Identification of Ankle Joint Stiffness Using Subspace Methods

Yong Zhao

Department of Biomedical Engineering
McGill University
Montreal, Quebec

August, 2009

A thesis submitted to McGill University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy.

© Yong Zhao, 2009
Abstract

Studying joint stiffness against a compliant load is a difficult problem because the intrinsic and reflex torques cannot be measured separately experimentally. Moreover, the joint stiffness is operated within a closed loop because the ankle torque is fed back through the load to change the ankle position. In this thesis, a state space model for ankle joint stiffness is developed. Then, a discrete-time, subspace-based method is used to estimate this state space model for overall stiffness. By using appropriate instrumental variables, the subspace method can estimate the state space model for joint stiffness in both open-loop and in closed-loop conditions. This thesis also presents a subspace method to identify state space models for biomedical systems with short transients or systems with time-varying behaviors, from ensembles of short transients. The simulation and experimental results demonstrate that those algorithms provide accurate estimates under their respective conditions.
Résumé

L'étude de la rigidité articulaire en réponse à une charge est un problème difficile car les couples réflexes et intrinsèques ne peuvent pas être mesurés séparément expérimentalement. En outre, la rigidité articulaire opère en boucle fermée car le couple de la cheville est réinjectée à travers la charge pour modifier la position de la cheville. Dans cette thèse, un modèle d'espace d'état pour la rigidité articulaire de la cheville est développé. Une méthode sous-espace à temps discret est ensuite utilisée pour estimer ce modèle d'espace d'état pour la rigidité globale. En considérant les variables instrumentales appropriées, la méthode sous-espace permet d'estimer le modèle espace d'état pour la rigidité articulaire en boucles ouverte et fermée. Cette thèse présente également une méthode sous-espace pour identifier les modèles d’espace d'état pour les systèmes biomédicaux ou les systèmes variant dans le temps caractérisés par des phénomènes transitoires de courte durée. Les simulations et les résultats expérimentaux démontrent que ces algorithmes fournissent des estimations précises en fonction de leurs conditions propres.
Acknowledgements

I would like to express my great appreciation to my supervisor, Dr. Robert Kearney, for his support and comments, not only as a thesis supervisor but also as a patient English teacher.

I would also like to thank Dr. David Westwick for his guidance and his collaboration during the years of my graduate study. I would also like to thank Daniel and Tanya for providing me the experimental data for validation purposes. My thanks also go to all my colleagues in the REKLAB and Dr. Ross Wagner as well as Ms Pina Sorrini, for their great attitude and help which made my time at McGill very enjoyable. I would also like to thank Dr. William H.A. Beaudot and Dr. Sunil Kukreja, for many late night discussions at the Thomson House, from the topics ranging from support vector machine to Myanmar Python.

I would like to thank my parents, Hunjian Zhao and Huizhen Ding, whom I owe a debt that I could never possibly pay. I would like to thank my wife, Hui, for not laughing at me too loudly for my switching between academia and industry. I would like to thank my daughter, Jiayi, who lights up my life, day and night, and makes me laugh and cry, at the same time.

This work was supported by grants from the Natural Science and Engineering Research Council of Canada (NSERC), and Canadian Institutes of Health Research (CIHR).
Contributions of Authors

The work done in Chapter 4 is almost completely my own. I created the algorithm, with some suggestions made by Dr. Kearney and Dr. Westwick. I carried out all the simulations studies that validated the method. I wrote the manuscript, with feedback from Dr. Kearney and Dr. Westwick.

The work in Chapters 5 and 6 is mostly my own. Dr. Kearney provided me with comments and suggestions to complete the work. I wrote the manuscript, implemented and validated the algorithms, with guidance from Dr. Kearney.
# Table of Contents

Abstract ......................................................................................................................... i  
Résumé ........................................................................................................................... ii  
Acknowledgements ......................................................................................................... iii  
Contributions of Authors ................................................................................................. iv  
Table of Contents ........................................................................................................... v  
Table of Figures ............................................................................................................. x  
Table of Tables ............................................................................................................. xii

1 Introduction ................................................................................................................ 1  
1.1 Joint Stiffness ............................................................................................................. 1  
1.2 Thesis Outline ............................................................................................................ 2

2 Background ................................................................................................................. 4  
2.1 Ankle Joint ................................................................................................................. 4  
2.2 Peripheral Neuromuscular System ............................................................................ 7  
2.2.1 Skeletal Muscle Structure .................................................................................... 7  
2.2.2 Neurons and Motor Units ..................................................................................... 14  
2.2.3 Control of Muscle Force ...................................................................................... 14  
2.2.4 Electromyography ............................................................................................... 18  
2.2.5 Peripheral Sensory Receptors .............................................................................. 18  
2.2.5.1 Muscle Spindles ............................................................................................ 20  
2.2.5.2 Golgi Tendon Organs ....................................................................................... 22  
2.2.6 Reflexes ................................................................................................................ 22  
2.2.6.1 Stretch reflex .................................................................................................. 22  
2.2.6.2 Inverse Stretch Reflex ................................................................................... 25  
2.2.6.3 Long loop Reflex ............................................................................................ 25  
2.2.6.4 Modulation of Reflex ...................................................................................... 25  
2.3 Joint Stiffness ............................................................................................................. 26  
2.3.1 Parallel Cascade Model ....................................................................................... 27  
2.3.2 EMG ..................................................................................................................... 31  
2.3.3 Separation of Intrinsic and Reflex Stiffnesses ..................................................... 31  
2.3.4 System Identification .......................................................................................... 32


2.3.5 Open-loop experiment VS Closed-loop experiment .................................................. 33
2.4 Thesis Rationale ................................................................................................................. 37

References .................................................................................................................................. 38

3 Subspace-based System Identification Methods ................................................................. 41
3.1 Introduction .......................................................................................................................... 41
3.2 Background ........................................................................................................................ 42
3.2.1 State Space Model ........................................................................................................ 42
3.2.2 Matrix Definitions ......................................................................................................... 46
3.2.2.1 Hankel Matrix ......................................................................................................... 46
3.2.2.2 Observability Matrix ............................................................................................... 47
3.2.2.3 Toeplitz Matrix of Impulse Response ...................................................................... 48
3.2.3 Orthogonal Projections ................................................................................................. 48
3.2.4 Singular Value Decomposition ..................................................................................... 51
3.2.5 Ergodic Algebraic Framework ..................................................................................... 52
3.2.6 Persistently Exciting Signals ......................................................................................... 53
3.3 The Ordinary MOESP Method ......................................................................................... 54
3.3.1 Estimate the Order of System and the Matrices A & C ............................................. 56
3.3.2 Estimate the Matrices B & D ...................................................................................... 62
3.3.3 Validation ..................................................................................................................... 63
3.4 Extended MOESP Algorithms ....................................................................................... 66

References .................................................................................................................................. 74

4 Identification of Human Ankle Joint Stiffness Using Subspace Methods

Part 1. Open-loop System Identification ................................................................................. 76
4.1 Abstract .............................................................................................................................. 78
4.2 Introduction ....................................................................................................................... 79
4.3 Subspace Method ............................................................................................................ 84
4.4 Estimation of Intrinsic and Reflex Torque ...................................................................... 88
4.4.1 State Space Model for Intrinsic Stiffness ..................................................................... 88
4.4.2 State Space Model for Reflex Stiffness ....................................................................... 89
4.4.3 State Space Model for Overall Stiffness ..................................................................... 93
4.4.4 Implementation ........................................................................................................... 95
4.4.5 Continuous Time Models ........................................................................................... 96
4.5 Simulation Studies ................................................................. 96
  4.5.1 Simulation Model................................................................. 97
  4.5.2 Identification Evaluation..................................................... 101
  4.5.3 Results.............................................................................. 102
    4.5.3.1 Additive Noise.............................................................. 102
    4.5.3.2 Reflex Gain................................................................. 104
    4.5.3.3 Data Length............................................................... 106
4.6 Experimental Studies ............................................................... 108
  4.6.1 Methods......................................................................... 108
  4.6.2 Results.......................................................................... 109
4.7 Discussion and Conclusion .......................................................... 114
  4.7.1 Constructed Inputs ............................................................ 114
  4.7.2 Comparison of Subspace Method and Parallel Cascade Method .... 115
  4.7.3 Application to Closed-loop Identification............................. 117
  4.7.4 Conclusion...................................................................... 118
Acknowledgements ........................................................................ 118
References....................................................................................... 120

5 Identification of Human Ankle Dynamics Using Subspace Based Methods.
Part 2: Closed-Loop System Identification........................................... 124
5.1 Abstract............................................................................. 126
5.2 Introduction....................................................................... 127
5.3 Identification of Ankle Stiffness In Closed-Loop.......................... 132
  5.3.1 Parametric Models for Ankle Stiffness................................. 132
  5.3.2 State space Model for Joint Stiffness..................................... 134
  5.3.3 Subspace Methods for Closed-Loop System Identification........ 136
  5.3.4 Implementation.............................................................. 140
5.4 Simulation Studies ................................................................. 141
  5.4.1 Simulation Methods............................................................ 141
  5.4.2 Identification and Validation............................................... 144
  5.4.3 Effects of Noise............................................................... 145
  5.4.4 Effects of Reflex Gain ..................................................... 147
5.5 Experimental Studies ..................................................................................... 149
5.6 Discussion and Conclusion ........................................................................... 153
  5.6.1 Parallel Cascade Method ........................................................................... 153
  5.6.2 Subspace Method ..................................................................................... 154
  5.6.3 Instrumental Variables ............................................................................. 155
  5.6.4 Conclusion ............................................................................................... 156

Acknowledgements ............................................................................................ 157

References: ............................................................................................................ 158

6 Identification of Biomedical Systems from Short Transients Using Subspace Methods ................................................................. 161
  6.1 Abstract ....................................................................................................... 163
  6.2 Introduction .................................................................................................. 164
  6.3 Identification of Systems from Short Transients ........................................ 166
    6.3.1 Identification Method ............................................................................. 166
    6.3.2 Simulation Study ..................................................................................... 172
  6.4 Identification of Time-Varying Systems ...................................................... 182
    6.4.1 Identification Method ............................................................................. 182
    6.4.2 Simulation Studies ................................................................................... 182
  6.5 Discussion and Conclusion ........................................................................... 187
    6.5.1 Effect of Noise ....................................................................................... 187
    6.5.2 Initial States ............................................................................................ 189
    6.5.3 Data Segment Length and Number of Trials ......................................... 190
    6.5.4 Different Data Segment Lengths ............................................................. 190
    6.5.5 Conclusion ............................................................................................... 191

Acknowledgements ............................................................................................ 192

Reference .............................................................................................................. 193

7 Discussion & Conclusions ................................................................................. 195
  7.1 Statement of Original Contributions ............................................................ 197
    7.1.1 Identification of Dynamic Ankle Stiffness in Open-Loop ....................... 198
    7.1.2 Identification of Ankle Joint Stiffness in Closed-loop ......................... 198
    7.1.3 Identification of Ankle Joint Stiffness from Short Transients ............... 200
  7.2 Future Work ................................................................................................. 200
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2.1</td>
<td>Basis Function</td>
<td>201</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Systems with Time-Varying Behavior</td>
<td>202</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Experimental Application</td>
<td>203</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Reference Input As Instrumental Variable for Closed-loop System</td>
<td>204</td>
</tr>
</tbody>
</table>

References: ............................................................................................................. 206

Appendix .......................................................................................................................... 208

A. Ethics Certificate ........................................................................................................ 208
B. Waivers From Co-Authors ........................................................................................... 209
# Table of Figures

Figure 2.1: Posterior view (left) and medial view (right) of ankle joint bones, tendons and muscles .......................... 5

Figure 2.2: Definitions of plantarflexion (pushing the foot down) and dorsiflexion (pulling the foot up) ................................................................. 6

Figure 2.3: Skeletal Muscle Structure .............................................. 9

Figure 2.4: Skeletal Muscle Structure Muscle length-tension relationship ................................................... 11

Figure 2.5: Muscle stretch velocity versus tension ................................................................. 13

Figure 2.6: Motor Unit .......................................................................................... 14

Figure 2.7: Effect of AP frequency on muscle force ................................................... 16

Figure 2.8: Successive recruitment ........................................................................ 17

Figure 2.9: Skeletal muscle and muscle sensory receptors .................................................. 19

Figure 2.10: Sensitivity of primary and secondary endings of muscle spindles ..................................................... 21

Figure 2.11: Pathways involved in the stretch reflex .................................................................. 24

Figure 2.12: Peripheral neuromuscular system contributors to joint dynamic stiffness ........ 29

Figure 2.13: Parallel cascade model for dynamic ankle stiffness ................................................... 30

Figure 2.14: Block diagram of the open-loop identification problem ........................................ 34

Figure 2.15: Simplified block diagram of the closed-loop identification problem .................. 36

Figure 3.1: Block diagram of a state space model ................................................................ 45

Figure 3.2: Input and output data from the simulation of a fifth-order system with no noise ..... 55

Figure 3.3: System order estimation for a simulated, fifth order system with no noise .......... 59

Figure 3.4: Comparison of estimated and true poles .............................................................. 61

Figure 3.5: Comparison of outputs from the simulated and estimated systems ...................... 65

Figure 3.6: Estimated systems order for a fifth order system with 20 dB measurement noise ... 68

Figure 3.7: Comparison of noise-free output from the simulated and system estimated with 20 dB measurement noise ................................................................. 69

Figure 4.1: Parallel cascade model for dynamic ankle stiffness ............................................ 80

Figure 4.2: Transformation of a SISO Hammerstein system to a MISO linear system .......... 90

Figure 4.3: Matlab Simulink model for ankle stiffness .......................................................... 98

Figure 4.4: Experimental and Simulated Torques .................................................................... 100

Figure 4.5: Estimation accuracy as a function of SNR ....................................................... 103

Figure 4.6: Estimation accuracy as a function of reflex gain ................................................ 105
Figure 4.7: Identification qualities as a function of data length. .................................................. 107
Figure 4.8: Estimated non-parametric models for intrinsic and reflex stiffness for subject from the subspace method. ........................................................................................................... 112
Figure 4.9: Estimated nonlinear output and reflex torque for a subject ........................................ 113
Figure 5.1: Parallel cascade structure of ankle dynamics. ............................................................. 128
Figure 5.2: Block diagram of the open-loop identification problem ............................................. 130
Figure 5.3: Simplified block diagram of the closed-loop identification problem ......................... 131
Figure 5.4: Matlab Simulink model of ankle dynamics................................................................... 143
Figure 5.5: Effects of SNR on %VAF between noise-free simulated and estimates ................... 146
Figure 5.6: Effects of Reflex Gain on %VAF between noise-free simulated and estimated ......... 148
Figure 5.7: Stiffness estimates for human subject ........................................................................... 152
Figure 6.1: Parallel-cascade model for dynamic ankle stiffness .................................................... 173
Figure 6.2: Matlab Simulink model for ankle stiffness ................................................................. 176
Figure 6.3: Simulated joint stiffness data ....................................................................................... 178
Figure 6.4: Simulated and estimated torques from the simulation study for system with short transients ............................................................. 179
Figure 6.5: VAFs between simulated and predicted torques as a function of the number of realizations ................................................................................................................................. 181
Figure 6.6: Matlab Simulink model of a second-order, time-varying system .............................. 184
Figure 6.7: Simulated and estimated outputs for one simulation trial of the second-order, time-varying system ...................................................................................................................... 185
Figure 6.8: Simulated and estimated gain as a function time for the second-order, time-varying simulation ........................................................................................................................... 186
Table of Tables

Table 4.1. Parameter values used to simulate ankle stiffness .................................................. 98
Table 4.2 Estimates from the subspace method and the parallel cascade method for experimental data................................................................. 110
Table 5.1. Intrinsic and reflex parameter values of simulated ankle dynamics ...................... 142
Table 5.2 Estimates from the subspace method and the parallel cascade method for experimental data................................................................. 151
1 Introduction

For humans, motor control provides the most fundamental means of interacting with the environment. The performance of any movement requires the control by a collection of rigid skeletal elements actuated by muscles acting about joints. The mechanisms that control human movement can be divided into central and peripheral. Central mechanisms involve the portion of the nervous system above the spinal cord. Peripheral mechanisms translate commands originating from central mechanisms into muscle forces.

1.1 Joint Stiffness

Dynamic joint stiffness plays an important role in the peripheral system. It defines the mechanic response to external perturbation and determines the force needed to move the body. In engineering terms, joint dynamics refers to the dynamic relation between the angular position of a joint and the torque acting about it. The study of joint dynamics requires a set of mathematical models to describe the dynamic relationship between the position and the torque. Such models may be formulated as differential (continuous time) or difference (discrete time) equations.

There are two ways to develop the mathematical models for joint dynamics. Physicists will be interested in constructing a model to carefully explain all the underlying mechanisms of the observed phenomena. Physical laws, such as Newton’s laws of motion, are typically used to construct such mathematical models. However, those methods require a complete knowledge of the system that is normally not available
in physiology.

An alternative approach is to use system identification to model the joint. System identification is a technique that constructs mathematical models analytically from measured input-output data. Compared with physical modeling, system identification provides a relatively easy way to obtain a model, and by reproducing the system’s response to relevant inputs, the model itself often provides insight into the underlying physiological behaviour.

1.2 Thesis Outline

This thesis focuses on developing subspace based system identification methods with specific applications to ankle joint dynamics. This thesis is developed as follows.

Chapter 2 reviews anatomical and physiological background including motor neurons, skeletal muscle physiology, the peripheral sensory systems and stretch reflexes. A review of previous identification methods for dynamic joint stiffness is also included.

Chapter 3 presents a theoretical review of an established, subspace method that is widely used for identification of linear systems. Subsequent chapters of the thesis will develop subspace-based method to estimate the ankle dynamics in open loop and in closed loop based on this analysis.

Chapter 4 presents a novel subspace method to estimate ankle dynamics in open loop. A state space model is developed to model the dynamic relationship between the position of the ankle and the torque acting about it. It also investigates how the method’s performance depends on measurement noise, the relative magnitude of the nonlinear
reflex contribution, and the length of the data set.

Chapter 5 describes an extension of the subspace method to estimate ankle dynamics from data acquired in closed loop. Simulation results demonstrate that the open loop, parallel cascade method provides biased estimates for the data acquired in closed loop. In contrast, the extended subspace method provides accurate estimates of ankle dynamics from closed loop data.

Chapter 6 formulates a version of the subspace method that estimates system dynamics from an ensemble of input-output measurements. It first focuses on solving the problem of estimating the system dynamics for biomedical systems for which only short transients can be measured. It then describes how the same ensemble approach can be applied to support the identification of time-varying systems.

Finally, Chapter 7 summarizes the contributions of the research described in the thesis and provides suggestion for further developments and improvements.
2 Background

This chapter reviews anatomical and physiological concepts related to the ankle joint and the peripheral neuromuscular system. Skeletal muscle structure is presented. Sensory stretch receptors and stretch reflexes are described followed by an introduction to dynamic joint stiffness.

2.1 Ankle Joint

The ankle joint is a complicated system consisting of bone, cartilage, muscle, and passive tissue. The ankle joint contains four major bones: the tibia and fibula of the lower leg which articulate with the talus and calcaneus of the foot [1]. The main bones and muscles of the ankle joint are shown in Figure 2.1. The main movement of the joint is in the sagittal plane and consists of plantarflexion (pushing the foot down) and dorsiflexion (pulling the foot up), as shown in Figure 2.2.

Dorsiflexion results from the contraction of the tibialis anterior muscle that originates at the tibia and inserts at the base of the first metatarsal bone via the tibialis anterior tendon [2]. The muscles primarily responsible for plantar flexion are the gastrocnemius and soleus. The gastrocnemius is attached to the femur and to the calcaneus via the calcaneal (Achilles) tendon, thus spanning the knee and ankle joints. The soleus is attached to the head of the fibula and a portion of the medial tibia [2].
Figure 2.1: Posterior view (left) and medial view (right) of ankle joint bones, tendons and muscles. Adapted from [2]
Figure 2.2: Definitions of plantarflexion (pushing the foot down) and dorsiflexion (pulling the foot up).
2.2 Peripheral Neuromuscular System

The nervous system can be divided into two parts: the central nervous system and the peripheral nervous system. The central nervous system comprises the brain and spinal cord. The peripheral nervous system consists of the nerves between the brain or spinal cord and the body’s muscles, glands and sense organs [3]. The peripheral nervous system can be divided into efferent and afferent divisions. Afferent neurons carry information from sensory organs to the central nervous system, while efferent neurons carry information from the central nervous system to muscles or glands. Efferent neurons are classified as autonomic or somatic. Autonomic neurons innervate smooth and cardiac muscle, glands, and the gastrointestinal tract. Neurons of the somatic division are motor neurons that innervate skeletal muscle. All motor neurons are excitatory; they stimulate muscle tissue to contract; muscle relaxation is accomplished by inhibiting motor neurons [3].

The peripheral nervous system delivers sensory and motor information through bundles of nerve fibres. A further division of the peripheral nervous system is made to differentiate between sensory information being brought to the central nervous system from receptors in peripheral tissues and organs, and motor commands being carried from the central nervous system to peripheral tissues [3].

2.2.1 Skeletal Muscle Structure

We use skeletal muscles to perform voluntary motion and maintain posture. Skeletal muscle contains many cylindrical muscle fibres, each having one elongated cell
with multiple nuclei. Muscle fibres in adults may be up to 20 cm in length and have diameters between 10 and 100 µm [3]. A muscle fibre consists of myofibrils lying parallel to one another that are striped in appearance due to the arrangement of their thick and thin protein filaments [3], as illustrated in Figure 2.3.

The repeating unit in the myofibril is called the sarcomere. A sarcomere consists of thin actin and thick myosin filaments and its length is between 1.6 and 2.6 µm [3]. Thick and thin filaments are typically 10-12 nm and 5-6 nm in diameter, and 1.6 µm and 1 µm in length, respectively [3].

The thick filaments consist primarily of myosin molecules. Myosin molecules have a long tail portion from which extends a globular head containing a binding site for actin [1]. Two intertwined helical chains of actin molecules form the backbone of the thin filaments. Secondary components in thin filaments include troponin and tropomyosin. Tropomyosin molecules block the myosin-binding site on each actin molecule and are held in place by troponin. As depicted in Figure 2.3, the actin filaments are anchored to a structure of interconnecting proteins known as the Z line [3].
Figure 2.3: Skeletal Muscle Structure. Adapted from [3]
Cross-bridges define the projections from myosin filaments that interact with the actin filaments. The sliding-filament theory describes how the action of the cross-bridges creates muscular contraction and force [4].

Muscles have both elastic and viscous properties. Elasticity is due to the structure of the sliding filaments. The total contractile force generated by the muscle depends on the amount of stretch of its filaments. The active tension depends on the degree of overlap between thin and thick filaments. The more overlap, the more force will be generated until an optimal length is reached. If the muscle is too long, the thin filaments will be too far away from the thick filaments to produce overlap and to generate force. Therefore, the force generated will increase with length until an optimal length is reached; after which the force will decrease, as shown in Figure 2.4. However, passive muscle tension increases continuously.
Figure 2.4: Skeletal Muscle Structure Muscle length-tension relationship. Adapted from [3].
On the other hand, viscosity of the muscle is originated from extracellular connective tissue. The force generated depends on the stretch rate of the filaments, as shown by Figure 2.5. As the load on a muscle increases, muscle shortening velocity decreases until a point is reached at which the load is too great for the muscle to shorten [3]. The tension at this point is known as the maximum isometric tension. Therefore, the response of skeletal muscles depends on both position and velocity.
Figure 2.5: Muscle stretch velocity versus tension. Adapted from [3].
2.2.2 Neurons and Motor Units

Skeletal muscle is activated by somatic efferent nerves called motor neurons. The cell bodies of these neurons are located in the brainstem or ventral horn of the spinal cord, but their axons can extend beyond one meter [4].

The central commands and sensory feedback converge at the motor neuron to control muscle contraction. Each muscle fibre is controlled by only one motor neuron, but each motor neuron innervates many muscle fibres. The motor neuron and muscle fibres it innervates are called a motor unit. Figure 2.6 shows two motor units. A motor unit is the smallest functional element of the motor system. All fibres in a motor unit contract each time the motor neuron fires.

Figure 2.6: Motor Unit. Adapted from [5]

2.2.3 Control of Muscle Force

The neuromuscular system uses two mechanisms to modulate muscle force. The first mechanism is rate coding, that is the adjustment of motor neuron firing rate. Active
tension produced by a muscle varies directly with the action potential firing rate of its nerve [3]. Low frequency stimuli evoke discrete and separate twitches, characterized by a rapid rise and slow decay. Force summation occurs when successive stimuli occur at intervals less than the decay time; force summation saturates at high rates of stimulation. At the high rates, individual stimuli do not produce discrete force fluctuations. Figure 2.7 illustrates how rate coding controls muscle force generation.

