988). "Grammatical category disambiguation by statistical optimization." In: *Computational Linguistics* 14.1, pp. 31–39.

Desmurget, Michel and Claude Prablanc (1997). "Postural control of three-dimensional prehension movements." In: *Journal of Neurophysiology* 77.1, pp. 452–464.

DeWeese, Michael R, Michael Wehr, and Anthony M Zador (2003). "Binary spiking in auditory cortex." In: *The Journal of Neuroscience* 23.21, pp. 7940–7949.

Di Pellegrino, Giuseppe et al. (1992). "Understanding motor events: a neurophysiological study." In: *Experimental Brain Research* 91.1, pp. 176–180.

Do, Minh N and Martin Vetterli (2005). "The contourlet transform: an efficient directional multiresolution image representation." In: *Image Processing, IEEE Transactions on* 14.12, pp. 2091–2106.

Domkin, Dmitry et al. (2005). "Joint angle variability in 3D bimanual pointing: uncontrolled manifold analysis." In: *Experimental Brain Research* 163.1, pp. 44–57.

Donoho, David L, Michael Elad, and Vladimir N Temlyakov (2006). "Stable recovery of sparse overcomplete representations in the presence of noise." In: *Information Theory, IEEE Transactions on* 52.1, pp. 6–18.

Efron, Bradley et al. (2004). "Least angle regression." In: *The Annals of Statistics* 32.2, pp. 407–499.

Ejaz, Naveed, Masashi Hamada, and Jörn Diedrichsen (2015). "Hand use predicts the structure of representations in sensorimotor cortex." In: *Nature Neuroscience* 18.7, 1034âĂŞ–1040.

Faisal, Aldo et al. (2010). "The manipulative complexity of Lower Paleolithic stone toolmaking." In: *PLoS One* 5.11, e13718.

Fara, Salvatore et al. (2013). "Robust, ultra low-cost mmg system with Brain-machine-interface applications." In: *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*. IEEE, pp. 723–726.

Farina, Dario et al. (2014). "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges." In: *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 22.4, pp. 797–809.

Févotte, Cédric, Nancy Bertin, and Jean-Louis Durrieu (2009). "Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis." In: *Neural Computation* 21.3, pp. 793–830.

Field, David (1994). "What is the goal of sensory coding?" In: *Neural Computation* 6.4, pp. 559–601.

Flash, T. and N. Hogan (1985). "The coordination of arm movements: an experimentally confirmed mathematical model." In: *The Journal of Neuroscience* 5.7, pp. 1688–1703.

Flombaum, Jonathan I and Laurie R Santos (2005). "Rhesus monkeys attribute perceptions to others." In: *Current Biology* 15.5, pp. 447–452.

Fugl-Meyer, Axel R et al. (1974). "The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance." In: *Scandinavian journal of rehabilitation medicine* 7.1, pp. 13–31.

Gavriel, Constantinos and A Aldo Faisal (2014). "A Comparison of Day-Long Recording Stability and Muscle Force Prediction between BSN-Based Mechanomyography and Electromyography." In: *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*. IEEE, pp. 69–74.

Gavriel, Constantinos and AA Faisal (2013). "Wireless kinematic body sensor network for low-cost neurotechnology applications âĂIJin-the-wildâĂİ." In: *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*. IEEE, pp. 1279–1282.

Gavriel, Constantinos et al. (2015a). "Kinematic body sensor networks and behaviourmetrics for objective efficacy measurements in neurodegenerative disease drug trials." In: *IEEE Body Sensor Networks Conf.* IEEE.

Gavriel, Constantinos et al. (2015b). "Towards neurobehavioral biomarkers for longitudinal monitoring of neurodegeneration with wearable body sensor networks." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 348–351.

Gentner, R. and J. Classen (2006). "Modular organization of finger movements by the human central nervous system." In: *Neuron* 52.4, pp. 731–742.

