University of Sussex

Chaotic Exploration and Learning of Locomotor Behaviours

Yoonsik Shim

Submitted for the degree of *Doctor of Philosophy*
University of Sussex
January 2012

Acknowledgements

After spending more than three decades of my life as a student, now I am merely standing on the doorstep to the realm of academia, holding this thesis as a key, although far from perfection. No thesis could have been finished without a great amount of help from many people.

Firstly, I deeply thank my supervisor Prof. Phil Husbands for his support and considerate supervision, having encouraged me to pursue my research topic as well as helping me to get direct/indirect financial support. Also thanks to Dr. Inman Harvey for being my second supervisor in my early D.Phil. years.

Many thanks to Dr. Andy Philippides and Prof. Rolf Pfeifer for agreeing to be my examiners and finding the time and effort to read and comment on this thesis.

Thanks to the UK government for granting me the Overseas Research Students Awards. Also thanks to the Department of Informatics for the Graduate Teaching Assistantship.

I am also indebted to the members of the Centre for Computational Neuroscience and Robotics for inspiring my study in all aspects. Good memories in the CCNR outpost.

Also thanks to the Sussex Korean society for being my friends and families during my days in Brighton, being my emotional recliners, drinking mates, and many more.

Most of all, I would like to thank my beloved family for supporting me over my D.Phil. course. Thanks to my parents for their unconditional love and support. Thanks to my bro Taesik for always being supportive of me. Finally, love and thanks to my wife Youngrang for all the things she has done for me, feeding me lovely food, cheering me up, and enriching my life in countless other ways.

Chaotic Exploration and Learning of Locomotor Behaviours

Yoonsik Shim

Summary

Recent developments in the embodied approach to understanding the generation of adaptive behaviour, suggests that the design of adaptive neural circuits for rhythmic motor patterns should not be done in isolation from an appreciation, and indeed exploitation, of neural-body-environment interactions. Utilising spontaneous mutual entrainment between neural systems and physical bodies provides a useful passage to the regions of phase space which are naturally structured by the neural-body-environmental interactions. A growing body of work has provided evidence that chaotic dynamics can be useful in allowing embodied systems to spontaneously explore potentially useful motor patterns. However, up until now there has been no general integrated neural system that allows goal-directed, online, realtime exploration and capture of motor patterns without recourse to external monitoring,evaluation or training methods. For the first time, we introduce such a system in the form of a fully dynamic neural system, exploiting intrinsic chaotic dynamics, for the exploration and learning of the possible locomotion patterns of an articulated robot of an arbitrary morphology in an unknown environment. The controller is modelled as a network of neural oscillators which are coupled only through physical embodiment, and goal directed exploration of coordinated motor patterns is achieved by a chaotic search using adaptive bifurcation. The phase space of the indirectly coupled neural-body-environment system contains multiple transient or permanent self-organised dynamics each of which is a candidate for a locomotion behaviour. The adaptive bifurcation enables the system orbit to wander through various phase-coordinated states using its intrinsic chaotic dynamics as a driving force and stabilises the system on to one of the states matching the given goal criteria. In order to improve the sustainability of useful transient patterns, sensory

homeostasis has been introduced which results in an increased diversity of motor outputs, thus achieving multi-scale exploration. A rhythmic pattern discovered by this process is memorised and sustained by changing the wiring between initially disconnected oscillators using an adaptive synchronisation method. The dynamical nature of the weak coupling through physical embodiment allows this adaptive weight learning to be easily integrated, thus forming a continuous exploration-learning system. Our result shows that the novel neuro-robotic system is able to create and learn a number of emergent locomotion behaviours for a wide range of body configurations and physical environment, and can re-adapt after sustaining damage. The implications and analyses of these results for investigating the generality and limitations of the proposed system are discussed.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Genesis

Properly coordinated rhythmic movements for locomotion are ubiquitous in animals. From insects to humans, locomotive ability is one of the basic requirements for survival. Biological locomotor systems (usually involving coordinated limb movements) evolved to be highly adaptable, dexterous and energy efficient. Consequently they are a major source of inspiration when designing robot locomotor systems. Indeed, there has been a major effort in autonomous robotics for finding ways to design neural controllers to produce coordinated rhythmic movements. Most biological locomotor systems involve neural networks acting as central pattern generators (CPGs) which are responsible for producing the basic rhythmic patterns for the oscillatory movement of limbs (Cohen et al., 1988; Stein et al., 1997). Understanding the subtleties of operation of such networks, and how to design artificial versions for robotic applications are ongoing challenges (Ekeberg, 1993; Kimura et al., 1999; Ijspeert, 2001; Ijspeert et al., 2007).

The majority of approaches to design CPG-based robotic locomotor systems have relied on optimisation and search methods, including evolutionary algorithms and other stochastic methods (Gallagher et al., 1996; Ijspeert, 2001; Kamimura et al., 2003; Itoh et al., 2004; Floreano et al., 2008), to find a suitable configuration of system parameters. These methods are usually very computationally expensive and often require a priori knowledge of the robot body and environment. Besides it is often very difficult to devise evaluation methods and metrics that can adequately

cover the enormous number of unexpected situations that a robot can encounter during its lifetime, such as environmental change or body defects. Hence there are many open issues in how to deal with unknown environments and adaptation to arbitrary or changed (e.g. damaged) body conditions.

This naturally led to efforts to develop methods for the on-line generation of motor control for the lifetime maintenance of locomotor functions. Among these, reinforcement learning (RL) (Matsubara et al., 2006; Nakamura et al., 2007; Endo et al., 2008) has been successfully applied to produce locomotion controllers. Heuristic optimisation algorithms such as conjugate direction methods (Sproewitz et al., 2008) have been adopted for faster on-line optimisation of motor behaviours. More systematic approaches such as continuous self-modelling (Bongard et al., 2006) used a number of stochastically optimised internal models to predict the robot's morphological changes using data collected from sensory-actuation causal relationship. Although these are powerful methods which utilise stochastic search algorithms in a more efficient way for on-line adaptation, they are still not free from the inherent difficulties of building-from-scratch and often need to incorporate a priori knowledge or to make use of a biased learning strategy in order to simplify and speed up the learning process.

An important consideration in understanding, and drawing inspiration from, biological motor behaviors, is the appreciation that studying the underlying neural circuitry in isolation ignores the considerable advantage that can be obtained from incorporating the physical body and its environment - i.e. exploiting the embodied nature of such behavior. In studies of animal development including human infant, many evidences support that sensory-motor information from motor activity has important effects on brain development (Rakic, 1988; Thelen and Smith, 1994; Crair, 1999; Johnson, 2005). From this perspective, much attention has been paid to the recent research paradigm called Embodied Artificial Intelligence, which puts a strong emphasis on the dynamical and reciprocal interaction across multiple time scales between and the brain, body, and environment of an agent (Brooks, 1999; Wheeler, 2005; Pfeifer and Bongard, 2007). In robotics this has led to efforts to exploit ready-made functionality provided by the physical properties of an embodied system by concentrating on brain-body-environment interactions.

Meanwhile, the great potential in exploiting intrinsic chaotic dynamics has attracted the attention of neurobiologists interested in how animals learn to coordin-

ate their limbs (Mpitsos et al., 1988; Kelso, 1995; Korn and Faure, 2003). Also a strand of topics in developmental robotics has been studied based on the growing body of observations of intrinsic chaotic dynamics in nervous systems (Rapp et al., 1985; Freeman and Viana Di Prisco, 1986; Terman and Rubin, 2007) to suggest that such dynamics can underpin crucial periods in animal development. Being joined with the concept of embodiment, roboticists strived to develop better and more life-like locomotion systems for articulated autonomous robots. Particularly, the movement patterns of the robots were generated by exploiting the mechanisms of self-organisation where the brain-body-environment dynamics are explored in a spontaneous way as part of the process of acquiring motor skills (Taga et al., 1999; Lungarella and Berthouze, 2002; Pfeifer and Bongard, 2007; Pfeifer et al., 2007). Together with these idea, a line of robotics studies have demonstrated that chaotic neural networks can indeed power the self-exploration of brain-body-environment dynamics in an embodied system, discovering stable patterns that can be incorporated into motor behaviors (Kuniyoshi and Suzuki, 2004; Pitti et al., 2005; Kuniyoshi and Sangawa, 2006; Pitti et al., 2010; Steingrube et al., 2010).

However, to date it has not been clear how to harness chaos in embodied robotic systems in an integrative way such that desired adaptive sensorimotor behaviors can be explored, captured and learnt. In this work we build on the essential concepts of the prior work outlined above, extending and generalising it as we attempt to develop a generally applicable methodology based around self-organisation through chaotic dynamics for neural-body-environment coupled systems. We present a prototypic study on the goal-directed on-line exploration of rhythmic motor patterns in an oscillator system coupled through physical embodiment, specifically generating forward locomotion behaviours without prior knowledge of the body morphology or its physical environment. In an important departure from the previous work outlined above, we employ the concept of Chaotic Mode Transition with External Feedback (Davis, 1990, 1998), which exploits the intrinsic chaoticity of a system orbit as a perturbation force to explore multiple synchronised states of the system, and stabilises the orbit by decreasing its chaoticity according to a feedback signal that evaluates the behaviour. This enabled the system to perform a 'deterministic search' guided by a global feedback signal from the physical system, which facilitates an active exploration toward a desired behaviour. Based on the initial idea, we further enhance our system by providing a coherent integration of a series of dynamic learning pro-

cesses, which attempts to build a realtime self-driven exploration-capture-learning system.

1.2 Thesis Organisation

The remainder of the thesis is structured as follows. Chapter 2 describes general background topics necessary for carrying out later simulations and experiments. The first section starts with the organisation of biological locomotor systems, putting emphasis on rhythmic pattern generation by central pattern generators. Next, there is a very brief description of optimisation based methods which have been widely used for designing robot locomotion controllers. Following that, two branches of previous research in which this work is deeply-rooted are introduced. The first one is about the spontaneous emergence of motor patterns from coupling through physical embodiment and the second work describes the method of chaotic search through adaptive bifurcation.

Chapter 3 describes the models and methods for the proposed exploration system. The general architecture of the exploration system for locomotor behaviours is outlined, followed by the brief description of each sub-component. The subsequent sections describe each component in more detail which include the CPG model, evaluation strategy, and neuromuscular system. Dynamical properties of both single and coupled CPG models are examined in terms of the applicability to the limbed robot movement, followed by how these CPGs are encapsulated as a unit module for a single joint. Then there is an explanation of how the CPGs are coupled indirectly through physical embodiment by local sensory information. After that, the robotic simulation model is addressed by describing the basic neuromuscular model which is used as the common component for building arbitrary bodies, as well as the simulation models of physical environments. Finally, the analysis method for the system behaviours are presented.

In chapter 4, the basic behaviour of the proposed exploration system is investigated. The analyses of the system behaviours are presented using a simple 4-fin aquatic swimmer placed in a simulated 2D hydrodynamic environment. Firstly, the system behaviours in its chaotic regime is investigated with different global sensor gains which determines the strength of functional coupling between neural elements

for the system to exhibit a well-structured dynamics from stable to chaotic patterns. To draw insights into the chaotic behaviour of the system, a Lyapunov analysis is performed for different sensor gains. Secondly, the stable movement behaviours of the robot which exist in the stable regime of the system are identified and categorised, and the result of the chaotic exploration for those patterns is presented together with the statistics of the appearance of each pattern. Also a report on the unwanted system deficiencies arising from the deterministic nature of the system are addressed.

Chapter 5 improves the basic framework of the system in the previous chapter by extending the system using two fully dynamic processes which are smoothly integrated with the basic system. Firstly, a biologically inspired process for calibrating incoming sensor signals is introduced, for which the neural system receives refined afferent input to deal with an unknown variety of sensor signals from an arbitrary robotic system, and to maintain an appropriate embodimental coupling between neural elements in realtime and continuous manner. Secondly the system is again extended by employing a fully dynamic learning method for CPGs in order to automatically wire initially disconnected neural elements, which is guided by the evaluation signal, so that the system can dynamically capture and memorise a locomotor pattern discovered by exploration process. In order to show the generality of the fully integrated system, it is tested using a 3D quadruped robot, which is modelled as a representative of terrestrial walking robots. A brief discussion about the comparative effect of the system on the two different robotic systems is presented. Finally, in the last chapter we summarise the result and discuss the generality and limitations of the system, as well as its biological relevance, all of which naturally lead to a series of future directions.

1.3 Contributions

A summary of the main achievements of this thesis are provided as follows:

- We significantly generalise and extend the two seemingly distinct previous concepts of the pattern emergence from embodiment and chaotic dynamics, and introduce a prototype of an integrated framework for robot locomotion, which is constructed as a single continuous dynamical system.

- We demonstrate that the presented system can exhibit reversible transitions between the stable regime where multiple stable states coexist and the unstable regime where the system orbit wanders quasiperiodically or chaotically by varying a single system parameter, and attempt to provide a quantitative analysis of system dynamics and its complexity over a range of dynamical regimes.

- We model a continuous self-evaluation strategy for robotic system and harness it to the bifurcation parameter of the system, achieving the selective entrainment of the system dynamics to one of the patterns by imposing goal directedness toward a higher performing behaviours.

- We provide homeostatic sensory adaptation for the system to generate flexible yet correlated activities to maintain a certain level of information exchange. This enhances the generality of control system to deal with an arbitrary robotic system and an arbitrary type of sensors, and also results in an increased diversity of limb motions by achieving multi-scale exploration.

- The system is further improved by incorporating an oscillator learning algorithm to capture and sustain useful transient patterns by rewiring the CPGs dynamically using an adaptive synchronisation process.

- The system is successfully demonstrated on a number of simulated robots, showing its generality and effectiveness in exploring, capturing and learning locomotion patterns and adapting in real time to damage and other radical changes.

1.4 Relation to Previous Publications

Parts of this thesis are based on previously published work. Particularly, Chapter 4 is based on Shim and Husbands (2010), with deeper analyses of underlying system dynamics as well as a few modifications on the neuromuscular model which is commonly used for robotic system. Due to the sensitivity of this work to the changes in models, a large part of the results in Chapter 4 has been newly generated. The description of the robotic simulation model and its environmental forces in Chapter 3

has been extracted and re-arranged from a number of sources (Shim and Kim, 2003; Shim et al., 2004b,a; Shim and Kim, 2006; Shim and Husbands, 2007b,a). Chapter 5 is based on Shim and Husbands (2012) with a few expansions of the results to emphasise the generality of this work.

Chapter 2

General Background

This chapter provides a general background on related work and theoretical concepts relevant to this work. A brief review is provided for each subject; the intention is not to give an exhaustive survey.

2.1 Vertebrate Locomotor System and CPGs

The vertebrate locomotor system in general can be functionally categorised into four components, which are for selection, initiation, maintenance, and execution of the motor task (Grillner et al., 2000; Grillner, 2003). Selecting a motor program is performed in the region in forebrain called the *basal ganglia*, involving multimodal sensory integration. The information of the selected motor task is communicated to the brainstem, which is responsible for initiating and maintaining locomotion. The brainstem system is organised serially from the *mesencephalic locomotor region* (MLR) and *diencephalic locomotor region* (DLR) to *reticulospinal neurons* in the hindbrain, which in turn project to locomotor neurons in the spinal cord (Figure 2.1). The basal ganglia determines when the motor program should be active. Non-selected motor programs in the brainstem are subject to a powerful tonic inhibition, such that locomotion is initiated only when the MLR is disinhibited (Grillner et al., 2005).

Locomotion is executed by the specialised neural circuits within the spinal cord called the *Central Pattern Generators* (CPG), which are responsible for generating rhythmic motor output. These circuits are located at the spinal level, and they

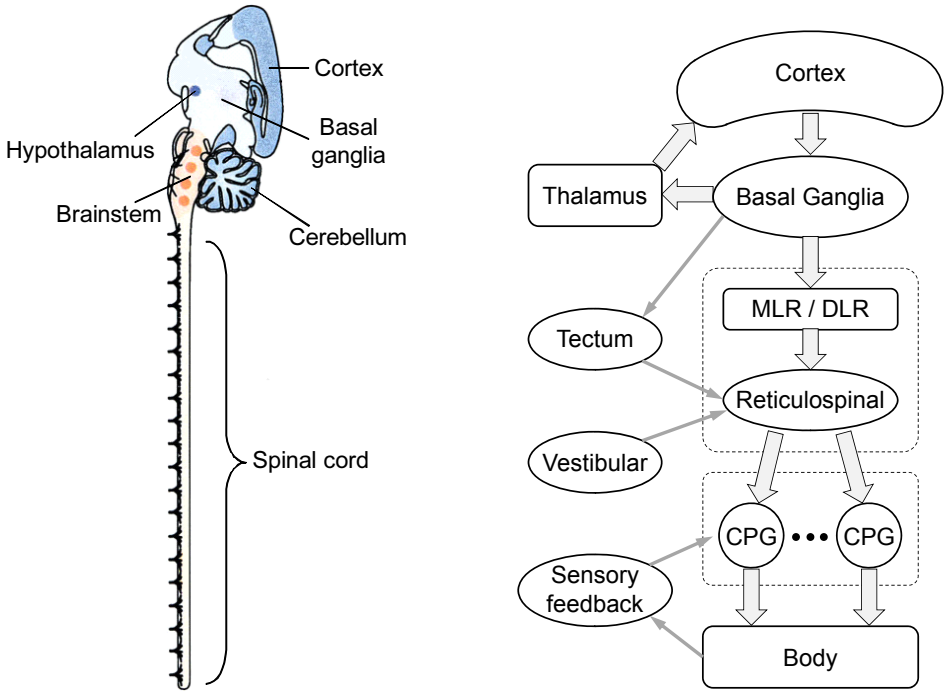

Figure 2.1: Organisation of the locomotor system in vertebrates. Basal ganglia (BG) receives inputs from the cortex and the thalamus, which performs the selection of a motor task. BG inhibits command centres in the DLR and the MLR during resting conditions. The visual, sensory, and vestibular information are integrated in the brainstem to control both steering and posture. The spinal CPGs are activated via reticulospinal neurons, and they are modulated by local sensory feedback. Images were arranged and reproduced from (Grillner et al., 2008; Goulding, 2009).

coordinate the basic propulsive movement synergy such as swimming, walking, and flying by sequentially activating the different motoneuron-muscle groups (Grillner and Wallen, 1985; Grillner, 2003; Rossignol et al., 2006). The spinal CPG is normally activated by brainstem reticulospinal cells that provide excitation along the spinal cord to the different spinal neurons and the motoneurons. The level of CPG activity is determined by the brainstem locomotor command (Orlovsky et al., 1999). Different levels of MLR stimulation can modulate the speed of locomotion by inducing the transition between completely different gait patterns. The walking, trotting or galloping in tetrapods such as cats arise from the different stimulation level (Shik

et al., 1966). Stimulation of the same region in a bird gives rise to walking and at higher strengths flapping movements of the wings, and finally in a fish or a lamprey swimming is initiated at progressively higher speeds (Stein et al., 1997). Although receiving only simple input signals, CPGs are capable of generating complex locomotor behaviors involving the switch between radically different gait patterns.

While the various descending inputs from the supraspinal level play a crucial role in initiating and shaping the rhythmic pattern of CPG, it can produce coordinated patterns of rhythmic activity without any rhythmic inputs from sensory feedback or from higher control centers. This has been demonstrated by decerebrating or spinalising an animal using a neuromuscular blockade or deafferentation of sensory feedback. When activated by a simple electrical or pharmaceutical stimulation, several vertebrates are able to produce patterns of activity (which is called fictive locomotion) after the isolation of spinal cord, which are very similar to normal locomotion (Grillner and Wallen, 1985; Cohen et al., 1988; Stein et al., 1997). Therefore it is evident that the CPG is the source of rhythmic activity. However, sensory feedback plays a very important role both in shaping and coordinating the rhythmic patterns of CPG. Many studies have shown that a decerebrated cat is able to restore locomotor function under treadmill training (Shik et al., 1966; Whelan, 1996; Rossignol, 2000). It is even possible to recover some locomotor activity in human patients suffering from a spinal cord injury through an intense training on a treadmill (Barbeau and Rossignol, 1994; Dietz et al., 1995).

The notable locomotor ability of decerebrated animals under the presence of local sensory feedback suggests that the CPG function is tightly coupled to the fast local sensory information to form a basic low-level circuit for rhythmic activity, which can be flexibly modulated by a few gradual supraspinal controls. This distributed nature of the vertebrate locomotor system provides advantages for designing machine locomotor systems; less time delays in coordinated rhythmic control and the reduced dimensionality of the descending control signals (Ijspeert, 2008).

It is widely accepted that the spinal motor circuitry is genetically hardwired. However, some biological evidences suggest that activity-dependent events can shape the locomotor network during development. A group of neural circuits and the motor neurons in the spinal cord are spontaneously active during embryonic development (Jean-Xavier et al., 2007). In rat experiments, the spontaneous motoneuronal activity can be mediated by the inhibitory neurotransmitter pathways in embryonic

neurons (Nishimura et al., 1996). The activity-dependent patterns are necessary for motor axon guidance (Hanson and Landmesser, 2004) or for shaping the topology of cutaneous sensory inputs (Schouenberg, 2004). Also, activity-dependent events can influence the modulation of neurotransmitter expression by the homeostatic mechanisms that maintain network excitability in the spinal cord at appropriate levels during development (Goulding, 2004). All these findings suggest that the spinal motor circuitry is not strictly hard-wired but can undergo activity-dependant changes during its lifetime.

2.2 Designing Robot Locomotor System

2.2.1 Mathematical Models of CPGs

Almost a century ago, a basic structure for the CPG network was first proposed by Brown (1914), which is called the half-centre model. It is comprised of two coupled neural populations (half-centres) with mutual inhibition, which produce the alternating rhythmic activity induced by a fatigue mechanism. The two half-centres can represent the flexors and extensors for a single limb, or any opposing motor pairs. This central concept became the core of modern CPG models, and a level of neurobiological details have been filled in for specific vertebrate systems of interest. To date, there are a number of CPG models in use from simple phase equations to detailed biophysical models. An excellent review about the various aspects of CPGs for designing robotic locomotion is available at Ijspeert (2008).

Among a variety of CPG models for designing robot locomotor system, the most detailed biophysical models of CPG are constructed using the Hodgkin-Huxley (HH) type neurons (Hodgkin and Huxley, 1952), a model developed by Alan Lloyd Hodgkin and Andrew Huxley in 1952 to explain the ionic mechanisms underlying the initiation and propagation of action potentials in the squid giant axon. The model is governed by the four dimensional differential equations based on the idea that the electrical properties of a segment of nerve membrane can be modelled by an equivalent electric circuit, which expresses the flows of current across the membrane associated with charging the membrane capacitance and the movement of specific types of ions. The ionic current is subdivided into three components, which are sodium current, a potassium current, and a small leakage current primarily carried

by chloride ions. Although HH model provides biologically valid neuronal dynamics, there are several parameters to be estimated for an accurate behaviour and is computationally intensive to simulate a large network of neurons.

One of the models that captures the core aspects of the dynamics with less equations is the FitzHugh-Nagumo (FN) model (Fitzhugh, 1961; Nagumo et al., 1962). It is also called Bonhoeffer-van der Pol (BVP, by Richard FitzHugh himself) oscillators, because it is reduced from the van der Pol equation for self-sustained oscillation whose behavior resembles that of the iron nitric acid model proposed by (Bonhoeffer, 1948). The FN model is a two-dimensional simplification of the Hodgkin-Huxley model, which can reproduce the behavior of pacemaker neurons qualitatively well. The model has been widely studied as models of pacemaking cells and interlimb coordination (Collins and Richmond, 1994; Golubitsky et al., 1999; Asai et al., 2000, 2003a,b; Sproewitz and Berthouze, 2005; Ohgane et al., 2009).

There are other kinds of approximated versions of the HH model which combine multiple models for compromising between computational complexity and biological plausibility, or being specialised for specific interests. The Morris-Lecar model (Morris and Lecar, 1981) combines HH and FN models into a voltage-gated calcium channel model with a delayed-rectifier potassium channel by using two dimensional differential equations. It is used to design quadruped (Buono, 2001) and biped (Pinto and Santos, 2011) robots. The Hindmarsh-Rose model (Hindmarsh and Rose, 1981) is based on the FN model but uses three coupled equations to allow a wider variety of dynamic behaviors including chaotic dynamics. For a comprehensive survey of more popular CPG models especially for designing robot locomotor system, see (Ijspeert, 2008).

2.2.2 Popular Design Methods for Robot Locomotion

2.2.2.1 Evolutionary robotics

The field of evolutionary robotics applies the techniques of evolutionary algorithms (EAs) to the design of robot controllers (Husbands and Harvey, 1992; Cliff et al., 1993; Husbands and Meyer, 1998; Nolfi and Floreano, 2001). Evolutionary algorithms are a group of techniques for automatic design which is inspired by natural evolutionary processes. They have been applied in a variety of domains including

combinatorial optimisation, adaptive control, and machine learning.