The second mechanism controlling force generation involves changing the number of motor units by recruitment/de-recruitment. Motor units are recruited in order according to size. In general, the smaller motor units are recruited first followed by stronger and larger motor units, as shown in Figure 2.8. The order of recruitment is highly correlated with the diameter and conduction velocity of the axons and the size of the motor neuron cell bodies. Since a cell’s electrical resistance is tied to its size, a smaller diameter cell will reach threshold levels before a larger neuron [3].
Figure 2.7: Effect of AP frequency on muscle force. Adapted from [3].
Figure 2.8: Successive recruitment of motor units is the primary means of adjusting muscle tension. Small motor units that generate small amounts of tension are activated first, followed by larger ones. Adapted from [6].
2.2.4 Electromyography

Electromyography (EMG) is the process of measuring the electrical activity of the muscle. When a motor neuron activates a muscle fibre, the depolarization of the muscle proceeds outwards in both directions along the fibre. The depolarization can be measured by an electrode. Since a single muscle fibre is rarely activated by itself, the voltage recorded is the result of the depolarization of many muscle fibres. A recording of the depolarizations from all the muscle fibres from one motor neuron is called the motor unit action potential (MUAP). Activation of one motor neuron produces a MUAP with a characteristic shape that depends on the geometry of the muscle and the location of the electrode [7]. EMG can be recorded with surface electrodes placed on the skin overlying the muscle with indwelling electrodes inserted into the muscle. Indwelling electrodes usually provide higher Signal to Noise Ratio (SNR) since they are closer to the muscle fibres.

2.2.5 Peripheral Sensory Receptors

Sensory receptors are stimulated by physical or chemical changes in the body or by changes in the external environment [8]. Proprioceptors are sensory receptors that perceive the position of the body in space by detecting position and force information. Two types of proprioceptors thought to be responsible for muscle regulation are: muscle spindles and Golgi tendon organs (GTO). Figure 2.9 demonstrates the position of spindle and GTO in muscle.
Figure 2.9: Skeletal muscle and muscle sensory receptors. Adapted from [9].
2.2.5.1 Muscle Spindles

Muscle spindles are small sensory receptors in the muscle, monitoring changes in muscle length and velocity. They comprise modified muscle fibres called intrafusal fibres, lying in parallel with the muscle. Therefore, changes in muscle length are mirrored in the intrafusal fibres.

Two types of nerves carry the afferent information from the spindle: primary (Group Ia) and secondary (Group II) sensory fibres [10]. Group Ia fibres are large-diameter (12-20 µm), myelinated nerves with unmyelinated endings that wrap around the central non-contractile portions of the intrafusal fibres. Group II fibres are small-diameter (4-12 µm), unmyelinated nerves that make connections onto the polar contractile endings of the intrafusal fibres. There are multiple Group II endings in each spindle but only one Ia ending for each spindle [8].

Fusimotor innervation involves efferent fibres called gamma motor neurons, which carry signals from the CNS (spinal cord) to innervate the spindle’s intrafusal fibres. Gamma motor neurons stimulate the two ends of the intrafusal fibres to shorten and thus maintain tension in the central region where the receptors are located. Previous research suggests that intrafusal fibres co-contract with extrafusal fibres to ensure information about muscle length is continuously available; this process is referred to as alpha-gamma coactivation [6].

Primary endings are more sensitive than secondary endings to changes in muscle length and velocity [8]. The frequency responses of both primary and secondary endings are shown in Figure 2.10.
Figure 2.10: Sensitivity of primary and secondary endings of muscle spindles to sinusoidal stretching [9].
2.2.5.2 Golgi Tendon Organs

Golgi tendon organs are sensory receptors located at the junction between muscle fibres and tendon. As opposed to spindles, GTOs are connected in series to a small number of muscle fibres, and in parallel with others. These sensors are very sensitive to changes in active muscle tension, but less sensitive to passive muscle tension.

GTOs are slender, encapsulated receptors, innervated by a single Group Ib axon. As the axon enters the capsule, it branches into fine endings intertwined into braided collagen fibres. Muscle tension will stretch the tendon organ, which straightens the collagen fibres and compresses the nerve endings, causing them to fire.

2.2.6 Reflexes

2.2.6.1 Stretch reflex

When a skeletal muscle is stretched, the change in muscle length is sensed by the muscle spindles. Group Ia afferent fibres from the spindles transmit a signal into the spinal cord where they synapse directly with the motor neurons controlling the muscle that was just stretched [8]. This monosynaptic pathway is known as the stretch reflex arc, and is illustrated by path A in Figure 2.11. Muscle contraction occurs approximately 20-50 ms after the stretch was initiated [1]. This resistance to elongation occurs in basically every extensor or flexor skeletal muscle in the body and is called the stretch reflex [9].

Group Ia afferent fibres also innervate motor neurons of agonistic muscles (path C in Figure 2.11), whose contraction would assist the intended motion of the stretched muscle. Conversely these afferent fibres synapse with interneurons (path B in Figure
2.11) that in turn inhibit motor neurons of antagonistic muscles, whose contraction would oppose the intended motion of the stretched muscle [9]. The activation of agonist muscles and simultaneous inhibition of antagonistic muscles is called reciprocal innervation and is present in many reflexes. The muscle length and velocity information is conveyed to higher centres of motor control in the brain via a final afferent pathway involving one or more interneurons, as shown by path D in Figure 2.11 [2].
Figure 2.11: Pathways involved in the stretch reflex. Adapted from [9].
2.2.6.2 Inverse Stretch Reflex

The inverse stretch reflex is a polysynaptic inhibitory reflex [8]. Stretch of a muscle will activate the Golgi tendon organs, which encodes muscle tension. The action potentials from the Golgi tendon organs travel along Group Ib afferents to the spinal cord where, through the actions of inhibitory interneurons, they tend to suppress activity of the alpha motorneurons of the same muscle [10]. The Group II projections may also form excitatory synapses with the motorneurons of the antagonistic muscles [10].

2.2.6.3 Long loop Reflex

In addition to the stretch reflex, there are also responses that occur at a latencies, between those expected for the short latency reflex and voluntary control. Those polysynaptic reflexes are called long-loop reflexes and believed to be mediated via the motor cortex. These reflex connections are more appropriate for tasks that require more fine control, while short-latency, spinal reflexes are more suited to gross movement tasks such as posture [8] [11].

2.2.6.4 Modulation of Reflex

The stretch reflex can be modulated in several ways. Alpha-gamma co-activation helps to maintain muscle spindle sensitivity during muscle contraction. But, gamma motor neurons can be controlled independently. This process is called fusimotor control and is used to fine tune muscle spindle sensitivity for specific tasks [8]. A second method of modulating the stretch reflex is through presynaptic inhibition [3]. This process involves interneurons that form a presynaptic synapse to Group Ia afferents and decrease
the amount of neurotransmitter released by the Ia afferent to the motorneuron [3]. A final method of stretch reflex modulation is through Renshaw cells. The Renshaw cell is a small inhibitory interneuron very close to the $\alpha$-motoneuron, from which it receives a collateral branch then gives it a postsynaptic inhibition (a negative feedback) [10]. In addition, the Renshaw cell also provides negative feedback to nearby synergistic motoneurons, other Renshaw cells, gamma-motoneurons, and Ia inhibitory interneurons.

2.3 Joint Stiffness

The concept of dynamic joint stiffness can be used to study the mechanical behavior of the mechanisms acting about the ankle. Joint stiffness may be defined as the dynamic relation between the angular position of a joint and the torque acting about it. Ankle joint stiffness plays an important role in the control of posture, where it defines the mechanical response to external perturbations, and of movement, where it determines the forces that must be generated to move the limb. Joint dynamics consists of three components: passive, intrinsic and reflex dynamics. Passive dynamics consist of two parts: limb dynamics and articular mechanics. Limb dynamics can be attributed to inertia if it is assumed that the limb is rigid and the joint rotates about a single, fixed axis. Thus, limb dynamics can be estimated between the second derivative of the joint position and the torque acting about it [12]. Intrinsic dynamics arise from intrinsic properties of active muscle fibers. Muscle mechanics consists of two components: contractile mechanics and activation dynamics. Contractile mechanics determine the forces generated in response to changes in length when the level of activation remains constant. Activation dynamics describe how electrical signals from the neuromuscular system relate to force generated
by the muscle. Activation dynamics are associated with the level of voluntary activation
and reflex responses. The stretch reflex dynamics arise from reflex activation dynamics
and muscle dynamics [13].

A flow diagram for the dynamics of a single joint is shown in Figure 2.13.

2.3.1 Parallel Cascade Model

Kearney et al. [14] found that a parallel cascade model could describe dynamic
joint stiffness well, as shown in Figure 2.13. For perturbations about an operating point,
intrinsic stiffness could be modeled well by a second-order, quasi-linear system with
transfer function [14]:

$$ H_{is}(s) = \frac{TQ_i(s)}{POS(s)} = Is^2 + Bs + K $$

(2.1)

where

$TQ_i$ is the intrinsic torque that is the output from intrinsic stiffness,

$POS$ is the position that is the input to intrinsic stiffness,

$I$, $B$ and $K$ are inertial, viscous, and elastic parameters which vary with the
operating point.

Reflex stiffness arises from muscle contraction in response to reflex activation
from stretch receptors in the muscle. At the ankle, reflex stiffness can be modeled with a
Linear-Nonlinear-Linear (LNL) block-structured model, comprising the series connection
of a differentiator, a delay of about 40 ms [12], a static non-linearity and a 2nd or 3rd
order low-pass system, whose transfer functions are given in Equations 2.2 and 2.3,
respectively.

\[
H_{RS}(s) = \frac{TQ_R(s)}{VL(s)} = \frac{g_R \omega_n^2}{(s^2 + 2\zeta \omega_n s + \omega_n^2)} e^{-\tau s} \quad (2.2)
\]

\[
H_{RS}(s) = \frac{TQ_R(s)}{VL(s)} = \frac{g_R \omega_n^2 \rho}{(s^2 + 2\zeta \omega_n s + \omega_n^2)(s + \rho)} e^{-\tau s} \quad (2.3)
\]

where

- \(TQ_R\) is the reflex torque,
- \(VL\) is the half-wave rectified joint angular velocity,
- \(g_R\) is the reflex gain,
- \(\omega_n\) is the 2nd order natural frequency,
- \(\zeta\) is the damping parameter,
- \(\rho\) is the 1st order cut-off frequency,
- \(\tau\) is the reflex delay.
Figure 2.12: Peripheral neuromuscular system contributors to joint dynamic stiffness
Adapted from [13]
Figure 2.13: Parallel cascade model for dynamic ankle stiffness. pos denotes the position. $tq_I$, $tq_R$ and $tq_N$ denote intrinsic torque, reflex torque and net torque. $z$ denotes the output from the static nonlinearity. Intrinsic stiffness is modeled as a linear system. Reflex stiffness is modeled as a delay, followed by a differentiator and a Hammerstein system comprising the series connection of a static nonlinearity and a linear system.
2.3.2 EMG

The simplest way to measure the stretch reflex is to record the EMG from a muscle in response to a stretch. EMG records of many stretch reflexes are rectified and then aligned on stretch onset and ensemble averaged. Since muscle activation is triggered by the stretch, the electrical activity of the muscle is synchronous. The stretch reflex response can be measured as the amplitude of the rectified and averaged EMG from a lag of 40-100 ms [8]. EMG recordings provide a good way to monitor muscle activity. However, the relationship between the EMG and muscle force is well understood only when the muscle force is generated during static, isometric contractions. Their relationship is not well understood when the muscle length changes and/or the contraction is dynamic [15]. Therefore, EMG recordings are not appropriate to study the mechanical behavior of the joint when the muscle changes dynamically to generate a dynamic force.

2.3.3 Separation of Intrinsic and Reflex Stiffnesses

Intrinsic and reflex torques cannot be measured separately experimentally; only their sum can be measured. Consequently, intrinsic and reflex stiffnesses cannot be estimated directly. To overcome this difficulty, a number of approaches have been developed to separate the torques due to each pathway.

One approach is to compare the responses before and after eliminating the reflexes experimentally by surgical differentiation [16], applying a neural block to [17] or electrically stimulating the primary afferents [18] [19]. The differences between the two responses are attributed to the contributions from the reflex stiffness. These methods
are only reliable if the operating points before and after eliminating reflexes are matched closely to ensure that intrinsic mechanics do not change. However, in practice, this is difficult to achieve since eliminating the reflex contribution by itself changes the system.

### 2.3.4 System Identification

Another approach is to use system identification-based methods to separate the intrinsic and reflex torques analytically. Thus, the parallel cascade method, developed in our laboratory [14] [20], estimates the intrinsic stiffness and reflex stiffness from the measured position and net torque using non-parametric, nonlinear identification methods. Intrinsic stiffness is estimated first, taking advantage of the stretch reflex delay to eliminate any contributions from reflex mechanisms. Then, reflex stiffness is estimated after the torques due to intrinsic stiffness have been removed. The algorithm repeats until it converges. Zhang and Rymer [21] estimated the intrinsic and reflex components of joint stiffness at the knee using a parametric method that involved solving a set of nonlinear delay differential equations describing the intrinsic and reflex stiffnesses. We developed a real time method [22] that achieves the same results in real-time by using a specially designed input sequence whose properties eliminate the correlation between intrinsic and reflex torques. Intrinsic stiffness is estimated from the cross-correlation between this perturbation and the measured torque. Then, reflex stiffness is estimated after the torques associated with intrinsic stiffness have been removed.

Using the parallel cascade method, Kearney et al. studied the effects of the properties of position perturbations on reflex stiffness [14]. They found that reflex stiffness was influenced strongly by the root-mean-square (RMS) velocity. The reflex
stiffness gain was highest with an the RMS velocity of 0.2 rad/s. Reflex stiffness Impulse Response Function (IRF) became smaller, shorter and more oscillatory as velocity increased [14].

Mirbagheri et al. used the parallel cascade method to study reflex stiffness in normal subjects at different operating points in a simple torque-matching task [23]. They found that intrinsic stiffness gain increased when the foot was moved from the mid-position to either plantarflexion or from dorsiflexion, in contrast, the reflex gain was significantly larger when the ankle was dorsiflexed as compared to plantarflexed.

Ludvig et al [22] used the real time algorithm to estimate the intrinsic and reflex stiffness in real time, and then investigate behavior of stretch reflex. They found that subjects could control stretch reflexes voluntarily and manipulate it independently of the intrinsic stiffness.

2.3.5 Open-loop experiment VS Closed-loop experiment

The parallel cascade method and real-time algorithm are correlation-based identification methods [24] that require the experimental data to be collected from an open-loop experiment. To meet this requirement, experiments were conducted by applying perturbations using an actuator operating in a position control mode as illustrated in Figure 2.14. In this mode, the actuator operates as a position servo and drives the ankle position to follow the command input. The actuator is much stiffer than the ankle, so that the torques developed at the ankle will not change the joint position. This effectively opens the feedback path from ankle torque to ankle position. Therefore, identifying the ankle dynamics can be viewed as an open-loop identification problem.
Figure 2.14: Block diagram of the open-loop identification problem showing intrinsic, $H_I(s)$, and reflex, $H_R(s)$, stiffness dynamics.
Such open-loop experiments have provided a good understanding of the behavior of intrinsic and reflex stiffness under these experimental conditions, which correspond to interacting with a very stiff environment. However, these conditions are far from realistic because, during most functional tasks and movements, subjects will interact with compliant loads and consequently the joint torque and position will be connected by feedback, as shown in Figure 2.15. Any change in torque will evoke a change in position and vice versa. Torques and positions measured under such conditions are therefore measured in closed-loop. It is well known that correlation-based identification methods, such as those we have used, will give biased results when applied to closed-loop data [24]. Therefore new tools are needed to estimate intrinsic and reflex stiffness while subjects interact with compliant environments.
Figure 2.15: Simplified block diagram of the closed-loop identification problem. $H_{IMP}(s)$ represents the dynamics of the impedance controller. $H_I(s)$ and $H_R(s)$ represent the intrinsic and reflex stiffness respectively. INP represents the input perturbation.
2.4 Thesis Rationale

Previous studies have shown that the intrinsic and reflex torques can be separated from the net torque analytically using system identification based technologies. The parallel cascade method [18] [19], real-time method [21] and nonlinear differentiation method [20] have been developed by different research groups to identify the intrinsic and reflex stiffness. However, all those methods are only suitable for open-loop system identification. The open-loop experiment is far from realistic because the limbs interact with compliant loads during most functional movements. Under these conditions, joint stiffness must be viewed as operating within a closed-loop because the ankle torque is fed back through the load to change the ankle position. Applying open-loop identification methods to such data will give biased results. In this thesis, I will show that open-loop methods, such as parallel cascade method, do provide biased estimate when used with data collected from closed-loop system.

One of the roadblocks to estimating the intrinsic and reflex stiffnesses in closed-loop has been the lack of analytic tools to identify nonlinear systems in closed-loop. Identification of nonlinear systems in closed-loop is a difficult topic because the feedback noise and nonlinearity will cause cross-correlation methods to fail. Therefore, there is a need to develop a new analytic tool to estimate the intrinsic and reflex stiffness in closed-loop. This thesis will present a subspace-based identification method to estimate the intrinsic and reflex stiffness in closed-loop.
References


3 Subspace-based System Identification Methods

3.1 Introduction

Subspace methods, and the Multivariable Output-Error State-space (MOESP) method in particular, will be used throughout this thesis. Consequently, this chapter will review subspace methods with particular attention to the MOESP approach.

Subspace methods are identification technologies that estimate a state space model of a system directly from input-output measurements. There have been three main implementations of the subspace method: De Moor’s N4SID (Numerical Algorithms State Space Subspace System Identification) [1], Verheagen’s MOESP [2], and Larimore’s CVA (Canonical Variate Analysis) [3]. A thorough comparison of these three subspace-based identification methods has been conducted in [4]. MOESP is the fastest in computation and can be extended to identify systems with noise and nonlinearities. Therefore, we opted to use MOESP in this thesis.

Compared with classical identification algorithms, such as the Prediction Error Method (PEM) [5], subspace-based methods have several advantages.

1. Model Structure

The subspace method estimates a state space model for a system directly from measured input and output data. State space models, whose form is shown in Equation 3.1, are good engineering models; they describe many systems efficiently, require few parameters and have good numerical properties. Moreover, many system design tools are available for controllers based on state space models [6] [7]. In addition, state space
models are well suited to Multiple Input and Multiple Output (MIMO) systems since their state space models have the same form as those for Single Input and Single Output (SISO) systems. This is particularly important to this thesis, since a MIMO state space model will be used extensively to model ankle joint stiffness. State space models and subspace methods have been used in biomedical engineering to model the behavior of a variety of physiological systems, including postural control [8] and joint stiffness [9] [10].

2. Fast Convergence

Subspace-based methods are computationally efficient and do not require iteration. Hence they are generally faster than classical iterative algorithms such as PEM, and do not have the associated convergence problems.

3. Order Selection

Subspace methods select the model order automatically in contrast to classical parametric algorithms that require the model structure to be chosen \textit{a priori}. For classical methods, model order selection typically requires two steps: first a high order, over-parameterized model is estimated; then a model order reduction procedure is applied to determine the optimal system order. Subspace methods determine the order automatically as will be described in Section 3.3.

3.2 Background

3.2.1 State Space Model

A state space model has the structure shown in the block diagram of Figure 3.1
and can be described by the equations:

\[
\begin{align*}
    x(k+1) &= Ax(k) + Bu(k) \\
    y(k) &= Cx(k) + Du(k) + w(k)
\end{align*}
\]

(3.1)

where

- \( u(k) \in \mathbb{R}^m \) and \( y(k) \in \mathbb{R}^l \), are vectors containing \( m \) inputs and \( l \) outputs of the process, at discrete time \( k \), respectively.
- \( w(k) \in \mathbb{R}^l \) is an additive measurement noise,
- \( x(k) \in \mathbb{R}^n \), is the state vector of the process at discrete time \( k \) containing the values of \( n \) states, where \( n \) is the order of the system,
- \( A \in \mathbb{R}^{n \times n} \) is the system matrix that describes the dynamics of the system,
- \( B \in \mathbb{R}^{n \times m} \) is the input matrix that defines how the deterministic inputs influence the states,
- \( C \in \mathbb{R}^{l \times n} \) is the output matrix that describes how the internal states are transformed to generate the output \( y(k) \),
- \( D \in \mathbb{R}^{l \times m} \) is the direct feed through term.

Given measurements of the input signal(s) \( u(k) \) and output signal(s) \( y(k) \), for \( k = 0,1,2,\cdots,N \), the subspace method estimates the order of the system \( n \), the matrices, \( A, B, C, D \) and the initial state, \( x(0) \). It should be noted that state space models for SISO systems and MIMO systems have the same form and consequently can be solved using the same methods. Therefore, the effort required to estimate a state space model for a MIMO system using subspace method is not much greater than that required to estimate a
model for a SISO system.
Figure 3.1: Block diagram of a state space model. Inputs, outputs and states are denoted as $u(k)$, $y(k)$ and $x(k)$, respectively. $w(k) \in \mathbb{R}^l$ is an additive noise. The state space model is described by the system matrices, $A$, $B$, $C$, $D$. The symbol delta represents a sample delay for this discrete time model.
3.2.2 Matrix Definitions

It is evident from Equation 3.1, that the output signals are dependent on the inputs, the system states and the noise. Subspace methods estimate the state space model of a system, by using geometrical projections to find the subspaces that contain information from only the states or the inputs. To achieve this, subspace methods manipulate the measured data using vectors and matrices with special forms. This section will review these forms as well as some mathematical tools and definitions that will be used frequently throughout this thesis.

3.2.2.1 Hankel Matrix

Hankel matrices have constant values on their anti-diagonals and are used by subspace methods to store and manipulate input and output data. For example, the Hankel matrix for the input signal, $u$, is

$$U_{i,j,N} = \begin{bmatrix}
  u(i) & u(i+1) & u(i+2) & \cdots & u(i+N) & u(i+N-1) \\
  u(i+1) & u(i+2) & \cdots & u(i+N) & \vdots \\
  u(i+2) & \ddots & \ddots & \vdots & \vdots \\
  \vdots & \ddots & \ddots & u(i+j+N) & \vdots \\
  u(i+j-2) & \cdots & u(i+j+N) & u(i+j+N-1) & \vdots \\
  u(i+j-1) & u(i+j) & \cdots & u(i+j+N) & u(i+j+N-1) & u(i+j+N-2)
\end{bmatrix}$$

(3.2)

where

- $i$ is the starting index,
- $j$ is the number of the rows,
- $N$ is the number of columns.
3.2.2.2 Observability Matrix.

The observability matrix for a state space model of the form given in Equation 3.1 is defined as \[11\]:

\[
\Gamma_i = \begin{bmatrix}
C \\
CA \\
\vdots \\
CA^{i-1}
\end{bmatrix}
\]  

(3.3)

Theoretically, the initial state of the system, \(x(0)\), can be obtained using the Observability matrix if the system is observable \[11\]. In particular, for a linear system the zero-input response is \(y(k) = CA^k x(0)\). Stacking the output samples at discrete times \(k = 1, 2, 3 \ldots\) into a column vector gives

\[
\begin{bmatrix}
y(0) \\
y(1) \\
\vdots \\
y(i-1)
\end{bmatrix} = \Gamma_i x(0)
\]  

(3.4)

If the system is observable, the Observability matrix, \(\Gamma_i\), will have full rank and so be invertible. Thus, the initial state of the system, \(x(0)\), can be estimated by solving the least square problem:

\[
x(0) = \Gamma_i^+ y
\]  

(3.5)

where
‡ indicates the Moore-Penrose pseudo-inverse [12].

When the number of samples exceeds the order of the system, \( i > n, \) \( \Gamma_i \) is called the extended Observability matrix. Subspace identification methods, such as MOESP, first estimate the extended Observability matrix \( \Gamma_i \) and then use it to estimate the \( A \) and \( C \) matrices.