Georgopoulos, Apostolos P, Andrew B Schwartz, and Ronald E Kettner (1986). "Neuronal population coding of movement direction." In: *Science* 233.4771, pp. 1416–1419.

Georgopoulos, Apostolos P et al. (1982). "On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex." In: *The Journal of Neuroscience* 2.11, pp. 1527–1537.

Gini, Corrado (1912). "Variabilità e mutabilità." In: *Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi* 1.

Giszter, S.F., F.A. Mussa-Ivaldi, and E. Bizzi (1993). "Convergent force fields organized in the frog's spinal cord." In: *The Journal of Neuroscience* 13.2, pp. 467–491.

Giuberti, Matteo et al. (2014). "Linking UPDRS scores and kinematic variables in the leg agility task of Parkinsonians." In: *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*. IEEE, pp. 115–120.

Goetz, Christopher G et al. (2007). "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan." In: *Movement Disorders* 22.1, pp. 41–47.

Graziano, Michael SA et al. (2002). "The cortical control of movement revisited." In: *Neuron* 36.3, pp. 349–362.

Grosse, Roger et al. (2007). "Shift-invariance sparse coding for audio classification." In: *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*.

Győrbíró, Norbert, Ákos Fábián, and Gergely Hományi (2009). "An activity recognition system for mobile phones." In: *Mobile Networks and Applications* 14.1, pp. 82–91.

Haber, David, Andreas AC Thomik, A Aldo Faisal, et al. (2014). "Unsupervised time series segmentation for high-dimensional body sensor network data streams." In: *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*. IEEE, pp. 121–126.

Hahnloser, Richard HR, Alexay A Kozhevnikov, and Michale S Fee (2002). "An ultra-sparse code underliesthe generation of neural sequences in a songbird." In: *Nature* 419.6902, pp. 65–70.

Hare, Brian, Josep Call, and Michael Tomasello (2001). "Do chimpanzees know what conspecifics know?" In: *Animal Behaviour* 61.1, pp. 139–151.

Hare, Brian et al. (2000). "Chimpanzees know what conspecifics do and do not see." In: *Animal Behaviour* 59.4, pp. 771–785.

Harris, Christopher M and Daniel M Wolpert (1998). "Signal-dependent noise determines motor planning." In: *Nature* 394.6695, pp. 780–784.

Hart, Corey B and Simon F Giszter (2010). "A neural basis for motor primitives in the spinal cord." In: *The Journal of Neuroscience* 30.4, pp. 1322–1336.

Hendrix, Claudia M, Carolyn R Mason, and Timothy J Ebner (2009). "Signaling of grasp dimension and grasp force in dorsal premotor cortex and primary motor cortex neurons during reach to grasp in the monkey." In: *Journal of Neurophysiology* 102.1, pp. 132–145.

Hepp-Reymond, M-C et al. (1994). "Force-related neuronal activity in two regions of the primate ventral premotor cortex." In: *Canadian journal of physiology and pharmacology* 72.5, pp. 571–579.

Honegger, Kyle S, Robert AA Campbell, and Glenn C Turner (2011). "Cellular-resolution population imaging reveals robust sparse coding in the Drosophila mushroom body." In: *The Journal of Neuroscience* 31.33, pp. 11772–11785.

Houweling, Arthur R and Michael Brecht (2007). "Behavioural report of single neuron stimulation in somatosensory cortex." In: *Nature* 450.7172.

Howard, Ian S et al. (2009). "Statistics of natural movements are reflected in motor errors." In: *Journal of Neurophysiology* 102.3, pp. 1902–1910.

Howard, Ian S et al. (2012). "Gone in 0.6 seconds: the encoding of motor memories depends on recent sensorimotor states." In: *The Journal of Neuroscience* 32.37, pp. 12756–12768.

Hoyer, Patrik O (2004). "Non-negative matrix factorization with sparseness constraints." In: *The Journal of Machine Learning Research* 5, pp. 1457–1469.