The algorithm operates on a population of individuals, each of which is expressed as a sequence of data that encodes the properties of the individual. This encoded data is called genotype (the set of genes that an organism carries). The EA procedure starts by generating initial population of random or seeded genotypes. The individuals are generated from genotype data and their fitness is evaluated by predesigned metric (fitness function). A new population is created by performing genetic operation such as crossover and mutation in such a way that individuals with higher fitness are more likely to pass genotype data to the next generation. This procedure is repeated until the fitness of the best individual (or of entire population) meets specified criteria. Designing a fitness function is crucial for the success of EA, which is often a hard problem.

Robots using EA have normally evolved the parameters of neural controller (Gallagher et al., 1996; Beer, 2003; Shim and Husbands, 2007b,a; Floreano et al., 2008), popularly using a connectionist model of neural system consists of continuous-time recurrent neural networks (Beer, 1995b). Also the robots have been generated by evolving both its controller and body simultaneously (Cliff et al., 1993; Sims, 1994; Shim and Kim, 2006; Bongard, 2010, 2011). In this case, more than merely a tool for optimisation, EA provides an interesting possibility that the extent of possible robot architectures produced by search can show the architecture beyond the human imagination, which can even possibly provide some insight back to biology.

Although EAs have been successfully applied to various robotic designs, there are several challenges to be addressed. Similar to other combinatoric optimisation methods, EA demands intensive computation power and is time consuming, and often require a priori knowledge of specific robotic system. Also they suffer from the difficulty of designing fitness functions that can cover the enormous number of unexpected situations that a robot can encounter during its lifetime, such as environmental change or body defects. Since EA normally uses computer simulations to deal with numerous individuals, an accurate simulation of physical system (which is often very difficult) is crucial to ensure a satisfactory level of reliability when the generated controller is applied to a real robot. This gap between the simulation and the real system raises the problem of 'transfer into reality', that the gap will widen as the complexity of robot system grows (Mataric and Cliff, 1996).

2.2.2.2 Reinforcement Learning

The reinforcement learning (RL) method (Barto et al., 1990; Sutton and Barto, 1998) is a general machine learning paradigm whereby an agent learns how to act to maximise the reward given as a consequence of its action. RL is suitable to solve sequential decision tasks such as Markov Decision Processes (MDPs) through trial and error interactions with the environment. The term reinforcement defines the desirability of a state of agent which is expressed in terms of rewards and punishments. RL is very general and highly extensible to a variety of problems, and it can be operated in on-line situations.

The RL agent interacts with its environment and receives a scalar reward signal for each action taken, and its goal is to learn the 'action policy' so as to maximise the cumulative reward it receives over time. The general RL process is modelled using a set of states (S), actions (A), and rewards. At each time t, an agent perceives its state $s_t \in S$ and the set of currently available actions $A(s_t)$. Then the agent selects one of possible actions $a_t \in A(s_t)$ and receives a new state s_{t+1} and a scalar reward r_{t+1}. The goal of an agent is to maximise the total reward over time ($R = \sum r_t$), by developing optimal action selection policy $\pi : S \rightarrow A$. The stochasticity intervines in the selection of action according to the learnt policy.

Concerning robotics, although RL was initially designed for discrete-step processes, a continuous version of the algorithm has been introduced and successfully applied to the learning of optimal control policies for simple mechanical systems such as pendulum and cart-pole swing-up tasks (Doya, 1996, 2000). For CPG based controllers, a line of studies has developed a hybrid RL architecture called the CPG-actor-critic model to control a CPG driven biped robot (Matsubara et al., 2006; Nakamura et al., 2007; Endo et al., 2008). The model basically uses hard-wired CPGs controlled by the descending commands from the higher center which is represented as a simple feed-forward neural network, and its parameter is subject to learning.

The major drawback of RL is the large number of runtime trials required, and it often requires a careful representation of states and actions. The combinatorial explosion problem is also an issue when the complexity of search space of possible actions increases.

2.3 Intrinsic Chaos in Neural Systems

A key influence on the current work is the growing body of observations of intrinsic chaotic dynamics in nervous systems (Guevara et al., 1983; Babloyantz and Salazar, 1985; Rapp et al., 1985; Freeman and Viana Di Prisco, 1986; Wright and Liley, 1996; Terman and Rubin, 2007). Freeman and colleagues were among the first to provide empirical evidence of chaotic brain dynamics in their analysis of multi-channel EEG recordings from the rabbit olfactory bulb (Freeman and Viana Di Prisco, 1986; Skarda and Freeman, 1987), showing that background neural activity in the area appears to be chaotic. These observations were backed up by mathematical models of the bulb that were able to replicate the electroencephalography (EEG) measured from living brain. Various authors have observed chaotic dynamics (low dimensional chaotic attractors) in human EEGs in various stages of sleep and wakefulness (e.g. (Babloyantz and Salazar, 1985)), in epileptic seizures and other disorders (Guevara et al., 1983; Babloyantz and Destexhe, 1986), as well as in other animal systems (Aihara, 2003). These observations of chaotic dynamics in both normal and pathological brain states, and at both global and microscopic scales (Wright and Liley, 1996), support the idea that chaotic dynamics, rather than being pathological, often plays a fundamental role in neural mechanisms (Freeman et al., 2001).

In addition to the examples cited above, many of which refer to perceptual states, there are also a number of studies indicating intrinsic chaotic dynamics in animal motor behaviours, both at the neural level (Rapp et al., 1985; Venaille et al., 2005; Terman and Rubin, 2007) and at the level of body and limb movement (Cordier et al., 1996; Riley and Turvey, 2002). These seem particularly prevalent during developmental and learning phases (e.g. when learning to coordinate limbs) (Ohgi et al., 2008).

Although the functional roles of chaotic dynamics in the nervous system are far from understood, a number of intriguing proposals have been put forward. Freeman and colleagues have hypothesised that chaotic background states in the rabbit olfactory system provides the system with "continued open-endedness and readiness to respond to completely novel as well as familiar input, without the requirements for an exhaustive memory search" (Skarda and Freeman, 1987). They postulate that such mechanisms underlie much of perception and play an important role in cognition in general (Freeman and Skarda, 1960; Freeman et al., 2001). A number

of neurobiological studies (Rakic, 1988; Thelen and Smith, 1994; Crair, 1999; Johnson, 2005) make the important suggestion that chaotic dynamics underpin crucial periods in the development of animal brains.

A growing number of computational models have demonstrated that the above kinds of functions do arise in chaotic neural networks. Many studies have shown interesting learning and self-organisation dynamics are possible in such networks, with the kind of open-ended properties postulated by Freeman and colleagues (Aihara et al., 1990; Hansel and Sompolinsky, 1992; Adachi and Aihara, 1997; Aihara, 2003; Huang and Cao, 2006; Andras and Lycett, 2007). Robotics studies also have demonstrated that chaotic networks can power the self-exploration of brain-body-environment dynamics in an embodied system, discovering spontaneously emerged patterns that can be incorporated into motor behaviours (Kuniyoshi and Suzuki, 2004; Pitti et al., 2005; Kinjo et al., 2008; Pitti et al., 2010).

2.4 Embodiment and Locomotion

2.4.1 Minimal Control by Exploiting Physical Embodiment

As has been increasingly pointed out over the past few years (Pfeifer and Iida, 2004; Pfeifer and Bongard, 2007; Pfeifer et al., 2007), studying neural circuitry underlying the generation of rhythmic motor behaviour in isolation ignores the considerable advantage that can be obtained from incorporating the physical body and its environment - i.e. exploiting the *embodied* nature of such behaviour. From a macroscopic perspective, even high-level cognition is inseparable from body-environment interactions (Wheeler, 2005) and any attempt to dissociate such interactions will destroy their natural function afforded by the information structure innately provided by physical embodiment (Kuniyoshi et al., 2007).

In robotics this has led to efforts to exploit ready-made functionality provided by the given physical properties of an embodied system for the automatic generation of motor movement. One such line of enquiry involves using a frequency adaptive oscillator whose frequency can be tuned to the resonant frequency of the mechanical system (Buchli et al., 2006; Righetti et al., 2006), and a systematic analysis of the importance of frequency scaling and its existence in the chaotic regime of a neuromechanical model have been presented in this context (Raftery et al., 2008).

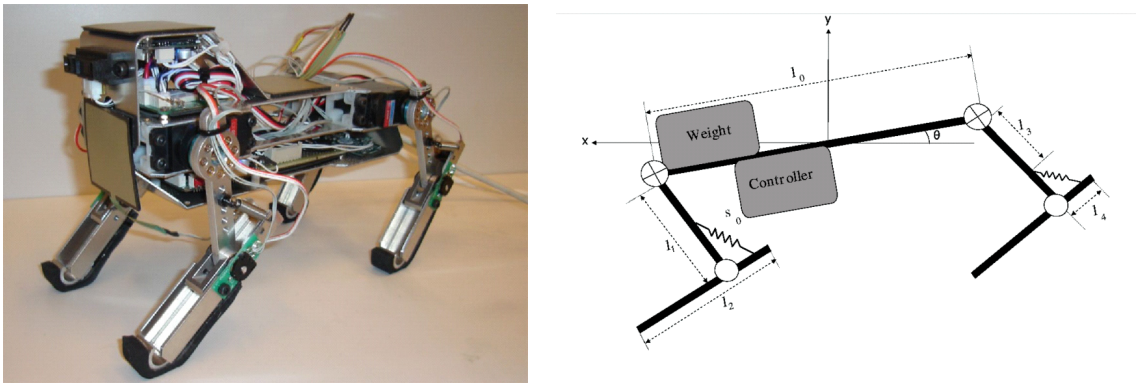

Figure 2.2: The photograph and schematic diagram of the quadruped robot "Puppy". Only four of eight joints are actuated, which are shown as the circles with a cross inside (right image). Images were taken from (Iida et al., 2005).

Pfeifer and his colleagues take a more active approach by using the concept 'morphological computation', where a considerable portion of the control or computation can be taken over by the dynamic interaction from morphological properties which include passive leg swing in walking, muscle compliance, and weight distribution. A realisation of this idea had been shown using a simple dog-like robot (Figure 2.2) (Iida et al., 2005), which could run robustly only with a purely feedforward controller using sinusoidal functions, although the parameters for the controller were found heuristically and set by hand.

2.4.2 Pattern emergence from Coupling through Physical Embodiment

One of the seminal works from the previously mentioned perspective is the exploration and acquisition of motor primitives, for a simple robot, using a mechanism which is embodied as a coupled chaotic field (Kuniyoshi and Suzuki, 2004; Pitti et al., 2005). Those groups modelled an extreme version of embodied coupling that had no electrical connection between neural units, with all neural coupling acting indirectly through body-environment dynamics. Therefore the neural circuit is 'embodimentally coupled' through the physical system as the agent interacts with the environment (we will use the term 'embodimental coupling' with this meaning

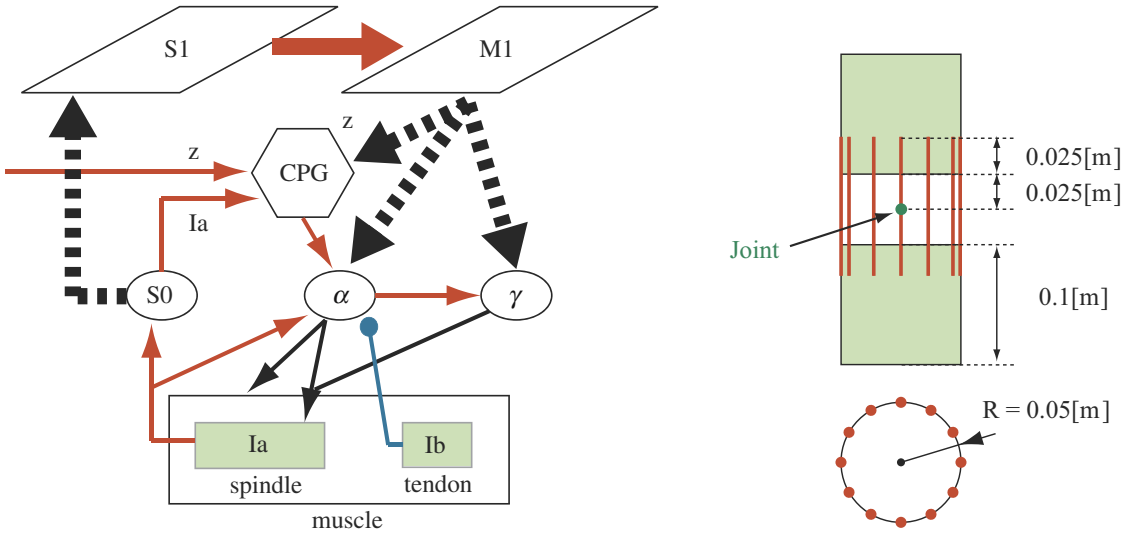

Figure 2.3: The schematic diagram of cortico-spinal-musculo-skeletal system (left) for the 1-joint link model (right) proposed by Kuniyoshi and Sangawa (2006). The model consists of, sensor map (S1), motor map (M1), α motor neuron (α), γ motor neuron (γ), and afferent interneuron (S0). Arrow and filled circles represent excitatory and inhibitory connections respectively, and thick broken lines represent all to all connections with plasticity. The physical body consists of two cylindrical rigid links connected with a ball-socket joint, actuated by 12 muscle fibers (red lines). Images were taken from (Kuniyoshi and Sangawa, 2006)

throughout the text). For instance, take the example of an insect walking, where the local neural circuits at the legs are coupled through the embodiment of the insect body and the interaction with the environment, as it pushes back with a leg. The neural element was implemented using a simple logistic map showing chaotic dynamics, such that the output of each element drives the muscle activation level and the sensor value from each actuator site is fed to the corresponding neural element. The system dynamics rapidly developed to a stable, coherent rhythmic motion by using mutual entrainment between the neural circuit and the body-environment interactions. The process was completely deterministic, not making use of any random search method.

More tractable systems (Pitti et al., 2009, 2010) have shown that a simple 2D

simulated biped controlled by indirectly coupled chaotic maps can generate stable locomotion when the coupling strength between controller and body was set in the specific regime of phase synchronisation. Phase synchronisation between chaotic controller and physical system allows the flexible self-assembly of motor patterns and adaptive frequency matching to the resonant frequency of the body. However, the motor patterns that emerge through phase synchronisation do not necessarily produce sustained locomotion behaviours unless the coupling strengths are properly set for a given neuromechanical system.

Another branch of work (Kuniyoshi and Sangawa, 2006) dealt with a more biologically plausible system in which a realistic musculo-skeletal model was employed and the neural control circuit consisted of a model CPG (Figure 2.3). This was embedded within a larger system involving cortical maps with interactions between the various components. The biomechanical system was modelled as a series of redundant muscles acting on a joint, and the information of the muscle combinations for any discovered coherent motor patterns were engraved on the model cortices as a sensorimotor representation. Later work (Kinjo et al., 2008) demonstrated the learning and replay of a motor pattern by adding a simple perceptron with a back-propagation learning on top of the previously learnt sensorimotor maps. The learning pattern was manually fed to the system during learning, and a simple 3-layer perceptron could acquire the movement by learning the relation between sensor and motor cortices. They showed that the representative power of the self-organised sensorimotor maps can greatly simplify the nontrivial sensorimotor learning problem into a simple mapping between the sensor and motor maps.

While these studies have developed detailed biological models that have significant implications for the understanding of motor development, concrete general methodologies for applying such techniques to the automatic generation of desired motor patterns for autonomous robots remains a challenge.

2.5 Chaos as a Source of Perturbation

2.5.1 Chaoticity as an External Perturbation

Conventional optimisation strategies generally use (external) stochastic perturbations on system parameters for search space exploration. However, a few studies

address the effectiveness of chaotic dynamics as behaving like a stochastic source (Parker and Chua, 1989; Ott et al., 1994). A number of studies have found that a deterministic chaos generator (usually using logistic map) outperforms a stochastic random explorer both in evolutionary algorithms (Determan and Foster, 1999; Zhang and Shao, 2001; Caponetto et al., 2003) and reinforcement learning (Morihiro et al., 2005, 2008). In these cases, the chaotic dynamics acts as an external module generating perturbations that cause system parameters to wander in parameter space.

2.5.2 Chaotic Search Using Adaptive Bifurcation

The adaptive chaotic search method presented here, using bifurcation to chaos, can directly drive the phase orbit of a embodied system for exploration because of the endogenous existence of chaotic dynamics in the system itself. The intrinsic chaotic dynamics of the system naturally power the search process without the need for external sources of noise. The general idea of applying a chaotic search method using adaptive parametric feedback control had been previously presented in the field of optical sciences (Aida and Davis, 1994; Davis, 1990, 1998; Liu and Davis, 1998) and for memory search (Nara and Davis, 1992). The idea couples the fitness response to the multimode system in such a way that bad response result in mode transitions, whereas good response suppress them.

It has been argued that this method should be generally applicable when the target device is capable of supporting a variety of stable modes, between which there exist chaotic transitions, which interacts with its environment and gives a feedback signal evaluating whether the mode is suitable or not. Chaotic transitions allow the system to try each of the modes sequentially, and the mode which is evaluated as suitable is selected and stabilised by changing a device parameter to take it into a multistable regime. The whole process can be thought as a controlled version of the concept of chaotic itinerancy (Kaneko, 2003), where the system wanders from one quasi-attractor to another, getting entrained in each of them for a while.

An indirectly coupled neural-body-environmental system, such as the one used in this paper, has the required characteristics of such a device, including multiple coordinated oscillation modes. It is known that a properly designed oscillator network can have multiple synchronised states which exhibit stable oscillations for both discrete (Feudel et al., 1996; Feudel and Grebogi, 1997) and continuous (Vadivasova

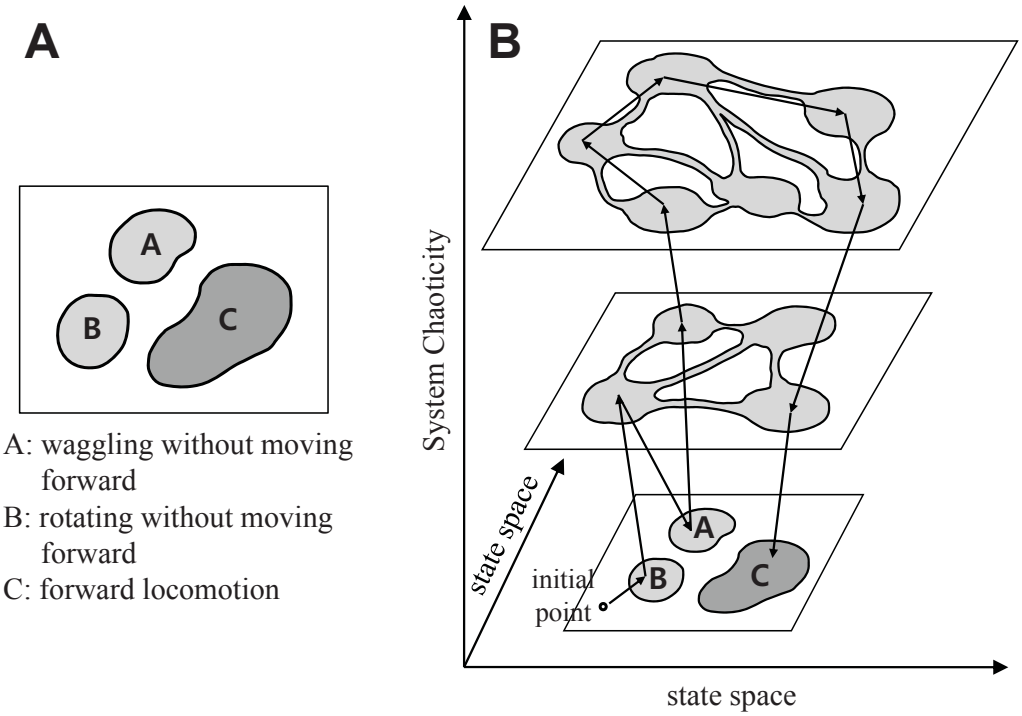

Figure 2.4: A conceptual illustration of chaotic exploration. (A) The state space of a neural-body-environment system coupled through physical embodiment, which consists of three basins of attraction (A,B,C) with different performances. (B) An exploration process to find the desired attractor, C, by varying the complexity of the state space landscape. Lump spaces and narrow passages in the landscapes of higher complexities represent quasi-attractors and itinerant pathways respectively.

et al., 1999; Olusola et al., 2010) systems, and the structure of emergent behaviour in these systems often reflects the spatial distribution of coupling strengths (Kaneko, 1994). Accordingly, a network of oscillators coupled through physical embodiment forms multiple synchronised states that reflect the body schema and its interactions with the environment, and each of them represents a potential candidate for 'meaningful' locomotion behaviour.

A conceptual description of the chaotic search process is briefly illustrated in Figure 2.4. The goal of the system can be regarded as finding and becoming entrained in the basin of a particular attractor which has high performance (denoted by C) while escaping from the low performing attractors (A and B) regardless of

the initial point in the state space. The idea is to 'open' a new pathway which connects those isolated basins through the use of an additional dimension afforded by changing the system dynamics through tuning the chaoticity according to the evaluation signal. The orbit will visit and evaluate each of the attractors (A,B,C) systematically yet chaotically by adaptively varying the bifurcation parameter of the system according to a feedback signal until it reaches the basin of the desired attractor. The method has some conceptual similarities with Ashby's idea of the ultrastable system (Ashby, 1952), although, not surprisingly given how long ago he conceived it, Ashby only envisaged simple stochastic perturbations. The process can be interpreted as a continuous and deterministic version of trial-and-error search which exploits the intrinsic chaotic behaviour of the system.

2.5.3 Chaotic Search of Embodied Rhythmic Movement

Inspired by the existence of the multistable solutions of rhythmic movements which spontaneously emerge in embodimentally coupled neuro-physical systems, our prior work (Shim and Husbands, 2010) has demonstrated how chaotic search can be applied to a neuro-robotic system for the goal-directed exploration of emergent movement patterns. Although the general purpose of our system is the same as that of the conventional parameter optimisation of robot controllers, there are radical differences between them. While parameter optimisation searches the parameters of the target system (e.g. connection weights of a neural network), chaotic search directly deals with the phase orbit of the dynamical system without altering the system parameters. It should be distinguished from chaotic optimisation, a subcategory of parameter optimisation methods which uses a chaotic number generator (ex. logistic map) to power the search space exploration. The system for chaotic search should exhibit multistability to be eligible for the method.

From the perspective of explorability of possible solutions, the parameter optimisation methods would generally be superior to chaotic search when the target system complexity is low or moderate. In the case of a highly complex embodied dynamical system, as used in this work, where there are many stable states in the phase space, the solution diversity would be equal or even higher than that of parameter optimisation. If the systems for both methods have rich enough sets of solutions, chaotic search has the advantage in that it does not need multiple trials and resets

and can be run continuously in realtime. Also the support of internal simulations to deal with the real world as in Bongard et al. (2006) is unnecessary. Assuming the final solution is a good stable rhythmic movement which is expressed in the globally stable attractor in state space, parameter optimisation changes (morphs) the landscape of the state space by optimising the parameters of the dynamical system (i.e. neural weights) to find the best state space landscape that satisfies the evaluator. The chaotic search method deals with a state space containing multiple attractors and this multistability stems from the embodimental coupling (no connection between neural elements) whose effective coupling strengths are time varying and nonlinear. The embodiment can be thought of as a coupling field for neural elements providing inherent flexibility against changes of the physical system.

From the viewpoint of developmental biology, it has been suggested that early motor development is influenced by the emergent patterns from the neuro-physical interaction, which is crucial for the organisation of neural circuitry (Granmo et al., 2008; Blankenship and Feller, 2010). This implies that the evolution of biological locomotion is deeply related to the body and the surrounding environment in which the specimen lives. Such movement patterns naturally reflect the information structure formed by the physical embodiment which is advantageous in terms of energy efficiency. In the case of generating biologically plausible locomotion behaviour of a robot with a given physical body and environment using a stochastic optimisation method, it is likely to be expected that the parameter set which represents the motor patterns similar to those that emerge from the embodimentally coupled system will lie in the neighbourhood around the highest peaks in the search space.

Previous works of the Kuniyoshi group on emergent movement from embodiment proposed a novel framework that serves as a base model for constructing a robotic system which can fully exploit information structures inherently formed by the body-environment interaction. Although this study builds on the previous idea, it differs from the previous works in a number of important aspects. Firstly, this work develops a *goal-directed* system which performs an active search of desired locomotion behaviour by adopting a chaotic search method using adaptive bifurcation, particularly for stable forward locomotion behaviours. The previous work did not focused on the directed search for long lasting forward locomotion. Rather, they focus on the embodied interaction between chaotic units (e.g. logistic maps), showing the possibility of emergent ordered patterns (including an example of locomotion

behaviour) with specifically pre-specified controller parameters which determine the controller chaoticity and feedback coupling strength (Kuniyoshi and Suzuki, 2004). Another work also dealt with transient systems by setting the neural controller to behave chaotically, promoting the system to explore possible temporal coordinations of multiple muscles acting on a joint, which are then stored in cortical maps as sensorimotor representation (but did not focus on the active search of interlimb coordination for locomotion) (Kuniyoshi and Sangawa, 2006). Although more recent works (Yamada et al., 2011a,b) were devoted to generating quadruped locomotion using non-chaotic controllers, the emergent movement of the robot spontaneously changed among forward/backward walking and jumping, and none of them dealt with the active search driven by external feedback to foster better performance.