### 3.2.2.3 Toeplitz Matrix of Impulse Response

Toeplitz matrices [11] have constant values on their diagonals. They are used in subspace methods to describe how the state-free response influences the system’s output. A Toeplitz matrix constructed from the system’s impulse response has the form:

\[
H_j = \begin{bmatrix}
D & 0 & \cdots & 0 \\
CB & D & 0 & \cdots \\
\vdots & CB & \ddots & \ddots \\
CA_iB & \cdots & 0 \\
CA_iB & CA_iB & CB & 0 \\
CA_iB & CA_iB & CA_iB & \cdots & CB & D
\end{bmatrix}
\]  

(3.6)

### 3.2.3 Orthogonal Projections

Subspace-based methods use geometric tools to find different subspaces from data matrices. These tools are based on geometric concepts, such as orthogonal projections, oblique projections, and intersections. In this thesis, we focus mainly on orthogonal projections [13].

For a matrix \( A \in \mathbb{R}^{m \times n}, (m \geq n) \) whose column-space is \( S_A \subseteq \mathbb{R}^n \), the orthogonal projection of matrix \( B \in \mathbb{R}^{m \times n} \) on to \( S_A \) is achieved by left multiplying \( B \) by a projection
matrix, denoted as $P_A$. If $A$ has full column rank, $P_A$ can be calculated as.

$$P_A = A(A^T A)^{-1} A^T$$ (3.7)

The matrix $P_A$ projects the columns of the matrix $B \in \mathbb{R}^{p \times n}$ orthogonally onto the subspace $S_A$, that is:

$$\text{column-space}(BP_A) \subseteq S_A$$ (3.8)

The orthogonal complement of the column-space of $A$ is given by $S_A^\perp = \mathbb{R}^m / S_A$. The projection matrix

$$P_A^\perp = I_m - P_A$$ (3.9)

projects the columns of $B$ orthogonally onto $S_A^\perp$ such that

$$BP_A^\perp \subseteq S_A^\perp$$ (3.10)

Similarly, the orthogonal projection of the rows of matrix $B \in \mathbb{R}^{m \times p}$ on the row-space of $A \in \mathbb{R}^{m \times n}$, $(m \leq n)$ is achieved by right multiplication with the matrix $\Pi_A$. If $A$ has full row rank, the projection matrix is calculated as

$$\Pi_A = A^T (AA^T)^{-1} A$$ (3.11)

The projection from $B$ onto the row space of $A$ is denoted as $\Pi_A B$ and has the property

$$\text{row-space}(\Pi_A B) \subseteq \text{row-space}(A)$$ (3.12)

The projection matrix
\[ \Pi_A^\perp = I_n - \Pi_A \]  

(3.13)

projects the rows of the matrix \( B \) onto the orthogonal complement of the row-space of \( A \), so that

\[
\text{row-space}(\Pi_A^\perp B) \subseteq \mathbb{R}^n/\text{row-space}(A)
\]

(3.14)

In practice, orthogonal projections are usually calculated using LQ factorization [13]. Thus, the LQ factorization of a matrix \( A \) is given by

\[ A = LQ \]

(3.15)

where

\( L \) is a lower triangular matrix,

\( Q \) is a matrix with orthogonal rows.

To calculate \( \Pi_A B \), LQ factorization is applied to the stacked \( A \) and \( B \) matrices as follows:

\[
\begin{bmatrix}
A \\
B
\end{bmatrix} = \begin{bmatrix}
L_{11} & 0 \\
L_{21} & L_{22}
\end{bmatrix} \begin{bmatrix}
Q_1 \\
Q_2
\end{bmatrix}
\]

(3.16)

where

\( Q_i \) are orthogonal, that is

\[
\begin{cases}
Q_i Q_j^T = I & \text{if } i = j \\
Q_i Q_j^T = 0 & \text{if } i \neq j
\end{cases}
\]

From Equation 3.16, it is evident that

\[
\begin{align*}
A &= L_{11} Q_1 \\
B &= L_{21} Q_1 + L_{22} Q_2
\end{align*}
\]

(3.17)

so that \( \Pi_A B \) can be seen to be:
\[ \Pi_A B = A^T (AA^T)^{-1} AB \]
\[ = Q_1^T L_{11}^T (L_{11} Q_1 Q_1^T L_{11})^{-1} L_{11} Q_1 (L_{21} Q_1 + L_{22} Q_2) \]
\[ = L_{21}^T L_{11}^T (L_{11} L_{11}^T)^{-1} L_{11} Q_1 \]
\[ = L_{21} Q_1 \] (3.18)

### 3.2.4 Singular Value Decomposition

Singular Value Decomposition (SVD) is another tool frequently used by subspace methods. The SVD of the matrix \( A \in \mathbb{R}^{m \times n} \) is:

\[ A = U S V^T \] (3.19)

where

- \( U \in \mathbb{R}^{m \times m} \) is an orthogonal matrix and \( U^T U = UU^T = I \),
- \( S \in \mathbb{R}^{m \times n} \) is a diagonal matrix with positive singular values on the diagonal sorted in decreasing order,
- \( V \in \mathbb{R}^{n \times n} \) is a second orthogonal matrix.

SVD is an important tool for determining low rank approximations [13]. The matrix \( A \) can be expanded in terms of its Singular Values as:

\[ A = u_1 \sigma_1 v_1^T + u_2 \sigma_2 v_2^T + \cdots + \sum_{i=1}^{\min(m,n)} u_i \sigma_i v_i^T \] (3.20)

where

- \( u_i \) is the \( i^{th} \) column of \( U \),
- \( v_i \) is the \( i^{th} \) column of \( V \),
- \( \sigma_i \) is the \( i^{th} \) diagonal element of \( S \).
This expansion is useful in determining a system’s order when generating low rank approximations. Thus, for example, the first $r$ columns of the $U$ matrix are the column base for the column-space of $A$ where $r$ is the number of non-zero singular values in $S$. Consequently, $A$ can be constructed from linear combinations of these $r$ bases. Thus the low rank approximation is described by:

$$A = USV^T$$

$$= [U_1 \ U_2] \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

$$= U_1 S_1 V_1^T$$

where

- $S_1$ contains the $r$ non-zero singular values in $S$,
- $U_1$ is the first $r$ columns in $U$,
- $V_1^T$ is the first $r$ rows in $V^T$.

Subspace methods also use SVD to estimate a system’s order and its Extended Observability matrix as described in Section 3.3.

### 3.2.5 Ergodic Algebraic Framework

In this thesis, we will assume that all signals are realizations of ergodic stochastic processes. For example, we assume that $u(j) \in \mathbb{R}^m$ and $v(k) \in \mathbb{R}^l$, are ergodic stochastic processes such that [14]:

$$E[u(j)v^T(k)] = \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} u(j+I-1)v^T(k+I-1)$$

(3.22)
An alternative way of expressing this limit is

\[
\frac{1}{N} \sum_{j=1}^{N} u(j + I - 1)v^T(k + I - 1) + O_N(\varepsilon) = E[u(j)v^T(k)]
\]  

(3.23)

where

\[ O_N(\varepsilon) \]

is a bounded matrix of appropriate dimensions with norm \( \varepsilon \), which vanishes as \( N \to \infty \).

In this thesis, we will frequently refer to sequences that are zero-mean, white noise; such a sequence \( \{w(k)\} \) with \( w(k) \in R^n \) satisfies:

\[
E[w(k)] = 0,
\]

\[
E[w(k)w^T(j)] = \begin{cases}
0, & \text{if } k \neq j, \\
\sigma_w^2 I, & \text{if } k = j,
\end{cases}
\]

(3.24)

### 3.2.6 Persistently Exciting Signals

System identification requires the input and output signals to be informative enough to permit the system dynamics to be estimated. Different identification methods may have different specific requirements for input signals. For subspace methods the requirement is that the input signal must be Persistently Exciting (PE) [15].

Let \( u \) be the input to a linear system and \( x \) be the state of the system. Construct the Hankel matrix \( U_{i,j,N} \) and vector \( X_{i,N} \) as follows:

\[
U_{i,j,N} = \begin{bmatrix}
u(i) & u(i+1) & \cdots & u(i+N-1) \\
u(i+1) & u(i+2) & \cdots & u(i+N) \\
\vdots & \vdots & \ddots & \vdots \\
u(i+j-1) & u(i+j) & \cdots & u(i+j+N-2)
\end{bmatrix}
\]

(3.25)
and

\[ X_{i,N} = [x(i) \ x(i+1) \ \ldots \ x(i+N+1)] \]  

(3.26)

Then the input, \( u \), is PE of order \( j \) if:

\[ \text{rank} \begin{bmatrix} U_{i,j,N} \\ X_{i,N} \end{bmatrix} = mj + n \]  

(3.27)

where

- \( m \) is the number of inputs,
- \( n \) is the number of the states.

### 3.3 The Ordinary MOESP Method

This section will introduce the ordinary MOESP algorithm [2] by solving an example problem step by step, explaining the mathematical concepts, and examining the results at each step. The data for the example were generated by simulating a fifth order, discrete-time, linear system [4] using Matlab [16].

\[ P(Z) = 10^{-2} \frac{-3.49Z^5 + 12Z^4 - 16.75Z^3 + 12.73Z^2 - 3.99Z - 0.37}{Z^5 - 4.0331Z^4 + 6.8927Z^3 - 6.2716Z^2 + 3.0211Z - 0.606} \]  

(3.28)

Fifty seconds of Gaussian white noise was used as the input. Figure 3.2 shows a 1 second segment of the simulated input/output data. The MOESP identification proceeds as follows.
Figure 3.2: Input and output data from the simulation of a fifth-order system without noise.
3.3.1 Estimate the Order of System and the Matrices A & C

For the sake of simplicity, no noise was considered in this example. Consequently, from Equation 3.1, the output, \( y \), can be written as:

\[
\begin{align*}
y(0) &= Cx(0) + Du(0) \\
y(1) &= Cx(1) + Du(1) \\
&= CAx(0) + CBu(0) + Du(1) \\
y(2) &= CA^2x(0) + CABu(0) + CBu(1) + Du(2) \\
&\vdots \\
y(k) &= CA^kx(0) + CA^{k-1}Bu(0) + \cdots + CBu(k-1) + Du(k)
\end{align*}
\]

(3.29)

This makes it clear that the output signal, \( y \), is a linear combination of the system states, \( x \), and the input signal, \( u \). Equation 3.29 can be rewritten in matrix form as:

\[
\begin{bmatrix}
y(0) & y(1) & \cdots & y(N-1) \\
y(1) & y(2) & \cdots & y(N) \\
\vdots & \vdots & \ddots & \vdots \\
y(i-1) & y(i) & \cdots & y(N+i-2)
\end{bmatrix}
= \begin{bmatrix}
C \\
CA \\
\vdots \\
CA^{i-1}
\end{bmatrix}
\begin{bmatrix}
x(0) \\
x(1) \\
\vdots \\
x(N-1)
\end{bmatrix}
+ \begin{bmatrix}
D & 0 & \cdots & 0 \\
CB & D & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
CA^{i-2}B & CA^{i-2}B & \cdots & D
\end{bmatrix}
\begin{bmatrix}
u(0) & u(1) & \cdots & u(N-1) \\
u(1) & u(2) & \cdots & u(N) \\
\vdots & \vdots & \ddots & \vdots \\
u(i-1) & u(i) & \cdots & u(N+i-2)
\end{bmatrix}
\]

(3.30)

The left hand side of Equation 3.30 is simply the Hankel matrix for the output \( y \), while the right-hand side comprises the Observability matrix, a vector of system states, the Hankel matrix for the input \( u \) and a Toeplitz matrix of the impulse response. Therefore, Equation 3.30 can be rewritten as:

\[
Y_{0,i,j} = \Gamma_{i}X_{0,N} + H_{i}U_{0,i,N}
\]

(3.31)

where
\[ X_{i,N} = \begin{bmatrix} x(i) & x(i+1) & \cdots & x(i+N-1) \end{bmatrix} \]

It is evident from Equation 3.31 that the output contains two parts: the first part is the zero-input response, \( \Gamma_i X_{0,N} \), due to the system’s internal states; the second part is the zero-state response, \( H_i U_{0,i,N} \), due to the external input. The first step of MOESP is to estimate a subspace containing information only from the zero-input response. This is achieved using orthogonal projection by LQ factorization. In particular, the input and output Hankel matrices are stacked into a tall matrix, and then LQ factorization is applied to give

\[
\begin{bmatrix} U_{0,i,N} \\ Y_{0,i,N} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}
\]

(3.32)

Therefore,

\[
U_{0,i,N} = L_{11} Q_1 \\
Y_{0,i,N} = L_{21} Q_1 + L_{22} Q_2
\]

(3.33)

Furthermore, \( Q_1 \) and \( Q_2 \) are orthogonal, so:

\[
Y_{0,i,N} Q_2^T = L_{21} Q_1 Q_2^T + L_{22} Q_2 Q_2^T = L_{22}
\]

(3.34)

Multiplying both sides of Equation 3.31 by \( Q_2^T \) gives:

\[
Y_{0,i,N} Q_2^T = \Gamma_i X_{0,N} Q_2^T + H_i U_{0,i,N} Q_2^T
\]

\[
= \Gamma_i X_{0,N} Q_2^T + H_i L_{11} Q_1 Q_2^T
\]

\[
= \Gamma_i X_{0,N} Q_2^T
\]

(3.35)

Combining Equations 3.34 and 3.35 gives:
This shows that the column space of $\mathbf{L}_{22}$ contains information only from the internal states. Applying SVD to $\mathbf{L}_{22}$ gives:

$$\mathbf{L}_{22} = \Gamma_{\mathbf{L}_0} \mathbf{X}_{\mathbf{L}_0} \mathbf{Q}_{\mathbf{L}_0}^T$$

(3.36)

The base of the column space of $\mathbf{L}_{22}$ can be determined from the $\mathbf{U}$ matrix, by determining the number of non-zero singular values in $\mathbf{S}$. This also determines the order of the system. However, because of computational errors and noise, the singular values will never be exactly zero. Consequently, the order of the system must be chosen by grouping the singular values into a set of “large” singular values, associated with the system modes, and a set of “small” singular values arising from noise and computational error. It is expected that there will be an obvious gap between these “large” and “small” singular value sets [2]. Figure 3.3 shows the singular values for the simulated sample system. A large gap is evident between the fifth and sixth singular values giving an estimated order for the system of five, which is correct.
**Figure 3.3:** System order estimation for a simulated, fifth order system with no noise. The x axis is the singular values number while the y axis is its value. Note the separation between the “large” eigenvalues (1-5) and “small” eigenvalues indicating that the system is of order 5.
Once the system order, $n$, is known, $\hat{\Gamma}$ is given by the column space of $U$ in Equation 3.37, that is by the first $n$ columns of $U$. So, using Matlab notation, it is

$$\hat{\Gamma} = U(:,1:n)$$

(3.38)

The $C$ matrix is given by the first $l$ rows of $\hat{\Gamma}$,

$$\hat{C} = \hat{\Gamma}(1:l,:)$$

(3.39)

where

$$l$$ is the number of the output signals.

The $\hat{A}$ matrix is given by

$$\hat{A} = U_1^\dagger U_2$$

(3.40)

where

$$U_1 = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-2} \end{bmatrix}$$ is the top $(i-1)l$ rows of $\hat{\Gamma}$

$$U_2 = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{bmatrix}$$ is the bottom $(i-1)l$ rows of $\hat{\Gamma}$.

The eigenvalues of the $\hat{A}$ matrix represent the poles of the estimated system. Therefore, they can be used to validate the estimate by comparing them to the actual poles of the simulated system. Figure 3.4 shows that the poles the simulated and estimated systems coincide almost exactly demonstrating that $\hat{A}$ is a good estimate of $A$. Thus, the
subspace method estimated the system’s poles accurately.

Figure 3.4: Comparison of the poles of the simulated system and those estimated using MOESP. The blue curve represents the unit circle.
3.3.2 Estimate the Matrices $B \& D$

The system output is linear in the $B \& D$ matrices

$$y(k) = CA^k x(0) + \sum_{\tau=0}^{k-1} CA^{k-\tau}Bu(\tau) + Du(k)$$  \hspace{1cm} (3.41)

This equation can be rewritten using the Kronecker product $\otimes$.

$$y(k) = CA^k x(0) + \left[ \sum_{\tau=0}^{k-1} u(\tau)^T \otimes CA^{k-\tau} \right] vec(B) + \left[ u(k)^T \otimes I_r \right] vec(D)$$  \hspace{1cm} (3.42)

where the $vec(\cdot)$ terms result from stacking the columns of the matrix $\cdot$ on top of each other. Equation 3.42 defines the output signal at time, $k$. Therefore constructing a vector containing the output signal $y(k)$ for samples 1 to $k$ gives:

$$Y_{0,N,1} = \begin{bmatrix} \Gamma_N & E_y & E_a \end{bmatrix} \begin{bmatrix} x_0 \\ vec(B) \\ vec(D) \end{bmatrix}$$  \hspace{1cm} (3.43)

where

$\Gamma_N$ is the Observability matrix,

$$Y_{0,N,1} = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix}$$  \hspace{1cm} (3.44)
\[
E_y = \begin{bmatrix}
0 \\
\mathbf{u}(0)^T \otimes \mathbf{C} \\
\vdots \\
\sum_{\tau=0}^{N-2} \mathbf{u}(\tau)^T \otimes \mathbf{CA}^{N-2-\tau}
\end{bmatrix} \tag{3.45}
\]

\[
E_u = \begin{bmatrix}
\mathbf{u}(0)^T \otimes \mathbf{I}_l \\
\mathbf{u}(1)^T \otimes \mathbf{I}_l \\
\vdots \\
\mathbf{u}(N-1)^T \otimes \mathbf{I}_l
\end{bmatrix} \tag{3.46}
\]

Consequently, the matrices \( \mathbf{B} \) & \( \mathbf{D} \) can be estimated by solving the least squares problem:

\[
\begin{bmatrix}
\mathbf{x}_0 \\
\text{vec}(\mathbf{B})
\end{bmatrix}
= \begin{bmatrix}
\mathbf{E}_y \\
\mathbf{E}_u
\end{bmatrix}
\begin{bmatrix}
\Gamma_N \\
\mathbf{y}_{0,N,l}
\end{bmatrix} \tag{3.47}
\]

The estimates obtained for \( \mathbf{B} \) and \( \mathbf{D} \) for the example system are:

\[
\hat{\mathbf{B}} = \begin{bmatrix}
0.23 \\
0.22 \\
0.04 \\
-0.07 \\
-0.02
\end{bmatrix}
\]

and

\[
\hat{\mathbf{D}} = [-0.04]
\]

### 3.3.3 Validation

We used the SMI toolbox [4] to estimate the system matrices. To compare the estimated simulated systems, we simulated the responses of the estimated and true systems to a new input and compared the outputs. Figure 3.5 shows a segment of
estimated and simulated outputs and indicates that they were very similar. Differences between the two signals were quantified in terms of the Variance Accounted For (\%VAF) defined as

\[
\%\text{VAF} = \left( 1 - \frac{\text{variance}(y - y_{\text{est}})}{\text{variance}(y)} \right) \times 100\%
\]

where

\( y \) is the simulated output,

\( y_{\text{est}} \) is the estimated output.

The \%VAF between the estimated and simulated output was 100\%.
Figure 3.5: Comparison of outputs from the simulated and estimated systems
3.4 Extended MOESP Algorithms

The previous section demonstrated the use of the ordinary MOESP algorithm to estimate a state space model for a linear system. However, the ordinary MOESP algorithm was developed to estimate systems from noise-free data, which only occurs for computer simulations. Experimental data will always have some noise so the internal states and output signals will contain both measurement and process noise. Therefore, Equation 3.31 becomes:

\[ Y_{0,i,j} = \Gamma_i X_{0,N} + H_i U_{0,i,N} + W_{i,j,N} \]  

(3.49)

where

\[ W_{i,j,N} \] is a Hankel matrix containing the output measurement noise.

Ordinary MOESP does not consider the presence of noise and so can be expected to provide biased results when noise is present. To examine these effects, we repeated the simulation of the example system, with the addition of white noise whose amplitude was selected to generate a Signal-to-Noise Ratio (SNR) of 20dB. The SNR is defined as:

\[ \text{SNR (dB)} = 20 \log_{10} \left( \frac{\text{RMS}_{\text{signal}}}{\text{RMS}_{\text{noise}}} \right) \]  

(3.50)

where

RMS is the root mean square amplitude.

The system matrices, \( A, B, C, D \), were estimated using the MOESP algorithm. Figure 3.6 (A) shows the estimated order of the system. The largest gap between the eigenvalues was between the second and the third. Therefore, ordinary MOESP would estimate the
system order to be two rather than the actual order of five. To evaluate the quality of this model prediction we compared the outputs from the simulation model and the estimated system for another realization of Gaussian, white input. No noise was added to the simulation in order to compare the estimated output with the noise-free simulated output. Figure 3.7A shows that there are evident differences between the two signals and the %VAF between them was 83%. Clearly, the ordinary MOESP does not provide accurate estimates in the presence of noise.
Figure 3.6: Estimated systems order for a fifth order system with 20 dB measurement noise. (A) Order estimated using the ordinary MOESP. (B) Order estimated using PI-MOESP.
Figure 3.7: Comparison of noise-free output from the simulated and system estimated with 20 dB measurement noise. The system was identified using noisy measurement but comparisons are from noise-free, simulated outputs. (A) Ordinary MOESP (B) PI-MOESP.
MOESP can be extended to eliminate the effects of noise and get unbiased estimate by using Instrumental Variables (IV). The IV, $\Theta$, must be uncorrelated with the noise [17].

$$\lim_{N\to\infty} \frac{1}{N} W_{i,j,N} \Theta^T = 0$$

(3.51)

but $\Theta$ must be correlated with the internal states so that

$$\lim_{N\to\infty} \frac{1}{N} \Gamma_i \Theta^T \text{ has full rank}$$

(3.52)

And $\hat{\Gamma}$ can be recovered. If these properties hold, projecting the input and output data onto the IV will eliminate the noise without disturbing the column space of $\Gamma_i$. For example, the past input can be used as IV to eliminate measurement noise [17]. In practice, the projection onto the IV is performed using LQ factorization. First, the signals are transformed into the Hankel matrices $U_{i,j,N}$ for input, $U_{0,i,N}$ for IV and $Y_{i,j,N}$ for output. Then a tall matrix is formed by stacking $U_{0,i,N}$ between $U_{i,j,N}$ and $Y_{i,j,N}$ and subjected to LQ factorization:

$$
\begin{bmatrix}
    U_{i,j,N} \\
    U_{0,i,N} \\
    Y_{i,j,N}
\end{bmatrix} =
\begin{bmatrix}
    L_{11} & 0 & 0 \\
    L_{21} & L_{22} & 0 \\
    L_{31} & L_{32} & L_{33}
\end{bmatrix}

\begin{bmatrix}
    Q_1 \\
    Q_2 \\
    Q_3
\end{bmatrix}
$$

(3.53)

where

$Q_1$ is the same size as $U_{i,j,N}$;

$Q_2$ is the same size as $U_{0,i,N}$;

$Q_3$ is the same size as $Y_{i,j,N}$.
$Q_i$ is orthogonal, that is
\[
\begin{cases} 
Q_i Q_i^T = I & \text{if } i = j \\
Q_i Q_j^T = 0 & \text{if } i \neq j
\end{cases}
\]

From Equation 3.53, we have:
\[
Y_{i,j,N} Q_2^T = L_{31} Q_i Q_2^T + L_{32} Q_2 Q_2^T + L_{33} Q_3 Q_2^T
\]
\[
= L_{32}
\]

(3.54)

Multiplying both sides of Equation 3.49 by $Q_2^T$ gives:
\[
Y_{i,j,N} Q_2^T = \Gamma_j X_{i,N} Q_2^T + H_j U_{i,j,N} Q_2^T + W_{i,j,N} Q_2^T
\]
\[
= \Gamma_j X_{i,N} Q_2^T + H_1 L_{11} Q_i Q_2^T + W_{i,j,N} Q_2^T
\]
\[
= \Gamma_j X_{i,N} Q_2^T + W_{i,j,N} Q_2^T
\]

(3.55)

Because $w(k)$ is uncorrelated with the input, we have:
\[
\lim_{N \to \infty} U_{i,j,N} W_{i,j,N}^T = \lim_{N \to \infty} L_{41} Q_1 W_{i,j,N}^T = 0
\]

(3.56)

Since the input is persistently exciting, the matrix $L_{11}$ is invertible. Multiplying both sides of Equation 3.56 with the inverse of $\lim_{N \to \infty} L_{11}$ gives:
\[
Q_1 W_{i,j,N}^T = 0
\]

(3.57)

Similarly, since the instrumental variable, $U_{0,i,N}$, is uncorrelated with $W_{i,j,N}$, we have:
\[
\lim_{N \to \infty} U_{0,i,N} W_{i,j,N}^T = \lim_{N \to \infty} (L_{21} Q_1 + L_{22} Q_2) W_{i,j,N}^T
\]
\[
= \lim_{N \to \infty} L_{22} Q_2 W_{i,j,N}^T
\]
\[
= 0
\]

(3.58)

Since $L_{22}$ is invertible, it is clear that
\[
Q_2 W_{i,j,N}^T = 0
\]

(3.59)
From Equations 3.54, 3.55 and 3.59, it can be concluded that the columns of $L_{32}$ span the subspace from the internal states.

$$L_{32} = \Gamma_j X_{\ell,n} Q^T_2$$  \hspace{1cm} (3.60)

Once this subspace has been estimated, the system matrices, $A, B, C, D$, can be obtained in the same way as for the noise-free case. To test the performance of the algorithm, we first examined the order of the system estimated in this way. Figure 3.6 (B) shows the eigenvalue; the largest difference now occurs after the fifth eigenvalue indicating, correctly, that it is of order five. Figure 3.7 (B) compares the simulated and estimated noise-free system outputs to a new realization of a white input signal; the two singles are very similar; the $\%$VAF between them was 99%. Evidently, PI-MOESP accurately estimates systems dynamics where there is white measurement noise.