Hromádka, Tomáš, Michael R DeWeese, Anthony M Zador, et al. (2008). "Sparse representation of sounds in the unanesthetized auditory cortex." In: *PLoS Biology* 6.1, e16.

Ingram, James N and Daniel M Wolpert (2011). "Chapter 1 - Naturalistic approaches to sensorimotor control." In: *Enhancing Performance for Action and Perception Multisensory Integration, Neuroplasticity and Neuroprosthetics, Part I*. ed. by John F. Kalaska Andrea M. Green C. Elaine Chapman and Franco Lepore. Vol. 191. Progress in Brain Research. Elsevier, pp. 3 –29.

Ingram, James N et al. (2008). "The statistics of natural hand movements." In: *Experimental Brain Research* 188.2, pp. 223–236.

Ivanenko, Yuri P et al. (2002). "Two-thirds power law in human locomotion: role of ground contact forces." In: *Neuroreport* 13.9, pp. 1171–1174.

Ivanenko, Yuri P et al. (2007). "Modular control of limb movements during human locomotion." In: *The Journal of Neuroscience* 27.41, pp. 11149–11161.

Jarrassé, Nathanaël et al. (2014). "Analysis of hand synergies in healthy subjects during bimanual manipulation of various objects." In: *Journal of Neuroengineering and Rehabilitation* 11.1, p. 113.

Jeannerod, Marc et al. (1995). "Grasping objects: the cortical mechanisms of visuomotor transformation." In: *Trends in Neurosciences* 18.7, pp. 314–320.

Johansson, Gunnar (1973). "Visual perception of biological motion and a model for its analysis." In: *Attention, Perception, & Psychophysics* 14.2, pp. 201–211.

Johnson, Leif and Dana Henry Ballard (2014). "Classifying movements using efficient kinematic codes." In: *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Cognitive Science Society, pp. 2447–2452.

Jortner, Ron A, S Sarah Farivar, and Gilles Laurent (2007). "A simple connectivity scheme for sparse coding in an olfactory system." In: *The Journal of Neuroscience* 27.7, pp. 1659–1669.

Kakei, Shinji, Donna S Hoffman, and Peter L Strick (1999). "Muscle and movement representations in the primary motor cortex." In: *Science* 285.5436, pp. 2136–2139.

Kalman, Rudolph Emil (1960). "A new approach to linear filtering and prediction problems." In: *Journal of Fluids Engineering* 82.1, pp. 35–45.

Kelso, JA Scott (1984). "Phase transitions and critical behavior in human bimanual coordination." In: *American Journal of Physiology* 246.6 Pt 2, R1000–R1004.

Kilbreath, Sharon L and Robert C Heard (2005). "Frequency of hand use in healthy older persons." In: *Australian Journal of Physiotherapy* 51.2, pp. 119–122.

Kingstone, Alan, Daniel Smilek, and John D Eastwood (2008). "Cognitive ethology: A new approach for studying human cognition." In: *British Journal of Psychology* 99.3, pp. 317–340.

Kock, Christiaan PJ de and Bert Sakmann (2009). "Spiking in primary somatosensory cortex during natural whisking in awake head-restrained rats is cell-type specific." In: *Proceedings of the National Academy of Sciences* 106.38, pp. 16446–16450.

Konorski, Jerzy (1967). *Integrative Activity of the Brain*. University of Chicago Press.

Krakowski, Aaron I et al. (2011). "The neurophysiology of human biological motion processing: a high-density electrical mapping study." In: *NeuroImage* 56.1, pp. 373–383.

Krasoulis, Agamemnon, Sethu Vijayakumar, and Kianoush Nazarpour (2015). "Evaluation of regression Methods for the continuous decoding of finger movement from

surface EMG and accelerometry." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 631–634.

Kreutz-Delgado, Kenneth et al. (2003). "Dictionary learning algorithms for sparse representation." In: *Neural Computation* 15.2, pp. 349–396.