We also provide a novel homeostatic sensory adaptation which can adaptively regulate incoming sensor signals. This enhances the generality of the control system enabling it to deal with arbitrary sensor types by ensuring the control units maintain a certain level of information exchange ensuring the system always generates flexible yet correlated activities even if the robotic system has a range of different types of sensors which give a variety of waveforms. The homeostatic process controls the slow variables for the amplitudes and offsets of incoming sensor signals, which are diversified in the chaotic phase of the search process. As a result, limb motions with different waveforms can give similar sensor signals to neural elements, achieving multi-scale exploration by the fast and slow dynamics of the system.

Finally, the system is further improved and differentiated from previous work by hybridising it with an oscillator learning algorithm to capture and sustain high performing transient patterns which are discovered during the search process by dynamically rewiring the neural elements using an adaptive synchronisation process. The learning process is integrated with the chaotic search into a single continuous dynamical system, which achieves a realisation of a fully dynamic resilient robotic agent. Such a system also provides a meaningful instance for embodied cognitive science, particularly being incorporated with dynamical systems theory, which is theorised as the situated, embodied, dynamical (SED) framework (Beer, 2011).

2.6 Concepts in Dynamical Systems Theory

A dynamical system is a mathematical abstraction that describes how the state of some system evolves over time. It is commonly expressed by a set of ordinary differential equations or iterated maps, but many other kinds of mathematical descriptions are available. Dynamical systems theory provides the appropriate theoretical language and tools for analysing autonomous agents such as our system. Using the language of dynamical systems theory, the behavior of the system can be qualitatively understood by identifying its invariant sets and the dependence on system parameters. Although dynamical systems theory is one of many hypotheses to explain the natural world, it provides the concepts, intuitions, and metaphors for the design of the systems we want, as well as for the interpretation of their results. This stance, called the dynamical perspective (Beer, 2011), is widely accepted in cognitive science, and it is one that this study also advocates. Since our system was designed and analysed from the dynamical systems perspective, let us briefly go over some of the concepts in dynamical systems theory for the terms which are used in this work. For a comprehensive introduction to dynamical systems theory see (Hirsch et al., 2012).

Autonomous/non-autonomous dynamical system

A dynamical system whose evolution depends only on its internal state is called autonomous, while one whose evolution also depends on external inputs is called non-autonomous. Characterising a dynamical system as autonomous or non-autonomous depends on how we define the boundary of the system of interest. Usually an agent and its environment are considered as two interacting non-autonomous systems. On the other hand, we can view the two coupled non-autonomous systems and their interaction as a single autonomous dynamical system whose state variables are the union of the state variables of the two non-autonomous systems (Beer, 1995a, 2000, 2008), which is the approach we take in this thesis.

Phase space

Phase space is an abstract mathematical space in which all possible states of a dynamical system are represented. Each possible state of the system is represented as a unique point in the phase space. The trace of state point evolving over time is

called an orbit or a trajectory.

Attractor

A set of points or a subspace in phase space toward which the orbit of dynamical system evolves over time. It is an attracting invariant set (a set in phase space is *invariant* if a trajectory that starts on the set remains on it forever). An attractor has its own geometry in phase space, which can be a point, a set of points, a curve, or a complicated fractal structure. In practice, classical attractors are equilibrium points, limit cycles, quasiperiodic or chaotic attractors. A dynamical system often has more than one attractor. In a phase space where multiple attractors coexist, each attractor has its basin of attraction which is a set of initial points from which all trajectories eventually approach that attractor.

Periodic orbit

A type of solution for a dynamical system whose orbit exhibits a periodic motion. A stable periodic orbit appears as an attractor in the phase space, whose shape is a closed curve. A periodic orbit on a two-dimensional manifold (on a plane) is called a *limit cycle*.

Quasiperiodicity

An oscillating solution consisting of two or more incommensurate frequencies (Two frequencies are said to be incommensurate if their ratio is irrational). A well-known instance of a quasiperiodic attractor is a *torus*, an invariant set which forms a surface of a doughnut in a three dimensional phase space. The circular motion of the orbit on the attractor is the sum of two oscillatory motions whose frequencies are incommensurate, one around the small radius and the other around the large radius. Accordingly, a hypertorus can exist in the phase space of higher dimensions, where the periodic motion has more than two frequency components.

Chaos

A type of motion that is sensitive to changes in initial conditions. A motion for which trajectories starting from infinitesimally close initial conditions diverge exponentially with time. It is usually identified by the motion exhibiting a positive *Lyapunov exponent*, a quantity of a dynamical system that characterises the rate of separation of slightly different initial points.

Bifurcation
A qualitative change in the dynamics or a change in the topological structure of a dynamical system by varying the system parameters. One of the most frequently covered types of bifurcation in the field of robot locomotion is the *Hopf bifurcation*, in which a stable fixed point of a dynamical system loses its stability and a stable periodic orbit arises.

Chaotic itinerancy
A type of transitory dynamics in which the system state wanders from one quasi-attractor to another through a high dimensional chaotic trajectory, getting entrained in each of them for a while. A *quasi-attractor* is a local region to which flows converge to an ordered periodic activity, while the flows diverge between the regions with disordered and chaotic activity. In chaotic itinerancy, an orbit successively itinerates over different ordered motions of small effective degrees of freedom, which are specifically called *attractor ruins*. The attractor ruin can be thought as having both stable and unstable subspaces, where the orbit is attracted through the stable trajectory and stays there for a while until it eventually leaves through the unstable trajectory.

2.7 Analysis Methods for Oscillation and Chaos

Understanding the behaviour of the exploration system as a whole might be best done by visual observation of the movement of the robot. Deeper understanding of the system which is unable to be captured by a simple visual observation requires a series of mathematical processes for analysing the accessible variables of the system. Although the neuro-robotic system has relatively small numbers of degrees of freedom, there is a lack of accessability to the internal variables of the physics engine used to simulate the articulated robots. Therefore, the analyses of system behaviours should be done using as many variables as possible which are observable from the outside of system. These observed (measured) data are generally the subject of time series analysis.

Time series is a sequence of measurements which follow non-random orders. Unlike the analyses of random samples of observations that are discussed in the context of most other statistics, the analysis of time series is based on the assumption

that successive values in the data file represent consecutive measurements taken at equally spaced time intervals. Often, many of the results produced by time series analysis methods are presented as an image whose data points for the whole time span are superposed to show qualitative or statistical features of the system behaviour. Since our system operates in realtime, the majority of our interest lies in how the system behaviour changes in time. In order to understand the system behaviour more intuitively, we prefer that the analysed time series data are presented and visualised as a function of time wherever it is possible.

Mainly a couple of aspects are considered for our system to investigate its time varying behaviour, namely the phase relationship of limbs and the chaoticity of the system. The interlimb coordination of multi-legged locomotion is often characterised by the phase relationship of the oscillating limbs or corresponding neural signals. In order to look over how the interlimb coordination changes over time, the instantaneous phase differences are calculated and plotted through time flow. The instantaneous phase of periodic time series data $x(t)$ can be obtained by calculating its Hilbert transformation $\hat{x}(t)$ which generates a 90 degrees phase-shifted signal of the original data. Hence the instantaneous (wrapped) phase becomes $\arctan(\hat{x(t)}/x(t))$. For measuring the chaoticity of a time series data, we use a characteristic quantity of dynamical systems called the maximum Lyapunov exponent, which is the rate of separation of initially close trajectories in the mostly separating direction. It can be analytically calculated from the differential equations describing the system dynamics, but if the complete set of the governing equations are unknown as in our system, it has to be numerically obtained from the observable time series data. In order to do this, a technique called state space reconstruction is employed to unfold the multidimensional structure of the system orbit with appropriately chosen embedding delay and embedding dimension. The following sections will present each method in brief detail.

2.7.1 Instantaneous Phase Relationship

Phase difference (phase shift) between two oscillating signals is the most frequently used method for describing the movement pattern of limbed robot. Phase difference is expressed in degrees or radian, between two waves having the same frequency and referenced to the same point in time. However, in order to observe how the phase

relationship between robot limbs varies throughout the exploration process, the time course of phase differences should be presented. The phase at a time instant can be expressed using *instantaneous phase*. Then we can obtain the instantaneous phase differences by subtracting one from another at a specific time.

The instantaneous phase of a two dimensional limit cycle orbit (x, y) can be calculated by $\theta = \arctan(y/x)$. When θ is constrained to an interval such as $(-\pi, \pi)$ or $(0, 2\pi)$ it is called the wrapped phase. Otherwise it is called unwrapped phase, which is seen as a continuous function of time. For the visualisation of phase differences vs. time, the unwrapped instantaneous phase is appropriate because wrapped phase difference exhibits discrete jumps. In case of relaxation oscillators, the speed of orbit on the limit cycle varies in time so the evolution of unwrapped instantaneous phase does not show linear increase, which results the oscillation in phase difference plot. This can be linearised by expressing the phase as the elapsed time normalised by the intrinsic period.

When only a single oscillatory signal is available, we need to generate the 90° shifted signal to calculate instantaneous phase. A simple method might work where the shifted signal is generated by the derivative of the original signal, but the derived signal often shows a highly distorted waveform compared to the original, which is inappropriate for our purpose. Instead, the Hilbert transformation (Bracewell, 1999; King, 2009) is frequently used for generating the shifted signal. Hilbert transform has been introduced by the German scientist David Hilbert (1862-1943) in the beginning of the 20th century. The Hilbert transform $H[f(t)]$ of a signal $f(t)$ is defined as

$$H[f(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau)}{t - \tau} d\tau. \tag{2.1}$$

We can see that $H[f(t)]$ is the convolution of $f(t)$ with the signal $1/\pi t$. However this integral is improper in that the integrand has a singularity and the limits of integration are infinite. Although the Hilbert transform is properly defined as the *Cauchy principal value* of the integral whenever this value exists, alternatively, the principal value integral can be written explicitly by changing variables as

$$H[f(t)] = -\frac{1}{\pi} \lim_{\epsilon \to \infty} \int_{\epsilon}^{\infty} \frac{f(t + \tau) - f(t - \tau)}{\tau} d\tau. \tag{2.2}$$

An original function $f(t)$ and its Hilbert transform $H[f(t)]$ are orthogonal to each other, hence the basic property of the Hilbert transform is $\pi/2$ phase-shift oper-

ator. The Hilbert transformation of signal $f(t)$ is written as $\hat{f}(t)$ as the common mathematical notation.

2.7.2 Hilbert-Huang Method

If the oscillating centre of a periodic signal deviates from zero, the centre of rotation of the two dimensional orbit $(f(t), \hat{f}(t))$ also deviates from the origin, which results in inaccurate or oscillating instantaneous phase. Therefore the time series data should be modified so that its centre of oscillation is aligned near zero. Often the measured data are nonstationary and nonlinear and they hardly show zero-offset oscillation. To address this problem, the Hilbert-Huang transformation (HHT) is used (Huang et al., 1998).

The fundamental part of the HHT is the empirical mode decomposition (EMD) method. EMD decomposes any complicated data set into a finite number of components, which is a collection of intrinsic mode functions (IMF). An IMF represents a generally simple oscillatory mode as a counterpart to the simple harmonic function. An IMF must satisfy the definition that it has the same number of extrema and zero crossings and its envelopes are symmetric with respect to zero. This definition guarantees a well-behaved Hilbert transform of the IMF. Since the decomposition is based on the local characteristic time scale of the data, it can be applied to nonlinear and nonstationary processes.

The extracting procedure of an IMF is called *sifting*. The sifting process is as follows:

0. Copy the original data set and name as I.

1. Identify all the local extrema in I.

2. Generate the upper envelope by connecting all the local maxima using a cubic spline.

3. Generate lower envelope by repeating the procedure 2 for the local minima.

4. Obtain a new data set I_N by subtracting the local mean values of upper and lower envelopes from I.

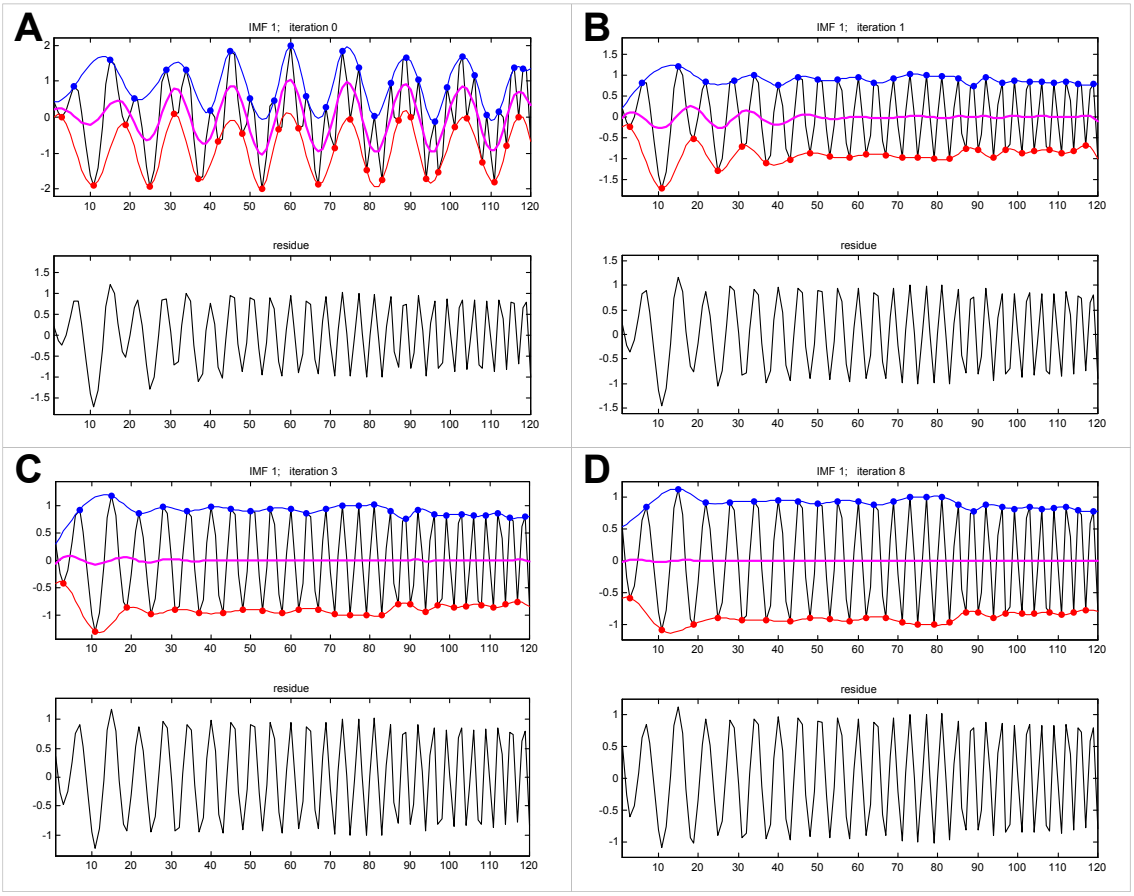

Figure 2.5: An example of iterations of the EMD procedure and the resulting residua after each iteration; (A) iteration 0, (A) iteration 1, (A) iteration 3, (A) iteration 8. Images were taken from Tolwinski (2007).

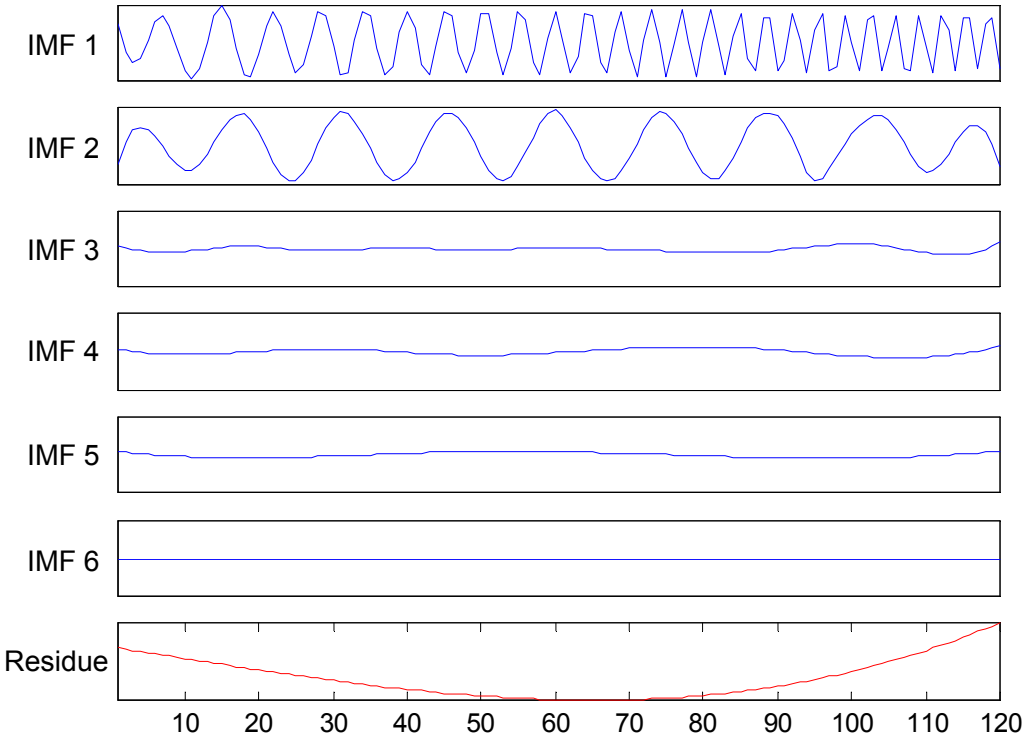

Figure 2.6: Decomposed signals by EMD procedure and the final residue. Images were taken from Tolwinski (2007).

5. If I_N has negligible local mean then terminate sifting, otherwise set $I = I_N$ and repeat 1 to 4.

An example of sifting process is shown in Figure 2.5. After the first IMF is produced by above loop it is subtracted from the original signal. This difference between the original data and the IMF is called *residue*. The construction of an IMF is now repeated for the residual signal as an outer loop. This EMD procedure is repeated until the residue is either constant or monotone. As a result, we have a series of IMF data and the final residue. The original signal can be reconstructed by summing up these decomposed data.

$$s(t) = \sum_i \text{IMF}_i + \text{residue} \tag{2.3}$$

Figure 2.6 shows the decomposed data. The remaining residue reveals any trend that exists in the data.

In this work, the periodic data measured from the robot movement is not decomposed into multiple components which have different frequencies. Instead, only sifting process is used for the purpose of well-behaving Hilbert transformation by generating the first satisfactory IMF of the original time series which is aligned around zero. This is because we want to preserve all the frequency components in order to generate instantaneous phase vs. time plot. However, generating IMF for highly variable periodic data such as chaotic signals often exhibit large fluctuation during the enveloping of extrema using cubic splines. This can be greatly relieved by slightly pre-smoothing the original data using low-pass filter, which removes the components out of our resolution of interest where the frequency is too high and the amplitude is too small.

2.7.3 Lyapunov Exponent from Time Series

Variable system dynamics of the chaotic exploration process can be quantitatively measured using the Lyapunov characteristic exponent (Lyapunov exponent). The Lyapunov exponent of a dynamical system is a quantity that characterises the rate of separation of infinitesimally close trajectories. In state space, two dynamical system are similar if two system orbits are close together. While the two infinitesimally close points change similarly when the systems are linear, those of nonlinear system may exhibit significant separation after being evolved through time. The rate of separation of neighbouring orbits usually increases exponentially, and this exponent is defined as Lyapunov exponent. Lyapunov exponents exist as many number as the system dimension, and estimation of the maximum Lyapunov exponent (MLE, usually denoted as λ_1) is often used in diagnosing chaos. The maximal Lyapunov exponent λ_1 can be defined using the relationship,

$$d(t) = d(0)e^{\lambda t} \tag{2.4}$$

where $d(0)$ is the initial separation of two neighbouring points and $d(t)$ is the divergence at time t.

Directly computing the Lyapunov exponent from the conventional mathematical definition is inappropriate for a finite experimental time series. Lyapunov exponents are only truly defined in the limit of infinite time which is impractical for real world problems. Using numerical simulation, inaccuracies from small numerical error will

accumulate over time until small perturbations in contracting direction eventually realign themselves along exponentially expanding directions. Besides, the mathematical definitions of lyapunov exponent can only be applied when the governing differential or difference equations are known. In many real world applications, these equations are either unknown or extremely difficult to derive, as in the case of our system.

2.7.4 State space reconstruction

The algorithms estimating maximum lyapunov exponent from experimental time series data are generally based on state space reconstruction procedures. From Taken's embedding theorem (Takens, 1981), one can reconstruct an appropriate state space from a single original time series and its time-delayed copies as:

$$\mathbf{x}(t) = [x(t), x(t+\tau), x(t+2\tau), ..., x(t+(m-1)\tau)]^T \tag{2.5}$$

where $\mathbf{x}(t)$ is the m dimensional state vector, and $x(t)$ is the original data. τ is the time delay, and m is the embedding dimension.

2.7.4.1 Choosing embedding delay

Choosing a good time delay is to find a delay large enough such that the information we get from measuring the value of variable $x(t)$ variable at time $x(t+\tau)$ is relevant and significantly different from the information we already have. At the same time, it should not be so large that they are completely independent statistically, not letting the system loose 'memory' of the previous state. The most widely used method to select proper time delay is to choose τ by taking the first minimum of the average mutual information (AMI) (Fraser and Swinney, 1986), which evaluates the amount of information shared between two datasets over a range of time delays. Given a time series of the form $(x_0, x_1, x_2, ..., x_t, ..., x_T)$, the difference between the maximum (x_{max}) and the minimum (x_{min}) of the sequence $(|x_{max} - x_{min}|)$ is partitioned into j equally sized intervals, where the bin size j should be a large enough integer number. Then the mutual information is expressed as:

$$I(\tau) = \sum_{h=1}^{j}\sum_{k=1}^{j} P_{h,k} \ln P_{h,k}(\tau) - 2\sum_{h=1}^{j} P_h \ln P_h \tag{2.6}$$

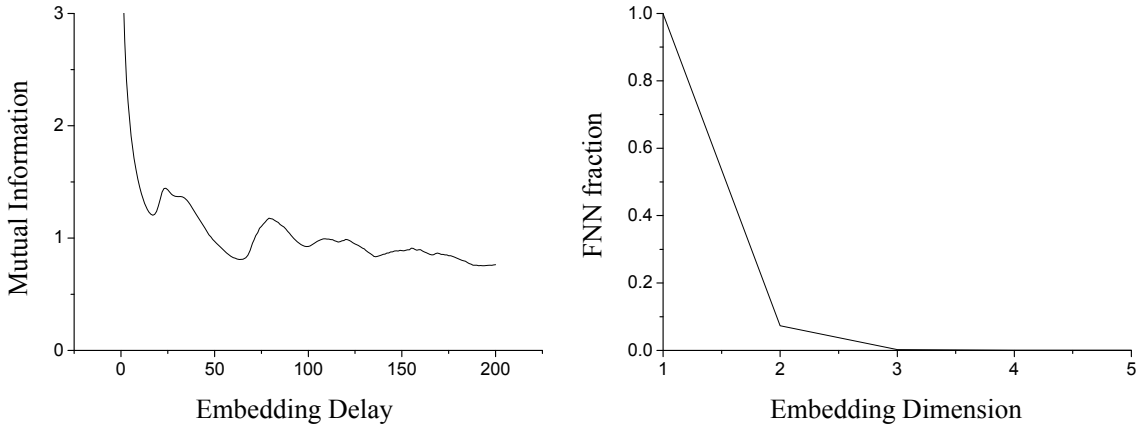

Figure 2.7: Mutual information and FNN calculated from the time series of Lorentz system. Embedding delay is chosen to $\tau = 17$ (bits) from the first minimum of mutual information. FNN graph shows the fraction of false nearest neighbour drops to zero at the embedding dimension $m = 3$.

where P_h and P_k are the probabilities each of which assumes a value inside the h-th and k-th bin. $P_{h,k}(\tau)$ is the joint probability that x_t is in bin h and $x_{t+\tau}$ is in bin k. This metric evaluates the amount of information shared between two datasets over a range of time delays. Choosing the first minimum of the mutual information provides adjacent delay coordinates with a minimum of redundancy. While it has been empirically shown that the first minimum of $I(\tau)$ often yields the optimal embedding delay, there is lack of a formal mathematical proof on this, and even it has been pointed out that there is no theoretical reason why a minimum of the mutual information exists (Kantz and Schreiber, 1997). Although it has often been proved reliable and has shown well performance for the given purpose, giving full credit to the determined embedding delay should be avoided if the reconstructed system shows somewhat doubtful results.