MOESP has been extended in a number of ways by using different signals as the instrumental variable. Thus:

- PI-MOESP uses the past input as the IV for to estimate systems containing both measurement and process noise [17].

- EIV-MOESP (Errors-In-Variable) [18] uses both the past input and the past output as IVs to identify systems operating in closed loop.

This thesis uses the instrumental variable formulation of MOESP extensively. Chapter 4 uses PI-MOESP to estimate a state space model for the ankle joint stiffness in open loop. In Chapter 5, EIV-MOESP is used to estimate ankle joint stiffness from data acquired in closed-loop. Finally, in Chapter 6 another MOESP [19] structure is used to
estimate a state space model of ankle stiffness from ensemble data.
References:


International Conference, New York, NY, pp 296--299, 2006


[16]. Matlab, Mathworks Inc.


In this chapter we present a state space model for ankle joint stiffness and a subspace method to estimate this state space model from position and torque measurements. We describe the methods behind the algorithm and present simulation and experimental studies that validate the algorithm. We summarize the contribution of this paper and discuss its applications.

Y. Zhao, D. T. Westwick, R. E. Kearney

To be submitted to:
IEEE Transactions on Biomedical Engineering
4.1 Abstract

Joint stiffness, the dynamic relationship between the angular position of a joint and the torque acting about it, describes the dynamic, mechanical behavior of a joint during posture and movement. Joint stiffness arises from both intrinsic and reflex mechanisms but their torques cannot be measured separately experimentally since they appear and change together. Therefore, the direct estimation of intrinsic and reflex stiffnesses is difficult. In this paper, we present a new, two-step procedure to estimate the intrinsic and reflex components of ankle stiffness. In the first step, a discrete-time, subspace-based method is used to estimate a state space model for overall stiffness from the measured overall torque and predict the intrinsic and reflex torques. In the second step, continuous-time models for intrinsic and reflex stiffnesses are estimated from these predicted intrinsic and reflex torques. Compared with current identification methods, the new subspace-based algorithm has three advantages: It does not require iteration, it provides better estimates under the most severe estimation conditions (e.g. high noise, short sample lengths) and it can be extended to closed-loop system identification problems. Simulations and experimental results demonstrate that the algorithm estimates the intrinsic and reflex stiffnesses accurately.

Keywords—ankle dynamics, subspace method, Parallel cascade structure, Hammerstein system identification.
4.2 Introduction

The concept of dynamic joint stiffness can be used to study the mechanical behavior of the mechanisms acting about the ankle. Ankle joint stiffness defines the dynamic relationship between the angular position of the ankle and the torque acting about it. Ankle joint stiffness plays an important role in the control of posture, where it defines the mechanical response to external perturbations, and of movement, where it determines the forces that must be generated to move the limb.

Ankle joint stiffness can be separated into intrinsic and reflex components and described by a parallel cascade model [1], as shown in Figure 4.1. Intrinsic stiffness arises from the mechanical properties of the joint, passive tissue, and active muscle fibers. For small perturbations about a fixed operating point, defined by the mean position and level of voluntary activation, intrinsic stiffness can be described by a linear, dynamic relationship that is modeled well by a second-order, quasi-linear system with transfer function:

\[
H_{IS}(s) = \frac{TQ_I(s)}{POS(s)} = Is^2 + Bs + K
\]  

(4.1)

where

- \( TQ_I \) is the intrinsic torque,
- \( POS \) is the position,
- \( I, B, \) and \( K \) are inertial, viscous, and elastic parameters that vary with the operating point.
Figure 4.1: Parallel cascade model for dynamic ankle stiffness. $pos$ denotes the position. $tq_I$, $tq_R$ and $tq_N$ denote intrinsic torque, reflex torque and net torque. $z$ denotes the output from the static nonlinearity. Intrinsic stiffness is modeled as a linear system. Reflex stiffness is modeled as a delay, followed by a differentiator and then a Hammerstein system comprising the series connection of a static nonlinearity and a linear system.
Reflex stiffness arises from muscle contraction in response to reflex activation from stretch receptors in the muscle. At the ankle, reflex stiffness can be modeled with a Linear-Nonlinear-Linear (LNL) block-structured model, comprising the series connection of a differentiator, a delay of about 40 ms [1], a static non-linearity and a 2nd or 3rd order low-pass system, whose transfer functions are given in Equations 4.2 and 4.3, respectively.

\[
H_{RS}(s) = \frac{TQ_{R}(s)}{VL(s)} = \frac{g_R \omega_n^2}{(s^2 + 2\zeta \omega_n s + \omega_n^2)} e^{-s\tau} \quad (4.2)
\]

\[
H_{RS}(s) = \frac{TQ_{R}(s)}{VL(s)} = \frac{g_R \omega_n^2 \rho}{(s^2 + 2\zeta \omega_n s + \omega_n^2)(s + \rho)} e^{-s\tau} \quad (4.3)
\]

where

- \(TQ_{R}\) is the reflex torque,
- \(VL\) is the half-wave rectified joint angular velocity,
- \(g_R\) is the reflex gain,
- \(\omega_n\) is the 2nd order natural frequency,
- \(\zeta\) is the damping parameter,
- \(\rho\) is the 1st order cut-off frequency,
- \(\tau\) is the reflex delay.

Intrinsic and reflex torques cannot be measured separately experimentally; only their sum can be measured. Consequently, intrinsic and reflex stiffnesses cannot be estimated directly. To overcome this difficulty, a number of approaches have been
developed to separate the torques due to each pathway.

One approach is to compare the responses before and after eliminating the reflexes experimentally by surgical differentiation [2], applying a neural block to [3], or electrically stimulating the primary afferents [4] [5]. The differences between the two responses are attributed to the contributions from the reflex stiffness. These methods are only reliable if the operating points before and after eliminating reflexes are matched closely to ensure that intrinsic mechanics do not change. However, in practice, this is difficult to achieve since eliminating the reflex contribution by itself changes the system.

Another approach is to use system identification methods to separate the intrinsic and reflex torques analytically. Thus, the parallel cascade method, developed in our laboratory [6] [7], estimates the intrinsic stiffness and reflex stiffness from the measured position and net torque using non-parametric, non-linear identification methods. Intrinsic stiffness is estimated first, taking advantage of the stretch reflex delay to eliminate any contributions from reflex mechanisms. Then, reflex stiffness is estimated after the torques due to intrinsic stiffness have been removed. The algorithm repeats until it converges. Zhang and Rymer [8] estimated the intrinsic and reflex components of joint stiffness at the knee using a parametric method that involved solving a set of nonlinear delay differential equations describing the intrinsic and reflex stiffnesses. We developed a method [9] that achieves the same end in real-time by using a specially designed input sequence whose properties eliminate the correlation between intrinsic and reflex torques. Intrinsic stiffness is estimated from the cross-correlation between this perturbation and the measured torque. Then, reflex stiffness is estimated after the torques
The parallel cascade [6] [7] and real-time [9] methods estimate the intrinsic and reflex stiffness efficiently under some circumstances, but they do have limitations. In particular, the parallel cascade method is an iterative algorithm that cannot be guaranteed to converge to the globally optimal solution. Furthermore, both the parallel cascade method and real-time algorithms use correlation-based identification methods that are suitable for open-loop system identification but cannot be applied or extended to closed-loop problems [10]. This is important because limbs interact with compliant loads during most functional movements. As a result, any torque change will be fed back to change the joint position. This means that position and torque are measured inside a closed-loop and as a result the open-loop methods used by the parallel cascade and real-time algorithms will give biased estimates [11]. Consequently, there is a need for an analytic method to estimate intrinsic and reflex stiffnesses that can be extended to closed-loop identification.

In this paper, we develop and validate a subspace-based method to estimate the intrinsic and reflex torques from open-loop measurements of overall torque. Compared with current identification methods, the new subspace-based algorithm has three advantages:

1. It does not require iteration.
2. It provides better estimates under the most severe estimation conditions (e.g. high noise, short sample lengths).
3. It can be extended to closed-loop system identification problems.

The paper is developed as follows: Section II reviews subspace algorithms for linear system identification. Section III develops a new subspace algorithm to estimate
intrinsic and reflex torque from the measured net torque. Section IV presents the method used to estimate continuous time models of intrinsic and reflex stiffness from the estimated intrinsic and reflex torques. Sections V and VI describe applications of the new method to simulated and experimental data, respectively. Section VII discusses the results, describes areas for future work, and summarizes the contributions of this research.

4.3 Subspace Method

The subspace method developed in this paper uses the PI-MOESP (Past Input as instrumental variable of Multivariable Output-Error State-space) algorithm [12] [13] which is reviewed briefly here. Consider systems of the form

\[
\begin{align*}
\mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{u}(k) \\
\mathbf{y}(k) &= C\mathbf{x}(k) + D\mathbf{u}(k) + \mathbf{w}(k)
\end{align*}
\]  

(4.4)

where

\(\mathbf{u}(k) \in \mathbb{R}^m\) and \(\mathbf{y}(k) \in \mathbb{R}^l\), are vectors containing measurements, at discrete time \(k\), of the \(m\) inputs and \(l\) outputs of the process.

\(\mathbf{w}(k) \in \mathbb{R}^l\) is an additive, zero mean, noise signal that is uncorrelated with the input.

\(\mathbf{x}(k) \in \mathbb{R}^n\), is the state vector of the process at discrete time \(k\) containing the values of \(n\) states, where \(n\) is the order of the system.

\(A \in \mathbb{R}^{n \times n}\), is the system matrix that describes the dynamics of the system.

\(B \in \mathbb{R}^{n \times m}\), is the input matrix that describes how deterministic inputs influence the
states.

\( C \in \mathbb{R}^{m \times n} \), is the output matrix that describes how the internal states are transformed to generate the output \( y_k \).

\( D \in \mathbb{R}^{b \times n} \), is the direct feed through term.

Subspace methods estimate the system matrices from measurements of inputs and outputs. Block Hankel matrices play an important role in MOESP; they have the general form, for a signal, \( r \), of:

\[
R_{i,j,N} = \begin{bmatrix}
  r(i) & r(i+1) & \cdots & r(i+N-1) \\
  r(i+1) & r(i+2) & \cdots & r(i+N) \\
  \vdots & \vdots & \ddots & \vdots \\
  r(i+j-1) & r(i+j) & \cdots & r(i+j+N-2)
\end{bmatrix}
\]

(4.5)

where

\( i \) is the left upper entry of the Hankel matrix,

\( j \) is the number of the rows,

\( N \) is the number of columns.

Equation 4.1 can be rewritten using Hankel matrices formed from samples of the system inputs and outputs as

\[
Y_{i,j,N} = \Gamma_j X_{i,N} + H_j U_{i,j,N} + W_{i,j,N}
\]

(4.6)

where

\( Y_{i,j,N} \) is the Hankel matrix constructed from output \( y(k) \),

\( U_{i,j,N} \) is the Hankel matrix constructed from output \( u(k) \),
\[
\Gamma_j = \begin{bmatrix}
  C \\
  CA \\
  \vdots \\
  CA^{j-1}
\end{bmatrix}
\]
is the extended Observability matrix \cite{14},
\[
\begin{bmatrix}
  D & 0 & 0 & 0 \\
  CB & D \\
  \vdots & \ddots & \ddots & \ddots \\
  CA^{j-2}B & CA^{j-3}B & \cdots & CB & D
\end{bmatrix}
\]
is a Toeplitz matrix of impulse response elements \cite{14},
\[
X_{i,N} \text{ contains the internal states } X_{i,N} = \begin{bmatrix} x(i) & x(i+1) & \cdots & x(i+N-1) \end{bmatrix},
\]
\[
W_{i,j,N} \text{ is the Hankel matrix constructed from the noise } w(k).
\]

It is evident from Equation 4.6 that the system’s output contains three parts. The first part is the zero-input response, \( \Gamma_j X_{i,N} \), which comes from the system’s internal states. The second part is the zero-state response, \( H_j U_{i,j,N} \), which comes from the external input. The last part is the noise term, \( W_{i,j,N} \). Thus, the extended Observability matrix, \( \Gamma_j \), can be estimated by finding the subspace of the output that contains only the zero-input response. To achieve this, PI-MOESP uses the past input as an instrumental variable \cite{12} to eliminate the noise term, \( W_{i,j,N} \), and then uses an orthogonal projection \cite{15} to remove the zero-state response, \( H_j U_{i,j,N} \).

Specifically, the matrix, \[
\begin{bmatrix}
  U_{i,j,N} \\
  U_{0,i,N} \\
  Y_{i,j,N}
\end{bmatrix},
\]
where \( U_{i,j,N} \), \( U_{0,i,N} \) and \( Y_{i,j,N} \) are Hankel
matrices, is constructed by stacking the instrumental variable, \( U_{0,i,N} \), between the input and the output Hankel matrices. Applying LQ factorization [15] to

\[
\begin{bmatrix}
U_{i,j,N} \\
U_{0,i,N} \\
Y_{i,j,N}
\end{bmatrix}
= \begin{bmatrix}
L_{11} & 0 & 0 \\
L_{21} & L_{21} & 0 \\
L_{31} & L_{32} & L_{33}
\end{bmatrix}
\begin{bmatrix}
Q_1 \\
Q_2 \\
Q_3
\end{bmatrix}
\]

(4.7)

where

\( Q_1 \) is the same size as \( U_{i,j,N} \),

\( Q_2 \) is the same size as \( U_{0,i,N} \),

\( Q_3 \) is the same size as \( Y_{i,j,N} \).

\( Q_i \) is orthogonal, that is

\[
Q_i Q_i^T = I \quad \text{if } i=j
\]

\[
Q_i Q_j^T = 0 \quad \text{if } i \neq j
\]

From Equation 4.7 it can be shown that [12]

\[
L_{32} = \Gamma_j X_{i,N} Q_2^T
\]

(4.8)

which demonstrates that columns of \( L_{32} \) span the subspace from the internal states [12].

The SVD of matrix \( L_{32} \) is

\[
L_{32} = \mu \xi \nu^T
\]

(4.9)

So \( \hat{\Gamma} \) is given by the column space of \( \mu \) in Equation 4.9, that is by the first \( n \) columns of \( \mu \) where \( n \) is the order of the system. Once \( \hat{\Gamma} \) is available, the system...
matrices, \( A B C D \), can be estimated using standard subspace methods [12].

### 4.4 Estimation of Intrinsic and Reflex Torque

#### 4.4.1 State Space Model for Intrinsic Stiffness

For perturbations about an operating point, intrinsic stiffness can be modeled by a linear dynamic relationship between position and torque that is described well by the second-order, quasi-linear system of Equation 4.4. This quasi-linear transfer function has no poles, so that the relationship between the position and torque can be described in the time domain as:

\[
\begin{pmatrix}
    t_q(t)
\end{pmatrix} =
\begin{pmatrix}
    p(t) & \text{vel}(t) & \text{acc}(t)
\end{pmatrix}
\begin{bmatrix}
    K \\
    B \\
    I
\end{bmatrix}
\]  

(4.10)

where

- \( t_q(t) \) is the intrinsic torque,
- \( p(t) \) is measured joint position,
- \( \text{vel}(t) \) is the differentiated position,
- \( \text{acc}(t) \) is second-order differentiated position.

A generic, discrete-time, state space model relating the inputs, \( U_i(k) = [p(k) \text{ vel}(k) \text{ acc}(k)] \), to the output, \( t_q_i(k) \), is

\[
\begin{align*}
    x_i(k+1) &= A_i x_i(k) + B_i U_i(k) \\
    t_q_i(k) &= C_i x_i(k) + D_i U_i(k)
\end{align*}
\]  

(4.11)

where
$tq_i(k)$ is the intrinsic torque at discrete time $k$,

$U_i(k)$ represents a set of constructed inputs for intrinsic stiffness given by

$U_i(k) = [p(k) \ vel(k) \ acc(k)]$.

$x_i(k)$ is the state vector.

The matrices, $A_i$, $B_i$ and $C_i$, will be zero, and so the model reduces to:

$tq_i(k) = D_i U_i(k) \quad (4.12)$

### 4.4.2 State Space Model for Reflex Stiffness

Reflex stiffness can be modeled with an LNL structure comprising a delay of 40 ms, a differentiator, a static non-linearity, and a linear dynamic system as shown in Figure 4.1. Alternatively, the reflex stiffness can be viewed as a Hammerstein system whose input is the delayed and differentiated position. This Hammerstein system can be identified using an extended subspace method [16] as follows.
Figure 4.2: Transformation of a SISO Hammerstein system to a MISO linear system. A) The SISO Hammerstein system B) The equivalent MISO linear system formed using the terms of the basis function $g_i(\bullet)$ to generate the set of constructed inputs, $U = \left[ g_1(u), \cdots, g_p(u) \right]^T$. 
Assume the static nonlinearity $n_{ns}()$ can be approximated by a basis expansion $g()$, so that the output of the nonlinearity, $z(k)$ is

$$z(k) = g(dvel(k), \alpha) = \sum_{i=1}^{p} \alpha_i g_i(dvel(k))$$

$$= [\alpha_1 \; \cdots \; \alpha_p] \begin{bmatrix} g_1(dvel(k)) \\ \vdots \\ g_p(dvel(k)) \end{bmatrix} \tag{4.13}$$

where

- $g_i(\cdot)$ are the basis function terms,
- $\alpha_i$ are the scale factors for each basis function term,
- $dvel(k)$ is the input to the nonlinearity, that is the delayed velocity.
- $z(k)$ is the output from the nonlinearity.

The linear dynamics of the reflex stiffness can be described by the state space model.

$$x_R(k+1) = A_R x_R(k) + B_R z(k)$$

$$tq_R(k) = C_R x_R(k) + D_R z(k) \tag{4.14}$$

where

- $tq_R(k)$ is the reflex torque at discrete time $k$,
- $z(k)$ is the output from the static nonlinearity,
- $x_R(k)$ is the state vector for the linear part of the reflex stiffness,
- $A_R, B_R, C_R, D_R$ are the system matrices for the linear dynamic block of the Hammerstein system.

Equation 4.13 makes it evident that the output from the nonlinearity is the product of a
row vector containing the parameters of the nonlinearity, \([\alpha_1 \ldots \alpha_p]\), and a column vector containing the terms of the basis function, \[
\begin{bmatrix}
g_1(dvel(k)) \\
\vdots \\
g_p(dvel(k))
\end{bmatrix}
\]. If we define:

\[
\hat{B} = \begin{bmatrix} B_R \alpha_1, & \ldots & B_R \alpha_p \end{bmatrix}
\]

\[
\hat{D} = \begin{bmatrix} D_R \alpha_1, & \ldots & D_R \tau_p \end{bmatrix}
\]

\[
U_R(k) = \begin{bmatrix} g_1(dvel(k)), & \ldots & g_p(dvel(k)) \end{bmatrix}^T
\]

then the Hammerstein system can be rewritten as

\[
x_R(k+1) = A_R x_R(k) + \hat{B} U_R(k)
\]

\[
tq_R(k) = C_R x_R(k) + D_R U_R(k)
\]

Thus, once a basis function has been chosen for the static nonlinearity, the SISO (Single Input and Single Output) Hammerstein system can be described by a MISO (Multiple Input and Single Output) state space model with inputs

\[
U_R(k) = \begin{bmatrix} g_1(dvel(k)), & \ldots & g_p(dvel(k)) \end{bmatrix}^T,
\]

as shown in Figure 4.2. The system matrices \([\hat{A}_R \hat{B} \hat{C}_R \hat{D}]\) can be estimated using subspace methods. We opted to use Chebyshev polynomials as the basis functions to describe the nonlinearity since they avoid the conditioning problems [17] associated with the higher order components of regular polynomials. The constructed input matrix becomes:

\[
U_R(k) = \begin{bmatrix} c_1(k) & c_2(k) & \ldots & c_p(k) \end{bmatrix}^T
\]

where the Chebyshev polynomials are given by:
\begin{align}
ch_1(k) &= 1 \\
ch_2(k) &= dvel(k) \\
&\vdots \\
ch_p(k) &= 2 \cdot dvel(k) \cdot ch_{p-1}(x) - ch_{p-2}(x)
\end{align} 
(4.18)

### 4.4.3 State Space Model for Overall Stiffness

The state space models for intrinsic and reflex stiffness cannot be estimated directly since their outputs, \(tq_I(k)\) and \(tq_R(k)\), cannot be measured separately. However, a state space model for the overall parallel cascade model of ankle dynamics can be estimated because the measured torque \(tq_N(k)\) is the sum of the intrinsic and reflex torques (i.e. \(tq_N(k) = tq_I(k) + tq_R(k)\)). Thus, combining Equations 4.11 and 4.16 gives:

\[
\begin{bmatrix}
\dot{x}_R(k+1) \\
tq_N(k)
\end{bmatrix} =
\begin{bmatrix}
A_R & B \\
C_R & D
\end{bmatrix}
\begin{bmatrix}
x_R(k) \\
U_I(k)
\end{bmatrix}
+ \begin{bmatrix}
0 & \hat{B} \\
0 & \hat{D}
\end{bmatrix}
\begin{bmatrix}
U_I(k) \\
U_R(k)
\end{bmatrix}
\] 
(4.19)

where

- \(tq_N(k)\) is the measured net torque,
- \(D_I\) is the system matrix for intrinsic stiffness from Equation 4.11,
- \(A_R, \hat{B}, C_R, \hat{D}\) are the system matrices for reflex stiffness from Equation 4.16,
- \(U_I(k)\) is the set of constructed inputs to intrinsic stiffness from Equation 4.11,
- \(U_R(k)\) is the set of constructed inputs to reflex stiffness from Equation 4.16.

The input, \(\begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix}\), for the overall state space model, can be constructed from the
measured input, while the output, \( t_{q_N}(k) \), is available directly from the measured data. Therefore the state space model of Equation 4.19, can be estimated from input, \[
\begin{bmatrix}
U_f(k) \\
U_R(k)
\end{bmatrix},
\]
to output, \( t_{q_N}(k) \).

The order of the system matrices of the state space model (Equation 4.19) must be determined before obtaining the system matrices \( A_R, \tilde{B}, C_R, \tilde{D} \). This is done by examining the eigenvalues of \( L_{32} \) [12], resulting from the SVD factorization in Equation 4.9. These are expected to have two groups: one group of eigenvalues, with “large” numerical values should be associated with the rank of matrix \( A \); the second group of “small” eigenvalues will arise from noise and computational error. Ideally, there should be an evident gap between these “large” and “small” groups of eigenvalues. However, this gap may not be obvious when the SNR is low. In this case, determining the order of the estimated system is more difficult and may require manual intervention. For our application, the rank of matrix \( A \) in Equation 4.19 will equal the order of the linear dynamic system in the reflex stiffness; previous work has shown this to be described well by a second-order or third-order linear low-pass filter. Thus, the order of the system will be either two or three, depending on which provides the better estimate.

Once the order of the system is determined, the state space model of Equation 4.19 can be estimated. Identifying this state space model does not estimate the intrinsic stiffness and the reflex stiffness directly. However, the estimated state space model can be used to estimate the torques from the intrinsic or reflex stiffness by simulating the model with appropriate inputs. Specifically, simulating the system with the input signal,
\[
\begin{bmatrix}
U_I(k) \\
0
\end{bmatrix}, \text{ will estimate the intrinsic torque, } tq_I(k), \text{ while the response to the input, }
\begin{bmatrix}
0 \\
U_R(k)
\end{bmatrix}, \text{ will estimate the reflex torque, } tq_R(k).
\]

### 4.4.4 Implementation

To summarize, the intrinsic and reflex torques are estimated as follows.