Krzanowski, WJ (1979). "Between-groups comparison of principal components." In: *Journal of the American Statistical Association* 74.367, pp. 703–707.

Kutch, Jason J et al. (2008). "Endpoint force fluctuations reveal flexible rather than synergistic patterns of muscle cooperation." In: *Journal of Neurophysiology* 100.5, pp. 2455–2471.

Lacquaniti, Francesco, Carlo Terzuolo, and Paolo Viviani (1983). "The law relating the kinematic and figural aspects of drawing movements." In: *Acta Psychologica* 54.1, pp. 115–130.

Land, Michael and Julia Horwood (1995). "Which parts of the road guide steering?" In: *Nature* 377.6547, pp. 339–340.

Land, Michael F and David N Lee (1994). "Where do we look when we steer." In: *Nature*.

Laughlin, Simon B (2001). "Energy as a constraint on the coding and processing of sensory information." In: *Current Opinion in Neurobiology* 11.4, pp. 475–480.

Le Pennec, Erwan and Stéphane Mallat (2005). "Sparse geometric image representations with bandelets." In: *Image Processing, IEEE Transactions on* 14.4, pp. 423–438.

Lesage, Sylvain et al. (2005). "Learning unions of orthonormal bases with thresholded singular value decomposition." In: *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*. Vol. 5. IEEE, pp. 293–296.

Levy, William B and Robert A Baxter (1996). "Energy efficient neural codes." In: *Neural Computation* 8.3, pp. 531–543.

Lewicki, Michael S (2002). "Efficient coding of natural sounds." In: *Nature Neuroscience* 5.4, pp. 356–363.

Lewicki, Michael S and Terrence Sejnowski (2000). "Learning overcomplete representations." In: *Neural Computation* 12.2, pp. 337–365.

Li, W., E. Todorov, and X. Pan (2004). "Hierarchical optimal control of redundant biomechanical systems." In: *Proceedings of the Annual International IEEE EMBS Conference*. Vol. 2. IEEE, pp. 4618–4621.

Mairal, Julien, Francis Bach, and Jean Ponce (2012). "Task-driven dictionary learning." In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 34.4, pp. 791–804.

Mairal, Julien, Michael Elad, and Guillermo Sapiro (2008). "Sparse representation for color image restoration." In: *Image Processing, IEEE Transactions on* 17.1, pp. 53–69.

Mairal, Julien et al. (2008). "Discriminative learned dictionaries for local image analysis." In: *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, pp. 1–8.

Mairal, Julien et al. (2009a). "Non-local sparse models for image restoration." In: *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, pp. 2272–2279.

Mairal, Julien et al. (2009b). "Supervised dictionary learning." In: *Advances in Neural Information Processing Systems*, pp. 1033–1040.

Mairal, Julien et al. (2010). "Online learning for matrix factorization and sparse coding." In: *The Journal of Machine Learning Research* 11, pp. 19–60.

Mallat, Stéphane G and Zhifeng Zhang (1993). "Matching pursuits with time-frequency dictionaries." In: *Signal Processing, IEEE Transactions on* 41.12, pp. 3397–3415.

Marzke, Mary W and Robert F Marzke (2000). "Evolution of the human hand: approaches to acquiring, analysing and interpreting the anatomical evidence." In: *Journal of anatomy* 197.1, pp. 121–140.

Mason, CR, JE Gomez, and TJ Ebner (2001). "Hand synergies during reach-to-grasp." In: *Journal of Neurophysiology* 86.6, pp. 2896–2910.

Massey, Joe T et al. (1992). "Three-dimensional drawings in isometric conditions: relation between geometry and kinematics." In: *Experimental Brain Research* 88.3, pp. 685–690.

Miller, Ruth and Wendy Brown (2004). "Steps and sitting in a working population." In: *International Journal of Behavioral Medicine* 11.4, pp. 219–224.

Milner, Theodore E (1992). "A model for the generation of movements requiring endpoint precision." In: *Neuroscience* 49.2, pp. 487–496.