2.7.4.2 Choosing embedding dimension

A valid state space must include a sufficient number of coordinates (m) in order to fully resolve the complex structure of the attractor. While Takens' embedding theorem guarantees a proper embedding for all sufficiently large $m > 2n$ (where n is the

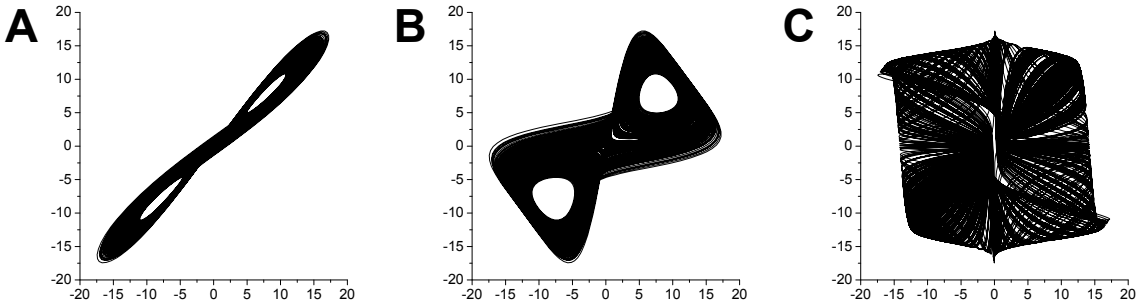

Figure 2.8: Reconstructed phase space ($x(t)$ versus $x(t+\tau)$) using different embedding delays and embedding dimensions. (A) $\tau = 3$, $m = 2$. (B) $\tau = 17$, $m = 3$. (C) $\tau = 100$, $m = 4$.

actual attractor dimension), it has been proven that attractors reconstructed using smaller values of m often yield reliable Lyapunov exponents, which suggests that the minimum number of dimension to describe the attractor is sufficient. Therefore the problem of finding good embedding dimension can be seen as an optimisation procedure yielding just the minimal required m.

The most widely used method to determine optimal embedding dimension is False Nearest Neighbors (FNN) algorithm (Kennel et al., 1992). The FNN assumes that the orbit of a deterministic system evolves smoothly with no sudden irregularities in the attractor structure. Two neighbouring points on the reconstructed embedding space have to stay sufficiently close during forward iteration of system. FNN algorithm compares the distances between neighboring trajectories at successively higher dimensions. 'False neighbors' occur when the points apparently lying close together due to projection are separated in higher embedding dimensions. A natural criterion for catching embedding errors is that the increase in distance between two neighboured points is large when going from dimension d to $d+1$.

When we denote the distance between two neighbouring points in the d-dimensional embedding space at time t as $R_d(t)$, the normalised distance R_i is calculated as:

$$R_i = \left(\frac{R_{d+1}^2(t) - R_d^2(t)}{R_d^2(t)} \right)^{\frac{1}{2}} = \frac{|x_{i+d\tau} - x_{t+d\tau}|}{\|\mathbf{p}(i) - \mathbf{p}(t)\|} \tag{2.7}$$

where τ is embedding delay, and $\mathbf{p}(t)$ and $\mathbf{p}(i)$ are a reference point in the d-dimensional embedding space and its nearest neighbour respectively. $\mathbf{p}(i)$ is chosen

so that $\|\mathbf{p}(i) - \mathbf{p}(i))\| < \epsilon$, where ϵ is a small constant. $\mathbf{p}(i)$ is designated as a false nearest neighbour of $\mathbf{p}(t)$ if R_i is larger than a given threshold R_{th}. This calculation is performed for the entire time series (for all t) for increasing d until the total percentage of false neighbors declines to near 0%. Kennel et al. (1992) suggested that $R_{th} = 10$ works well for most cases, although other values should be explored if the reconstructed system does not show convincing result since there is no concrete mathematical proof for choosing optimal R_{th}.

Figure 2.7 shows the mutual information and FNN calculated from the time series data obtained from Lorentz system whose differential equations,

$$\dot{x} = \sigma(y - x) \tag{2.8}$$

$$\dot{y} = rx - y - xz \tag{2.9}$$

$$\dot{z} = xy - bz \tag{2.10}$$

with the parameters $\sigma = 10$, $r = 25$, $b = 8/3$, which exhibits chaotic behaviour. By solving the above equations, the time series data was obtained from the variable x sampled at a time interval of 0.01s for 400 seconds. The 'optimal' embedding parameters were estimated as $\tau = 17$ for embedding delay and $m = 3$ for embedding dimension. The reconstructed attractors with different embedding parameters are shown in Figure 2.8, suggesting that the shape of reconstructed attractor using optimal embedding parameters seems to make good agreement with the original system.

2.7.5 Wolf's Method

As the maximum Lyapunov exponent is often of greatest interest in diagnosing chaos and because of the inherent computational difficulties in estimating the full Lyapunov spectrum from a single time series data, a number of methods have been proposed to compute just the maximum Lyapunov exponent. Perhaps the most well-known method is Wolf's algorithm (Wolf et al., 1985) which is conceptually simple and implements the theory in a direct fashion.

The algorithm monitors the long-term evolution of a single pair of initially nearby orbits. Given some initial point $\mathbf{p}(t)$ in reconstructed state space and its nearest neighbor $\mathbf{p}(i)$ which is not on the same trajectory, the initial Euclidean distance between the two points is denoted $\epsilon_t = \|\mathbf{p}(i) - \mathbf{p}(t)\|$. After iterating both points

forward in time for a fixed evolution time h, the initial length will have evolved to length $\epsilon_{t+h} = \|\mathbf{p}(i+h) - \mathbf{p}(t+h)\|$.

Assuming that the system is chaotic, the distance after the evolved time ϵ_{t+h} will be larger than ϵ_t. After each forward evolution, the replacement step is performed that a new nearest neighbor point to the original reference trajectory is selected that minimises both the replacement length and the orientation change between the reference and neighboring points. At each replacement step, a new neighbouring point $\mathbf{p}(j)$ is selected so that both the distance to the evolved point $\mathbf{p}(t+h)$ and the angular separation between the line segments which is formed by the points ($\mathbf{p}(t+h)$,$\mathbf{p}(i+h)$) and ($\mathbf{p}(t+h)$,$\mathbf{p}(j)$) are minimised. This procedure is repeated until the original reference trajectory has traversed the entire dataset. Then the maximal Lyapunov exponent λ_1 is calculated as

$$\lambda_1 = \frac{1}{Mh} \sum_{i=1}^{M} \ln \frac{\epsilon_{t+h}}{\epsilon_t} \tag{2.11}$$

where M is the total number of replacement steps. The result of Wolf's algorithm for the Lorentz system with optimal embedding parameters is shown in Figure 2.9 (left). Practical implementation of the algorithm must carefully choose parameters for the replacement step. Maximum allowed replacement distance and orientation errors should be minimised while one can still find a neighbouring point within that limits. The fixed evolution time (h) also have to be tuned so that it is maximised in order to save computation cost and to reduce the number of replacement step which cause the length and orientation errors to be accumulated. Also the reconstructed attractor should be densely populated with a sufficient number of data points in order to achieve successful replacement steps. Although the algorithm implements the theoretical concept of Lyapunov exponent straightforwardly, it is generally accepted that the algorithm suffers from the unreliability for small datasets as well as the sensitivity of result to the parameter choice.

2.7.6 Rosenstein's Method

The alternative approaches which are more robust for small datasets were proposed by Rosenstein (Rosenstein et al., 1993) and Kantz (Kantz and Schreiber, 1997). Both algorithms are based on the basic definition of the maximum Lyapunov exponent

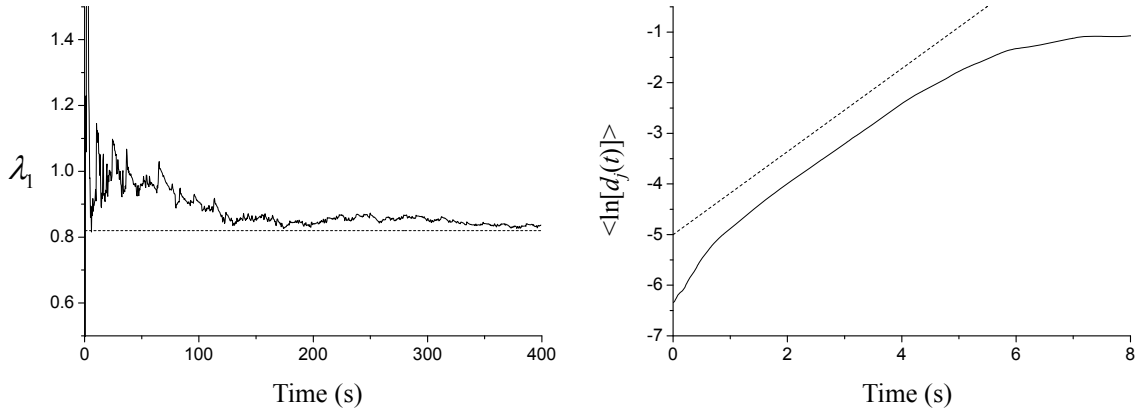

Figure 2.9: Maximal Lyapunov exponents calculated from the time series of Lorentz system with $\tau = 17$ and $m = 3$. Dashed lines indicate the maximal Lyapunov exponent calculated using the differential equations of Lorentz system ($\lambda_1 = 0.82$). (Left) Result using Wolf's method. (Right) Result using Rosenstein's method.

from the equation 2.4. Since both methods are very similar (differs in that using a single neighbour or a cluster of neighbours) we use Rosenstein's method in this work. From the definition, the jth pair of nearest neighbors in the reconstructed system diverge approximately at a rate given by the maximal Lyapunov exponent:

$$d_j(i) \approx d_j(0) e^{\lambda_1 (i \triangle t)}, \tag{2.12}$$

where $d_j(i)$ is the mean Euclidean distance between the jth pair of nearest neighbors after i discrete time steps $i \triangle t$ seconds), and $d_j(0)$ is the initial separation. By taking the logarithm of both sides of above equation,

$$\ln[d_j(i)] \approx \ln[d_j(0)] + \lambda_1 i \triangle t, \tag{2.13}$$

The above equation yields a set of approximately parallel lines (for each j), each of slope roughly proportional to λ_1. Then the maximal Lyapunov exponent can be robustly estimated from the slope of a least-squares fit to the 'average' log divergence curve defined by:

$$S(i) = \frac{1}{\triangle t} \langle \ln[d_j(i)] \rangle \tag{2.14}$$

where $\langle \cdot \rangle$ represents the average over all values of j. Figure 2.9 (right) shows the divergence vs. time curve generated from the Lorentz system data. We can see that the slope of the linear region of the curve agrees well with the expected value (the slope of dashed line: 0.82).

Chapter 3

Models and Methods

This chapter describes the models and methods for the proposed exploration system. The chapter includes the model of the neural part for limb control, the simulation model of the common parts for articulated bodies, and its interaction with the physical environments for 2D aquatic and 3D terrestrial robots. Before going into the detailed description of simulation models, we start with outlining the general architecture of the exploration system for locomotor behaviours, followed by brief explanation of each sub-component. The subsequent sections describe each component in more detail which include central pattern generator (CPG) model and simulated neuromuscular system. Dynamical properties of both single and coupled CPG models are examined in terms of the applicability to the limbed robot movement, particularly choosing appropriate common fixed parameters of CPGs which are suitable for the chaotic exploration process. After that, the CPG equations are incorporated with local sensory information and encapsulated as a unit module in order to be applied to the robot with arbitrary degrees of freedom. Finally the robotic model is addressed by describing the basic neuromuscular model for the joint-motor of a limb which is used as the common component independent of the body shape and the physical environment. Integration of mechanical and neural parts is also presented by modelling a simple reflex loop using muscle proprioceptors. Finally, some methods for analysing the system behaviours are presented.

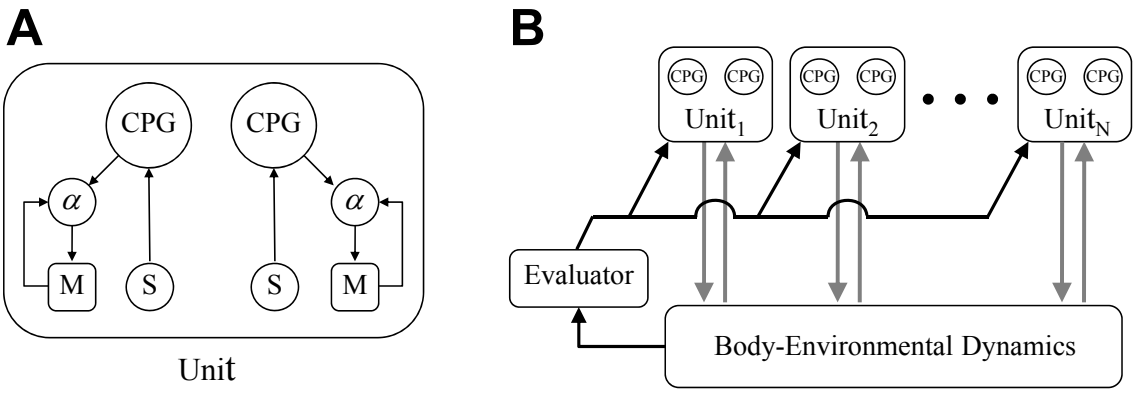

Figure 3.1: Schematic diagram of the general exploration system. (A) A *motor unit* for a single degree of freedom in the muscle-joint system. A unit consists of two electrically disconnected oscillators, which receive integrated information of other oscillators in the system from the sensor (S), via body-environment interactions, and give a control signal to the muscle (M) via an α motor neuron. (B) Schematic diagram of the exploration system whose body has N degrees of freedom. The chaoticities of all units are altered according to the global control signal from the evaluator.

3.1 General Scheme

The basic architecture of the neural part of the system developed in this study is based on Kuniyoshi and Sangawa's (Kuniyoshi and Sangawa, 2006) *Medulla Oblongata* inspired model, but with a more compact and modular configuration for each joint of the limbed robot. It is intended to be applicable to a wide range of robotic systems. The architecture consists of a number of identical control modules connected to each of the body parts in their environment. Each neuromuscular system for a joint which receives afferent sensory input and gives motor output to a antagonistic muscle pair can be encapsulated as a single *motor unit*, and the whole system consists of identical motor units whose number is the same as the number of degrees of freedom of the robot (Fig. 3.1). This configuration eliminates muscle redundancies by constraining joint-motors to be operated only by an antagonistic muscle pair, thus giving more weight to interlimb interactions. Although it is expected that our system will likely be able to deal with multiple muscles acting on a single joint, we

constrain the system complexity to a pair of muscles per joint in order to focus on the interlimb coordination. The signal from the sensors of a motor unit (in most case the mechanosensory information from haptic sensors or muscle afferent) is fed to the corresponding oscillators that each motor unit contains. A muscle is activated by an efferent (α) motoneuron and the motoneuron has an afferent connection from the homonymous muscle spindle, forming a simple stretch reflex loop.

The overall architecture of the system is illustrated in the Figure 3.1A. The neural architecture generalises and extends that presented in (Kuniyoshi and Sangawa, 2006). The CPG neurons are all disconnected to each other but they form functional coupling through physical embodiment. The CPG neurons receive sensory signals that integrate information from the body-environment interaction dynamics experienced by the system. Hence, while the direct connections are absent, any coupling between the oscillators will be indirect via bodily and environmental interactions. The network of oscillators, coupled through physical embodiment, has multiple synchronised states (modes) in the state space of neuro-body-environment dynamics that reflect the body schema and its interactions with the environment, each of which can be regarded as a potential candidate for 'meaningful' motor behavior. The exploration process, powered by adaptive bifurcation through the feedback evaluation signal, allows the system to become entrained in these modes, one at a time, until one is found that is sufficiently stable and high performing for the bifurcation parameter to reduce to zero and the system to fully stabilise. In the following sections, each component will be described in more detail.

3.2 CPG Model

The control signals for the basic motor patterns are generated by central pattern generators (CPGs), which are composed of a collection of neurons that produce an oscillatory signal for various locomotor patterns by synchronisation with the movement of the physical systems. The model consists of a coupled Bonhoeffer-van der Pol (BVP, or Fitzhugh-Nagumo) oscillators (Fitzhugh, 1961; Nagumo et al., 1962), which are widely studied as models of pacemaking cells and interlimb coordination. BVP model is a two-dimensional simplification of the Hodgkin-Huxley model of spike generation in squid giant axons (Hodgkin and Huxley, 1952), and it repro-

duces the behavior of pacemaker neurons qualitatively well. The model is called Bonhoeffer-van der Pol (by Fitzhugh) because it is reduced from the van der Pol equation for self-sustained oscillation and its behavior resembles that of the iron nitric acid model proposed by (Bonhoeffer, 1948).

A particularly interesting feature of coupled BVP equations, that allows adjustment of the complexity of the system orbit, had been presented by (Asai et al., 2000, 2003a,b). A pair of coupled BVP oscillators generates a stable limit cycle when the two control inputs are the same, but a quasiperiodic/chaotic orbit otherwise. Another interesting feature of the BVP model is flexible phase locking (Sproewitz and Berthouze, 2005; Ohgane et al., 2009), where the phase relationship between CPG activity and body motion can be flexibly locked according to a loop delay. This is a beneficial feature for covering a range of sensorimotor delays originating from different body-environment configurations.

3.2.1 Bonhoeffer-van der Pol Oscillator

Before going into the detailed application of the BVP model to the chaotic exploration system, some dynamical properties of the basic model of BVP equations are addressed. The model consists of a voltage-like variable (u) having cubic nonlinearity that allows regenerative self-excitation via a positive feedback, and a recovery variable (v) having a linear dynamics that provides a slower negative feedback. The equations of a single BVP oscillator are:

$$\dot{u} = c(u - \frac{u^3}{3} - v + z) \tag{3.1}$$

$$\dot{v} = \frac{1}{c}(u - bv + a), \tag{3.2}$$

where a, b, and c are the fixed parameters of the oscillator. z is the control parameter from the higher centre. The solution of equations 3.1 and 3.2 mimics neuronal activity if parameters a, b, and c are set as following criteria.

$$1 - \frac{2b}{3} < a < 1, \quad 0 < b < 1, \quad b < c^2 \tag{3.3}$$

In this range of parameters, the equation 3.1 and 3.2 exhibit Hopf bifurcation as the bifurcation parameter z varies. The analytically derived critical values of z (Nomura

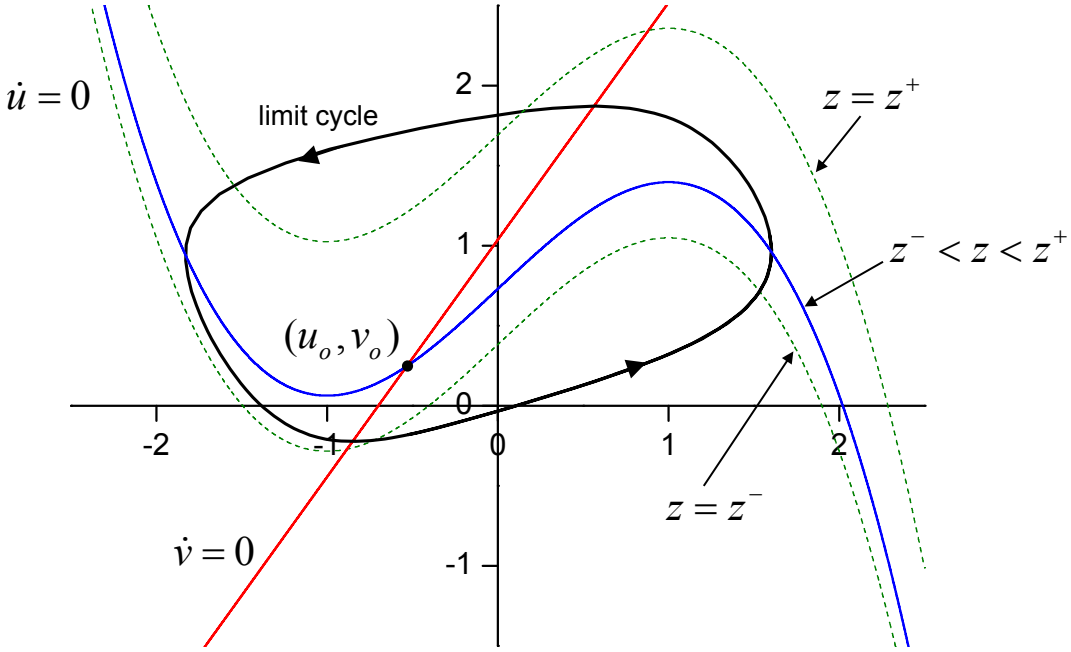

Figure 3.2: Nullclines and limit cycle of BVP model of the equations 3.1 and 3.2.

et al., 1993) where Hopf bifurcations take place are

$$z^{\pm} = \frac{a}{b} \pm \sqrt{1 - \frac{b}{c^2}} \left(\frac{1}{b} - \frac{2}{3} - \frac{b}{3c^2} \right). \tag{3.4}$$

Between these values, i.e., $z^- < z < z^+$ there exist an unstable equilibrium point (u_o, v_o) and an asymptotically stable limit cycle.

The fixed parameters of BVP model were set $a = 0.7$, $b = 0.675$, and $c = 1.75$ throughout this work, which were chosen based on the previous work (Asai et al., 2003a) whose parameter setting was rigorously studied in terms of exhibiting various dynamical regime from stable to chaos which meets the purpose of our work. They were also verified by later studies using robotic platform (Kuniyoshi and Sangawa, 2006; Kinjo et al., 2008). Therefore, the critical values of z are $z- = 0.38247$ and $z^+ = 1.6916$, where subcritical bifurcations take place. The asymptotically stable limit cycle rotates counterclockwise around the equilibrium point, and the model acts as a pacemaker neuron. The magnitudes of vector fields (the right-hand sides of the equations 3.1 and 3.2) near the right and left branch of the N-shaped nullcline (i.e., $\dot{u} = 0$, blue line in Figure 3.2) are small, so the state variables move slowly

there. In contrast, vector fields in the horizontal direction along the bottom or the top of the limit cycle are large, and the state point moves quickly.

3.2.2 Variable Chaoticity of Coupled BVP Oscillators

Let us consider two coupled oscillators of equation 3.1 and 3.2. By denoting the state variables of each oscillator as (x_1, y_1) and (x_2, y_2), they are mutually coupled by excitatory-to-all connections to form the CPG model as follows (Asai et al., 2000, 2003b):

$$\dot{x}_1 = c(x_1 - \frac{x_1^3}{3} - y_1 + z_1) + \delta(x_2 - x_1) \tag{3.5}$$

$$\dot{y}_1 = \frac{1}{c}(x_1 - by_1 + a) + \varepsilon x_2 \tag{3.6}$$

$$\dot{x}_2 = c(x_2 - \frac{x_2^3}{3} - y_2 + z_2) + \delta(x_1 - x_2) \tag{3.7}$$

$$\dot{y}_2 = \frac{1}{c}(x_2 - by_2 + a) + \varepsilon x_1 \tag{3.8}$$

where two coupling constants are $\delta = 0.013$ and $\varepsilon = 0.022$ respectively. Since the refractoriness variable y_i $(i = 1, 2)$ inhibits activity of x_i, the connection by ε from contralateral output (x_j, where $j \neq i$) to y_i may result in inhibition of x_i, resulting in the above CPG model becoming a half-centre model with reciprocal inhibition.