1. Calculate \(vel(k)\) and \(acc(k)\) by numerically differentiating [18] the measured position. Construct the input signals to the intrinsic stiffness, 
\[
U_I(k) = \begin{bmatrix} p(k) & vel(k) & acc(k) \end{bmatrix}.
\]

2. Construct the input signal to the reflex stiffness using the basis functions
\[
U_R(k) = \begin{bmatrix}
ch_1(k) & ch_2(k) & \cdots & ch_p(k)
\end{bmatrix}^T, \text{ defined by Equations 4.12 and 4.15.}
\]

3. Use the PI-MOESP algorithm from the SMI 2.0 Toolbox [19] to estimate the system order and the matrices, \(A, B, C, D\), from the input 
\[
\begin{bmatrix}
U_I(k) \\
U_R(k)
\end{bmatrix}
\]
and the measured overall torque, \(tq_N(k)\).

4. Simulate the estimated system with the input, 
\[
\begin{bmatrix}
U_I(k) \\
0
\end{bmatrix}
\]
to estimate the intrinsic torque, \(tq_I(k)\).

5. Simulate the estimated system with the input, 
\[
\begin{bmatrix}
0 \\
U_R(k)
\end{bmatrix}
\]
to estimate the reflex torque, \(tq_R(k)\).
4.4.5 Continuous Time Models

The subspace method estimates the joint stiffness as a MISO state space model relating the constructed inputs \( \begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix} \) to the measured net torque. This state space model can be used to estimate the intrinsic and reflex torques but provides little physical insight since it is in discrete-time and its structure and parameters are not related directly to underlying neuromuscular mechanisms. Consequently, we thought it useful to estimate continuous-time stiffness models from the intrinsic and reflex torques predicted by the state space model. This is straightforward once the intrinsic and reflex torques have been separated.

Thus, intrinsic stiffness can be described by the second-order, quasi-linear model of Equation 4.4 that relates the intrinsic torque to the position input and operating point dependent inertial \((I)\), viscous \((B)\), and elastic \((K)\) parameters. Kearney et al [1] found that the intrinsic stiffness could be described well by a two-sided, impulse response function (IRF). Therefore, we opted to estimate a two-sided IRF from the measured position and estimated intrinsic torque using the Nonlinear Identification Toolbox [20].

A continuous-time Hammerstein system can be estimated between the velocity and the reflex torque provided by the state space model. To do so, we used a Separable Least Square (SLS) [21] method to estimate the static nonlinearity and the IRF for \( H_{rs}(s) \) from the velocity and the predicted reflex torque.

4.5 Simulation Studies
4.5.1 Simulation Model

To evaluate the performance of the subspace method, we simulated the parallel cascade model of ankle stiffness (Figure 4.1) using Matlab’s Simulink [22] and the model defined by the diagram shown in Figure 4.3. The position input was filtered by a 2nd order, low-pass, Bessel filter with a cutoff frequency of 30 Hz to simulate the actuator dynamics in a real experiment. Intrinsic stiffness torques were simulated by multiplying the position, the velocity and the acceleration by the elastic, viscous and inertial parameters respectively. Reflex torque was simulated by differentiating the position input, delaying it for 40 ms, half-wave rectifying it, and finally filtering it by a second-order, low-pass filter (Equation 4.2). Parameter values used for the simulation were based on those identified experimentally [23] during a torque-matching task and are given in Table 4.1.

The Simulink model produced overall torques that were similar to those observed experimentally. Figure 4.4 shows the position input (Figure 4.4A), and overall torque (Figure 4.4B) recorded in a typical experimental trial. Figure 4.4C shows the torque generated by the simulation using the input of Figure 4.4A. Clearly, the amplitudes and waveforms of the experimental and simulated torques are very similar.
Figure 4.3: Matlab Simulink model for ankle stiffness. A realization of Gaussian white noise was added to the net torque as measurement noise. The noise-free intrinsic and reflex torques are recorded in the simulation for comparison with the estimates from the algorithm.
### Table 4.1. Parameter values used to simulate ankle stiffness

<table>
<thead>
<tr>
<th>Intrinsic parameters</th>
<th>Reflex parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$ (Nm/rad/s$^2$)</td>
<td>$B$ (Nm/rad/s)</td>
</tr>
<tr>
<td>0.0172</td>
<td>1.31</td>
</tr>
</tbody>
</table>
Figure 4.4: Experimental and Simulated Torques. A) Experiment position B) Experimental torque C) Simulated torque generated by the Simulink model shown in Figure 4.3.
Monte-Carlo studies were used to examine the performance of the algorithm. Each Monte-Carlo study comprised 100 simulation trials using Pseudorandom Binary Sequence (PRBS) inputs, each with a peak-to-peak amplitude 0.04 rad and a switching interval of 80 ms; these parameters were similar to those used in previous experimental studies of ankle dynamics [6]. Each simulation trial lasted 50 seconds and used a different realization of the PRBS input. The model was simulated at 1 KHz, and the signals decimated to 100 Hz for analysis purposes. After decimation, a different realization of Gaussian, white noise was added to the output as the measurement noise to generate the required Signal-to-Noise Ratio (SNR) defined as:

\[ SNR(dB) = 20 \log_{10} \left( \frac{RMS_{signal}}{RMS_{noise}} \right) \]  

(4.20)

where

\( RMS \) is the root mean square amplitude.

### 4.5.2 Identification Evaluation

Stiffness dynamics were estimated for each simulation trial using both the subspace and parallel cascade methods. The simulations used a parametric, continuous time model, the parallel cascade method estimates a non-parametric, continuous time model, and the subspace method estimates a discrete time, state space model. Since it is not possible to compare directly these three model structures we chose to evaluate their accuracy in terms of their ability to predict the simulated, noise-free intrinsic, reflex and total torques. This was quantified in terms of the percentage Variance Accounted For (%VAF):
\[
\% \text{VAF} = \left(1 - \frac{\text{var}(tq_s - tq_E)}{\text{var}(tq_s)}\right) \times 100
\]  
(4.21)

where

\(tq_s\) is the simulated, noise free torque

\(tq_E\) is the estimated, noise free torque.

### 4.5.3 Results

#### 4.5.3.1 Additive Noise.

We first examined the robustness of the algorithm in the presence of additive noise by carrying out Monte-Carlo studies using SNRs that varied from -5 db to 20db. The reflex gain was chosen to be 25 so that the reflex torque accounted for 40% of the net torque variance. Six sets of Monte-Carlo studies were conducted, each set using an additive noise with the same SNR.

Figure 4.5 shows the mean %VAF, bracketed by the standard deviation, for the net, intrinsic and reflex torques estimated by the subspace and parallel cascade methods. At high SNRs (>10) both methods performed very well with mean %VAFs greater than 95% for all three signals. The results from the subspace method were slightly more reliable as evidenced by the somewhat higher mean %VAFs and lower standard deviations for the predicted intrinsic and net torques. The estimates of the reflex torques were similar for both methods, as evidenced by the similar values for VAF fit and standard deviations (variance < 0.5%). However, at low SNRs, the subspace method was more reliable. Indeed, at a SNR of -5 dB, only the subspace method provided good results.
consistently; the parallel cascade method failed to converge nearly half the time (45/100) and so provided no useful results.

Figure 4.5: Estimation accuracy as a function of SNR. Each point represents the mean value bracketed by its standard deviation of 100 Monte-Carlo trials. (A) Intrinsic torque; (B) Reflex torque; (C) Net torque.
4.5.3.2 Reflex Gain

Previous studies [24] have shown that the relative contribution of the reflex stiffness to joint stiffness varies substantially from subject to subject. Thus, in one study, the percentage of the net torque variance accounted for by reflex stiffness in different subjects ranged from as little as 5% to as much as 45% [24]. Moreover, in spastic subjects the reflex gain was much greater than it in normal subjects [24]. The nonlinear behavior of ankle stiffness becomes more marked as the reflex gain increases; this could be expected to make the identification more difficult. Consequently, we thought it important to determine how the identification accuracy varied with reflex gain.

Consequently, we carried out Monte-Carlo simulations using a fixed intrinsic stiffness and reflex gains ranging from 5 to 50, with the corresponding reflex torques accounting for 5% to 70% of the total torque. The SNR was set to 13dB corresponding to that seen experimentally.

Figure 4.6 shows the results of these simulations. Both the subspace and parallel cascade methods performed well for low values of reflex gain. Indeed, for reflex gains less than 15, the parallel cascade method performed slightly better than the subspace method, as shown by a slightly higher %VAF and smaller standard deviation. However, the subspace method results became progressively better than the parallel cascade’s results as the reflex gain increased. The difference was most apparent in the reflex torque estimates. The %VAF of the parallel cascade estimate dropped with increasing reflex gain while that for the subspace method stayed almost constant.
Figure 4.6: Estimation accuracy as a function of reflex gain. Each point represents the mean value bracketed by the standard deviation of the 100 Monte-Carlo trials (A) Intrinsic torque; (B) Reflex torque; (C) Net torque.
4.5.3.3 Data Length

Next, we examined how the quality of the identification varied with the length of data set. Ten Monte-Carlo studies were carried out using data lengths ranging from 4 to 11.2 second in 0.8 second increments and the parameter values given in Table 4.1. A Gaussian, white noise, scaled to produce a SNR of 13 dB, was added to the output. Figure 4.7 shows the results for the subspace (left column) and parallel cascade (right column) methods. Both behaved similarly as the record length decreased; the %VAF decreased while the variance, indicated by the size of the error bars, increased. However, the overall performance of the subspace method was superior. The %VAF of its torque predictions was higher at all record lengths and their variance was substantially lower.
**Figure 4.7:** Identification qualities as a function of data length. Each point represents the mean value bracketed by the standard deviation of the 100 Monte-Carlo trials (A) Intrinsic torque; (B) Reflex torque; (C) Net torque for the subspace method. (D) Intrinsic torque; (E) Reflex torque; (F) Net torque for the parallel cascade method.
4.6 Experimental Studies

Next, experimental data were used to validate the algorithm and compare results with those obtained using the parallel cascade algorithm.

4.6.1 Methods

One subject with no history of neuromuscular disorders participated in the study. The subject lay supine with the left foot attached to an electrohydraulic actuator by a custom-made fiberglass boot. Movement of the foot was restricted to plantarflexion and dorsiflexion. The left leg was immobilized with a leather strap and supported with two sand bags under the knee. Two safety stops, one physical and one hydraulic, limited the actuator movement to the subjects’ voluntary range of motion. Ankle position was measured with a potentiometer (BI Technologies 6273), and torque with a torque transducer (Lebow 2110-5K). Both position and torque were defined to be positive for dorsiflexion and negative for plantarflexion.

Position and torque were sampled at 1 kHz using an NI-447 data acquisition card. Anti-aliasing was performed by the card which over-sampled the data and then applied a digital low-pass filter with a 486.3 Hz cutoff frequency. Data were then stored on the Host PC (AMD Athlon 1.33 GHz, 1 GB RAM). The controller and the real-time identification were implemented using xPC Target on the Target PC (AMD Athlon 1.6 GHz, 256 MB RAM). Position and torque were sampled on this machine for actuator control and real-time identification of intrinsic stiffness and reflex gain. Sampling was done at 1 kHz using a ComputerBoards PCIM-DAS1602/16 Analog-to-Digital Card.
Prior to sampling, data were filtered with 8-pole, 6-zero, linear phase, constant delay, low-pass filter with a cutoff of 400Hz (Frequency Devices 9064). The control signal was output with a ComputerBoards PCIM-DDA06/16 Digital-to-Analog Card to the servo-valve which controlled the position of the actuator.

During the experiment, the subject was asked to relax and let the actuator move the ankle. A Pseudorandom Binary Sequence (PRBS) signal was used as the position signal to drive the actuator. The experiment lasted for 60 seconds. The position and torque signals were sampled at 1 kHz and then decimated to 100 Hz. Figure 4.4 shows a segment of the experimental data.

### 4.6.2 Results

The intrinsic torque, reflex torque and net torque were estimated for the subject using both the subspace and parallel cascade methods. Since the intrinsic and reflex torques cannot be measured experimentally, it was not possible to use %VAF as the criterion to compare the estimated and true intrinsic or reflex torques. Therefore, we first computed the %VAF between the estimated and true net torques. Then, the %VAF between the estimate of intrinsic and reflex torque and the measured net torque was computed to quantify the contributions of the intrinsic and reflex torques to the net torque. Table 4.2 shows the %VAF for net torque by the two estimates as well as the %VAF for the intrinsic and reflex results. Note that with the PRBS input signal used in this experiment the intrinsic and reflex torques were not completely uncorrelated. As a result the total %VAF is not equal to the sum of the %VAFs for the intrinsic and reflex torques.
Table 4.2 Estimates from the subspace method and the parallel cascade method for experimental data

<table>
<thead>
<tr>
<th>Method</th>
<th>Intrinsic torque</th>
<th>Reflex torque</th>
<th>Net torque</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subspace</td>
<td>46%</td>
<td>45%</td>
<td>81%</td>
</tr>
<tr>
<td>Parallel cascade</td>
<td>34%</td>
<td>42%</td>
<td>80%</td>
</tr>
</tbody>
</table>
Figure 4.8 shows the intrinsic stiffness, static nonlinearity and reflex stiffness estimated using the two methods. The intrinsic stiffness IRFs estimated by the two methods were very similar (Figure 4.8 A). Superficially, the static nonlinearities (Figure 4.8 B) and the linear part of the reflex stiffness (Figure 4.8 C), identified by the subspace and parallel cascade methods appear quite different. However, in reality, these differences are not significant since the difference is due to differences in the way that the overall gain is distributed between the linear and nonlinear blocks in the two models. In particular, the estimated output of the nonlinearity and the reflex torques from the subspace and parallel cascade methods are shown in Figure 4.9. The output from subspace nonlinearity was larger than that from parallel cascade method. Correspondingly, the IRF of linear part of the reflex stiffness had a smaller magnitude than that from the parallel cascade method. Moreover, the estimates of the reflex stiffness from both methods have similar shapes and the same overall contributions from the velocity to reflex torque. Therefore, estimates from the two methods are both correct.

In summary, the subspace method provides realistic estimates of intrinsic and reflex stiffness when applied to experimental data. The estimates are generally consistent with those obtained by applying the parallel cascade method to the same data set although there are minor differences in the results given by the two methods. Given that the “true” dynamics are not known it is not possible to determine which method gives the correct results. However, based on the simulation results, it appears that the state space method is likely to be more robust to the effects of noise and hence more likely to be correct.
Figure 4.8: Estimated non-parametric models for intrinsic and reflex stiffness for subject from the subspace method; (solid line) and from the parallel cascade method (dashed line). (A) Intrinsic stiffness; (B) Static nonlinear element of reflex stiffness; (C) Linear impulse response of reflex stiffness.
Figure 4.9: Estimated nonlinear output and reflex torque for a subject. The solid line is the estimate from subspace method; the dashed line is the estimate form the parallel cascade method. (A) Output from nonlinearity; (B) Reflex torque
4.7 Discussion and Conclusion

We have presented a method for the identification of a state space model of dynamic ankle stiffness based on the MOESP subspace identification algorithm. Using a set of static, nonlinear basis functions, the nonlinear, SISO ankle stiffness is transformed to a linear, MISO state-space model in discrete time. Then, the PI-MOESP subspace algorithm outlined in Section II is applied to estimate the state space model. Simulating the estimated system with the appropriate inputs gives estimates of the intrinsic and reflex torques. The simulation studies reported in this paper demonstrate that the estimates obtained with the subspace method are accurate and robust even in presence of substantial output noise. Furthermore, the underlying continuous-time model can be recovered readily from these estimated torques. The experimental study demonstrated that estimates from the subspace method were similar to those obtained by applying the parallel cascade method to the same data set.

4.7.1 Constructed Inputs

The subspace method proposed in this paper used constructed inputs to estimate intrinsic and reflex stiffness, but the method used to choose the constructed inputs for each stiffness is different. For the intrinsic stiffness, the position and its first and second order derivatives are used to as the constructed inputs. This quasi-linear system can be simplified to a constant gain vector applied to the constructed inputs. Thus the estimated $A$ matrix of the MISO system contains system information only from the reflex stiffness. The system dynamics of intrinsic stiffness are estimated only in the $D$ matrix.
For the reflex stiffness, the constructed inputs are produced using a nonlinear expansion basis. A one-dimensional input is transformed to a high dimensional input vector which contains the coefficients of the expansion of \( g(\cdot) \) in the basis. Thus the SISO Hammerstein system is transformed to a MISO linear system. The number of constructed inputs depends on the complexity of the nonlinearity. Considering the half-wave rectifier in the ankle reflex stiffness, a fifth order of Chebyshev polynomial was used. However, choosing the optimal dimension for this expansion before the identification is impossible. A practical way to choose the order of basis functions would be to select a relatively high order at first, and then proceed with the identification and examine the estimated parameters of the nonlinearity. The scale factor for each basis function term, \( \alpha_i \), can obtained using the method listed in [16], once the reflex torque is estimated. Basis function terms having scale factors with small numeric values are likely to have no significant contribution to the reflex torque and so can be excluded without influencing the reflex torque. Therefore, a low order basis function is obtained.

4.7.2 Comparison of Subspace Method and Parallel Cascade Method

The parallel cascade method uses non-parametric identification technology. Thus it requires little prior information. The intrinsic stiffness is estimated using an impulse response function with the memory length less than the reflex delay. The computation must be repeated until the estimated net torque fails to improve the %VAF. Thus, it is intense computationally. Moreover, although the computation usually converges in practice, there is no proof that this will always occur. Consequently, the iteration may not converge to a global minimal solution if the initial guess for the parameters of the
nonlinearity is far from the global minimal solution or if the SNR is low. Indeed, as illustrated in the simulation studies reported in Figure 4.7, there are situations in which the algorithm will fail at the first iteration and so fail to provide any useful information. In contrast, the subspace method provides accurate estimates without iteration. The state space model is estimated directly from the input and output data. Thus, subspace methods are fast computationally and do not have convergence problems.

In general, both subspace and parallel cascade methods estimated the intrinsic, reflex and net torques accurately in the presence of noise. However, the subspace method performed better than the parallel cascade method under the most demanding situations associated with low SNR, high reflex gain and short data records. Thus, as Figure 4.6 shows, the subspace method provided a better estimate with a higher VAF and a smaller standard deviation, when significant noise was present. Indeed, with a SNR of -5 dB, the parallel cascade failed to provide consistently useful results while the subspace method provided good estimates.

The subspace method also performed better than the parallel cascade method when the reflex gain was relatively large. The parallel cascade method treats the reflex torque as noise when estimating the intrinsic stiffness. Therefore, for the same set of data, the effective SNR for the parallel cascade method is lower than that for subspace method because the noise contains both system noise and reflex torque. Thus with a low SNR, the parallel cascade method will provide a less accurate estimate than the subspace method. Thus, we recommend using the subspace method rather than the parallel cascade method to study spastic subjects who have been reported to have a much greater reflex stiffness.
Finally, the subspace method provided more reliable estimates from short data records. When the data is shorter than 10 seconds, the standard deviation of its predictions was much smaller than that from the parallel cascade method. The parallel cascade method is an iterative algorithm. If the initial “guess” of the nonlinear parameters is far from the global minimum, the algorithm could get stuck in the local minimum, and then provides a low fit with %VAF less than 50%, which leads to a high standard deviation in the parameter uncertainty.

4.7.3 Application to Closed-loop Identification

The parallel cascade method is a correlation-based algorithm that is only suitable for open loop system identification. In our laboratory, we collect data from an open loop experiment by using the actuator as a position servo driving the subject’s ankle position to follow the command input. Since the actuator is much stiffer and faster than the ankle, the effects of torques generated by the subject on the actuator position are insignificant. Therefore, the ankle can be viewed as being in open loop because changes in torque will not affect the position. This makes it feasible to use open loop identification algorithms to estimate dynamic ankle joint stiffness accurately. However, these experimental conditions are far from realistic since, during most functional movements, the ankle moves against a compliant load. Consequently, any changes in ankle torque resulting from position perturbations will in turn produce changes in ankle position leading to a closed loop situation. The subspace method for estimating for ankle stiffness presented in this paper can be extended to deal with closed loop data by using an appropriate set of
instrumental variables to eliminate the feedback noise. We have described this extension in several conferences papers [25] [26] and it is the topic of a companion paper [27].

4.7.4 Conclusion

This paper presents a novel method to identify the intrinsic and reflex components of ankle joint stiffness. A state space model for the joint stiffness is estimated, using a subspace method, from a set of inputs constructed from the position signal and the measured torque. Simulating this state space model with appropriate inputs gives estimates of the intrinsic and reflex torques. Our results show that the new subspace method provides results similar to those of the parallel cascade method under most normal estimation conditions. However, the subspace method is not an iterative method as the parallel cascade method is. Therefore, the subspace method is more efficient in computation. In addition, simulation studies show that the subspace method provides better estimates under the most severe estimation conditions (e.g. high noise, short sample lengths). Moreover, the subspace method can be extended to the closed-loop system identification problems, while the parallel cascade method can only estimate systems operating in open loop.

Though this method was developed specifically for ankle joint stiffness, it should be applicable to study the mechanical behavior of other human joints with parallel cascade structure, such as the back [24] and the elbow [28].

Acknowledgements

Supported by grants from the Natural Sciences & Engineering Research Council of
Canada, and the Canadian Institutes of Health Research. The authors wish to thank D. Ludvig for providing the experimental data.
References


5 Identification of Human Ankle Dynamics Using Subspace Based Methods. Part 2: Closed-Loop System Identification

In this chapter we present a subspace method to estimate ankle joint stiffness using data acquired from closed loop experiments. We describe the method and provide simulation and experimental studies to validate the algorithm. Finally, we summarize the contribution of this paper and discuss further development.
Identification of Human Ankle Dynamics Using Subspace Based Methods. Part 2: Closed-Loop System Identification

Y. Zhao and R. E. Kearney

To be submitted to:
IEEE Transactions on Biomedical Engineering
5.1 Abstract

Dynamic joint stiffness, the dynamic relationship between the angular position of a joint and the torque acting about it, can be used to describe the mechanical behavior of the human ankle during posture and movement. Joint stiffness can be separated into intrinsic stiffness and reflex stiffness, and modeled as a linear system and a Linear-Nonlinear-Linear (LNL) system respectively [1]. Previously, we have used a single-input single-output (SISO), nonlinear, parallel cascade identification method to study joint stiffness under open-loop conditions. However, in most functional tasks, the ankle will interact with a compliant load and the joint stiffness must be viewed as operating within a closed-loop because the ankle torque is fed back through the load to change the ankle position. Applying open-loop identification methods to such data will give biased results. In this paper, we describe a new method to estimate the intrinsic and reflex stiffness from data measured in closed-loop. The SISO nonlinear stiffness model is first transformed into a multiple-input/multiple output, linear, state space model. Then, a MOESP (Multivariable Output-Error State space) subspace system identification method is used to estimate the dynamics of each pathway directly from measured data. Bias errors due to feedback noise are eliminated through the use of the past input and past output signals as instrumental variables. Simulation and experimental studies demonstrate that the method produces accurate results.

Keywords—ankle dynamics, closed-loop system identification, subspace method, MOESP algorithm
5.2 Introduction

The concept of dynamic joint stiffness is used to study the mechanical behavior of the mechanisms acting about a joint. It defines the dynamic relationship between a joint’s angular position and the torque acting about it. It can be separated into two components: intrinsic stiffness and reflex stiffness. The intrinsic component is due to the mechanical properties of the ankle, passive tissue, and active muscle fibers; the reflex component is due to muscular contraction in response to activation by stretch receptors in the muscle. Previous work from our laboratory [2] demonstrated that the parallel cascade model, shown in Figure 5.1, describes dynamic joint stiffness at the ankle very well.

Intrinsic and reflex torques arise and change together; only their sum can be measured experimentally. Consequently, it is not possible to estimate intrinsic and reflex stiffnesses directly. Previous work from our laboratory addressed this problem using a systems identification approach. Thus, we developed a parallel cascade method [2][3], a real-time algorithm [4] and a subspace method [5] to separate the intrinsic and reflex components from the overall torque and estimate intrinsic and reflex stiffness.
Figure 5.1: Parallel cascade structure of ankle dynamics. $TQ_i$ and $TQ_R$ denote intrinsic torque and reflex torque.
The parallel cascade method and real-time algorithm are correlation-based identification methods [6] that require the experimental data to be collected from an open-loop experiment. To meet this requirement, experiments were conducted by applying perturbations using an actuator operating in a position control mode as illustrated in Figure 5.2. In this mode, the actuator operates as a position servo and drives the ankle position to follow the command input. The actuator is much stiffer than the ankle, so that the torques developed at the ankle will not change the joint position. This effectively opens the feedback path from ankle torque to ankle position. Therefore, identifying the ankle dynamics can be viewed as an open-loop identification problem.