Mollazadeh, Mohsen et al. (2014). "Principal components of hand kinematics and neurophysiological signals in motor cortex during reach to grasp movements." In: *Journal of Neurophysiology* 112.8, pp. 1857–1870.

Moran, Daniel W and Andrew B Schwartz (1999). "Motor cortical representation of speed and direction during reaching." In: *Journal of Neurophysiology* 82.5, pp. 2676–2692.

Morasso, P and FA Mussa Ivaldi (1982). "Trajectory formation and handwriting: a computational model." In: *Biological Cybernetics* 45.2, pp. 131–142.

Morasso, Pietro (1981). "Spatial control of arm movements." In: *Experimental Brain Research* 42.2, pp. 223–227.

Morasso, Pietro (1983). "Three dimensional arm trajectories." In: *Biological Cybernetics* 48.3, pp. 187–194.

Mukherjee, Sayan and Chandra Sekhar Seelamantula (2014). "A split-and-merge dictionary learning algorithm for sparse representation: Application to image denoising." In: *Digital Signal Processing (DSP), 2014 19th International Conference on*. IEEE, pp. 310–315.

Murata, Akira et al. (1996). "Parietal neurons related to memory-guided hand manipulation." In: *Journal of Neurophysiology* 75.5, pp. 2180–2186.

Murata, Akira et al. (2000). "Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP." in: *Journal of Neurophysiology* 83.5, pp. 2580–2601.

Napier, John (1962). "Fossil hand bones from Olduvai Gorge." In: *Nature* 196, pp. 409–411.

O'Connor, Daniel H et al. (2010). "Neural activity in barrel cortex underlying vibrissa-based object localization in mice." In: *Neuron* 67.6, pp. 1048–1061.

Okada, Masato (1996). "Notions of associative memory and sparse coding." In: *Neural Networks* 9.8, pp. 1429–1458.

Olree, Kenneth S and Christopher L Vaughan (1995). "Fundamental patterns of bilateral muscle activity in human locomotion." In: *Biological Cybernetics* 73.5, pp. 409–414.

Olshausen, Bruno A and David J Field (1996). "Natural image statistics and efficient coding*." In: *Network: Computation in Neural Systems* 7.2, pp. 333–339.

Olshausen, Bruno A and David J Field (1997). "Sparse coding with an overcomplete basis set: A strategy employed by V1?" In: *Vision Research* 37.23, pp. 3311–3325.

Olshausen, Bruno A et al. (1996). "Emergence of simple-cell receptive field properties by learning a sparse code for natural images." In: *Nature* 381.6583, pp. 607–609.

Overduin, Simon A et al. (2015). "Representation of Muscle Synergies in the Primate Brain." In: *The Journal of Neuroscience* 35.37, pp. 12615–12624.

Palm, Günther (2013). "Neural associative memories and sparse coding." In: *Neural Networks* 37, pp. 165–171.

Paninski, Liam et al. (2004). "Spatiotemporal tuning of motor cortical neurons for hand position and velocity." In: *Journal of neurophysiology* 91.1, pp. 515–532.

Pati, Yagyensh Chandra, Ramin Rezaiifar, and PS Krishnaprasad (1993). "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet

decomposition." In: *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, pp. 40–44.

Perez-Orive, Javier et al. (2002). "Oscillations and sparsening of odor representations in the mushroom body." In: *Science* 297.5580, pp. 359–365.

Peyré, Gabriel (2009). "Sparse modeling of textures." In: *Journal of Mathematical Imaging and Vision* 34.1, pp. 17–31.

Poo, Cindy and Jeffry S Isaacson (2009). "Odor representations in olfactory cortex:"sparse" coding, global inhibition, and oscillations." In: *Neuron* 62.6, pp. 850–861.

Quiroga, R Quian et al. (2005). "Invariant visual representation by single neurons in the human brain." In: *Nature* 435.7045, pp. 1102–1107.