This model originates from clinical experiments to evaluate severity of symptoms in Parkinson's Disease (PD) patients by looking at coordination between lower limbs during pedaling exercise (Abe et al., 2002). Subjects were instructed to pedal a special kind of pedalling machine that allows its left and right pedals to rotate independently. Coordination between lower limbs during the pedaling exercise was characterised by the velocity and relative phase profiles to see whether the symptoms on one side of the body are severer than that of the other side. The velocity profile in normal subjects was periodic with almost constant amplitude for every cycle with the relative phase locked at about 180 degree. Experiment in PD patients exhibited disordered coordination that shows the amplitude modulation of the velocity profiles and/or drift of the relative phase. This left-right asymmetries could be reproduced using a coupled BVP equations by giving different control inputs which models a flow of descending signals from higher motor centres down to each half-centre of the CPG. When these two control inputs are identical ($z_1 = z_2$), the model has

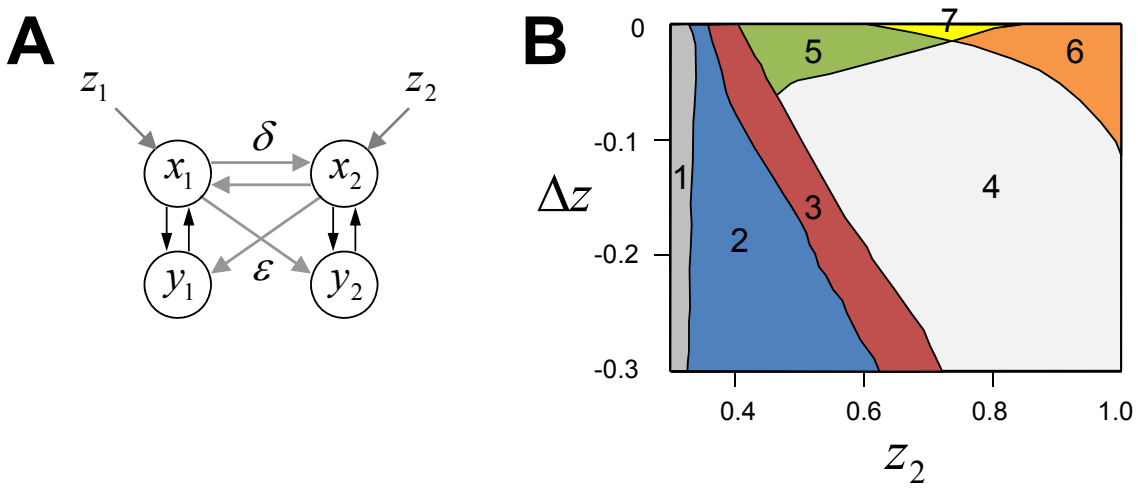

Figure 3.3: Coupled BVP oscillators and the classification of its dynamics. (A) Schematic diagram of coupled BVP oscillators as coupled half-centre system. (B) Classification of the dynamics of coupled BVP system which is reproduced from (Asai et al., 2003a). Numbered regions indicate qualitatively different dynamics of the system which exhibit limit cycle, quasiperiodic, and chaotic orbits. Their brief behaviours correspond to those of figure 3.4. For instance, the region 5 and 6 correspond to A and B, region 4 to C and D, and region 3 to F and G (Refer (Asai et al., 2003a) for more details). The bistable phase-locking of both in-phase and anti-phase occurs in region 7, which is suitable parameter range to promote the diversity of emergent behaviours for the exploration system.

left-right symmetry. If two control inputs unbalance ($z_1 \neq z_2$), the model's symmetry is broken.

Figure 3.3B describes the bifurcation map of the qualitative dynamics of two coupled oscillators reproduced from (Asai et al., 2003a). Each numbered region indicates the range of two parameters z_2 and $\triangle z = z_1 - z_2$ where the coupled system exhibits distinct qualitative behaviour. In region 1, the system shows no oscillatory dynamics. In region 2, x_1 and x_2 oscillates with different amplitude even when the system is driven by the identical tonic signals ($\triangle z = 0$). The relative phases in 2 are locked at some value between 0° and 180°. However, the amplitude of the oscillator receiving smaller control input (x_1) is usually too small, which may

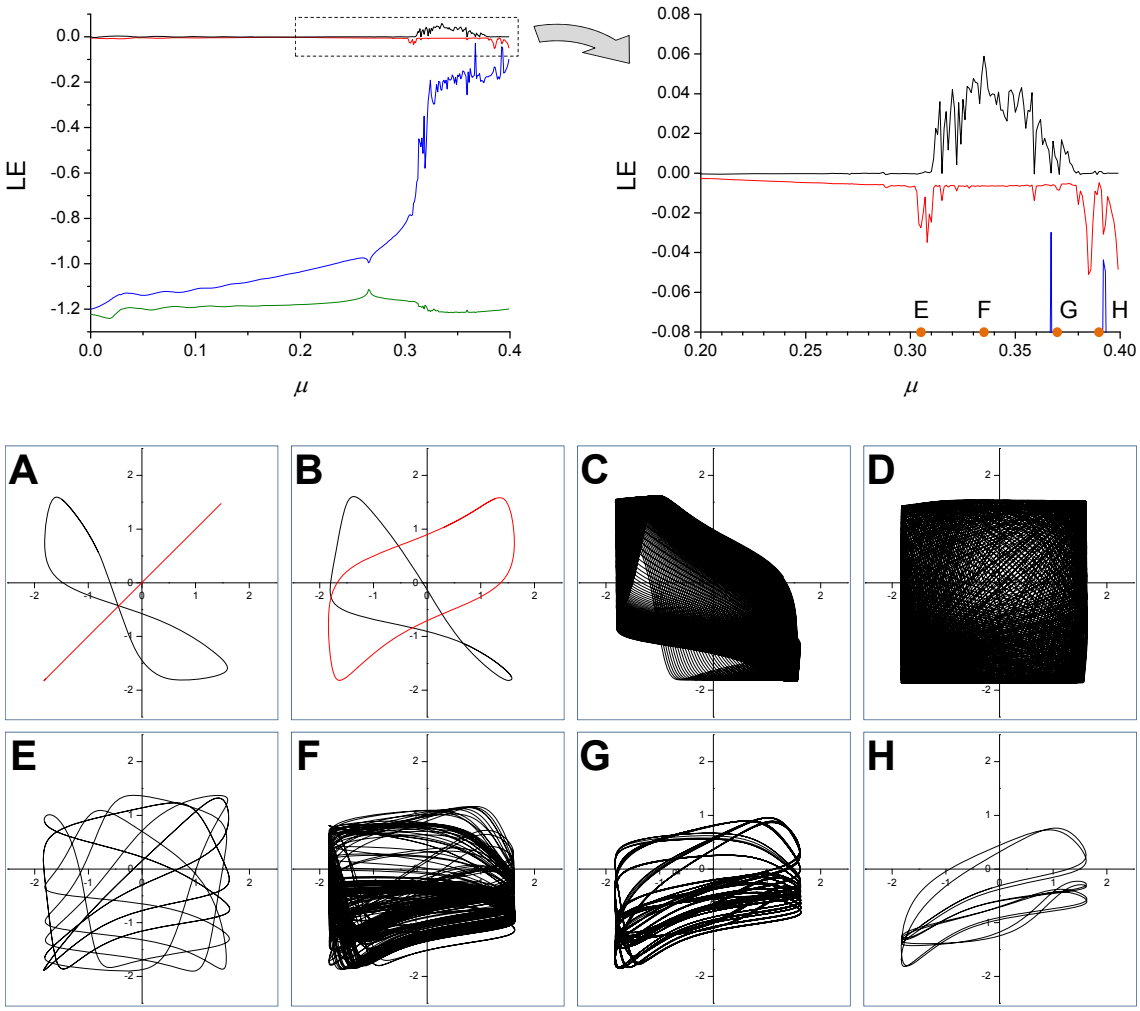

Figure 3.4: Behaviour of two coupled BVP equations. Two upper graphs show the lyapunov exponents (λ_1: black, λ_2: red, λ_3: blue, λ_4: green) of coupled system, where μ is the difference between control inputs ($\mu = z_2 - z_1$) with $z_2 = 0.73$. The region around the chaotic dynamics of system is shown as an enlarged graph on the right side. Lower eight images depict the evolution of the state variable (x_2 vs. x_1) with different μ; A:0.0, B:0.01, C:0.02, D:0.15, E:0.305, F:0.335, G:0.37, and H:0.39. Superposed plots in A and B show two different phase-locked dynamics from different initial states, which indicates the bi-stability of system in its stable regime. C and D show slow and fast drift of relative phases in quasiperiodic regime. E shows the oscillation with multiple period. In F, chaotic oscillation occurs. As the control input of the first oscillator (z_1) crosses its critical value (G and H) the amplitude of x_1 slowly decreases.

not be suitable for the purpose of robot control where the limb is commanded by an antagonistic oscillator outputs, because the movement of limb will depend mostly on one side of motor output. In region 5, the oscillation is locked nearly anti-phase, and the amplitudes of x_1 and x_2 are almost or completely identical. In region 6, the oscillation is almost in-phase, with almost or completely identical amplitudes of both outputs. Region 7 shows the bi-stability in phase-locked oscillation, which exhibits either the dynamics of region 5 or those of region 6, depending on the initial condition. Above mentioned parameter regions 2, 5, and 6 exhibit periodic oscillations of period one. In case of region 3, the model exhibits periodic oscillations with multiple periods, or chaotic oscillations via a sequence of bifurcations. In the largest region 4, the model shows quasi-periodic oscillations in which the relative phases drift from 0° to 360°.

Some examples of the above qualitative dynamics are described in Figure 3.4. Let us denote the difference between two control inputs $\triangle z = z_2 - z_1$ as μ, which will act as the control parameter of coupled BVP system. Upper images in Figure 3.4 depict full Lyapunov spectrum according to the control parameter μ with $z_2 = 0.73$ where the system shows bi-stability. The orbits of A and B clearly show the bi-stable solution of the system which corresponds to the region 7 in Figure 3.3. Slight increase of control parameter in B results in the deformation of orbits, which indicates the variation of relative phases between x_1 and x_2. When the control parameter enters into quasiperiodic regime, the relative phases of x_1 and x_2 drift from 0° to 360° as seen in C and D where the orbit eventually covers the whole rectangular region in the $x_2 - x_1$ plane. In E, the orbit loses its quasiperiodicity and exhibits stable oscillations with finite period. The parameter region in upper right graph where the second lyapunov exponent (red line) drops just before chaotic regime indicates this stage. F shows the chaotic orbits which occupies most part of the region 3 in Figure 3.3. In G and H, the control input (z_1) of the first oscillator is below its critical value, entering into the region 2 in Figure 3.3.

3.3 Bodily Coupled CPGs through Local Sensory Information

We have looked at the dynamical properties of the basic CPG model and could draw the proper range of some of its parameters in terms of generating oscillatory dynamics for robot limb control as well as the applicability to the chaotic exploration process. Given the previously described CPG model and its parameters, we build a motor unit as an encapsulated modular component for robot control.

For practical use of the oscillators for the activation of muscles and to deal with the oscillator learning which will be addressed in later sections, we locate its centre of rotation at near the origin in the phase space such that $x \leftarrow (x + 0.2138)$ and $y \leftarrow (y - 0.7202)$. The centre of rotation was determined by averaging each pair of variables of a limit cycle for a sufficiently long duration. By introducing the sensory input and the coupling from other oscillators, a pair of oscillators in a motor unit m is expressed as follows:

$$\tau \dot{x}_l = c(x_l - \frac{x_l^3}{3} - y_l + z_1) + \delta F_x(s_l, x_l) \tag{3.9}$$

$$\tau \dot{y}_l = \frac{1}{c}(x_l - b y_l + a) + \varepsilon F_y(s_l, x_l) \tag{3.10}$$

$$\tau \dot{x}_r = c(x_r - \frac{x_r^3}{3} - y_r + z_2) + \delta F_x(s_r, x_r) \tag{3.11}$$

$$\tau \dot{y}_r = \frac{1}{c}(x_r - b y_r + a) + \varepsilon F_y(s_r, x_r) \tag{3.12}$$

where $l = 2m-1$ and $r = 2m$, which indicates each consecutive pair in the set of $2N$ oscillators are sequentially allocated to each motor unit (we will also use expressions such as ${}^m x_l$ and ${}^m x_r$ to refer to the mth motor unit where it avoids confusion).

To adjust the chaoticity of the motor unit, the difference of control parameters ($\mu = z_2 - z_1$) changes identically in all motor units as a function of the evaluation signal, which will act as the global bifurcation parameter for the chaotic search with adaptive feedback. The two constants δ=0.013 and ε=0.022 are the coupling strengths for afferent input functions $F(s, x)$. The input functions F_x and F_y are normally a linear function of the raw sensor signal s and the oscillator output x. Note that we need to preserve the topology of indirect couplings between oscillators close to that of a coupled BVP model (Figure 3.5) in equations 3.5-3.8, so their equations

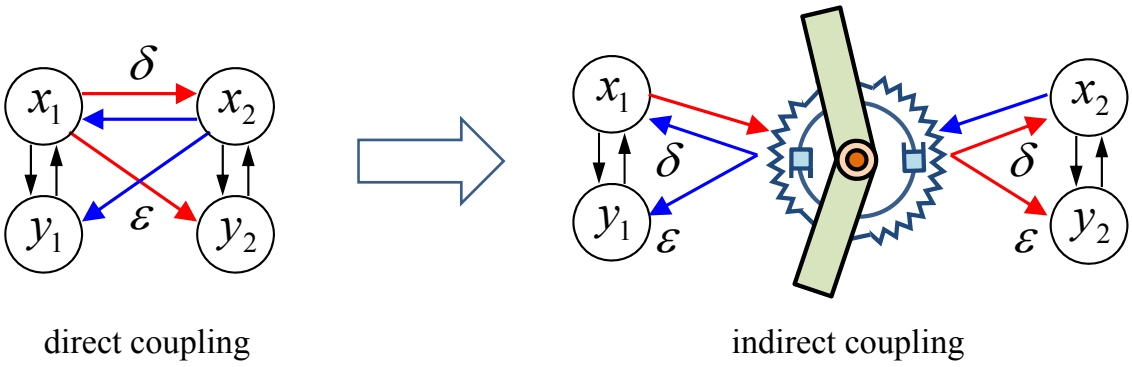

Figure 3.5: Coupling through physical embodiment. Indirect coupling via local sensor should have the same structure of information flow as that of direct coupling. Red and blue arrows show the equivalent information flow of direct and indirect coupling between two oscillators.

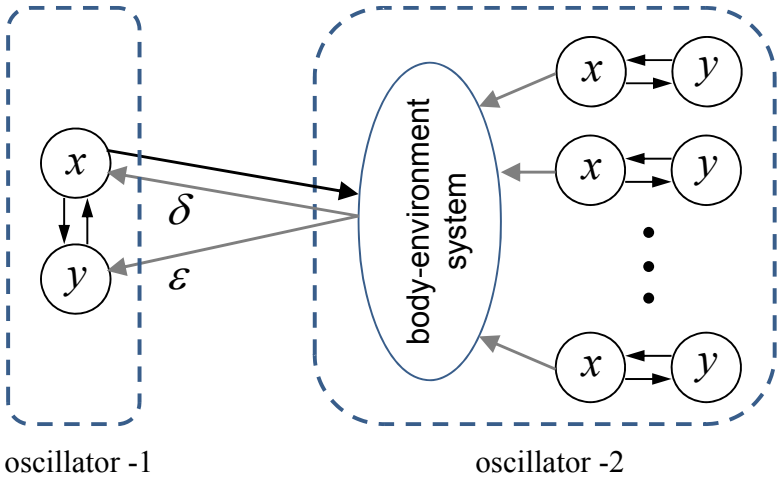

Figure 3.6: Conceptual illustration of the scale-free equivalence between a pair of coupled oscillators and the embodimentally coupled system of arbitrary size. Regardless of the system size, any oscillator can be seen as the first oscillator which receives the afferent input from (and gives output to) another (imaginary) oscillator whose output reflects the composite effect of neuro-body-environmental interaction.

may show slight variance according to sensor designs, which will be shown in later chapters. When the oscillators are coupled through physical embodiment via local sensory information, the whole system can be viewed as a pair of coupled oscillators, that is, any oscillator can be thought of as interacting with another imaginary oscillator which is a complex of the interaction among body, environment, and the rest of oscillators (Figure 3.6). This is useful for both determining the appropriate sensor function $F(s, x)$ and applying global bifurcation control to adjust the chaoticity of system by which an arbitrary neuro-body-environment system can be treated as two coupled oscillators whose parameters and dynamics are well-understood.

From the analyses of qualitative categorisation of two coupled BVP equations in the previous section, the stable regime of a pair of weakly coupled BVP oscillators where the two control parameters are symmetric exhibit bistable phase locking of their oscillations in a parameter range of $0.6 < z_1 = z_2 < 0.88$ (region 7 in Figure 3.3). In this range, observation of a number of experiments on the oscillator dynamics suggests that its medium point $z_2 = 0.73$ ensures a higher probability of multistability of the system in its stable regime. Although multiple strategies are possible to design the moving trajectory for z_1 and z_2 on the $\triangle z - z$ space in Figure 3.3, we will fix $z_2 = 0.73$ and vary z_1 using the global bifurcation parameter $\mu = z_2 - z_1$ throughout this work.

3.4 Evaluation and Feedback

The coherent integration of a performance evaluation signal that is able to control the chaoticity of the system is an essential feature of the exploration process. During exploration, the bifurcation parameter continuously drives the system between stable and chaotic regimes as a function of the evaluation signal. The evaluation signal is determined by a ratio of the actual performance (for instance, forward speed) to the desired performance. The desired performance can be set explicitly, or it can adaptively vary according to the actual performance. When the actual performance is low compared to the desired performance, the global bifurcation parameter increases to destabilise the system that the orbit wanders chaotically to escape from the low performing state. If the performance reaches the desired performance, the bifurcation parameter decreases to zero and the system stabilises.

3.4.1 Locomotion Performance and Feedback Bifurcation

In the experiments to be described throughout this work, the performance evaluation signal E is measured by the forward speed of the robot. Since the system has no prior knowledge of the body morphology of the robot, it does not have direct access to the direction of movement nor of information on body orientation. In order to facilitate steady movement in one direction without gyrating in a small radius, a temporal integration of the velocity of the centre of mass is formulated as an evaluation function. The centre of mass velocity of a robot is continuously averaged over a certain time window and its magnitude was used as the performance of the system. The performance signal E at any time instance can be calculated by applying a leaky integrator equation to the velocity vector as

$$E(t) = \|\bar{\mathbf{v}}\|, \quad \tau_E \frac{d\bar{\mathbf{v}}}{dt} = -\bar{\mathbf{v}} + \mathbf{v}. \tag{3.13}$$

τ_E is the time scale of integration which is larger than that of an oscillator, but typically not exceeding it by more than an order of magnitude.

A global feedback signal determines the degree of chaoticity of motor units. The bifurcation parameter for feedback control is continuously modified by an amount governed by the evaluation signal. If the current entrained state is not satisfactory, parameter μ is increased to where the orbit will follow quasiperiodic or chaotic dynamics, and when the performance increases, μ is decreased so that the system attempts to stabilise the orbit at the vicinity of the satisfactory mode. The time course of the bifurcation parameter μ $(= z2 - z1)$ is given as follows:

$$\tau_\mu \frac{d\mu}{dt} = -\mu + \mu_c G(\frac{E}{E_d}) \tag{3.14}$$

$$G(x) = 1/(1 + e^{16x-8}) \tag{3.15}$$

As described in the last section, z_2 (Equation 3.11) is fixed, hence z_1 (Equation 3.9) varies as μ changes. $G(x)$ is a function of the ratio of the actual performance E to the desired performance E_d (Figure 3.7). The function implements a decreasing sigmoid function which maps monotonically from (0,1) to (1,0). $16x - 8$ shapes the sigmoid function so that the boundary value at $x = 1$ and its derivative ($[dG(x)/dx]_{x=1}$) become almost 0 so as to make the function smoothly vanish to zero. We automatically set μ to zero when it falls below a small threshold ($\approx$0.0001) since the bifurcation

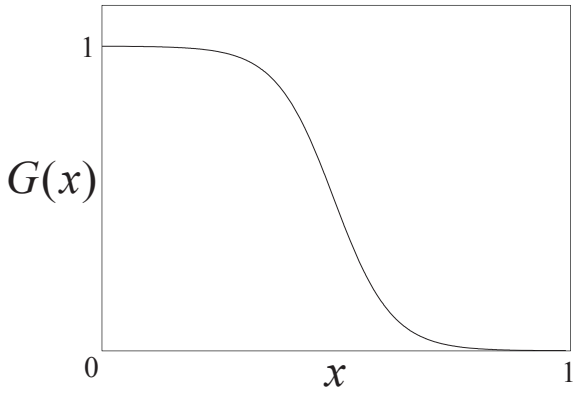

Figure 3.7: Sigmoid function $G(x)$.

parameter μ should be zero in order to make the system completely stable. This also allows some margin for the system to stay in the stable regime ($\mu = 0$) despite the small oscillation of E/Ed near unity.

τ_μ determines the time scale of the change of μ and is normally set faster ($\tau_\mu < T$) than the oscillation period (T) of the controller. If its value is too high, stabilisation of the system dynamics is significantly delayed which results in a partition mismatch (Aida and Davis, 1994). If it is too low, μ fluctuates too much according to the undulation of the robot movement which acts as a disturbance for stabilisation, or the system can become locked in a ring of undesirable patterns in a regime of intermediate chaoticity.

The level of chaoticity is adjusted by varying μ as a sigmoid function of E/E_d in the range [0,μ_c] where μ_c is the maximum level of chaoticity of the system. Momentum was given by using a first order differential equation to reduce the fast oscillation of μ due to the gyration of the robot body. Using the sigmoidal function not only provides a smooth boundary for the variation of μ but also has a uniformalising effect on the area occupied by the different dynamical regimes of the system (stable, quasiperiodic, and chaotic) which are defined by the level of μ. As illustrated by region 4 of Figure 3.3, the range of μ ($= \triangle z$ where $z_2 = 0.73$ in this work) exhibiting a quasiperiodic regime is much wider than those of other regimes. Operating μ linearly will diminish the performance of exploration by ensuring the system stays mostly in the quasiperiodic regime during the exploration process. The sigmoidal control enables μ to move fast in the middle (quasiperiodic) region and

slow in both the saturated regions (stable and chaotic) of the sigmoid function, so that the system can generate each type of dynamical regime in a less biased way.

From the analysis of a single BVP oscillator in Section 3.2 where the Hopf bifurcation occurs with an increase of the parameter z, it can be known that an analytically estimated critical (minimum) value of z is $z_c \approx 0.38247$, which indicates that the maximum possible value of μ_c is $\mu_{c,max} = 0.73 - 0.38247 = 0.34753$. However, because the situation is different from the dynamics of a single oscillator, the actual experiments on the robotic systems with indirectly coupled multiple oscillators reveal that the actual behavioural criticality of z in the integrated system varies slightly among different body and environmental settings. Determining the exact range of the variance of z_c in the general case requires testing for all possible robotic bodies. Even in the same robot body it also varies over different initial conditions. One way to determine the system specific criticality of the control parameter is to simply observe the dynamics of the system with fixed $\mu = \mu_c$. If the system is beyond its critical state, one of the oscillators in the motor unit will generate near zero amplitude by crossing the Hopf bifurcation point. Normally we chose μ_c to be slightly less than its maximum observed value, taking into consideration the saturating region of the sigmoidal function $G(x)$, so that it does not stay near the critical value for an unnecessarily long time when the oscillation amplitude becomes small.

3.4.2 Adaptive Goal Setting Strategy

E_d indicates the desired locomotion performance of a given robot. Since the robotic system is arbitrary, we do not have prior knowledge of what level of performance it can achieve. Hence E_d is modelled as another dependent variable which varies adaptively according to the system performance. Using the concepts of a goal setting strategy (Barlas and Yasarcan, 2006) and the Rescorla-Wagner model (Rescorla and Wagner, 1972), the dynamics of the desired performance are modelled as a temporal average of the actual performance, such that the expectation of a desired goal is influenced by the history of the actual performance experienced in the past. We can express this strategy in terms of simple continuous dynamics for E_d, which slowly decays toward the current performance:

$$\tau_d \frac{dE_d}{dt} = -E_d + E \tag{3.16}$$

where τ_d is set sufficiently large so that E_d does not follow E too fast. Since E_d continuously decays toward E, the changing speed of control parameter μ depends both on τ_μ and τ_d.

Although this evaluation strategy does not explicitly impose a bias for continuously striving for higher performing behaviours (because of the dynamics of E_d), an implicit bias towards better performing behaviours is partially imposed on the system by the way in which the bifurcation parameter μ behaves as a function of E/Ed (Equation 3.14). Once the system has been stabilised to some behaviour, the speed of system destabilisation, for a given amount of behaviour degeneration, depends on the performance level of the initially stabilised behaviour. In the quasiperiodic regime which occupies a large portion of the entire system dynamics (μ in the lower saturation part near zero and middle part of the sigmoid function, $G(x)$) the phase relationships of ongoing patterns shifts slowly, while fast and catastrophic change occurs in the chaotic regime where μ is located around the upper saturation part (near μ_c) of $G(x)$. When the actual performance E of a stabilised behaviour decreases by a given amount, a low performing behaviour is destroyed more quickly, because E_d will be relatively small, while a high performing behaviour is smoothly degenerated, giving it much more chance of being sustained or re-entrained to itself. In this way, in practice, the system fully stabilises onto behaviours that exhibit stable relatively high performance.

3.5 Robot Simulation

This section describes physically simulated limbed robots and their environmental forces. The following sections start with a brief description of the articulated rigid body simulation and its basic settings, followed by a general framework for a neuromuscular system which incorporates a simple spinal reflex. Detailed description of its subcomponents including muscle and proprioceptor models are described. In order to investigate the generality of the exploration system coping with an arbitrary morphology under unknown physical environment, two kinds of robotic simulation models for different environment are developed; a 2D aquatic swimmer and a 3D terrestrial walker. Generic simulation models for robotic systems in each environment are presented.