Such open-loop experiments have provided a good understanding of the behavior of intrinsic and reflex stiffness under these experimental conditions [2] [7], which correspond to interacting with a very stiff environment. However, these conditions are far from realistic because, during most functional tasks and movements, subjects will interact with compliant loads and consequently the joint torque and position will be connected by feedback, as shown in Figure 5.3. Any change in torque will evoke a change in position and vice versa. Torques and positions measured under such conditions are therefore measured in closed-loop. It is well known that correlation-based identification methods, such as those we have used, will give biased results when applied to closed-loop data [6]. Therefore new tools are needed to estimate intrinsic and reflex stiffness while subjects interact with compliant environments.
Figure 5.2: Block diagram of the open-loop identification problem showing intrinsic, $H_i(s)$, and reflex, $H_r(s)$, stiffness dynamics.
**Figure 5.3:** Simplified block diagram of the closed-loop identification problem. $H_{\text{IMP}}(s)$ represents the dynamics of the impedance controller. $H_I(s)$ and $H_R(s)$ represent the intrinsic and reflex stiffness respectively. INP represents the input perturbation.
For linear systems, issues related to closed-loop identification are well understood and methods have been developed that use instrumental variables to provide accurate estimates of system dynamics [6]. For nonlinear systems, such as ankle stiffness, the effects of feedback are more complex, less well understood and identification methods are not well developed [8].

This paper investigates these issues for the specific case of ankle stiffness and is developed as follows. Section II presents a novel identification method that estimates ankle stiffness accurately from closed-loop data. To achieve this, the single-input/single-output (SISO), nonlinear, parallel cascade model of ankle dynamics is first transformed into a multiple-input/single-output, linear state space model as described previously [5] [9]. Then, we develop a subspace method that uses instrumental variables to estimate stiffness dynamics accurately from closed-loop data. Section III describes a simulation study that validates the new algorithm and compares its performance to that of the parallel cascade algorithm. It demonstrates that estimates obtained with the parallel cascade method are badly biased when applied to data acquired with a compliant load. Moreover, the estimates are biased even without measurement noise due to the effects of reflex torque. In contrast, we show that our new algorithm provides accurate estimates of intrinsic and reflex stiffness under the most demanding conditions. Section IV presents the results of an experimental study that demonstrates the utility of the new method. Section V summarizes the important results in this paper and discusses their significance.

5.3 Identification of Ankle Stiffness In Closed-Loop

5.3.1 Parametric Models for Ankle Stiffness
Previous work [1] showed that the intrinsic stiffness can be modeled well by a second-order, quasi-linear system with transfer function:

\[ H_{\text{Is}}(s) = \frac{TQ_{\text{I}}(s)}{\text{POS}(s)} = Is^2 + Bs + K \]  

(5.1)

where

- \( TQ_{\text{I}} \) is the intrinsic torque that is the output from intrinsic stiffness,
- \( \text{POS} \) is the position that is the input to intrinsic stiffness,

\[ I, B \text{ and } K \] are inertial, viscous, and elastic parameters which vary with the operating point.

Reflex stiffness can be modeled with a LNL structure, a series connection of differentiator, a static non-linearity and a 2nd- (or in some cases 3rd-) order low-pass system in series with a delay. Equation 5.2 shows the transfer function of the 2nd order low-pass filter for reflex stiffness.

\[ H_{\text{Rs}}(s) = \frac{TQ_{\text{R}}(s)}{\text{VL}(s)} = \frac{g_{\text{R}}\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)} e^{-st} \]  

(5.2)

where

- \( TQ_{\text{R}} \) is the reflex torque, the output from reflex stiffness,
- \( \text{VL} \) is the delayed, half-wave rectified joint angular velocity, the input to reflex stiffness,
- \( g_{\text{R}} \) is the reflex gain,
- \( \omega_n \) is the 2nd order natural frequency,
ξ is the damping parameter,

τ is the reflex delay.

5.3.2 State space Model for Joint Stiffness

In our previous paper [5] [9], we showed that the single-input/single-output (SISO), non-linear, parallel cascade model of joint stiffness could be transformed into a multiple-input/single-output (MISO), linear, state space model. Our new algorithm is based on this formulation and consequently it will be reviewed here.

The state space model describing intrinsic stiffness is:

\[ t_{q_i}(k) = D_i U_i(k) \] (5.3)

where

\( t_{q_i}(k) \) is the intrinsic torque at discrete time \( k \),

\( U_i(k) = [p(k) \ vel(k) \ acc(k)] \) is a set of constructed inputs for intrinsic stiffness where \( p(k) \) is measured position, \( vel(k) \) is the differentiated position,

\( acc(k) \) is the second differential of position

\( x_i(k) \) is the state vector.

Reflex stiffness is modeled as a MISO state space model defined by:

\[
\begin{align*}
    x_R(k + 1) &= A_R x_R(k) + B_R U_R(k) \\
    t_{q_R}(k) &= C_R x_R(k) + D_R U_R(k) 
\end{align*}
\] (5.4)

where

\( t_{q_R}(k) \) is the reflex torque,
\[ U_R(k) = \left[ g_1(dvel(k)), \ldots, g_p(dvel(k)) \right]^T, \] is a set of constructed inputs for reflex stiffness obtained by transforming \( dvel(k) \) by a set of basis functions \( g_i \), \( dvel(k) \) is the delayed and differentiated position.

\( x_R(k) \) is the state vector for the linear part of the reflex stiffness,

\( A_R, C_R \) are the system matrices for the linear part of the Hammerstein system,

\[ \tilde{B} = \left[ B_R \alpha_1, \ldots, B_R \alpha_p \right] \text{ and } \tilde{D} = \left[ D_R \alpha_1, \ldots, D_R \alpha_p \right], \] are the system matrices for reflex stiffness.

\( \alpha_i \) are the scale factors for the basis function terms,

The state space model for the overall parallel cascade model is obtained by combining the state space models for intrinsic (Equation 5.3) and reflex (Equation 5.4) stiffness to give:

\[
x_R(k+1) = A_R x_R(k) + \begin{bmatrix} 0 & \tilde{B} \end{bmatrix} \begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix} \\
tq_N(k) = C_R x_R(k) + \begin{bmatrix} D_I & \tilde{D} \end{bmatrix} \begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix}
\]

where

\( tq_N(k) \) is the measured net torque,

Once the state space model for overall ankle joint stiffness has been estimated from the measured position and torque, the torques from the intrinsic and reflex stiffness can be predicted by simulating the estimated system with the appropriate inputs.
Specifically, the output from the simulation with the input signal \[
\begin{bmatrix}
U_i(k) \\
0
\end{bmatrix}
\]
predicts the intrinsic torque, \( t(q_i(k)) \); the response to the input \[
\begin{bmatrix}
0 \\
U_r(k)
\end{bmatrix}
\]
predicts the torque from the reflex stiffness, \( t(q_r(k)) \). Continuous-time models for intrinsic and reflex stiffness can be estimated from these predictions as described in [5], which also provides details about the choice of basis functions for reflex stiffness.

### 5.3.3 Subspace Methods for Closed-Loop System Identification

The difficulty with closed-loop system identification is feedback noise. Subspace methods use instrumental variables to eliminate the effects of noise. The instrumental variable must be uncorrelated with the noise, but correlated with the states. In particular, the instrument variable, \( \Theta \), must be uncorrelated with the noise, \( V_{i,j,N} \) and \( W_{i,j,N} \), as described by the first and two terms of Equation 5.6. Moreover, the instrumental variable, \( \Theta \), must be correlated with the states, as indicated by the third term of Equation 5.6. The requirements for the instrumental variables are as follows.

\[
\lim_{N \to \infty} \frac{1}{N} V_{i,j,N} \Theta^T = 0
\]

\[
\lim_{N \to \infty} \frac{1}{N} W_{i,j,N} \Theta^T = 0
\]  \hspace{1cm} (5.6)

\[
\lim_{N \to \infty} \frac{1}{N} X_{i,N} \Theta^T \text{ has full rank}
\]

where

\( \Theta \) is the instrumental variable,
$X_{i,N} = [x_i \ x_{i+1} \ \cdots \ x_{i+N-1}]$ contains the internal states,

$V_{i,j,N}$ and $W_{i,j,N}$ are Hankel matrices for process noise and measurement noise respectively.

Note that MOESP algorithm family organizes data using block Hankel matrices where the Hankel matrix for a signal $r$ is defined as:

$$
\begin{bmatrix}
  r(i) & r(i+1) & \cdots & r(i+n-1) \\
  r(i+1) & r(i+2) & \cdots & r(i+n) \\
  \vdots & \vdots & & \vdots \\
  r(i+j-1) & r(i+j) & \cdots & r(i+j+n-2)
\end{bmatrix}
$$

(5.7)

where

- $i$ is the left upper entry of the Hankel matrix,
- $j$ is the number of the rows,
- $n$ is the number of columns.

Chou and Verhaegen [10] developed an EIV-MOESP (Errors-In-Variables Multivariable Output-Error State space) identification method for closed-loop systems. EIV-MOESP assumes that the noise processes of a system are statistically independent of past input and past output as well as their own past values. Thus, the first two terms of Equation 5.6 are satisfied. Moreover, it has been shown in [11] [12] that $\lim_{N \to \infty} \frac{1}{N} X_{i,N} \Theta$ has full rank if the input signal and the instrumental variable are persistently exciting of system order $n$. A signal is persistently exciting of order $n$, if its two-sided power spectrum is non-zero at more than $n$ points [6]. The third term of Equation 5.6 ensures that multiplication by the IV does not change the rank of the equation. Therefore, the
system’s order and its system matrices can be estimated correctly. Therefore, the third term of Equation 5.6 is satisfied.

EIV-MOESP can be used to estimate the state space model for overall joint stiffness in closed-loop since the past position and torque signals satisfy the requirements for instrumental variables. Specially, there are two noise terms in the experiment: measurement noise and torque changes due to variations in central commands. The measurement noise is assumed to be Gaussian white, and so will be uncorrelated with the past position, past torque and its own past values. It will be assumed that the variations in voluntary commands are generated randomly by the subject and so are not correlated with the past position or past torque. However, due to the low-pass filter properties of muscle contraction the voluntary torque will be correlated with some of its past values. Therefore the offset between the “past” IV and the “current” data must be at least twice the system order and large enough to eliminate any correlation between the voluntary torque and its past values. Moreover, the perturbation signal can be chosen to be persistently exciting. Consequently, the past position and torque meet all the requirements for instrumental variables.

The algorithm eliminates feedback noise by projecting the observed signals onto the instrumental variables. To do so, the Hankel matrices, $U_{i,j,N}$ and $Y_{i,j,N}$, are constructed from the input and output signals. Hankel matrices for the instrumental variables are constructed by shifting the input and output matrices, $U_{i,j,N}$ and $Y_{i,j,N}$, to the past by $p$ samples, to give $U_{i-p,j,N}$ and $Y_{i-p,j,N}$. In practice, the value of $p$ is chosen to be the same as number of columns in the input and output Hankel matrices and is
normally chosen at least twice of the order of the system. Then, the projection is implemented using the LQ factorization [13] as

\[
\begin{bmatrix}
U_{i,j,N} \\
Y_{i,j,N}
\end{bmatrix}
\begin{bmatrix}
U_{i-p,j,N} \\
Y_{i-p,j,N}
\end{bmatrix}^T = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}
\]

(5.8)

\( L_{22} \) is the subspace that spans the internal states after the effects of noise are eliminated by the projection. The singular value decomposition (SVD) [13] of the matrix \( L_{22} \) is

\[
L_{22} = \mu \xi \nu^T
\]

(5.9)

The column space of \( L_{22} \) is captured by the matrix \( \mu \), so, the \( C \) matrix is given by the first \( l \) rows of \( \mu \),

\[
\hat{C} = \mu(1:l,:)
\]

(5.10)

where

\( l \) is the number of the output signals.

The \( A \) matrix is given by

\[
\hat{A} = \mu_1 \hat{\mu}_2
\]

(5.11)

where

\[
\mu_1 = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-2} \end{bmatrix}
\]
is the top \((i-1)l\) rows of \( \mu \).
\[ \mathbf{\mu}_2 = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^i \end{bmatrix} \] is the bottom \((i-1)l\) rows of \(\mathbf{\mu}\).

‡ indicates the Moore-Penrose pseudo-inverse [14].

Once \(\hat{A}\) and \(\hat{C}\) are available, the system output is linear in the \(\mathbf{B}\) & \(\mathbf{D}\) matrices

\[ y(k) = \mathbf{CA}^k x(0) + \sum_{\tau=0}^{k-1} \mathbf{CA}^{k-\tau} \mathbf{Bu}(\tau) + \mathbf{Du}(k) \quad (5.12) \]

Therefore, \(\hat{B}\) and \(\hat{D}\) can be estimated in a least-square sense. Details of the proof and derivation are given in [15] [16].

5.3.4 Implementation

The identification algorithm is summarized as follows:

1. Calculate \(vel(k)\) and \(acc(k)\) by differentiating the measured position [17]. Construct the input signal to intrinsic stiffness, \(\mathbf{U}_I(k) = [p(k) \ vel(k) \ acc(k)]\).

2. Construct the input signal to reflex stiffness using the basis function,

\[ \mathbf{U}_R(k) = \begin{bmatrix} g_1(dvel(k)) & \cdots & g_p(dvel(k)) \end{bmatrix}^T \], where \(dvel(k)\) is the delayed velocity.

3. Construct \(\mathbf{U}_{i-p,N}, \mathbf{Y}_{i-p,N}, \mathbf{U}_{i,j,N}\) and \(\mathbf{Y}_{i,j,N}\) from the constructed inputs and measured torque. Use EIV-MOESP from the SMI toolbox [18] to estimate the system matrices \(\hat{A}, \hat{B}, \hat{C}, \hat{D}\) from the input \(\begin{bmatrix} \mathbf{U}_I(k) \\ \mathbf{U}_R(k) \end{bmatrix}\).
4. Simulate the estimated system with the input, $\begin{bmatrix} U_f(k) \\ 0 \end{bmatrix}$ to estimate the intrinsic torque.

5. Simulate the estimated system with the input, $\begin{bmatrix} 0 \\ U_h(k) \end{bmatrix}$ to estimate the reflex torque.

### 5.4 Simulation Studies

#### 5.4.1 Simulation Methods

To investigate the performance of our algorithm, we developed a simulation model of the experimental paradigm used in our laboratory. Thus, we modeled the behavior expected from the subject whose ankle was connected to the pedal of an electro-hydraulic actuator using a position-based impedance controller [19] to generate a compliant load. The simulation was implemented using Matlab’s Simulink [20] with the model diagram shown in Figure 5.4.

Intrinsic and reflex stiffnesses were based on the parallel cascade model of ankle stiffness shown in Figure 5.1. Intrinsic stiffness was modeled by multiplying the position, the velocity and the acceleration by the elastic, viscous and inertial parameters respectively. Reflex torque was simulated by delaying the velocity signal, half-wave rectifying it, and then filtering with the second-order, linear system defined in Figure 5.2. Parameter values were based on those identified experimentally from a typical subject [7] and are given in Table 5.1. The impedance controller was modeled as a second-order, low-pass, Bessel filter with a 10 Hz cut-off frequency. The actuator dynamics were
modeled as a 2nd order, low-pass, Bessel filter with a 30 Hz cut-off frequency.

Table 5.1. Intrinsic and reflex parameter values of simulated ankle dynamics

<table>
<thead>
<tr>
<th>$h_{IS}$ Parameters</th>
<th>$h_{RS}$ Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$ $(Nm/rad/s^2)$</td>
<td>$B$ $(Nm/rad/s)$</td>
</tr>
<tr>
<td>0.0172</td>
<td>1.31</td>
</tr>
<tr>
<td>$K$ $(Nm/rad)$</td>
<td>$\omega_n$ (rad/s)</td>
</tr>
<tr>
<td>200</td>
<td>26</td>
</tr>
<tr>
<td>$\xi$ $(rad/s)$</td>
<td>$G_R$ $(Nm/rad/s)$</td>
</tr>
<tr>
<td>0.98</td>
<td>10</td>
</tr>
</tbody>
</table>
Figure 5.4: Matlab Simulink model of ankle dynamics. A Gaussian white noise was added to the net torque as measurement noise. An independent Gaussian white noise signal was filtered by a second order low-pass system and added to the net torque as voluntary torque. The noise-free intrinsic and reflex torques are recorded in the simulation for comparison with the estimates from the algorithm.
The external input was a Pseudorandom Binary Sequence Signal (PRBS), with peak-to-peak amplitude of 0.04 rad and a switching interval of 100 ms, similar to that used experimentally. Gaussian white noise was added to the simulated output to model measurement noise. Another Gaussian white noise signal was filtered by a second-order, low-pass filter with a 2 Hz cut-off frequency, and added to the feedback loop to model variations in the voluntary torque due to changes in voluntary drive. The voluntary torque and measurement noise were added and considered as the total noise, as shown in Figure 5.4. The magnitudes of the voluntary torque and measurement noise were chosen so that they contributed equally to the total noise variance.

We used this simulation model to systematically examine how the algorithm’s performance varied with the measurement noise and the reflex gain. For each set of parameters we carried out a Monte-Carlo study comprising 100 simulation trials, each lasting 50 seconds with the same parameter set but having different realizations of the input and noise signals. The position and the torque signals were simulated at 1 KHz and then decimated to 100 Hz.

### 5.4.2 Identification and Validation

The subspace method was used to estimate stiffness models for each simulation trial. In particular, the inputs for identification were constructed from the measured position, while the output was the recorded torque. We chose the delay between the current data and IV to be 15 samples, to fulfill the requirement of the IV (larger than twice of the system order, 4 data points at 100 Hz). The parallel cascade method was also used to estimate the stiffness from the same data as those used by subspace method.
However, the two methods used quite different model structures and so cannot be compared directly. Consequently, we opted to evaluate the methods’ performance in terms of how well they predicted the noise-free torques from the simulation. Thus, the estimated models were used to predict the intrinsic, reflex and net torques and these were compared to the simulated signals in terms of the percentage variance accounted for (%VAF) defined as:

\[
\text{%VAF} = \left(1 - \frac{\text{var}(tq_s - tq_E)}{\text{var}(tq_s)}\right) \times 100
\]

where

\( tq_s \) is the simulated, noise free torque,

\( tq_E \) is the estimated torque.

5.4.3 Effects of Noise

We first examined the effect of measurement noise on accuracy of estimation. Ten sets of Monte-Carlo studies were carried out with Signal to Noise Ratios (SNR) ranging from 11 dB to 30 dB. The reflex gain was set to 10, a representative intermediate value [7]. Figure 5.5 shows the %VAFs between the simulated and estimated torques, as a function of SNR, for the subspace and parallel cascade methods. At high SNRs (>20dB) both methods gave good results although the subspace estimates were consistently better that the parallel cascade estimates, However, as the SNR decreased the performance of the parallel cascade became worse; the %VAF decreased and variability increased as reflected by the size of the standard deviation bars. The drop is particularly
evident in the intrinsic and total torques. In contrast the subspace method continued to give accurate estimates for all torque signals.

**Figure 5.5:** Effects of SNR on %VAF between noise-free simulated and estimates (A) Intrinsic (B) Reflex and (C) Net torque. Each point is the mean value bracketed by its standard deviation from the 100 Monte-Carlo trials.
5.4.4 Effects of Reflex Gain

The preceding section demonstrated that, when used with closed-loop data, the parallel cascade algorithm gave biased results in the presence of measurement noise. In view of the structure of the parallel cascade algorithm, which iterates between estimating the intrinsic and reflex pathways, we felt that closed-loop data might yield biased results even without measurement noise. Thus, reflex torques might act as noise when estimating intrinsic stiffness and, conversely, intrinsic torques as measurement noise when estimating reflex stiffness. Consequently, we investigated the performance of the two algorithms as a function of reflex gain with no measurement noise. Six Monte-Carlo studies were conducted, in which the reflex gain was varied from 5 to 30 while the intrinsic stiffness gain was held constant.

Figure 5.6 summarizes the results of these Monte-Carlo studies. When the reflex gain was low, both methods gave high quality estimates for all torques. However, the performance of the parallel cascade method decreased as the reflex gain increased. Intrinsic torques were estimated reasonably well but the estimates of the reflex torque became progressively worse as the reflex gain increased; indeed they dropped below 20% at the highest reflex gains. In contrast, the subspace method gave high quality estimates for all reflex gain values. Indeed, the quality of the reflex and net torque predictions remained high regardless of the value of reflex gain.
Figure 5.6: Effects of Reflex Gain on %VAF between noise-free simulated and estimated (A) Intrinsic (B) Reflex and (C) Net torque. Each point is the mean value bracketed by its standard deviation from the 100 Monte-Carlo trials.
5.5 Experimental Studies

Next, we undertook to evaluate the two algorithms with experimental data. To this end, we examined one subject with no history of neuromuscular disease. Details of the experimental methods are given in [7]. The subject lay supine with the left foot attached to the pedal of an electrohydraulic actuator operating under impedance control to simulate a lightly damped mass-spring system. The experimental paradigm was a simple torque matching task; subject was provided with a display of their average torque on a LCD monitor and asked to maintain it at a fixed level. A position perturbation comprising a PRBS with a peak-to-peak value of 0.4 rad and switching time of 120 ms was applied continuously. Experimental data were collected at 1 kHz for 65 seconds and then decimated to 100 Hz prior to analysis.

As with the simulation studies, stiffness models were estimated using both the parallel cascade and subspace methods and used to predict the intrinsic, reflex and total torques. Table 5.2 shows the %VAF between predicted and experimentally observed total torque. Note that these %VAFs cannot be compared directly to those for the simulation studies since the true reflex and intrinsic torques were not available and the net torque measurement contained noise. Therefore, we compared the observed net torque to those predicted by the models estimated with the two methods. The magnitudes of intrinsic and reflex torques relative to the net torque were evaluated by using %VAF. Note that the sum of %VAFs of intrinsic and reflex torques to the net torque was not equal to that between the estimated and measured net torque, because the intrinsic and reflex torque were correlated. Nevertheless, it is evident that the subspace method provided better
results than the parallel cascade method did. Thus, the %VAF for the net torque, was 19% higher for the subspace method than for the parallel cascade method indicating that the overall model was much better. The %VAFs for the intrinsic and reflex torque were both higher for the subspace estimates than those from the parallel cascade method suggesting that estimates of both pathways were improved.

To compare the two estimates further, we predicted the intrinsic and reflex torques from the state space model estimates and then used the Nonlinear Identification Toolbox [21] to estimate the intrinsic stiffness IRF and the Hammerstein system for reflex stiffness. Figure 5.7 shows the non-parametric models estimated with the torques predicted by the state space model with those for the parallel cascade method. It is evident that the subspace method estimates were better than those from the parallel cascade methods. Thus, the static nonlinearity predicted by the subspace method resembles a half wave rectifier as expected while that from the parallel cascade method appears to be linear. In addition, the reflex IRF from the subspace method estimates is smooth while that from the parallel cascade data is very noisy. Thus, the subspace method estimates not only account for more variance than the parallel cascade estimates but yield results more similar to those expected from previous, open-loop experiment.
Table 5.2 Estimates from the subspace method and the parallel cascade method for experimental data

<table>
<thead>
<tr>
<th></th>
<th>Subject (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intrinsic Torque</strong></td>
<td></td>
</tr>
<tr>
<td>39% Subspace</td>
<td></td>
</tr>
<tr>
<td>15% Parallel cascade</td>
<td></td>
</tr>
<tr>
<td><strong>Reflex Torque</strong></td>
<td></td>
</tr>
<tr>
<td>48% Subspace</td>
<td></td>
</tr>
<tr>
<td>45% Parallel cascade</td>
<td></td>
</tr>
<tr>
<td><strong>Net Torque</strong></td>
<td></td>
</tr>
<tr>
<td>86% Subspace</td>
<td></td>
</tr>
<tr>
<td>67% Parallel cascade</td>
<td></td>
</tr>
</tbody>
</table>
**Figure 5.7:** Stiffness estimates for human subject obtain using the subspace (left column) and parallel cascade (right column methods). (A&D) Intrinsic stiffness impulse response function, (B&E) Static nonlinearity for the reflex Hammerstein system (C&F) Linear impulse response function for the reflex Hammerstein system.
5.6 Discussion and Conclusion

5.6.1 Parallel Cascade Method

Our simulation results demonstrate that the parallel cascade method will yield biased estimates of joint stiffness when applied to data acquired in closed-loop. This occurs both with and without measurement noise. The finding that measurement noise results in biased estimates with closed-loop data is expected since the parallel cascade method uses a least-squares based system identification algorithm that is known to give biased estimates for closed-loop data [6]. This is because in closed-loop any measurement noise is fed back to the input giving rise to a correlation between the input and the measurement noise. As a result, least-squares estimates will be biased.