Quiroga, R Quian et al. (2008). "Sparse but not âĂŸgrandmother-cellâĂŹcoding in the medial temporal lobe." In: *Trends in Cognitive Sciences* 12.3, pp. 87–91.

Raina, Rajat et al. (2007). "Self-taught learning: transfer learning from unlabeled data." In: *Proceedings of the 24th International Conference on Machine learning*. ACM, pp. 759–766.

Raos, Vassilis et al. (2004). "Functional properties of grasping-related neurons in the dorsal premotor area F2 of the macaque monkey." In: *Journal of Neurophysiology* 92.4, pp. 1990–2002.

Raos, Vassilis et al. (2006). "Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey." In: *Journal of Neurophysiology* 95.2, pp. 709–729.

Rinberg, Dmitry, Alex Koulakov, and Alan Gelperin (2006). "Sparse odor coding in awake behaving mice." In: *The Journal of Neuroscience* 26.34, pp. 8857–8865.

Rizzolatti, Giscomo and Maurizio Gentilucci (1988). "Motor and visual-motor functions of the premotor cortex." In: *Neurobiology of Neocortex* 42, pp. 269–284.

Rohrer, Brandon and Neville Hogan (2003). "Avoiding spurious submovement decompositions: a globally optimal algorithm." In: *Biological Cybernetics* 89.3, pp. 190–199.

Rohrer, Brandon and Neville Hogan (2006). "Avoiding spurious submovement decompositions II: a scattershot algorithm." In: *Biological Cybernetics* 94.5, pp. 409–414.

Rolls, Edmund T and Alessandro Treves (1990). "The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain." In: *Network: Computation in Neural Systems* 1.4, pp. 407–421.

Rolls, Edmund T, Simon Yaxley, and Zenon J Sienkiewicz (1990). "Gustatory responses of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey." In: *Journal of Neurophysiology* 64.4, pp. 1055–1066.

Santello, M., M. Flanders, and J.F. Soechting (1998). "Postural hand synergies for tool use." In: *The Journal of Neuroscience* 18.23, pp. 10105–10115.

Santello, Marco, Gabriel Baud-Bovy, and Henrik Jörntell (2013). "Neural bases of hand synergies." In: *Frontiers in Computational Neuroscience* 7.

Santello, Marco, Martha Flanders, and John F Soechting (2002). "Patterns of hand motion during grasping and the influence of sensory guidance." In: *The Journal of Neuroscience* 22.4, pp. 1426–1435.

Schaal, Stefan and Dagmar Sternad (2001). "Origins and violations of the 2/3 power law in rhythmic three-dimensional arm movements." In: *Experimental Brain Research* 136.1, pp. 60–72.

Schaffelhofer, Stefan, Andres Agudelo-Toro, and Hansjörg Scherberger (2015). "Decoding a wide range of hand configurations from macaque Motor, Premotor, and Parietal Cortices." In: *The Journal of Neuroscience* 35.3, pp. 1068–1081.

Schieber, Marc H and Lyndon S Hibbard (1993). "How somatotopic is the motor cortex hand area?" In: *Science* 261.5120, pp. 489–492.

Schieber, Marc H and Marco Santello (2004). "Hand function: peripheral and central constraints on performance." In: *Journal of Applied Physiology* 96.6, pp. 2293–2300.

Schlich, Robert and Kay W Axhausen (2003). "Habitual travel behaviour: Evidence from a six-week travel diary." In: *Transportation* 30.1, pp. 13–36.

Schmitz, Anne et al. (2015). "The measurement of in vivo joint angles during a squat using a single camera markerless motion capture system as compared to a marker based system." In: *Gait & Posture* 41.2, pp. 694–698.

Schmitz-Hübsch, Tanja et al. (2008). "SCA Functional Index A useful compound performance measure for spinocerebellar ataxia." In: *Neurology* 71.7, pp. 486–492.