3.5.1 Articulated Robot Simulation

A robot is simulated using an articulated rigid body dynamics. The problem of simulating articulated rigid bodies is a long-standing problem in its own right in that a good simulation algorithms need to meet multiple conflicting demands such as speed, stability and accuracy. There are several commercial/free application program interfaces (APIs) for physics simulation which are used for a variety of purposes from computer animation to industrial use. Because it exhibits a good compromise between accuracy and speed, this work has chosen Open Dynamics Engine (ODE) developed by Russell Smith (Smith, 1998), an open source middleware physics simulation library. It is platform independent and relatively stable compared to other simulators. The numerical integrator used in ODE is a first order semi-implicit integrator (Stewart and Trinkle, 1996), where the constraint forces are implicit, and the external forces are explicit. A medium speed ($O(n^3)$, where n is the number of links) Lagrange multiplier method (Anitescu and Potra, 1997) is used to calculate the articulated body dynamics. Refer http://ode.org/ode.html for more detail. While the joint constraints for linking limbs of robot and the ground contact with friction are provided by physics engine, the actuation of limbs by simulated muscles and the hydrodynamic forces exerted on the robot should be explicitly modelled as will be presented later.

3.5.2 Neuromuscular System

The neuromuscular model for a single joint controlled by the corresponding motor unit comprises a pair of nonlinear torsional muscles and a neural reflex loop (Figure 3.8). The neural reflex loop includes an α motor neuron and muscle proprioceptors which are a muscle spindle and tension sensor.

3.5.2.1 Muscle Model

A muscle is activated by the (α) motoneuron and produces forces that cause movements of the body. We employed a simple yet biologically relevant actuation model proposed by (Ekeberg, 1993; Wadden and Ekeberg, 1998). In this model the muscles are located on the two (opposite) sides of the linked bodies with their axes of contraction being mainly perpendicular to the joint axis. It is assumed that the length

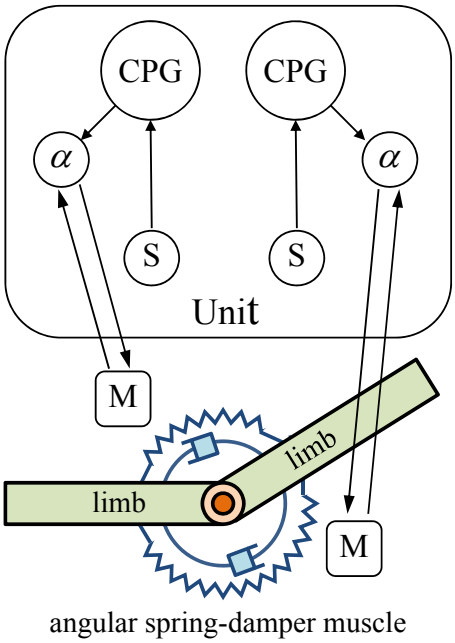

Figure 3.8: Neuromuscular system with motor unit. A pair of torsional muscle is modelled as an angular damped spring and controlled by each α motor neuron from the corresponding motor unit (M). Each muscle gives a proprioceptive feedback to the homonymous motor neuron. Sensor input (S) to CPGs is not necessarily a proprioceptive feedback, but it rather depends on the sensor design (for example, angle sensor or touch sensor).

of the muscle fibers varies linearly with the local curvature of the body. Therefore, this model can be regarded as a torsional muscle whose stretch length is expressed as bending angle as described in Figure 3.8. Accordingly, the *force* produced by muscle corresponds to the torque on the relevant joint. These muscles are modelled by including both an elastic and a viscous component. The elastic component can be viewed as an angular spring where the spring constant is used to set the torque produced in a steady-state condition (in its resting angle). The viscous component of the force is proportional to the angular speed of the movement.

The motoneuron output linearly controls the forces generated by the muscles, and this corresponds to a linear relationship between motoneuronal activity and the muscular spring constant. The linear relationships were preserved by modelling a

muscle as a damped torsional spring making it possible to express the torque acting at a particular joint as a linear function of the two motoneuronal activities in motor unit. The torque exerted on a joint by a pair of symmetric muscles is

$$T = \alpha(\sigma_l - \sigma_r) - \beta(\sigma_l + \sigma_r + \gamma)\theta - \delta\dot{\theta}, \tag{3.17}$$

where σ_l and σ_r are the output signals from the two motoneurons and θ is the joint angle. α is the muscle activation gain, β is the stiffness gain, γ is the tonic stiffness, and δ is the damping coefficient. Although they can be chosen arbitrarily from a range of values, we set those parameters in order that the musculo-skeletal system at rest behaves close to a critically damped system. The above arrangement enables the neural signal not only to control the muscle torque but also to control the stiffness of the joint.

3.5.2.2 Reflex Loop

Each motoneuronal output σ is contributed by CPG ouput (x) and a simplified muscle stretch reflex (s_m) according to the following canonical formulae based on the literature (Prochazka, 1999; Yakovenko et al., 2004).

$$\sigma = \tanh(x + k_m s_m) \tag{3.18}$$

$$s_m = \theta/\Phi + R(\dot{\theta}/\Phi) \tag{3.19}$$

$$R(v) = sgn(v)\sqrt{|v|} \tag{3.20}$$

$k_m = 0.2$ is a constant, and Φ is a denominator which normalises the angle and the angular velocity of torsional muscle by the unit of its resting angle, and it is set to the maximum available range of the muscle, by assuming that the angle of the torsional muscle is stretched twice as much as its resting angle when the joint is at its neutral position.

Although several types of proprioceptive feedback mechanisms including Group Ia, Ib, II, and cutaneous afferents operate on the spinal reflex system and their collective interaction accounts for the regulation of ongoing locomotor activities (Grillner and Wallen, 1985; Hiebert and Pearson, 1999; Pearson et al., 2006; Rossignol et al., 2006), it would be sufficient to support the mechanical stability of musculoskeletal system using a minimal model for a basic reflex loop (only with muscle spindle) since the Group Ia pathway is the most sensitive of all. From the viewpoint

of the global system, even the muscular-motoneuronal reflex loop can be broadly considered as a part of intact anatomical properties that may vary across the different robotic designs which should be covered by the exploration process.

The next sections will introduce two class of robot architectures both of which are controlled by the above basic neuromuscular system. Class specific modelling of the mechanical system and the interaction with environment will be presented in detail.

3.5.3 2D Aquatic Swimmers

This section describes the general model of an aquatic swimmer robot moving on a 2D hydrodynamic environment. The swimmer robot consists of a central body (trunk) and multiple limbs attached to it. The actual model was constructed using a 3D rigid body simulator, but it was constrained to move only on the x-y plane, so that it effectively undergoes 2D dynamics (Figure 3.9). Each limb has a rectangular-shaped rigid 'fin' which is located at its end in order to interact with hydrodynamic environment. Fins are joined by a nonlinear torsional spring with damper and the hydrodynamic forces are exerted exclusively on them, making the robot generate propulsion by beating its limbs. The use of passive fins under a fluidic environment is intended to produce smooth and continual reaction forces in order to provide continuous sinusoidal sensor signals to the neural system, which is advantageous both for the neural synchrony through embodimental coupling and the tractability of model analysis.

With this basic architecture one can build a 2D swimmer with an arbitrary number of limbs. The preferred shape of a swimmer robot in the context of the exploration system is radially symmetric. The radially symmetric shape in a 2D underwater environment is interesting because it makes generating continuous asymmetric propulsion forces challenging; in other words forward locomotion is non-trivial. The agent will not be able to move in a single direction unless the movements of all four arms are successfully synchronised with appropriate phase differences.

3.5.3.1 Passive Fin Model

A nonlinear damped torsional spring for passive fin bending is simulated using a second order differential equation. The spring is tuned to be critically damped so

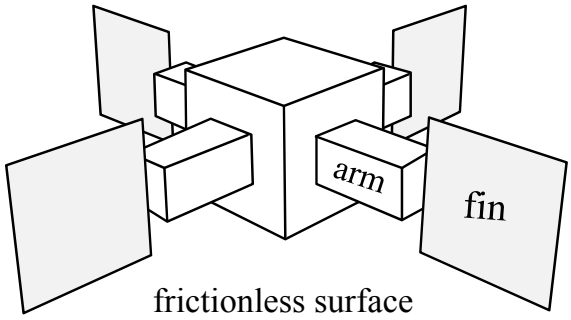

Figure 3.9: An example of 2D aquatic swimmer which is implemented using 3D physical simulator.

that the bend angle smoothly moves toward the equilibrium position between the hydrodynamic torque and the spring torque. At each time step, the bend angle and the angular speed of the fin receiving hydrodynamic force (Figure 3.10A) can be calculated by the relationships:

$$\mathbf{T} = \mathbf{r} \times \mathbf{F}_N \tag{3.21}$$

$$T = |\mathbf{T}| = I\ddot{\phi} + c\dot{\phi} + k\phi. \tag{3.22}$$

$\mathbf{F}_N$ is the component of net hydrodynamic force in the direction of fin surface normal. Thus the torque exerted to the fin joint can be calculated by the cross product of the vector $\mathbf{r}$ and $\mathbf{F}_N$. The relationship between damping constant c, spring constant k, and the moment of inertia of the fin is set to $c = 2\sqrt{Ik}$ so that the free fin movement (when not driven by T) shows critical damped motion.

The net hydrodynamic force exerted on the centre of mass of the fin are then calculated from the velocity of the centre of mass of the moving fin. Figure 3.10B shows the linear and angular velocities of the limb-fin movement. Given that the linear and angular velocity of the limb at any time instance is $\mathbf{v}_l$ and $\boldsymbol{\omega}_l$ and the rate of the bending angle of fin is $\boldsymbol{\omega}_f$, the velocity of the centre of mass of fin ($\mathbf{v}$) can be written as:

$$\mathbf{v} = \mathbf{v}_u - \mathbf{v}_f \tag{3.23}$$

$$\mathbf{v}_u = -\{\mathbf{v}_l + (\boldsymbol{\omega}_l \times \mathbf{R})\} \tag{3.24}$$

$$\mathbf{v}_f = \boldsymbol{\omega}_f \times \mathbf{r} \tag{3.25}$$

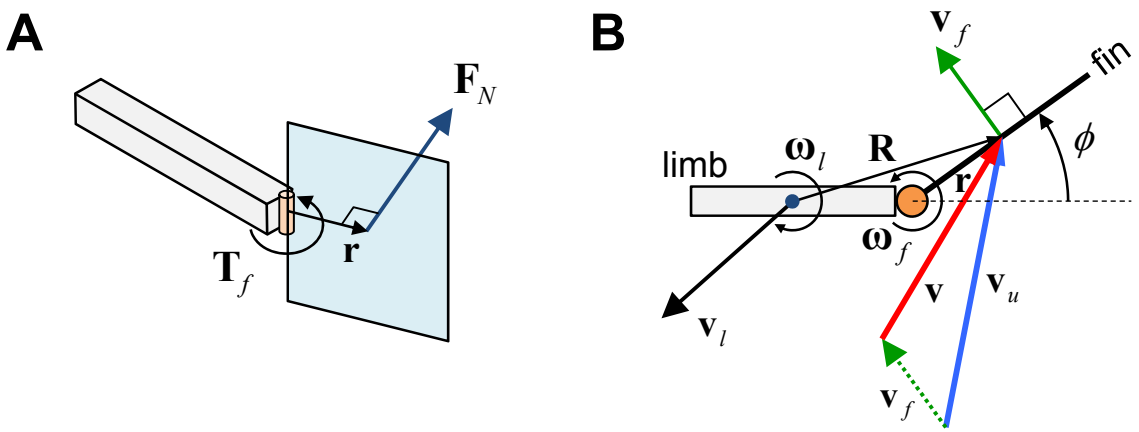

Figure 3.10: Passive fin dynamics. (A) The fin is attached at the end of a limb joined by a damped angular spring. The bending angle of fin depends on the torque exerted to the spring which is produced by the hydrodynamic force. (B) At each time step, the velocity vector of fluid stream ($\mathbf{v}$) is calculated from the linear and angular velocities of the limb and fin (see text for details).

where $\mathbf{R}$ is the vector from the centre of mass of limb to the centre of mass of fin, and $\mathbf{r}$ is the vector pointing the centre of mass of fin from the fin joint. With the incoming stream velocity $\mathbf{v}$ is known, we can calculate resultant hydrodynamic force on the fin using the equations which will be described in next section. The new bend angle and its angular speed from previous bending torque are used again as the parts of a parameter set for calculating new hydrodynamic forces and bending torque in a circular manner.

3.5.3.2 Hydrodynamic Forces

Each fin of the swimmer robot receives hydrodynamic forces which are calculated by simplified aerodynamics. The forces acting on a fin surface depend on its area and the angle of attack with respect to the velocity of the stream of fluid. Assuming that the flow is laminar which has low Reynolds number, it is reasonable to use a conventional aerodynamics to derive hydrodynamic forces acting on an immersed flat plate. The calculation of forces is based on the blade element theory (BET) with a quasi-steady assumption (Weis-Fogh and Jensen, 1956) as used in the simplified

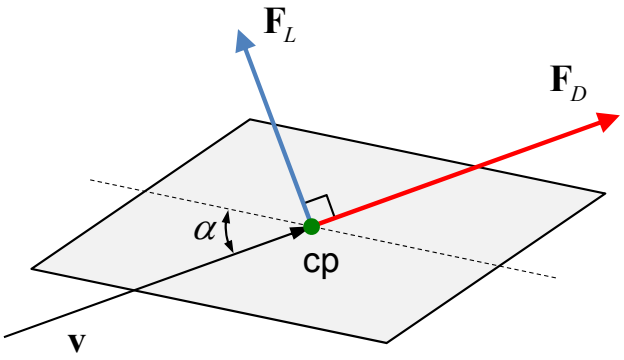

Figure 3.11: Hydrodynamic forces on a surface. All forces act on the centre of pressure (cp) of the surface. The direction of drag ($\mathbf{F}_D$) is colinear with the incoming stream velocity ($\mathbf{v}$), while the lift ($\mathbf{F}_L$) direction is perpendicular to that of drag.

aerodynamics. The calculated forces are considered to be exerted on the centre of mass of each fin surface. The stream velocity for each surface is calculated by inverting the sign of the vector representing the velocity of the centre of mass of the surface.

The lift and drag forces can then be calculated using the lift coefficient C_L and the drag coefficient C_D. The lift coefficient is proportional to the angle of incidence (α) of the surface and, from thin airfoil theory (Goldstein, 1942), its slope is 2π for a 2D wing of infinite span. However, Prandtl's lifting-line theory (Prandtl, 1918a,b) shows that for a finite-span, the slope depends on the aspect ratio of the wing because of the effect of induced velocity. In Prandtl's model, the drag coefficient C_D is composed of induced and parasite components, and yields an airfoil polar equation of the form

$$C_D = C_{D_0} + \frac{C_L^2}{\pi\varepsilon(AR)} \tag{3.26}$$

where the lift coefficient is expressed as

$$C_L = 2\pi \left(\frac{AR}{AR+2} \right) \alpha. \tag{3.27}$$

C_{D_0} is the parasite drag coefficient due to viscosity, and is assumed to be 0.05, ε is the Munk span efficiency, which is normally slightly less than 1 (this work set it to 1), and AR is the aspect ratio of surface.

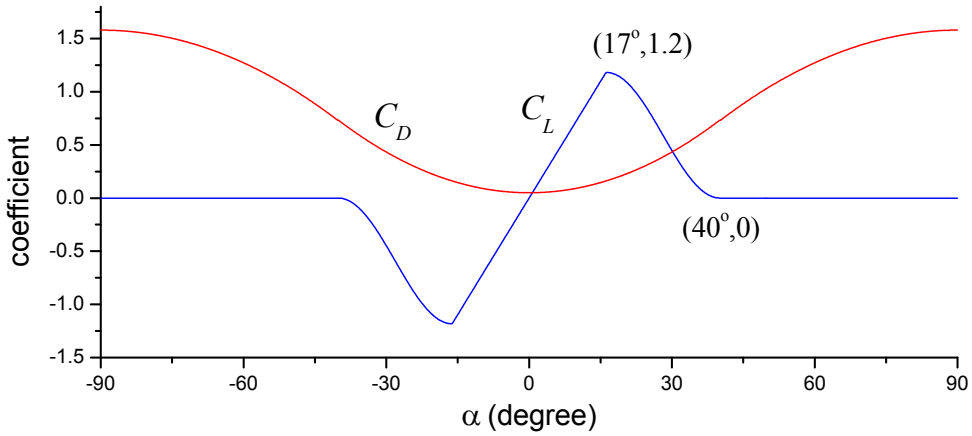

Figure 3.12: Hydrodynamic coefficients. Lift (C_L) and drag (C_D) coefficients are depicted as a function of angle of attack (α).

However, the lift cannot be very large because flow separates and the wing stalls at high angles of attack. In reality, if the angle of attack increases and passes the stall angle of attack (approximately at the angle somewhere between 15° and 17°), at one point all lift will be lost while the drag continues to increase. However this discontinuity may cause the numerical simulation to become unstable. In order to avoid sudden change, the shape of the lift coefficients is modelled as shown in Figure 3.12, by synthesising functions to return lift coefficients for any given angle (Shim and Kim, 2006). The drag coefficient is generated over the 'active' range (±40°) of the lift coefficient, and gradually flattened near ±90 by quadratic functions.

Drag acts in the direction of the airstream velocity. The direction of lift is perpendicular to that of the drag, so a patch produces not only lift but also thrust tangential to its surface (Figure 3.11). As the lift and drag coefficient for a surface at a specified angle of attack is known, the lift and drag forces produced for specific flow conditions can be determined using the following vector equation (Anderson, 2004):

$$\mathbf{F}_L = \frac{1}{2}\rho C_L(\alpha) A |\mathbf{v}|^2 \mathbf{e}_L \tag{3.28}$$

$$\mathbf{F}_D = \frac{1}{2}\rho C_D(\alpha) A |\mathbf{v}|^2 \mathbf{e}_D \tag{3.29}$$

where A is the area of surface, and ρ is the density of the fluid. $\mathbf{v}$ is the stream

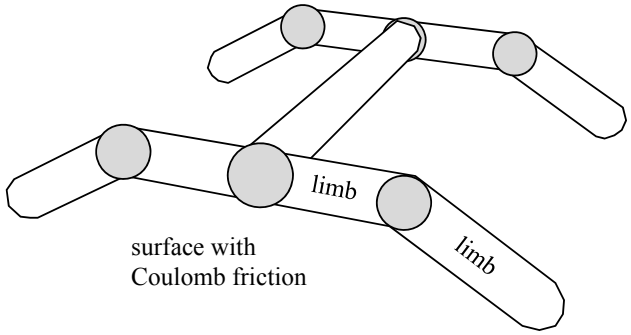

Figure 3.13: An example of 3D terrestrial walker.

velocity calculated from Equation 3.23. $\mathbf{e}_L$ and $\mathbf{e}_D$ are the unit vectors in the direction of the lift and the drag respectively. After the forces on each fin have been calculated, they are transformed to the force and the torque acting on the corresponding arms of the swimmer robot and are sent to the simulator as the environmental forces.

3.5.4 3D Terrestrial Walkers

Walking robots are placed in 3D terrestrial environment. Ground friction is simulated by the Coulomb friction model provided by the ODE engine. The Coulomb friction model defines a simple relationship between the normal and tangential forces present at a contact point.

$$|\mathbf{F}_T| \leq \mu |\mathbf{F}_N| \tag{3.30}$$

$\mathbf{F}_N$ and $\mathbf{F}_T$ are the normal and tangential force vectors respectively, and μ is the friction coefficient. This relationship defines a 'friction cone'; a cone with $\mathbf{F}_T$ as the axis and the contact point as the vertex. If the total friction force vector is within the cone then the contact point is influenced by static friction force, where the friction force is enough to prevent the contacting surfaces from moving with respect to each other. If the force vector is on the surface of the cone then the contact is in dynamic (kinetic) friction mode, and the friction force is typically not large enough to prevent the contacting surfaces from sliding. The parameter mu thus specifies the maximum ratio of tangential to normal force.

ODE approximates the friction cone by a friction pyramid aligned with the first and second friction directions which are automatically calculated by the simulator at every timestep. First it computes the normal forces assuming that all the contacts are frictionless. Then it computes the maximum limits F_m for the tangential forces from $F_m = \mu|\mathbf{F}_N|$ and then applies the force to the contact point within this limit.

A leg of robot may have various shapes, but for less simulation time, the capped cylinder shape was used in all legs (Figure 3.13). A capped cylinder is like a normal cylinder except it has half-sphere caps at its ends. This feature makes the internal collision detection code particularly fast and accurate because the leg of this shape produces only a single contact point at a time.

Chapter 4

Experiments with the Basic Exploration System

4.1 Introduction

In this chapter, we apply the previously described methods to a simulated robot and investigate the behaviour of our exploration system. The initial experiment involves integrating the previously modelled exploration system with a simple 4-fin aquatic swimmer placed in a simulated 2D hydrodynamic environment, and the analyses of the system behaviours are presented mainly in this setting. Firstly, the behaviours of a fully embodied system in its chaotic regime is investigated with different global coupling parameter which determines the strength of embodimental coupling between neural elements. It is essential for the system dynamics to have a well-structured dynamical range from stable to chaotic patterns. Since the chaoticity of system should be strong enough to satisfactorily destabilise the ongoing synchronisation between subparts, Lyapunov analysis is performed to characterise the chaotic dynamics of the system. Secondly, the permanently stable movement patterns (behaviours) existing in the stable regime of the system are identified and categorised. These existing stable behaviours are the targets to be searched by the exploration process, each of which is a possible candidate for efficient locomotion behaviour. Based on those target behaviours, the chaotic exploration is tested, and the statistics of the appearance of each pattern is investigated, followed by a report on the exploration deficiencies which are unwanted system behaviours to be addressed.

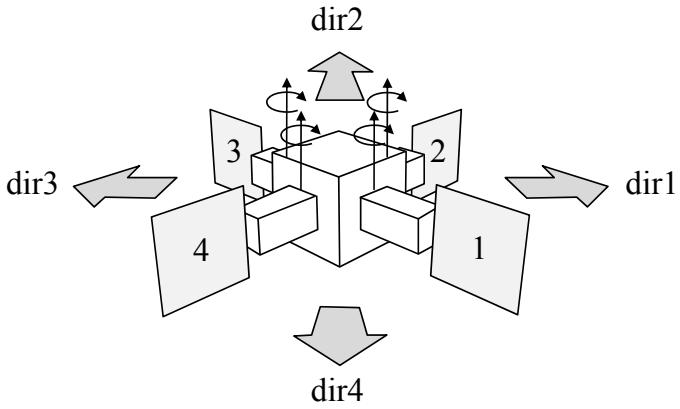

Figure 4.1: 4-Fin Swimmer robot in 2D hydrodynamic environment. Four body orientations are shown as direction 1-4.

4.2 4-Fin Swimmer Preparation

Initial experiments with the framework described in the previous chapter used the 4-Fin Swimmer (Figure 4.1) which has four fins, each at the end of a separate arm, placed in a simulated hydrodynamic planar (2D) environment. Each arm is attached to the torso by a hinge joint and a pair of torsional muscles actuates the arm as the neuromuscular system presented in section 3.5.2. The detailed physical parameters of 4-Fin Swimmer are presented in Table A.1 Each muscle pair is controlled by a motor unit, so there are four motor units (eight CPGs in total). The robot's radially symmetric shape in a 2D underwater environment is interesting because it makes generating continuous asymmetric propulsion forces challenging; forward locomotion is non-trivial. The agent will not be able to move in a single direction unless the movements of all four arms are successfully coordinated with appropriate phase differences.

The functional structure of embodimental coupling between motor units is formed by the transmission of hydraulic reaction forces from one arm to the others through articulation of the body. Since the information transfer between CPGs is mediated by sensory information, the information structure of physical embodiment is considerably influenced by the design and choice of sensory systems. While it is possible to use composite sensory information from multiple sensors (e.g. the linear combination of the input from fin sensor and muscle receptor), we use only a single fin

angle sensor for a motor unit for simplicity. The reaction forces are sensed by the fin bending angle (ϕ) and each sensor signal is fed to the corresponding motor unit. The fin angle implements the stretch receptor at each side of fin, so the afferent input s in Equations 3.9 and 3.11 were defined as: $s_l = k\phi$ and $s_r = -k\phi$ where k is sensor gain.

This results in only a single sensor signal with opposite sign provided for a couple of CPGs in a motor unit. Assuming that a fin angle sensor reflects the output difference of the oscillator pair in the corresponding motor unit (i.e. $s_{l,r} = f(x_{r,l} - x_{l,r})$ where $s_r = -s_l$), let us take the simplest form of the sensor functions F_x and F_y in equations 3.9 and 3.10. From equations 3.5 and 3.6 we mimic the functional coupling structure of a pair of coupled BVP model by denoting $s_{l,r} = x_{r,l} - x_{l,r}$. Then the sensor functions are written as

$$F_x(s_{l,r}, x_{l,r}) = s_{l,r} \tag{4.1}$$
$$F_y(s_{l,r}, x_{l,r}) = s_{l,r} + x_{l,r}, \tag{4.2}$$

which is equivalent of replacing the coupling terms in equations 3.5 and 3.6 by $\delta(x_2 - x_1) = \delta s$ and $\varepsilon x_2 = \varepsilon(s + x_1)$. Thus, together with the global bifurcation parameter μ, the reformulated CPG equations for a motor unit of 4-Fin Swimmer can be written as:

$$\tau \dot{x}_l = c(x_l - \frac{x_l^3}{3} - y_l + z - \mu) + \delta s_l \tag{4.3}$$
$$\tau \dot{y}_l = \frac{1}{c}(x_l - by_l + a) + \varepsilon(s_l + x_l) \tag{4.4}$$
$$\tau \dot{x}_r = c(x_r - \frac{x_r^3}{3} - y_r + z) + \delta s_r \tag{4.5}$$
$$\tau \dot{y}_r = \frac{1}{c}(x_r - by_r + a) + \varepsilon(s_r + x_r). \tag{4.6}$$

The time constant used was $\tau = 0.8$ and all other CPG parameters are as defined in section 3.2. The control input parameter z was set to 0.73. For convenience, we will use the notation ${}^i x_l$ and ${}^i x_r$ to express the anogist (left: l) or the antagonist (right: r) CPG output of the ith motor unit. In the 4-Fin swimmer robot we have four motor units; $i = 1, 2, 3, 4$.