However, our simulation results also showed that the parallel cascade method gave biased results with closed-loop data even when there was no measurement noise. We believe this is because of the iterative nature of the algorithm. Thus, in effect, the parallel cascade method treats the reflex torque, \( t_{rf} \), as part of the measurement noise, when estimating the impulse response function (IRF) for the intrinsic stiffness. This method assumes that the input signal is uncorrelated with the measurement noise. However, with closed-loop data this requirement is violated. Thus, when the load is compliant, any reflex torque is fed back via the load to change the position. Thus, the position signal is correlated with the reflex torque, meaning that the input signal is correlated with the measurement noise. The parallel cascade method will provide biased IRF estimates for intrinsic stiffness, which in turn leads to biased estimates for joint
stiffness.

5.6.2 Subspace Method

The subspace method uses instrumental variables to eliminate the effects of feedback noise. For closed-loop system identification, EIV-MOESP was implemented using the past-input and past-output as the instrumental variables. Noise effects were eliminated by projecting the data onto these instrumental variables. Indeed the simulation studies demonstrated that the subspace method estimated the intrinsic, reflex and net torques accurately with SNRs as low as 10 dB.

The reflex stiffness has a nonlinear, Hammerstein system structure. Thus, a large reflex gain will result in a larger contribution to the net torque making the identification problem more difficult. The subspace method uses a basis function to transform the SISO, nonlinear, Hammerstein system into a MISO linear system. This linearization converts the difficult nonlinear identification problem into a classic linear system identification problem. In the simulation studies, we showed that the subspace method can estimate the ankle joint stiffness regardless of the value of the reflex gain. Even when the reflex gain was large (>25), the subspace method still provided reliable results.

The experimental results were consistent with simulation results and indicated that the subspace method provides better estimates than the parallel cascade method. The estimated net torque from the subspace method accounted for more of the observed torque than did that from the parallel cascade method. Furthermore, the estimated IRF and Hammerstein models associated with the subspace method were more consistent with previous results than were those from the parallel cascade method.
5.6.3 Instrumental Variables

The EIV subspace method uses past input and past output as the IV to eliminate the effects of feedback noise and thus provide accurate estimates of system dynamics in closed-loop. The requirements of the IV are described by Figure 5.6. The IV must be uncorrelated with the noise, but correlated with the states. In particular, the past input and past output must be uncorrelated with the noise. However, any measurement of the input and output signals will contain some noise. In the simulation studies of ankle joint stiffness, the noise was modeled as containing a measurement considering to be Gaussian and white, and a component due to voluntary torques, which was assumed to be uncorrelated with the past, noise-free position and torque signals. Therefore, the autocorrelation of the measurement noise and voluntary must be analyzed to choose the correct offset between the current data and the IV.

Since the measurement noise comprised realizations of white noise, it can be shown easily that its auto-correlation is close to zero, if there is at least 1 sample of offset, because of the independent property of white noise. The voluntary torque is not white, because it is generated by filtering a white noise via a linear, second-order, low-pass filter, whose state space model is:

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + B\mathbf{v}(k)$$
$$\omega(k) = C\mathbf{x}(k) + D\mathbf{v}(k)$$

(5.14)

where

$\mathbf{v}(k)$ and $\omega(k)$, are vectors containing input and output measurements, at discrete time $k$, 
\( A, B, C, D \) are system matrices.

Therefore, \( \omega(k) \) can be obtained from \( \omega(k - n) \) by:

\[
\omega(k) = CA^n x(k - n) + CA^{n-1} B \nu(k - n) + CA^{n-2} B \nu(k - n + 1) + \cdots + CB \nu(k - 1) + D \nu(k)
\]

where

\( n \) is the offset,

\( A^n \) is the \( n \)th power of \( A \)

Therefore, only the terms \( CA^n x(k - n) \) and \( CA^{n-1} B \nu(k - n) \) may be correlated with \( \omega(k - n) \). However, if the system is stable, \( A^n \) will be close to zero if \( n \) is large.

Thus, \( \omega(k) \) is uncorrelated with \( \omega(k - n) \). In practice, the subspace method requires that the offset \( n \) must be at least twice of the order. In this paper, 15 samples were used as the offset, \( n \), which was large enough to eliminate the correlation between the noise in the current data and in the IV. Therefore, the past input and past output can be used as the IV to provide estimates of the ankle joint stiffness in closed-loop.

### 5.6.4 Conclusion

In this paper, we first demonstrated that using the open-loop, parallel cascade method to identify the ankle stiffness from closed-loop data will provide biased estimates. Then, we presented a subspace-based method to estimate ankle dynamics with compliant loads. The nonlinear, parallel cascade model of ankle dynamics is transformed into a linear, MISO state space model with constructed inputs from the measured position signal and terms of the basis function. The EIV-MOESP algorithm was implemented to estimate the MISO state space model for ankle dynamic. Simulating the estimated model
with the appropriate inputs gives estimates of the intrinsic and reflex torque. Simulation and experiment studies demonstrated that the subspace method provides high quality estimates for ankle stiffness from closed-loop data.

The ability to estimate intrinsic and reflex stiffness from data acquired while subjects interact with compliant loads is essential to the understanding of neuromuscular control. It is the most common functional situation since it prevails during all voluntary movements. Moreover, as a result of recent advances of chronic recording technology and the interest in brain-machine-interfaces data describing motor behavior under natural conditions is increasingly available. Our results demonstrate that the correct analysis of these data requires the use of appropriate algorithms, such as that presented in this paper.

Acknowledgements

Supported by grants from the Natural Sciences & Engineering Research Council of Canada, and the Canadian Institutes of Health Research.
References:


6 Identification of Biomedical Systems from Short Transients

Using Subspace Methods

In this chapter we present a subspace method to estimate systems with short transients. We also show that the algorithm can be used to estimate time-varying systems. Simulation studies demonstrate the performance of the method when applied to both the systems with short transients and time-varying systems.
Identification of Biomedical Systems from Short Transients Using Subspace Methods

Y. Zhao, and R. E. Kearney,

To be submitted to:
IEEE Transactions on Biomedical Engineering
6.1 Abstract

Most system identification methods are designed to work with a single data record acquired under stationary conditions. In general, the accuracy of the resulting estimates increases with the length of the data record. However, for many biomedical problems, it is difficult or impossible to acquire long records from a system in the same condition. Nevertheless, it may be possible to obtain ensembles of short records. This paper presents a subspace method to identify Multiple Input/Multiple Output (MIMO) state space models for biomedical systems from ensembles of short transients. Two applications of the method are demonstrated with simulated data sets. The first simulation demonstrates that the algorithm can estimate the intrinsic and reflex components of dynamic ankle stiffness from ensembles of records as short as 20 times the system order. The second simulation shows how the algorithm can be used to estimate time-varying systems from an ensemble of trials having the same time-varying behavior.

Keywords—biomedical system with short transients, time-varying systems, ensemble system identification, subspace method.
6.2 Introduction

Subspace methods are identification technologies that estimate state space models of dynamic systems directly from measurements of inputs and outputs. Although subspace methods originated in control engineering, these methods can be used to model physiological systems in biomedical engineering. Indeed, our laboratory recently developed a subspace method that separates intrinsic and reflex stiffness to study the dynamic, mechanical behavior of the human ankle joint during posture and movement. Compared with correlation-based methods, such as the parallel-cascade method [1], the subspace method is fast and can yield accurate estimates in the presence of noise through the use of instrumental variables (IV) [2] [3]. Moreover, by choosing appropriate IVs, the subspace method can identify joint stiffness for both open-loop [4] and closed-loop [5] [6] data.

As with other methods, subspace methods may require long records, acquired under stationary conditions, to yield accurate estimates. However, in many biomedical experiments it is not possible to acquire long data records. In some cases, such as with the vestibular-ocular system [7], the system dynamics may switch rapidly between two or more modes so that only short transients are available from any mode. In other cases, the system dynamics may vary rapidly with time so that the data may be considered to be stationary for only short segments. Thus, for example, joint stiffness varies rapidly throughout voluntary movements [8]. The direct identification of systems from such data is difficult with current methods.

A number of identification methods have been developed to identify system
dynamics from ensemble data. Kearney et al. [8] proposed a numerical scheme to identify a time-varying Finite Impulse Response (FIR) model from an ensemble of data. Verhaegen et al. [9] extended the subspace method to identify a time-varying state space model from ensemble data. However, these methods did not estimate the initial states that play an important role in a system’s transient response. For example, consider the continuous state space system:

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t)
\end{align*}
\] (6.1)

(6.2)

where

\[ u(t) \in \mathbb{R}^m \] and \[ y(t) \in \mathbb{R}^l \] are vectors containing measurements, at time \( t \), of the process’s \( m \) inputs and \( l \) outputs.

\[ x(t) \in \mathbb{R}^n \] is the state vector of the process at time \( t \).

\[ A \in \mathbb{R}^{nxn} \] is the system matrix that describes the system dynamics.

\[ B \in \mathbb{R}^{nxm} \] is the input matrix defining the linear transformation that describes how the deterministic inputs influence the states.

\[ C \in \mathbb{R}^{lxn} \] is the output matrix that describes how the internal states are transformed to generate the output.

\[ D \in \mathbb{R}^{lxm} \] is the direct feed through term.

The state vector at any time \( t \) can be calculated given the state vector at the starting time \( t_0 \) and the input signal between \( t_0 \) and \( t \) as follows:
\[ x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)\,d\tau \]  

(6.3)

If the system is stable, \( e^{A(t-t_0)}x(t_0) \) will tend to zero as \( t \) becomes large and the influence of the initial states will be insignificant. However, for systems with short transients, \( t \) will be small and the influence of initial states will be significant. Consequently, the same input applied to a system with different initial states may give rise to completely different responses. Thus, the initial states of the system response must be estimated.

This paper presents a subspace identification method to identify systems using ensembles of short transients collected from repeated experiments. It estimates a state space model and the associated initial conditions directly from the ensemble data. This paper is structured as follows. Section II describes the ensemble-based subspace method and demonstrates its application with data from a simulation of dynamic ankle stiffness. Section III extends the method for use with time-varying systems and demonstrates its application to a simulated second-order system whose gain varies with time. Section IV summarizes this paper’s contributions, discusses its advantages and limitations, and describes areas for future work.

6.3 Identification of Systems from Short Transients

6.3.1 Identification Method

This section presents an algorithm to estimate the system matrices and the associated initial states of a state space model from an ensemble of input-output records. The method used to estimate the system matrices is based on that in [9]. The initial states
are estimated by modifying the standard subspace method to utilize ensemble data.

It will be assumed that the experiment can be repeated so that multiple trials are available; the problem is to estimate the state space model:

\[
\begin{align*}
  x_j(k+1) &= Ax_j(k) + Bu_j(k) \\
  y_j(k) &= Cx_j(k) + Du_j(k)
\end{align*}
\] (6.4)

where

\[
\begin{align*}
  u_j(k) &\in \mathbb{R}^m \text{ and } y_j(k) &\in \mathbb{R}^l \text{ are vectors containing measurements, at discrete time } k \text{ of trial } j, \text{ of the system’s } m \text{ inputs and } l \text{ outputs.} \\
  x_j(k) &\in \mathbb{R}^n, \text{ is the state vector of the process at discrete time } k \text{ of trial } j.
\end{align*}
\]

In particular, for a trial with length of \( h+1 \), the input and output data from the \( j \)th experimental trial are related as follows:

\[
\begin{align*}
  y_j(0) &= Cx_j(0) + Du_j(0) \\
  y_j(1) &= Cx_j(1) + Du_j(1) \\
  &= CAx_j(0) + CBu_j(1) + Du_j(1) \\
  y_j(2) &= CA^2x_j(0) + CABu_j(0) + CBu_j(1) + Du_j(2) \\
  &\vdots \\
  y_j(h) &= CA^hx_j(0) + CA^{h-1}Bu_j(0) + \cdots + CBu_j(h-1) + Du_j(h)
\end{align*}
\] (6.5)

If the length of all trials is the same, the ensemble of the input and output data can be formulated as the matrix equation:
Equation 6.6 can be rewritten as

\[ Y = X_j \Gamma_h^T + UH \quad (6.7) \]

with

\[
Y = \begin{bmatrix} y_1(0) & y_1(1) & \cdots & y_1(h) \\ y_2(0) & y_2(1) & \cdots & y_2(h) \\ \vdots & \vdots & & \vdots \\ y_p(0) & y_p(1) & \cdots & y_p(h) \end{bmatrix},
\]

\[
U = \begin{bmatrix} u_1(0) & u_1(1) & \cdots & u_1(h) \\ u_2(0) & u_2(1) & \cdots & u_2(h) \\ \vdots & \vdots & & \vdots \\ u_p(0) & u_p(1) & \cdots & u_p(h) \end{bmatrix},
\]

\[
\Gamma_h = \begin{bmatrix} C & CA & \cdots & CA^h \end{bmatrix}.
\]

\[
X_j = \begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix}
\]
contains the internal state at the first sample for all trials,

\[
H = \begin{bmatrix} D & CB & \cdots & CA^{h-1}B \\ 0 & D & \cdots & CA^{h-2}B \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & D \end{bmatrix}.
\]

To estimate the system matrices, the subspace, \( X_j \Gamma_h^T \), that contains information
only from the internal states, must be estimated. To achieve this, generate the matrix

\[ [U \quad Y] \text{ and then use QR factorization [10] to get:} \]

\[
[U \quad Y] = [Q_1 \quad Q_2] \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} \tag{6.8}
\]

where

- \( Q_1 \) is the same size as \( U \),
- \( Q_2 \) is of the same size as \( Y \),
- \( Q_1 \) and \( Q_2 \) are orthogonal, that is \( Q_1 Q_2^T = 0, \ Q_1 Q_1^T = I, \ Q_2 Q_2^T = I \).

From Equation 6.8, we have

\[
Q_2^T Y = Q_1^T Q_1 R_{12} + Q_2^T Q_2 R_{22}
= R_{22} \tag{6.9}
\]

Multiplying both sides of Equation 6.7 by \( Q_2^T \) gives

\[
Q_2^T Y = Q_2^T X \eta_N^T + Q_2^T U H
= Q_2^T X \eta_N^T + Q_2^T Q_1 R_{11} H
= Q_2^T X \eta_N^T \tag{6.10}
\]

From Equations 6.9 and 6.10, it can be concluded that \( R_{22} \) is the subspace from internal states.

\[
R_{22} = Q_2^T X \Gamma_N^T \tag{6.11}
\]

The Singular Value Decomposition (SVD) [10] of \( R_{22} \) is
\[ R_{22} = \mu \zeta v^T \]  

(6.12)

So \( \hat{\Gamma} \) is given by the first \( n \) left singular values of \( v \), where \( n \) is the order of the system. Once \( \hat{\Gamma} \) is available, \( C \) is given by the first \( l \) rows of \( \hat{\Gamma} \), where \( l \) is the number of the outputs.

\[ \hat{C} = \hat{\Gamma}(1:l,:) \]  

(6.13)

Then, \( A \) can be estimated by solving the least-squares problem:

\[ \hat{A} = \Gamma_1^T \Gamma_2 \]  

(6.14)

where

\[
\begin{bmatrix}
C \\
CA \\
\vdots \\
CA^{k-2}
\end{bmatrix}
\]

is the top \((i-1)l\) rows of \( \hat{\Gamma} \),

\[
\begin{bmatrix}
CA \\
CA^2 \\
\vdots \\
CA^{i-1}
\end{bmatrix}
\]

is the bottom \((i-1)l\) rows of \( \hat{\Gamma} \),

\( \dagger \) indicates the Moore-Penrose pseudo-inverse [11].

Once \( \hat{A} \) and \( \hat{C} \) are known, the standard subspace method can be used to estimate \( x_j(0) \) and the system matrices \( B \) and \( D \). In particular, the output is a linear function of:

\[ B \] and \( D \)

\[ y_j(k) = CA^k x(0) + \sum_{\tau=N}^{N+k-1} CA^{k-\tau} Bu_j(\tau) + Du_j(k) \]  

(6.15)
Then the initial states, \( x_j(0) \), and the system matrices \( B \) and \( D \) can be estimated in a least square sense since \( y_j(N+k) \), \( u_j(\tau) \), \( C \) and \( A \) are known. Equation 6.15 can be rewritten using the Kronecker product \( \otimes \).

\[
y(k) = CA^k x(0) + \left[ \sum_{\tau=0}^{k-1} u(\tau)^T \otimes CA^{k-1-\tau} \right] vec(B) + \left[ u(k)^T \otimes I_j \right] vec(D) \tag{6.16}
\]

where the \( vec(\cdot) \) terms result from stacking the columns of the matrix \( \cdot \) on top of each other.

Equation 6.16 defines the output signal at time, \( k \). Therefore constructing a vector containing the output signal \( y(k) \) for samples 1 to \( k \) gives:

\[
Y = \begin{bmatrix} \Gamma_N & E_y & E_u \end{bmatrix} \begin{bmatrix} x_0 \\ vec(B) \\ vec(D) \end{bmatrix} \tag{6.17}
\]

where

\[
\Gamma_N \text{ is the Observability matrix,}
\]

\[
Y = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(h-1) \end{bmatrix} \tag{6.18}
\]

\[
E_y = \begin{bmatrix} 0 \\ u(0)^T \otimes C \\ \vdots \\ \sum_{\tau=0}^{h-2} u(\tau)^T \otimes CA^{h-2-\tau} \end{bmatrix} \tag{6.19}
\]
\[
E_u = \begin{bmatrix}
  u(0)^T \otimes I_f \\
  u(1)^T \otimes I_f \\
  \vdots \\
  u(h-1)^T \otimes I_f
\end{bmatrix}
\]

Consequently, the matrices \(B, D\) can be estimated by solving the least squares problem:

\[
\begin{bmatrix}
  x_0 \\
  \text{vec}(B) \\
  \text{vec}(D)
\end{bmatrix} = \begin{bmatrix}
  \Gamma_N & E_x & E_u
\end{bmatrix}^T \mathbf{Y}
\]

### 6.3.2 Simulation Study

We tested and validated the algorithm by simulating a model of ankle joint stiffness. Dynamic joint stiffness is a complex biological system that is important in neuromuscular control [12]. A parallel cascade model [1] has been developed that describes the dynamic behavior of ankle joint stiffness. Therefore, we opted to use the parallel cascade model to test the ensemble subspace method.

Joint stiffness describes the dynamic relationship between the angular position of a joint and the torque acting about it. Joint stiffness can be separated into intrinsic stiffness and reflex stiffness. The intrinsic component, modeled as a linear system, is due to the mechanical properties of the joint, passive tissue, and active muscle fibers. The reflex component is due to muscle activation in response to the activation of stretch receptors in the muscle and can be modeled as a Hammerstein system. A two-pathway, parallel cascade model [1], with the intrinsic stiffness in one pathway and the reflex stiffness in the other, can be used to describe joint stiffness, as shown in Figure 6.1.
Figure 6.1: Parallel-cascade model for dynamic ankle stiffness. *pos* denotes the position. $tq_I$, $tq_R$ and $tq_N$ denote intrinsic torque, reflex torque and net torque. The intrinsic stiffness is modeled as a linear system. The reflex stiffness is modeled as a delay, followed by a differentiator and a Hammerstein system comprised of a series connection of a static nonlinearity and a linear system.
Previously, we developed a state space model \[4\] \[5\] for the overall parallel cascade model of ankle joint stiffness as:

\[
x_R(k+1) = A_R x_R(k) + \begin{bmatrix} 0 & \hat{B} \end{bmatrix} \begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix}
\]

\[
tq_N(k) = C_R x_R(k) + \begin{bmatrix} D_I & \hat{D} \end{bmatrix} \begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix}
\]

(6.22)

where

\( tq_N(k) \) is the measured net torque,

\( A_R, \hat{B}, C_R, \hat{D} \) are the system matrices for reflex stiffness,

\( D_I \) is the system matrices for intrinsic stiffness,

\( U_I(k) = [p(k) \text{ vel}(k) \text{ acc}(k)] \) is a set of constructed inputs to intrinsic stiffness, where \( p(k) \) is measured position, \( \text{vel}(k) \) is the differentiated position, \( \text{acc}(k) \) is second-order differentiated position,

\( U_R(k) = [ch_1(k) \text{ ch}_2(k) \cdots \text{ ch}_p(k)]^T \) is a set of constructed inputs to reflex stiffness, where \( \text{ch}_p(k) \) are the terms of the basis function describing the static nonlinearity in the Hammerstein system.

For this state space model, the input \( \begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix} \) can be constructed from the measured input, and the output \( tq_N(k) \) can be obtained directly from the measured data.

Therefore, the state space model for ankle joint stiffness can be estimated from \( \begin{bmatrix} U_I(k) \\ U_R(k) \end{bmatrix} \) to \( tq_N(k) \). Simulating the estimated system with the appropriate inputs permits the
torque from the intrinsic and reflex stiffness to be predicted. Specifically, the output from
the simulation with the input signal $\begin{bmatrix} U_I(k) \\ 0 \end{bmatrix}$ is an estimate of the intrinsic torque.

Similarly, the response to the input $\begin{bmatrix} 0 \\ U_R(k) \end{bmatrix}$ is an estimate of the torque from the reflex stiffness.

Simulated data were generated using Matlab’s Simulink [13] with the model
diagram shown in Figure 6.2. Intrinsic stiffness was modeled as a second-order, quasi-
linear system with transfer function:

$$TQ_I(s) = \frac{0.015s^2 + 0.8s + 150}{\theta(s)}$$

(6.23)

where

$\theta$ is the joint angle,

$TQ_I$ is the torque from the intrinsic stiffness.

Reflex stiffness was modeled by a series connection of a differentiator, a delay for
40ms [1], a half-wave rectifier and a second order low-pass filter as

$$\frac{TQ_R(s)}{V_R(s)} = \frac{3200}{s^2 + 80s + 1600}$$

(6.24)

where

$TQ_R$ is the reflex torque,

$V_R$ is the half-wave rectified and delayed joint angular velocity.
Figure 6.2: Matlab Simulink model for ankle stiffness. The noise-free intrinsic and reflex torques are recorded in the simulation for comparison with the estimates from the algorithm.
The net torque was calculated by adding the intrinsic and reflex torques as:

\[ TQ_N = TQ_I + TQ_R + w \]  \hspace{1cm} (6.25)

A white noise, \( w \), with SNR as 40dB was added to the net torque to model measurement noise. A Pseudorandom Binary Sequence Signal (PRBS) was used as the external position input. Each simulation trial lasted for 10 seconds; 1000 trials were generated. The input and output data were sampled at 1 KHz and then decimated to 100 Hz.

We generated an input-output ensemble from these data where each realization consisted of 50 points, from 5s to 5.5s. Figure 6.3 shows the simulated position and torque signals for a subset of this ensemble. The transient algorithm was applied to this ensemble to estimate a state space model of joint stiffness. The quality of the resulting estimate was evaluated by using it to predict a full trial lasting for 10s. Figure 6.4 shows the total, intrinsic and reflex torques from the simulation with those predicted from the model estimate superimposed. Clearly, the simulated and estimated torques are very similar.
Figure 6.3: Simulated joint stiffness data. Five realizations of (A) position and (B) Net torque output used to identify ankle stiffness. Note that there were 1,000 realizations in the ensemble and each ensemble was only 500 ms long (50 points at 100 Hz sampling rate).
Figure 6.4: Simulated and estimated torques from the simulation study for system with short transients. (A) Net torque; (B) Intrinsic torque; (C) Reflex torque.
Next, we investigated how many trials were required for the subspace method to provide a good estimate. The percentage Variance Accounted For (%VAF) was used to measure how well the estimated model predicted the true torque. The VAF between the true torque, $y$, and estimated torque, $y_{est}$, was calculated as:

$$% \text{VAF} = \left(1 - \frac{\text{variance}(y - y_{est})}{\text{variance}(y)}\right) \times 100\% \tag{6.26}$$

Figure 6.5 shows the performance of the algorithm as a function of the number of realizations. The VAFs of reflex and net torque were low when the number of realizations was small but increased monotonically as the number realizations increased. The estimates of intrinsic torque were accurate even when the numbers of realization were small. A possible explanation of this phenomenon is that it results from the simple structure of the intrinsic stiffness. The dynamics for intrinsic stiffness in the state space model (Equation 6.22) comprise only a constant gain matrix, $D_j$. Therefore, the intrinsic stillness can be viewed as a static system while the reflex stiffness has a more complicated MISO dynamic system that requires more data to estimate accurately.
Figure 6.5: VAFs between simulated and predicted torques as a function of the number of realizations
6.4 Identification of Time-Varying Systems

6.4.1 Identification Method

The subspace method can also be used to identify time-varying systems from an input-output ensemble. In particular, if it is assumed that:

1. The time-varying behavior of the system can be treated as time-invariant provided the length of each segment is short enough.
2. An ensemble of input and output records having the same time-varying behavior, is available.