Scott, Stephen H (2004). "Optimal feedback control and the neural basis of volitional motor control." In: *Nature Reviews Neuroscience* 5.7, pp. 532–546.

Shadmehr, Reza and Ferdinando A Mussa-Ivaldi (1994). "Adaptive representation of dynamics during learning of a motor task." In: *The Journal of Neuroscience* 14.5, pp. 3208–3224.

Sherrington, Charles Scott (1910). "Flexion-reflex of the limb, crossed extension-reflex, and reflex stepping and standing." In: *The Journal of Physiology* 40.1–2, p. 28.

Shoemake, Ken (1985). "Animating rotation with quaternion curves." In: *ACM SIGGRAPH Computer graphics*. Vol. 19. 3. ACM, pp. 245–254.

Simoncelli, Eero P et al. (1992). "Shiftable multiscale transforms." In: *Information Theory, IEEE Transactions on* 38.2, pp. 587–607.

Sternad, Dagmar and Stefan Schaal (1999). "Segmentation of endpoint trajectories does not imply segmented control." In: *Experimental Brain Research* 124.1, pp. 118–136.

Swain, Michael J and Dana H Ballard (1991). "Color indexing." In: *International Journal of Computer Vision* 7.1, pp. 11–32.

Tarkka, IM and Mark Hallett (1990). "Cortical topography of premotor and motor potentials preceding self-paced, voluntary movement of dominant and non-dominant hands." In: *Electroencephalography and Clinical Neurophysiology* 75.1, pp. 36–43.

Teichmann, Daniel et al. (2014). "A bendable and wearable cardiorespiratory monitoring device fusing two noncontact sensor principles." In: *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*. IEEE, pp. 58–63.

Thakur, Pramodsingh H, Amy J Bastian, and Steven S Hsiao (2008). "Multidigit movement synergies of the human hand in an unconstrained haptic exploration task." In: *The Journal of Neuroscience* 28.6, pp. 1271–1281.

Thomik, Andreas AC, Sonja Fenske, and A Aldo Faisal (2015). "Towards sparse coding of natural movements for neuroprosthetics and brain-machine interfaces." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 938–941.

Thomik, Andreas AC, David Haber, A Aldo Faisal, et al. (2013). "Real-time movement prediction for improved control of neuroprosthetic devices." In: *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*. IEEE, pp. 625–628.

Thompson, LT and PJ Best (1989). "Place cells and silent cells in the hippocampus of freely-behaving rats." In: *The Journal of Neuroscience* 9.7, pp. 2382–2390.

Tibshirani, Robert (1996). "Regression shrinkage and selection via the lasso." In: *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288.

Todorov, E. and M.I. Jordan (2002). "Optimal feedback control as a theory of motor coordination." In: *Nature Neuroscience* 5.11, pp. 1226–1235.

Todorov, E., W. Li, and X. Pan (2005). "From task parameters to motor synergies: A hierarchical framework for approximately optimal control of redundant manipulators." In: *Journal of Robotic Systems* 22.11, pp. 691–710.

Todorov, Emanuel (2008). "General duality between optimal control and estimation." In: *Decision and Control, 2008. 47th IEEE Conference on*. IEEE, pp. 4286–4292.

Todorov, Emanuel and Zoubin Ghahramani (2004). "Analysis of the synergies underlying complex hand manipulation." In: *Engineering in Medicine and Biology Society,*

2004. IEMBS'04. 26th Annual International Conference of the IEEE. vol. 2. IEEE, pp. 4637–4640.

Torre, Fernando De la et al. (2009). "Guide to the carnegie mellon university multimodal activity (cmu-mmac) database." In: *Robotics Institute, Carnegie Mellon University*.

Torres-Oviedo, Gelsy, Jane M Macpherson, and Lena H Ting (2006). "Muscle synergy organization is robust across a variety of postural perturbations." In: *Journal of Neurophysiology* 96.3, pp. 1530–1546.