4.3 Global Strength of Embodimental Coupling

While the sensory coupling constants δ and ε in equations 4.3-4.6 are presented, these parameters originate from the model of directly connected oscillators. In our neuro-robotic system where the CPGs are indirectly coupled by local sensory information, these parameters only define the ratio of input strengths for the first and second variables of an oscillator (as shown in Figure 3.5). Since the robotic system and its physical environment are arbitrary, the actual intensity of afferent input is highly non-stationary and it depends on the overall effect from the body-environmental interaction. With the parameters preliminarily given from the previous chapter, let us investigate how the sensory input strength affects the overall system behaviour. The sensory input signal is typically scaled by an appropriate gain before it is fed to the controller. This sensor gain determines the overall strength of the coupling through physical embodiment, which is analogous to the global coupling strength in the coupled oscillator network model.

4.3.1 Effect of Global Coupling Parameter in the Chaotic Regime

One of the key issues of the behaviour of the embodimentally coupled oscillator network in our exploration system is whether the system is able to exhibit satisfactory behaviour spanning from stable to chaotic dynamics. The neuro-body-environment system should realise a weakly coupled interaction between CPGs where its mode of dynamics is sufficiently flexible while not losing coherency among individual parts. In fact, this is already the case in the previously introduced two coupled BVP model, which will be the reference for our system. As it is previously presented in section 3.3 (Figure 3.6), the neuro-body-environment system having arbitrary degrees of freedom can be equivalently treated as a pair of coupled oscillators such that every single oscillator views the rest of system as another (imaginary) oscillator. Thus the proper intensity of sensor signals for a CPG is expected to be same as that of the output of a single BVP oscillator.

In order to investigate how the system behaves in its chaotic regime, a test was performed by setting the global bifurcation parameter μ to be fixed at some value in the chaotic regime. The bifurcation parameter was set to $\mu = 0.335$ where the

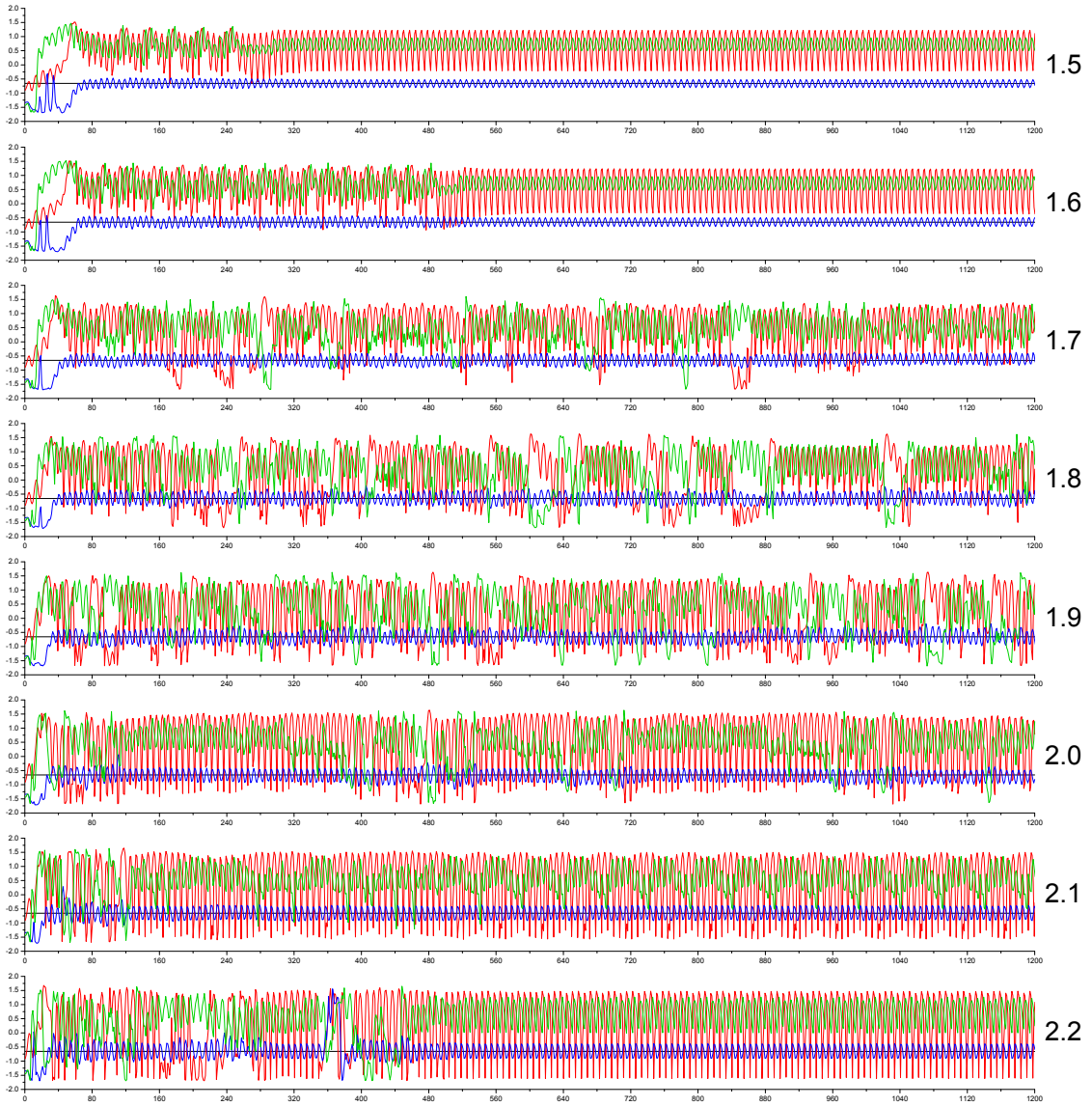

Figure 4.2: Neural output vs. time (oscillation cycle) for different global coupling parameters (g=[1.5,2.2]). The agonist signals of three motor units (^2x_l:red, ^3x_l:blue, ^4x_l:green) are plotted whenever ^1x_l crosses its singular point (i.e. unstable equilibrium point, shown as black line) of the periodic orbit of CPG.

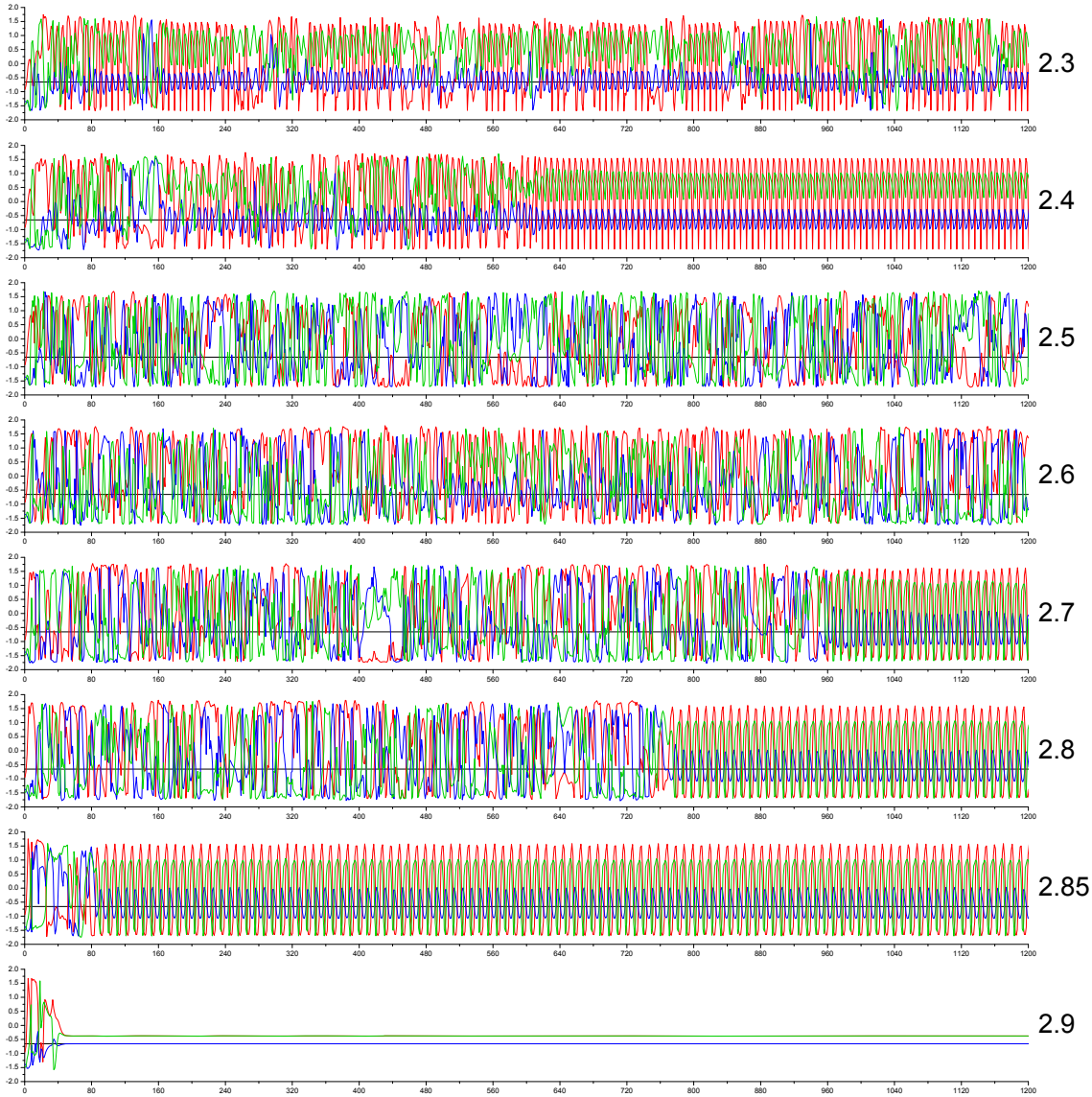

Figure 4.3: Time plot of neural output for different global coupling parameters (g=[2.3,2.9]).

two coupled BVP models exhibits the highest chaoticity (the highest maximal lyapunov exponent) as presented in Figure 3.4, and the system behaviour was observed for different sensor gain k. As it will be described later in this chapter, we will address the processing of raw sensor signal by an adaptive sensor mechanism. The adaptive sensor mechanism is designed to give continuously varying output values in an unknown range (which can have both positive and negative signs) by numerically integrating a set of ordinary differential equations, which will determine the sensor gain. In order to make sure that the gain always has positive sign and varies smoothly, let us denote the sensor gain as $k = e^g$ where g is the output from the adaptive sensor mechanism. We will use g as the global coupling parameter. Initial investigation was made by visually observing the complicatedness of the neural signal of embodied system for different coupling parameter g. Figure 4.2 and 4.3 depict the realtime plot of Poincare-like map. The graph was generated by plotting the output of selected CPGs at particular instant of time so as to best visualise the time varying dynamics of the system. The agonist signals of three motor units (${}^2x_l$, ${}^3x_l$, ${}^4x_l$) are plotted whenever ${}^1x_l$ crosses its singular point (i.e. unstable equilibrium point) of the periodic orbit of CPG. The coupling parameter g was increased by 0.1 in the range $1.5 \leq g \leq 2.9$. Each behaviour was depicted approximately for 1200 oscillator cycles.

In our neuro-body-environment system, observation suggests that the result can be roughly classified into three consecutive stages that show qualitatively distinct behaviours. Different from the globally coupled map (GCM) model whose individual elements exhibit chaotic dynamics regardless of the existence of coupling, our embodied system consists of limit cycle oscillators, that generate chaoticity only when connected to each other with appropriate coupling strength. For the weak coupling strength (small gain; $1.5 \leq g < 1.7$ in Figure 4.2) the system exhibits synchronised movement instead of chaotic motion. The output of CPGs initially fluctuates in the early stage of simulation, then eventually all elements are synchronised with a certain phase relationship which is guided by the given neuro-body-environmental interaction. The stabilised motion exhibits the stable oscillation of multiple periods.

As the coupling strength increases in this regime, the duration of initial fluctuations increases, which indicates that 'coupled chaoticity' starts to take effect. In the range of $1.7 \leq g < 2.7$ (Figure 4.2 and 4.3), both chaotic and synchronised motions are observed together. While weak or moderate chaotic dynamics dominates throughout

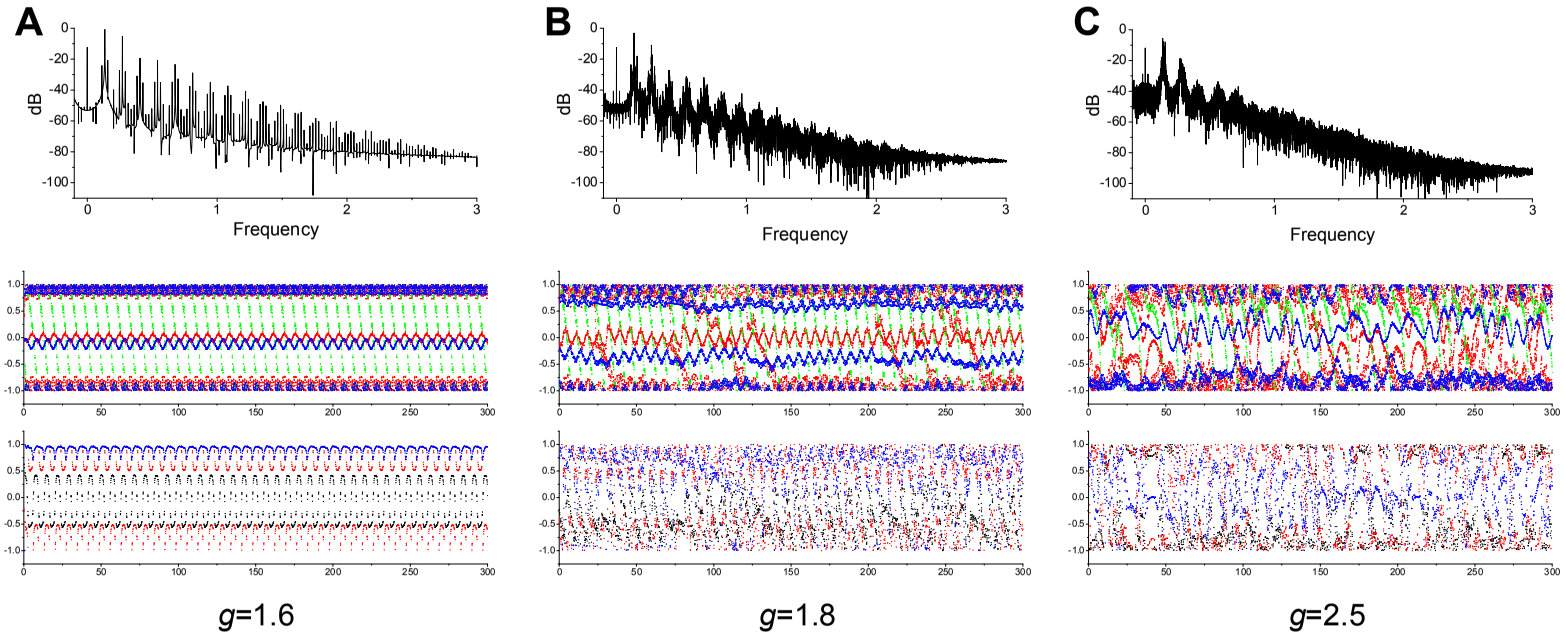

Figure 4.4: Frequency spectrums and phase relationships (πrad) vs. time (oscillator cycle) in chaotic regime for selected gain g, where (A) $g = 1.6$, (B) $g = 1.8$, and (C) $g = 2.5$. Plots in second and third rows depict the phase relationships of neural output and arm movements respectively. For neural phase relationship, each colour represents the phase differences for three groups without specifying individual members, where red plots are for the phase difference between left CPG outputs; $({}^1x_l - {}^2x_l)$, $({}^1x_l - {}^3x_l)$, $({}^1x_l - {}^4x_l)$, blue for the difference between right CPG outputs $({}^1x_r - {}^2x_r)$, $({}^1x_r - {}^3x_r)$, $({}^1x_r - {}^4x_r)$, and green for the difference between the left-right ouput of motor unit 1; $({}^1x_l - {}^1x_r)$. For arm movements, black for (arm1-arm2), red (1-3), and blue (1-4).

the range of gain parameter, some parameter regime showed a few distinct synchronised patterns intermittently. Because the presented body-environmental interaction is formed by the symmetric body shape emersed in the dense hydrodynamic force field, most of the synchronised patterns commonly exhibits strong coherence between the diagonal CPGs, that is, ${}^1x_l$ and ${}^1x_3$ tend to move nearly in-phase and ${}^1x_l$ and the rest move anti-phase. This tendency is dominant in the weak chaotic regime, and the frequency of escaping from this tendency increases in more chaotic regime. This can be observed in Figure 4.4, where the phase relationships between agonist (${}^ix_l$) CPG ouputs (red plots of the graphs in first column) show occasional or frequent changes between 0 and 180 degrees (Figure 4.4B,C). However, the stability of this tendency was so strong that the escaped state quickly re-attracted to it within a hundred cycles. Phase differences between antagonistic (${}^ix_r$, blue) CPGs rarely exhibit significant exchange of phase relationship while the fluctuation of phase differences increases with increased chaoticity. It is because they have identical control input z, indicating the change of phase relationship between arm movements is mostly influenced by agonist CPGs.

If coupling strength increases beyond $g = 2.7$, the dynamics begins to synchronise again into another multi-period oscillation. In contrast to to the regime of weak coupling strength, the duration of initial fluctuation decreases as the coupling strength increases. Finally, when g is over 2.9, the dynamics are tightly locked to the periodic orbit of period one which is a completely stable form of the diagonal in-phase motion, that we call 'bound antiphase' (BA), which is the most stable pattern emerged from the 4-Fin swimmer system.

4.3.2 Analysis with Lyapunov Exponents

Quantification of the above observation was performed by calculating the maximal Lyapunov exponent which is the indication of chaos. The analysis was done by estimating the maximal Lyapunov exponents from two time series corresponding to arm-1 of the robot, which are the neural output (${}^1x_l$) and the arm movement angle. For both time series, 10^5 data points are sampled at 40Hz ($\triangle t$=0.025 sec), and these time series data were generated for different g. The state space reconstruction was done by obtaining appropriate embedding delay and embedding dimension using the average mutual information (AMI) and the false nearest neighbours (FNN) methods.

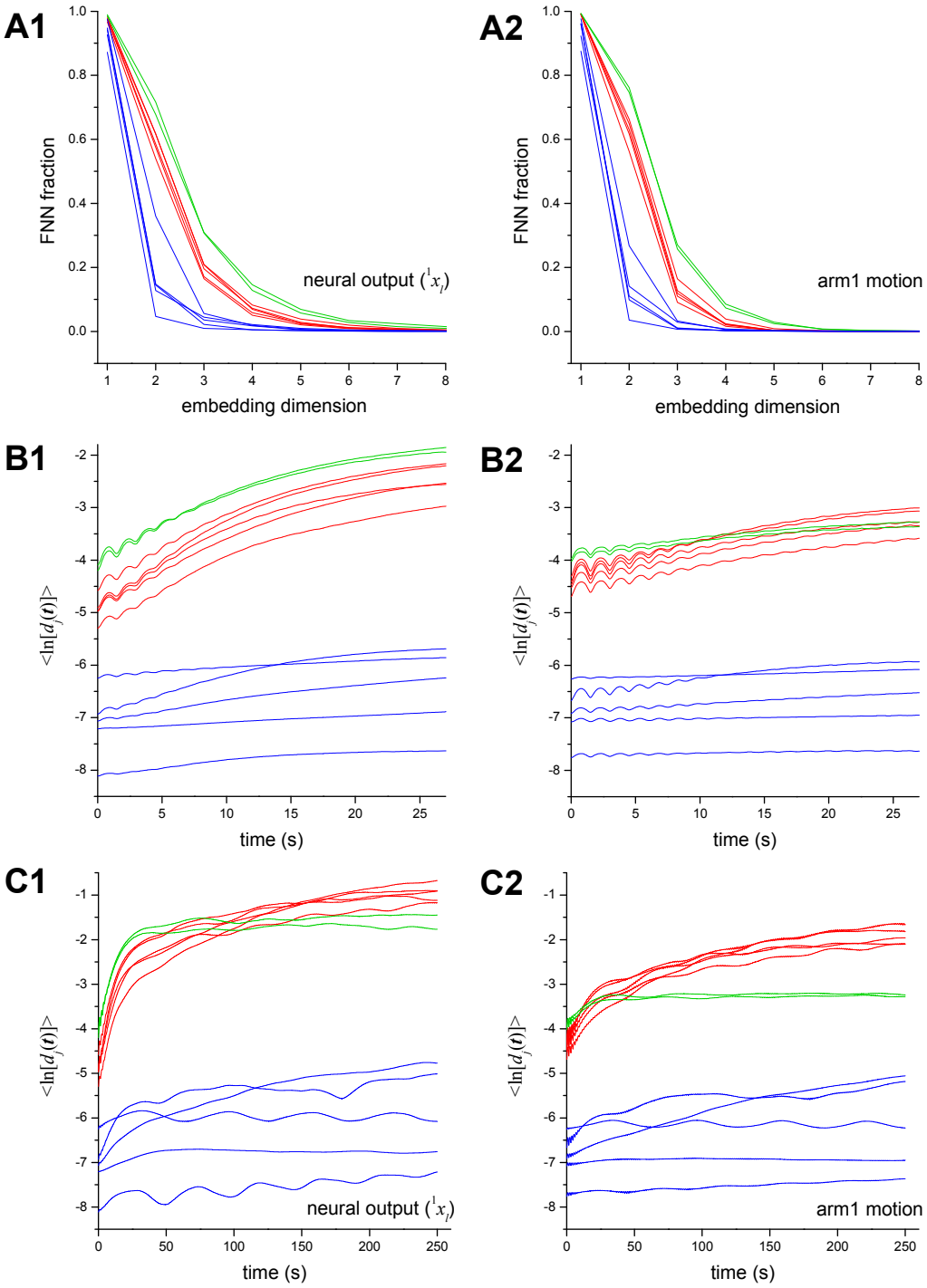

Figure 4.5: Lyapunov analysis of system in chaotic regime. (A1,A2) False nearest neighbours analysis. (B1,B2) Logarithmic divergence curves iterated for 25 sec. (C1,C2) Long range curves iterated for 250 sec. Three distinct groups are plotted as: blue (g =1.6, 2.1, 2.2, 2.4, 2.8), green (g =2.5, 2.6), and red (g =1.7, 1.8, 1.9, 2.0, 2.3). The rest values of g (= 1.5, 2.7, 2.85, 2.9) which are easily distinguished from others by observing the plots in Figure 4.2 and 4.3, are omitted for convenience.

For all data, the first minimum of the AMI ranged from 56 to 66 bits, which is approximately τ=60 bits (1.5 sec) in average. By assuming that they share the same dynamical system, $\tau = 60$ was used throughout the analysis.

Figure 4.5A shows the analyses of FNN and the logarithmic divergence curves generated using Rosenstein's algorithm. The graphs show clear distinction which is analogous to the previously observed neural dynamics shown in Figure 4.2 and 4.3. FNN plots for the gains whose patterns exhibit stable oscillation (blue) show relatively fast convergence of FNN fraction to zero, which indicates the low effective dimension caused by tight synchronisation between elements. The non-stable patterns (red and green) resulted in slower convergence of FNN, and it reached minimum plateau values (indication of an appropriate minimum required embedding dimension) at around m=8 for ${}^1x_l$ and m=6 for arm angle.