Then, a state space model, at each time sample, can be estimated from short segments of this ensemble using the method presented in Section II. Thus, the state space model at time $k$, given by Equation 6.27, can be identified using the ensemble subspace method.

$$
\begin{align*}
\mathbf{x}_j(k+1) &= A(k)\mathbf{x}_j(k) + B(k)\mathbf{u}_j(k) \\
\mathbf{y}_j(k) &= C(k)\mathbf{x}_j(k) + D(k)\mathbf{u}_j(k)
\end{align*}
$$

(6.27)

where

$$
A(k) \in \mathbb{R}^{n \times n}, \quad B(k) \in \mathbb{R}^{n \times m}, \quad C(k) \in \mathbb{R}^{l \times n} \quad \text{and} \quad D(k) \in \mathbb{R}^{l \times m}
$$

are the system matrices at sample time $k$.

Then, by repeatedly applying the algorithm to different segments of the ensemble data, the matrices $A(k)$, $B(k)$, $C(k)$ and $D(k)$ can be estimated at each sample time.

6.4.2 Simulation Studies

To demonstrate the time-varying application of the method we simulated a linear,
time-varying system by multiplying the second-order, linear, low-pass filter of Equation 6.14, by a time-varying gain, $g(t)$, using the Simulink model shown in Figure 6.6. The time-varying gain followed the ramp-and-hold trajectory shown by the blue curve in Figure 6.8. Each simulation trial used a different Gaussian noise as the input and lasted for 1 second; 1000 trials were generated each with a different input realization. There was no additive noise. The simulated data was sampled at 100 Hz.

The time-varying, state space model identified from the ensemble data comprised a separate state space model (Equation 6.27) estimated at each time sample, $k$, using the transient method with a segment length of 5 samples. Thus, the state space model at sample 1 was estimated using the ensemble data constructed from the 1st to the 6th points in each of the 1000 simulation trials. The state space model at sample 2 was estimated using the ensemble data constructed from the 2nd to 7th data points in each trial. Therefore, a time-varying state space model (Equation 6.27) was estimated.

Once the estimated time-varying system was available, its predicted output was obtained by simulating the estimated time-varying system with the input from simulation. Figure 6.7 shows the simulated and estimated output from one trial. The estimated and simulated outputs are almost identical indicating that the subspace method identified the time-varying system accurately. To estimate the time-varying values of $k$, the estimated time-varying state space model was first transformed to a time-varying transfer function [14]. Then, the value of $k$ was obtained by calculating the static gain of the transfer function by changing the transfer function to a zero, pole, state gain format. Figure 6.8 shows the estimated and simulated value of $k$. It is evident from these results that the identification algorithm correctly tracked the time-varying properties of the simulated
system.

Figure 6.6: Matlab Simulink model of a second-order, time-varying system.
Figure 6.7: Simulated (blue) and estimated (red) outputs for one simulation trial of the second-order, time-varying system.
Figure 6.8: Simulated and estimated gain ($k$) as a function of time for the second-order, time-varying simulation.
6.5 Discussion and Conclusion

6.5.1 Effect of Noise

This paper demonstrates how the ensemble subspace method can be used to estimate state space models of time-invariant systems from input-output ensembles of short transients. Furthermore, if all members of the input-output ensemble have the same time-varying behavior the same approach can be used to estimate the dynamics of time-varying systems. Simulations demonstrated the application of both approaches. However, it should be noted that these two simulations treated measurement noise rather differently. Thus, the first simulation study of time-invariant systems modeled measurement noise as an additive Gaussian white noise with a SNR as 40 dB. The results demonstrated that the algorithm provided accurate estimate with the presence of noise. In the second simulation study of time-varying systems, no noise was added because the time-varying formulation does not yet handle noise properly.

In general, subspace methods use instrumental variables (IV) to eliminate the effects of noise. In our laboratory, we have implemented several different IV structures to estimate ankle joint stiffness from noisy measurements with the choice of the IV depending on the experimental situation. Thus, the past input was used as the IV to estimate the joint stiffness in open loop [4]; while both the past input and past output were used as IVs to estimate the joint stiffness in closed loop [5]. Both algorithms provide accurate estimates even with significant noise. However, both algorithms use previous or “past” data as the IVs. In practice, the number of data points between the “past” IV and current data must be chosen to be at least twice order of the system; larger
values may be necessary if the noise is not white. The state space model for ankle joint stiffness (Equation 6.22) is a second order or sometimes a third order system. Therefore, at least 5 data points are required between the IV and current data. Moreover, the column numbers of Hankel matrices, $U$ and $Y$, in Equation 6.7 must be at least twice the system order. Therefore, at least 10 points of data, 5 points for the current data and an additional 5 points for the IV, are required to estimate the ankle joint stiffness. In the first simulation study for ensemble subspace method, the ensemble data was constructed using 50 data points from each trial. Therefore, each trial of ensemble data was long enough to permit the use of IVs to eliminate noise effects.

However, implementing the IV structure for time-varying system is more difficult. The time-varying system is treated as a time-invariant system provided the segment is short enough. If the segment is chosen to be too long, the time-varying behavior of system, especially for systems whose time-varying behaviors which change rapidly may not be estimated. Therefore, the segment of data chosen should be as short as possible. The subspace method requires that one segment of data must be at least twice the system order. Provided that the system is a second order system, the ensemble data was constructed by using 5 points of data from each trial. However, at least 10 points of data are required to implement the IV structure and estimate the system dynamics. Thus, the data length is too short to implement the IV structure.

The classic IV structures of subspace method may be applied to time-varying system identification, provided that the system’s behavior does not vary to quickly. Thus a segment of data can be viewed as time-invariant and implemented the IV structure to estimate system dynamics from noisy measurement. However, a novel IV structure needs
to be developed to identify the time-varying systems whose dynamics change fast with time. A further development of this research topic would be to find another IV appropriate for time-varying system identification using the ensemble subspace method. In general, the IV is required to be uncorrelated with the noise, but correlated with the states. A possible solution might be to use the input data from another trial as the IV. In particular, if the noise is Gaussian, the noise in one trial should be uncorrelated with the input signal from another trial. Besides, if the input is persistently exciting, the input signal is correlated with the states [15]. Therefore, the data from another trial can be used as IV. The noise effects could be eliminated by projecting the data from one trial onto the one from another trial.

### 6.5.2 Initial States

Compared with other ensemble identification method, such as non-parametric method [8], a significant contribution of this paper is to provide a method to estimate the initial states of the system from short transients. Initial state estimation may not be important if the system has long response, since the zero-input response, $CA^t x_0$, will decay to zero if system is stable. However, initial states are important for short transient data sets because the $CA^t x_0$ term can have a large influence if the time constant, $k$, is not large enough. Therefore, the initial state must be estimated to predict the response accurately. This paper shows that the initial states of the system can be estimated with the system matrices. This makes it possible to identify biomedical systems with short transients and switching between different operating points, such as Vestibulo-Ocular Reflex (VOR) [7].
6.5.3 Data Segment Length and Number of Trials

The subspace method requires that the length of the data segment to be at least twice of the order of the system to ensure identifiability [14]. However, it is likely that with the ensemble algorithm the segment length must be longer to permit estimation of the initial conditions. Therefore, the length of the segment in each trial, $h$, should be more than twice the order, so that the ensemble subspace can estimate a model of the system. However, it is still unknown how many points the subspace method needs to estimate the initial states. The simulation study showed that the subspace method could estimate the initial state with only 50 points of data when the SNR was 40 dB. However, a more noisy measurement may require more points of data to estimate the initial states. We also observed that this ensemble-based subspace method required a relatively large number of the experimental trials to get an accurate estimate. However, we expect the number of trials required will depend on the length of each trial and the noise level. In principle, for a given noise level, the number of required trials should decrease as the length of each trial increases. Conversely, for a fixed trial length, the number of trials required to obtain an accurate identification should increase with the noise level. This behavior needs further investigation since each additional trial not only increases the data available for identification but also increases the complexity of the model due to the need to estimate the initial states.

6.5.4 Different Data Segment Lengths

As formulated, the algorithm requires the length of the segment of each experiment trial to be the same. Ideally, this condition would be met by collecting the
same length of data for all trials. When this is not feasible, because of the nature of the experiment or unexpected interruption, the data length for the identification, $h$, can be chosen as the shortest collected data length from all the trials. One topic of further research would be to extend the algorithm to work with data ensembles where the trials had different lengths. One approach, applicable when most trials are of the same length and a few trials are shorter, would be to construct Equation 6.2 by filling the missing samples with zeros. This can be viewed as adding extra process and measurement noise to the system. Subspace methods use instrumental variables to eliminate the noise. Unbiased estimates of the system can be obtained by choosing different instrumental variables. Therefore, choosing appropriate instrumental variable may solve the problem.

### 6.5.5 Conclusion

In this paper, we presented a subspace method to estimate biomedical systems from short transients. This is usually a difficult problem because the input and output records are not long enough to carry sufficient information to identify the system. The method requires multiple experiments, each input and output sequence is recorded while the system displays the same dynamics. The simulation study shows that the algorithm provides an accurate estimate of the system dynamic from the ensemble data.

The idea of using ensemble data to estimate system dynamics can be extended to estimate systems with time-varying behavior. A state space model for each sample can be estimated from the ensemble data using subspace method. In this paper, we showed that a time-varying, state space model can be estimated using ensemble data. The simulation study demonstrated that the algorithm provided accurate estimate for a linear, time-
varying system.

**Acknowledgements**

Supported by grants from the Natural Sciences & Engineering Research Council of Canada, and the Canadian Institutes of Health Research.
Reference


7 Discussion & Conclusions

The concept of dynamic joint stiffness can be used to study the mechanical behavior of the mechanisms acting about the ankle. Ankle joint stiffness plays an important role in the control of posture, where it defines the mechanical response to external perturbations, and of movement, where it determines the forces that must be generated to move the limb. Ankle joint stiffness comprises two components: intrinsic stiffness and reflex stiffness [1]. Intrinsic stiffness arises from the mechanical properties of the joint, passive tissue, and active muscle fibers. Reflex stiffness arises from muscle contraction in response to reflex activation from stretch receptors in the muscle.

Examining intrinsic and reflex stiffness is difficult because intrinsic and reflex torques cannot be measured separately experimentally; only their sum can be measured. Consequently, intrinsic and reflex stiffnesses cannot be studied separately directly. One approach is to use system identification methods to separate the intrinsic and reflex torques analytically. Several methods, including the non-parametric parallel cascade method [1], parametric method [2] and real-time method [3], have been developed to separate intrinsic and reflex torques from the net torque, and thus study the intrinsic and reflex stiffness individually.

These methods require the experimental data to be collected from an open-loop experiment, where the torques developed at the ankle will not change the joint position. Such open-loop experiments have provided a good understanding of the behavior of intrinsic and reflex stiffness under these experimental conditions [1] [4], which correspond to interacting with a very stiff environment. However, these conditions are far
from realistic because, during most functional tasks and movements, subjects will interact with compliant loads and consequently the joint torque and position will be connected by feedback. Any change in torque will evoke a change in position and vice versa. Torques and positions measured under such conditions are therefore measured in closed-loop.

It is known that correlation-based identification methods, such as those we have used, will give biased results when applied to closed-loop data [5]. Therefore new tools are needed to estimate intrinsic and reflex stiffness while subjects interact with compliant environments. In particular, a new model structure for joint ankle stiffness needs to be developed. After that, identification methods must be developed to estimate the new model structure in closed-loop.

To achieve this objective, we developed a state space model for dynamic ankle stiffness in this thesis. Chapter 4 presented this novel state space model and showed that by using subspace methods, the model can be estimated directly from measurements of position and torque. To validate the state space model and subspace identification method, simulation and experimental studies were carried that demonstrated that the new method provided accurate estimates of ankle joint stiffness in open-loop. Chapter 5 considered the identification of ankle joint stiffness in closed-loop. Additional simulation and experimental studies showed, as expected, that the open-loop algorithm would provide biased estimates when used with data collected from closed-loop experiments. Consequently, we modified the subspace method to estimate ankle joint stiffness in closed-loop. The simulation and experimental studies demonstrated that the modified subspace method provides accurate estimates of intrinsic, reflex and net torques using
data collected from closed loop.

We also observed that the subspace method might need a large amount of data to estimate system dynamics. However, for many biomedical problems, it is difficult or impossible to acquire long records from a system in the same condition. Nevertheless, it may be possible to obtain ensembles of short records. Chapter 6 presents a subspace method to identify state space models for biomedical systems from such ensembles of short transients. The second half of Chapter 6 shows how the ensemble subspace method can be used to identify the systems with time-varying behavior.

The state space model was developed specifically for ankle joint stiffness, but can be implemented to estimate other systems with parallel cascade structures. The subspace method can be used to identify the systems in either open-loop or closed-loop. The ensemble subspace method can only estimate time-varying systems whose behavior changes slowly with time because the algorithm treats the time-varying system as a time-invariant system, for short data segments. The minimal length of the segment must be at least twice of the system order. Therefore, if the time-varying behavior changes during this small segment, the subspace method won’t be able to estimate the time-varying behavior.

7.1 Statement of Original Contributions

The original contributions of this thesis comprise the development and validation of subspace based methods for the identification of dynamic ankle stiffness under a number of different conditions. These include:
7.1.1 Identification of Dynamic Ankle Stiffness in Open-Loop.

We developed a novel state space model for dynamic ankle joint stiffness in which the nonlinear, SISO, parallel cascade model of ankle dynamics was transformed into a linear, MISO state space model with constructed inputs from the measured position signal and terms of a nonlinear basis function. This formulation is an original contribution and is significant since it makes it possible to use state space identification methods to study joint stiffness. Thus, once this new state space model for ankle joint stiffness was formulated, we were able to use the PI-MOESP algorithm to estimate its parameters. Simulating this state space model with appropriate inputs gives estimates of the intrinsic and reflex torques. Simulation and experimental studies showed that the new subspace method provides results similar to those of the parallel cascade method under most normal estimation conditions. However, it was also shown that the subspace method provides better estimates for the most severe estimation conditions (e.g. high noise, short sample lengths). Moreover, in contrast to the parallel cascade identification method, the state space model does not require iteration and so is computationally more efficient and has better convergence properties.

7.1.2 Identification of Ankle Joint Stiffness in Closed-loop.

The results of the research described in this thesis make it evident that identifying ankle joint stiffness when the joint interacts with a compliant load is a closed-loop identification problem. This is difficult since two problems must be solved at the same time. First, the output from each parallel pathway must be separated from the overall measurement. Second, an identification algorithm must be chosen that eliminates the
effects of feedback noise. This thesis, contributes to our understanding of this problem by
demonstrating that: (1) Correlation-based methods will give biased results when applied
to closed loop data; (2) Accurate estimates of ankle joint stiffness can be obtained from
closed-loop data by applying the EIV-MOESP method to the novel MISO state space
model for ankle dynamics.

Thus, we first transform the SISO, nonlinear parallel cascade model to a MISO
linear model. Then, we use the EIV-MOESP algorithm to eliminate the feedback noise
and estimate this MISO state space model. Simulating the estimated model with the
proper inputs permits the outputs from each pathway to be estimated.

Simulation and experiment studies demonstrated that when used with closed-loop
data, the parallel cascade method produces biased estimates of ankle stiffness. In contrast,
the subspace-based method provided accurate, unbiased estimates of ankle dynamics
using the same data.

The MISO state space model was developed specifically to study ankle joint
stiffness, however the basic idea can be extended to any nonlinear system with parallel
cascade structure. Indeed, if the structure of the nonlinear system with parallel cascade
structure is known, this nonlinear system can be described as a linear MIMO state space
model, using appropriate constructed inputs. Therefore, by using the EIV-MOESP
method feedback noise can be eliminated and the state space model estimated accurately.
However, the structure of the system must be known a priori. Therefore, this method is
not suitable for estimating “black box” systems, whose dynamics and structure are
completely unknown before the identification is conducted.
7.1.3 Identification of Ankle Joint Stiffness from Short Transients.

We presented a subspace method to estimate biomedical systems from short transients. The algorithm is based on the ensemble subspace method described in [6], but was extended to estimate the system matrices and initial states. The method needs multiple trials, each containing an input and output sequence recorded while the system displays the same dynamic behavior. Simulation results showed that the algorithm provides an accurate estimate of the system dynamics from ensemble data.

Time varying systems can be also identified using the ensemble subspace method by acquiring a series of realizations in which the systems undergoes the same time-varying behavior. The time-varying system is treated as a time-invariant system for a short segment of data and identified using the ensemble method. Separate models are identified in this way for segments at different times to characterize the time-varying behavior. The algorithm accurately estimated the dynamics of a simple, linear, time-varying system. However, a simulation study demonstrated that the segment must be at least twice of the system order in order to use subspace methods. This will determine the nature of the time-varying behavior that the method can be applied to identify time-varying system whose dynamics changes very fast with time.

7.2 Future Work

As with all scientific research, progress brings new questions. Much work remains to be done. In this section, we discuss the new problems we found during our research and propose possible developments.
7.2.1 Basis Function

A set of nonlinear basis functions is used frequently in this thesis to transform the SISO, nonlinear, Hammerstein system into a MISO, linear, state space model. Based on previous results from the parallel cascade method [1], we opted to use Chebyshev polynomials as the basis function. The order of this Chebyshev polynomial was chosen iteratively. At first, a high order (>7) Chebyshev polynomial was used as the basis function and the identification procedure was executed. Then estimated parameters for the nonlinearity were examined. The scale factor for each basis function term is available once the reflex torque has been estimated. Basis function terms having scale factors with small numeric values are likely to make no significant contribution to the reflex torque and so can be excluded without influencing the reflex torque. Thus, by discarding such terms, we determined that a fifth-order basis function was adequate. However, this method may not be the optimal solution. It is possible the terms with insignificant numeric contribution to the reflex torque may contain valuable information about the system dynamics. However, how to choose the optimal basis function and its order *a priori* is unknown. Further development of this topic is needed to provide an analysis tool to obtain a basis function with an optimal order to identify a nonlinear system with a cascade structure, such as a Hammerstein system or a Wiener system. An interesting solution for this problem might be to optimize the order of basis function and its scale factors using Least Square Support Vector Machine (LS-SVM) [8], which can estimate the order of basis function from the input and output measurements, using linear optimization technologies. In particular, a set of least squares support vector machines (LS-SVMs) regression problems can be solved to estimate the linear model and static
nonlinearities using a low-rank approximation.

7.2.2 Systems with Time-Varying Behavior

Chapter 6 presents a subspace method that estimates a state space model from an ensemble of input and output trials. We used this method to identify time-varying systems by treating them as a series of time-invariant systems using a series of short segments. The time-varying system was treated as a time-invariant system, provided the segment length in the data ensemble was short enough. The simulation study showed that this method provided accurate estimates for systems with slowly varying variations. In the example used, the segment of data was chosen to be the shortest length needed to implement the subspace. As a result, the segment was too short to apply the classic instrumental variable approach to eliminate noise. Therefore, no noise was added to the simulation study. Thus, this algorithm is not yet at a stage appropriate to solve “real world” problems, where the signal measurements contain significant noise. To improve this algorithm, we propose to extend the algorithm to two different areas.

First, we plan to expand the algorithm to estimate a time-varying system whose time-varying behavior can be modulated by a linear relationship. In particular, the time varying system matrices, \( A(k), B(k), C(k), D(k) \), will be described by a set of linear autoregressive functions, \( f(k), g(k), h(k), j(k) \), respectively, so that the time-varying system as

\[
\begin{align*}
x(k + 1) &= A(k)x(k) + B(k)u(k) \\
y(k) &= C(k)x(k) + D(k)u(k)
\end{align*}
\]  

(7.1)

can be transformed to
In this formulation, the segment length does not need to be short to be treated as time-invariant system, because the time varying system matrices are modeled by linear autoregressive models with extra parameters. Rather, a longer segment can be used since each data point can be described by the system matrices and the linear functions. Therefore, the IV structure can be implemented to estimate the system matrices from noisy measurement. Once the system matrices are estimated, the linear functions $f(k), g(k), h(k), j(k)$ can be obtained by linear factorization.

Another approach will be to use the data from different segments as the IV, $\Theta$. The IV must be uncorrelated with the noise, but correlated with the states. If the noise is white, the data from another segment, the input signal for example, will be a good candidate for use as the IV. In particular, the noise from one segment will be uncorrelated with the input from another segment because of the independent property of white noise. Moreover, if the input from another segment is persistently exciting, $\lim_{N \to \infty} \frac{1}{N} X_{i,N} \Theta$ will have full rank as discussed in chapter 5. Therefore, the system matrices can be estimated from the noisy measurement.

### 7.2.3 Experimental Application

In this thesis, we presented a state space model and a subspace method to identify the ankle joint stiffness against a compliant load. This is an important development since it enables us to study reflex stiffness when the subject performs tasks which are closer to...
those encountered in everyday life. Currently, experiments for studying the ankle joint stiffness are mainly conducted in open-loop experiment, where the subject cannot move his/her ankle freely. The subspace method enables us to study the ankle dynamics in closed-loop, where the ankle is against a compliant load and the position and torque can change together. We will conduct more experiments to study the ankle joint stiffness in closed-loop and hopefully be able to discovery some more interesting results.

7.2.4 Reference Input As Instrumental Variable for Closed-loop System Identification

The past input and past output are used as the IV to estimate the state space model of ankle joint stiffness against a compliant load. The follow matrix multiplication is required before the LQ factorization is conducted.

\[
\begin{bmatrix}
    U_{i,j,N} & Y_{i,j,N}
\end{bmatrix}
\begin{bmatrix}
    U_{i-p,j,N} & Y_{i-p,j,N}
\end{bmatrix}^T =
\begin{bmatrix}
    L_{11} & 0 \\
    L_{21} & L_{22}
\end{bmatrix}
\begin{bmatrix}
    Q_1 \\
    Q_2
\end{bmatrix}
\]  

(7.3)

However, considering the input and the IV are constructed inputs for intrinsic and reflex stiffness, \( U_{i,j,N} \) is large. In particular, the number of rows is equal to the product of the number of inputs and times the value of \( j \). Multiplication of two such large matrices is not efficient computationally and may lead to serious numerical error. Therefore, we investigated another IV structure, which is uses the external position perturbation as the IV, to estimate the ankle joint stiffness in closed-loop. The LQ factorization is applied a matrix as follows:
\[
\begin{bmatrix}
U_{i,j,N} \\
R_{1-p,i,N} \\
Y_{i,j,N}
\end{bmatrix} =
\begin{bmatrix}
L_{11} & 0 & 0 \\
L_{21} & L_{21} & 0 \\
L_{31} & L_{32} & L_{33}
\end{bmatrix}
\begin{bmatrix}
Q_1 \\
Q_2 \\
Q_3
\end{bmatrix}
\]  

(7.3)

where

\[U_{i,j,N}, R_{1-p,i,N} \text{ and } Y_{i,j,N}\] are Hankel matrices for current input, IV and current output.

In particular, the IV, the external perturbation, contains only one position input. Therefore, the dimension of \(R_{1-p,i,N}\) is much lower than that of \(U_{i-p,j,N}\), so that it should be more efficient in computation. The algorithm and results are presented in [9]. Though this new IV structure provided reasonable result for simulation studies, the results for experimental studies from IV and EIV methods were rather different. A further development would be to conduct more experimental studies to compare those two results and find out which algorithm is more suitable to identify MIMO system, with large numbers of inputs and outputs in closed-loop.
References:


Appendix

A. Ethics Certificate

Please find attached our Ethics Certificate dated May 11, 2009, valid until April 12, 2010.
B. Waivers From Co-Authors

Please find attached a waiver from Dr. David Westwick allowing the manuscript titled “Identification of Human Ankle Joint Stiffness Using Subspace Methods Part 1. Open-loop System identification”, which he co-authored to be included in this thesis.