Torres-Oviedo, Gelsy and Lena H Ting (2007). "Muscle synergies characterizing human postural responses." In: *Journal of Neurophysiology* 98.4, pp. 2144–2156.

Tresch, Matthew C and Anthony Jarc (2009). "The case for and against muscle synergies." In: *Current Opinion in Neurobiology* 19.6, pp. 601–607.

Uno, Yoji, Mitsuo Kawato, and Rika Suzuki (1989). "Formation and control of optimal trajectory in human multijoint arm movement." In: *Biological Cybernetics* 61.2, pp. 89–101.

Valero-Cuevas, Francisco J, Madhusudhan Venkadesan, and Emanuel Todorov (2009). "Structured variability of muscle activations supports the minimal intervention principle of motor control." In: *Journal of Neurophysiology* 102.1, pp. 59–68.

Vargas-Irwin, Carlos E et al. (2010). "Decoding complete reach and grasp actions from local primary motor cortex populations." In: *The Journal of Neuroscience* 30.29, pp. 9659–9669.

Vella, I. (2012). "Towards the Human Ethome: Human Kinematics Study in Daily Life Environments." MA thesis. Imperial College London.

Vidal, Rene, Yi Ma, and Shankar Sastry (2005). "Generalized principal component analysis (GPCA)." in: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27.12, pp. 1945–1959.

Vinje, William E and Jack L Gallant (2000). "Sparse coding and decorrelation in primary visual cortex during natural vision." In: *Science* 287.5456, pp. 1273–1276.

Vinje, William E and Jack L Gallant (2002). "Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1." In: *The Journal of Neuroscience* 22.7, pp. 2904–2915.

Viviani, Paolo and Marco Cenzato (1985). "Segmentation and coupling in complex movements." In: *Journal of Experimental Psychology: Human Perception and Performance* 11.6, p. 828.

Volkmann, J et al. (1998). "Handedness and asymmetry of hand representation in human motor cortex." In: *Journal of Neurophysiology* 79.4, pp. 2149–2154.

Vonderschen, Katrin and Maurice J Chacron (2011). "Sparse and dense coding of natural stimuli by distinct midbrain neuron subpopulations in weakly electric fish." In: *Journal of Neurophysiology* 106.6, pp. 3102–3118.

Šarić, Marin (2011). *LibHand: A Library for Hand Articulation*. Version 0.9.

Wann, John, Ian Nimmo-Smith, and Alan M Wing (1988). "Relation between velocity and curvature in movement: equivalence and divergence between a power law and a minimum-jerk model." In: *Journal of Experimental Psychology: Human Perception and Performance* 14.4, p. 622.

Weiss, Erica J and Martha Flanders (2004). "Muscular and postural synergies of the human hand." In: *Journal of Neurophysiology* 92.1, pp. 523–535.

Wohlberg, Brendt (2003). "Noise sensitivity of sparse signal representations: Reconstruction error bounds for the inverse problem." In: *IEEE Transactions on Signal Processing* 51.12, pp. 3053–3060.

Wu, Ge et al. (2002). "ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motionâĂŤpart I: ankle, hip, and spine." In: *Journal of Biomechanics* 35.4, pp. 543–548.

Wu, Ge et al. (2005). "ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motionâĂŤPart II: shoulder, elbow, wrist and hand." In: *Journal of Biomechanics* 38.5, pp. 981–992.

Xiloyannis, Michele et al. (2015a). "Dynamic forward prediction for prosthetic hand control by integration of EMG, MMG and kinematic signals." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 611–614.

Xiloyannis, Michele et al. (2015b). "Gaussian Process Regression for accurate prediction of prosthetic limb movements from the natural kinematics of intact limbs." In: *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*. IEEE, pp. 659–662.

Yang, Jianchao et al. (2009). "Linear spatial pyramid matching using sparse coding for image classification." In: *Computer Vision and Pattern Recognition, 2009. IEEE Conference on*. IEEE, pp. 1794–1801.