The lyapunov analysis was done using m=8 and d=1.5 sec for all data. The divergence curve generated by Rosenstein's method tells us that the time series data is chaotic if there exists reasonably linear region over relatively long time scales. From Figure 4.5B, the estimated maximal lyapunov exponents of red neural group ranged approximately $0.13 < \lambda_1 < 0.18$. The same group for arm movement showed smaller value. Although all data are generated by the simulation, the divergence curve did not show clear linear lines while the curves are positively diverging. This could be from a multiple reasons that the number of data point does not fulfil the Eckmann-Ruelle requirement (Eckmann and Ruelle, 1992), which suggests the minimum required data size is $N > 10^D$ (for our case 10^8 points) where D is the dimension of attractor. Another possibility is that the true system dimension is far larger than 8, and the hidden variables in higher dimensions act as a stochastic noise to the reconstructed attractor. The sampled data set is the result of complex calculation including physics simulation. The physics simulation uses a number of discrete processes which cannot be expressed by a smooth differential equations, which might invalidate the assumption of the determinism of system. Despite all the possible glitches, we are able to generally understand the sensitivity of orbit to initial conditions by determining the divergence curves. The curves in the figure also show similar differences to the FNN analysis. The long term divergence curves in Figure 4.5C show that the group showing synchronised motions (blue) exhibit flat or fluctuating curves, which is the typical behaviour of n-torus system (Rosenstein et al., 1993). The non-synchronising groups normally result positive curves that

show initial steep rising followed by gentle slopes.

In all images, the two green lines are the data analysed for g=2.5 and 2.6 which show somewhat distinguished results. They show the most complex looking motion in the observation of the plots in Figure 4.2 and 4.3, and also their FNN analyses converge slower than others and result in slightly higher minimal embedding dimensions. Estimated λ_1 for neural output were similar to other chaotic data, but those for arm motion indicate they are less chaotic than others. Long term divergence curves (especially the data from arm movement) also show slight deviation from others in that the slope of the curves are flat most of the time. It can be thought that although there are frequent changes of phase relationships between subparts, the actual movement of the robot may exhibit less variability, or there may exist some periodicity in this variability.

4.4 Movement Patterns in the Stable Regime

The observation of the system dynamics in the chaotic regime has provided us with insights into how to provide the system with an appropriate degree of embodimental coupling represented by the global coupling parameter that has to be pre-assigned before running the exploration process. Now let us investigate the emergent patterns of the system in its stable regime which are the candidates for high-performing locomotion behaviours to be searched. The global bifurcation parameter μ was fixed to zero, and the stable patterns which spontaneously emerge from the system were observed and categorised. The sensor gain was chosen as $k = e^{1.8} \approx 6$ from previous experiments.

An 'existing stable pattern' was designated by observing the movement behaviour of the robot for a reasonably long duration (200-300 oscillation cycles). The number and stability of stable patterns greatly varies over different body-environment systems, each of which will form different 'landscape' of the state space of neuro-body-environment dynamical system. The state space might have a small number of strongly attracting states if its landscape has a few 'deep' basins of attraction. Gently undulating landscape may result in a stream of slowly degenerating patterns which might be misinterpreted as stable patterns. In the case of 4-Fin aquatic swimmer, it is anticipated that there will be a few highly stable movement patterns for

Table 4.1: Categorised emergent behaviours existing in the stable regime. The average locomotion performance Avrg E was measured by calculating running average of actual performance using a leaky integrator equation with the time constant of $5\tau_E$, where τ_E is the time constant of performance evaluation (as shown in Equation 3.13).

Pattern	# of variations	Avrg E
1. straight (ST)	4 (each dir)	0.48
2. circular (STC)	8 (4×(CW,CCW))	0.4
3. rotate slow (R-slow)	8 (4×(CW,CCW))	0.052
4. rotate fast (R-fast)	2 (CW,CCW)	0.046
5. vibration (VB)	2 (dir 1-3 & 2-4)	0.034
6. bound antiphase (BA)	1	0.0

which is easy to determine whether or not the pattern is stable.

4.4.1 Categorised Emergent Behaviours

After running a reasonably large number of simulation (more than 500 trials), the stable movement behaviours in the stable regime of the system were identified. If a robot behaviour was observed as being permanently sustaining, it was identified as an individual behaviour. Basic movement behaviours of the swimmer were categorised into motion in four directions (along the body axes dir1, dir2, dir3 and dir4 as shown in Figure 4.1) which met expectations given the symmetric shape of the swimmer. Taking the directional symmetry into account, six different basic behaviours were observed and classified according to the locomotion performance as shown in Table 4.1. The shape of the 4-Fin Swimmer robot is radially symmetric, so different synchronised pairs of joints (variations) can exist for a single behaviour. For example, the straight swimming behaviour has four different combinations of synchronised joint pairs, all of which show the same frog-like swimming behaviour. As shown in Table 4.1 there are a total of 25 different arm coordinations when including all variations. Careful viewing reveals that the circling movement (STC) can show slightly different circling radii resulting from small differences in passive fin tilting, but these are too small to be considered as separate distinguishable beha-

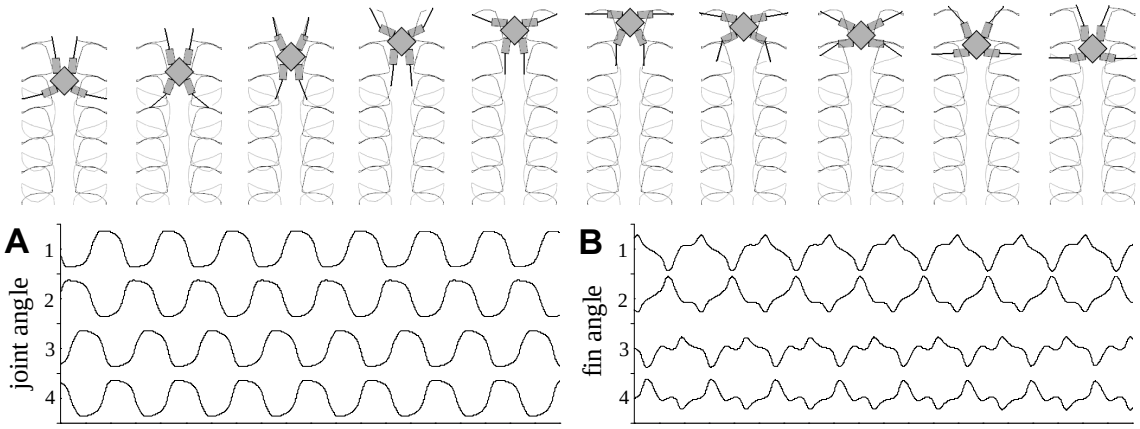

Figure 4.6: (Upper) Snapshots of the straight swimming (ST-dir3) behaviour of 4-Fin Swimmer. Images were taken every 1/10 gait cycle. The tip trajectories of the fore (black) and rear (grey) fins are shown. (Lower) (A) Joint angles and (B) fin bending angles of the behaviour. Each segment along the vertical axis indicates the range [-1,1] rad.

viours. In order to keep the analysis clearer these kinds of variations are not counted as different behaviours.

Figure 4.7 shows the instantaneous phase differences vs. time of each behaviour as well as the agonist neural patterns. Because of the strong tendency of the anti-phase synchronisation between four radially symmetric arms, most of the movement patterns of robot arms are similar each other, which is basically the variations of the frog-like swimming. However, the slight asymmetry of the hydrodynamic forces on the contralateral arms, or improper phase locking between ipsilateral arms resulted in significant differences in forward locomotion performance.

The forward locomotion behaviour involves straight movements (ST), or moving in circles (STC). Straight locomotion is a frog-like swimming action which has the highest performance (Figure 4.6). Since the robot shape is symmetric, the straight swimming has four variations into the body directions 1-4. Circular swimming (STC) moves on a circle by a slight asymmetry between contralateral arms, and their movements can be either clockwise or counterclockwise, having total eight variations. Non-locomotion behaviours were categorised into bound antiphase (BA), vibration (VB) and rotation (R-slow/fast) (Figure 4.9). BA results in no net movement of

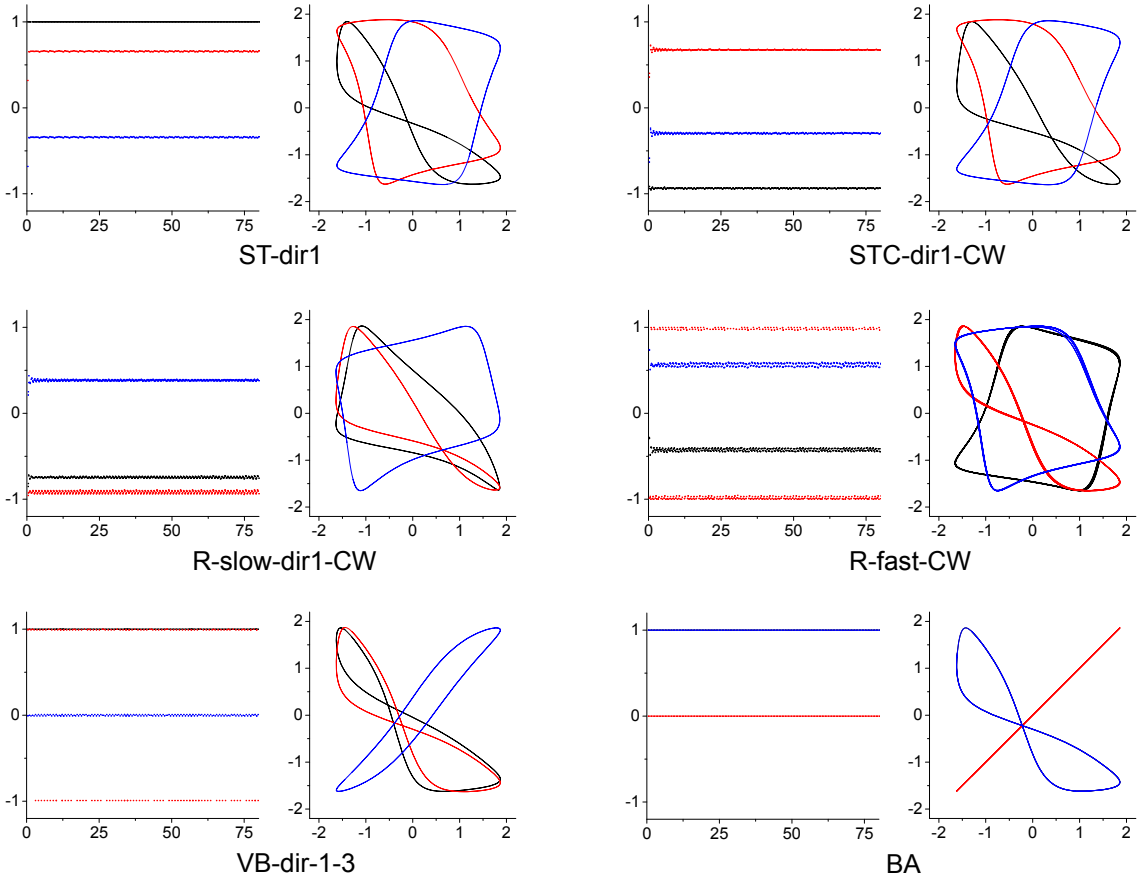

Figure 4.7: Existing behaviours of the 4-Fin Swimmer selected from Table 4.1. For each behaviour, the left plot depicts the instantaneous phase differences vs. time (oscillator cycle), and the right plot shows the agonist neural output of arm2, arm3, arm4 vs. arm1. Each colour of the phase difference plot indicates the phase difference between arm1-arm2 (black), arm1-arm3 (red), and arm1-arm4 (blue). Colours of the neural output plots are the same as the phase difference plots.

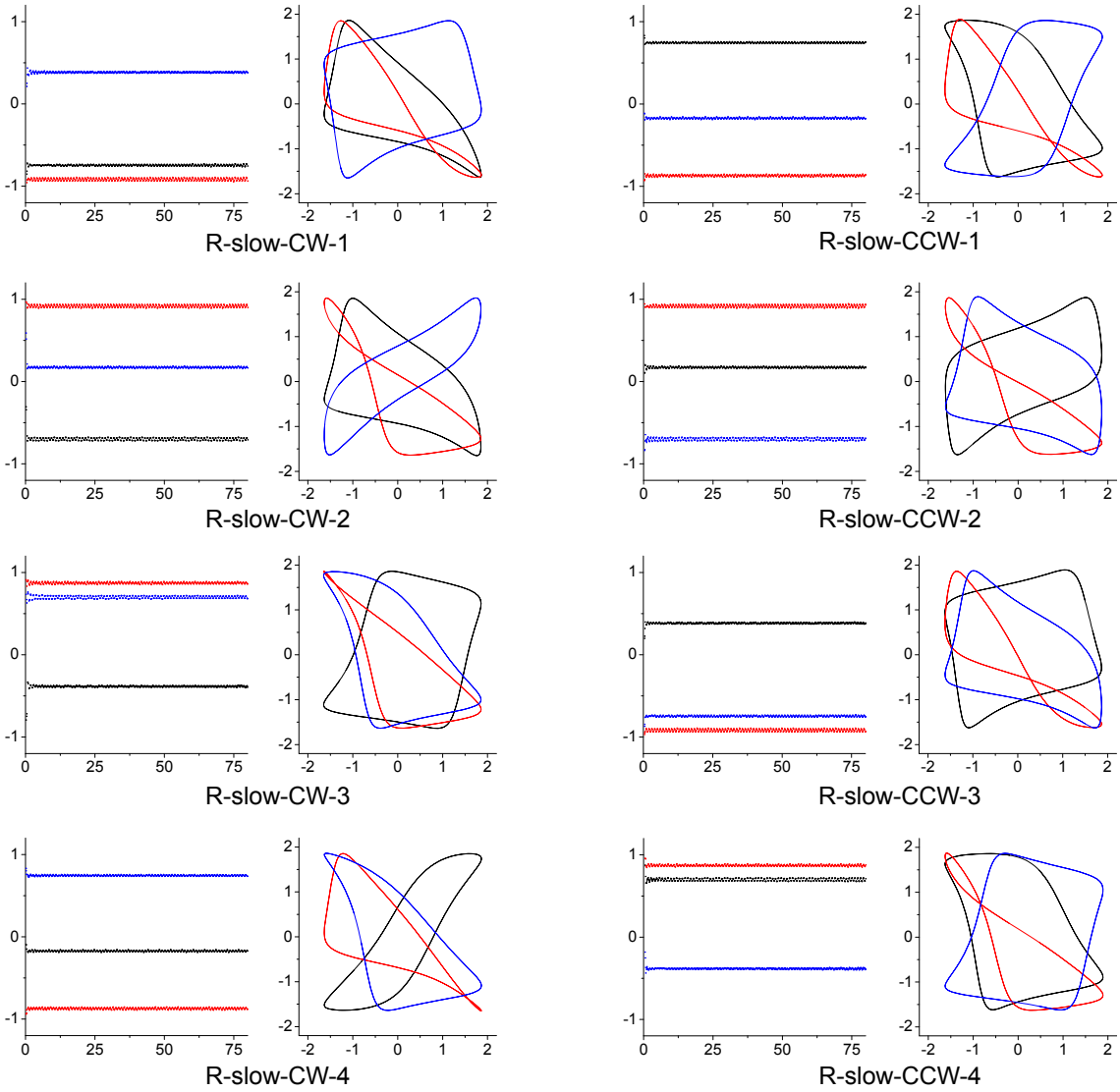

Figure 4.8: Phase relationships and neural outputs of all eight variations of Rotation behaviours. The variations of patterns reflect the radial symmetry of robot shape.

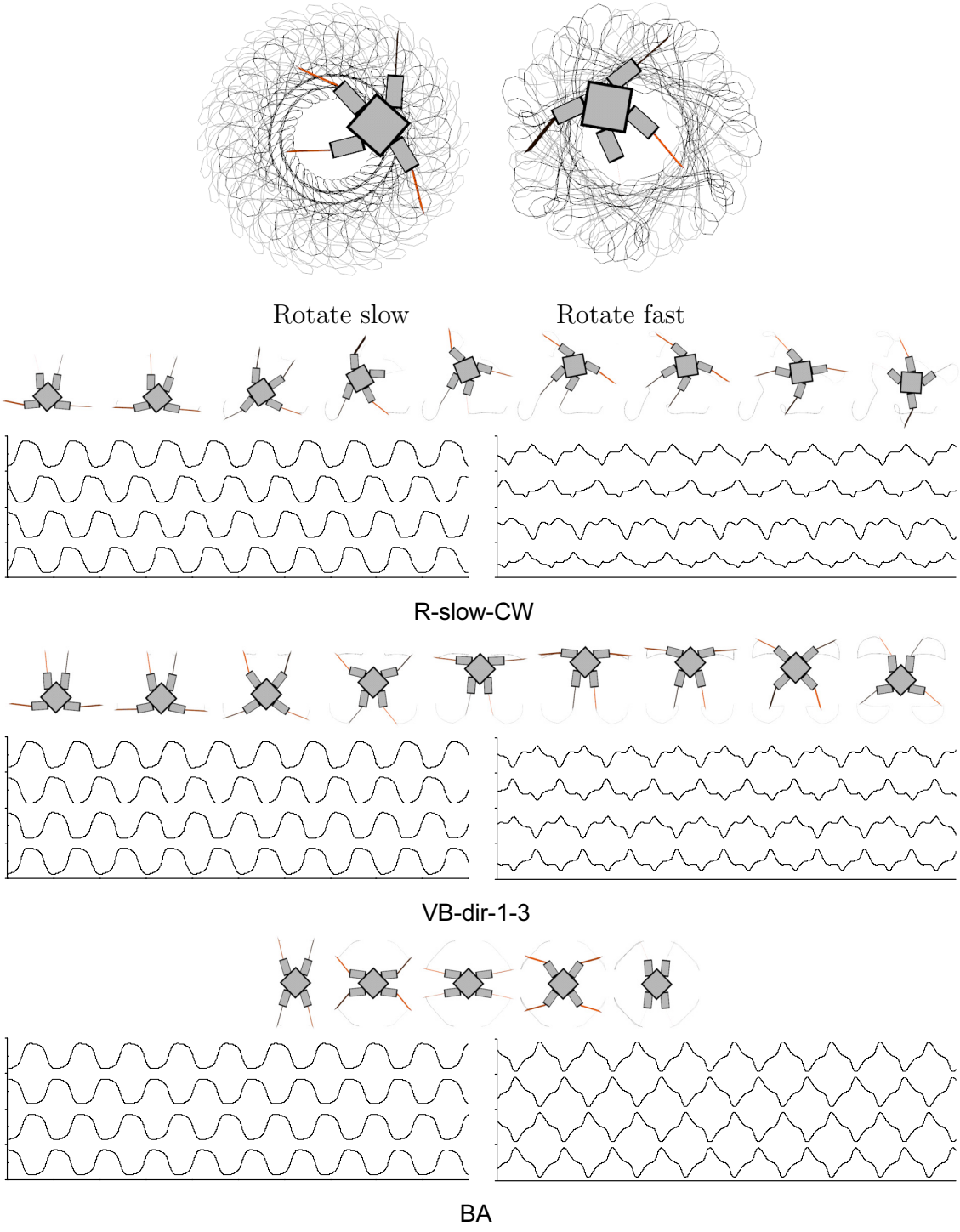

Figure 4.9: The uppermost image shows the continuous trajectories of fin tips of Rotation motions. The lower three images show the movements of R-slow-CW, VB, and BA as presented in Figure 4.6. The snapshots of BA was captured every 1/5 oscillation cycle.

the robot due to anti-phase locking between adjacent pairs of arms. The phase relationship of VB is qualitatively the same as BA with different arm combination. The arm movements of VB are contralaterally antiphase and ipsilaterally in-phase based on the vibrating axis, and it has two variations for the vibrating directions 1-3 and 2-4. It is a failed version of straight swimming whose ipsilateral arms move out of phase with an appropriate phase difference. Rotation behaviours show that the robot rotates clockwise or counterclockwise without moving forward. It is essentially another version of circular swimming, which moves on a circle with very small radius. There were slow and fast rotations. The slow rotation has eight variations just like STC motion (Figure 4.8), whereas the moving radius of fast rotation becomes negligible and regarded as having only two (CW,CCW) variations.

4.4.2 Statistics of Stable Behaviours

Now let us investigate the statistics of the stable patterns that appear when the system is in its stable regime. The system is run until a stable behaviour is observed, and the statistics of the appearance of behaviours are done by a number of trials. Before going into the test, a simple strategy in order to evenly designate the initial conditions of the system was considered. Since the system is completely deterministic, and is run by a numerical simulation, the final pattern depends solely on the initial condition. Therefore, in order to check all possible patterns of the system, the initial conditions should be evenly and densely distributed in the state space of the system. However, due to limited computational power, it is impossible to explore all possible initial conditions. Although the presented neuro-robotic system has a relatively small number of degrees of freedom, the number of actual independent variables of the entire neuro-body-environment system far exceeds that of the robotic body. If we consider the neural part only, the 8 oscillators have 16 variables, which will result n^{16} initial conditions when the initial point for each variable is picked from the discretised n equally spaced regions. Even if we pick only a single initial value located in one of the quadrants in an 2D oscillator space ($n = 2$; two values for one variable with different signs), the number of total initial points becomes 65536, which still requires an enormous amount of time for running the simulations. This means that only a limited (much less) number of trials compared to the number of all possible combinations of initial conditions are available, so running the system

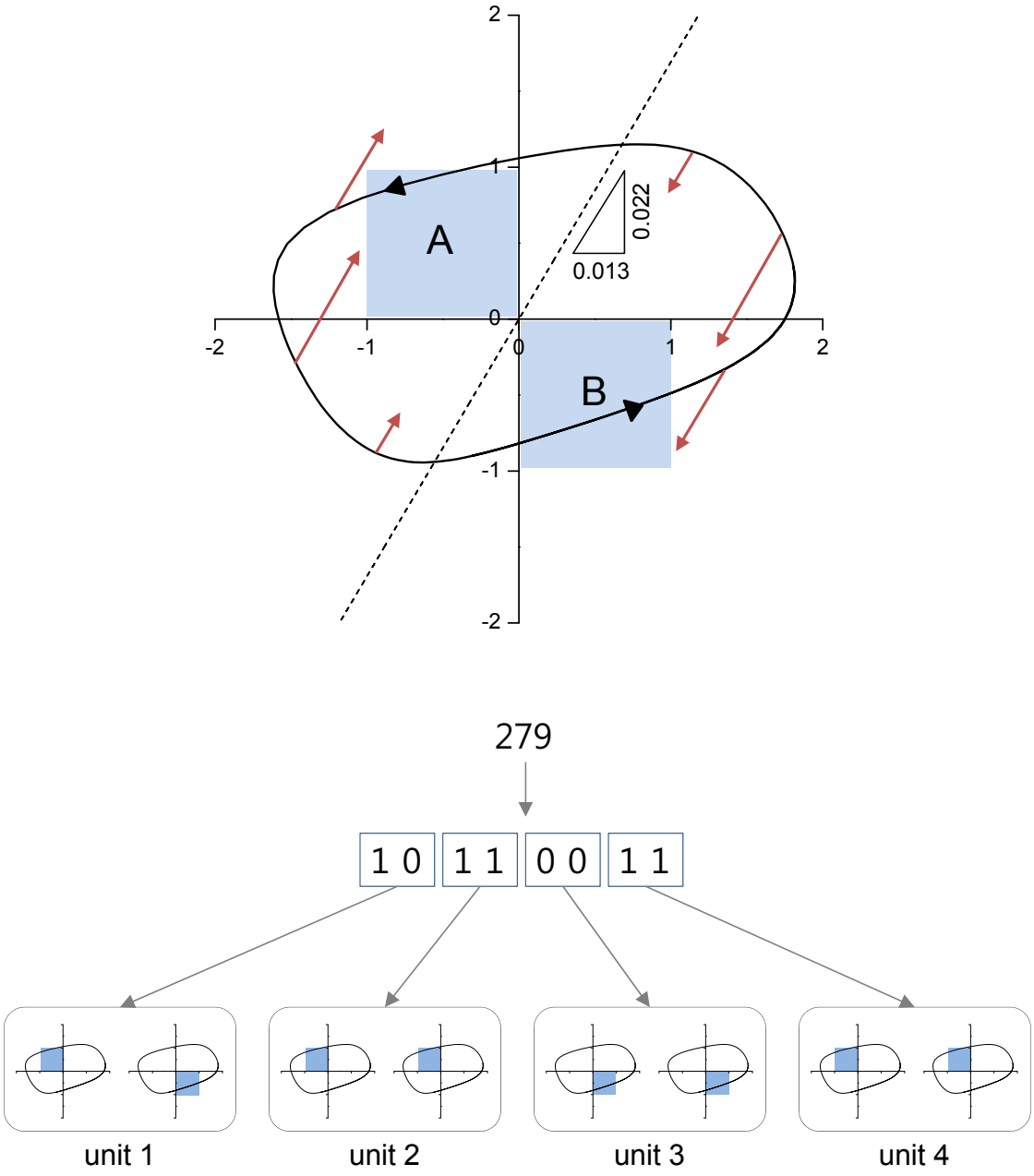

Figure 4.10: (Upper) Initial point of an oscillator for each simulation trial is selected in one of the regions (A or B). (Lower) The regions for choosing initial points of all 8 oscillators (shaded regions) for 279th simulation trial.

from purely randomly generated initial conditions may result in biased statistics when the initial conditions are not evenly distributed in the state space.

Instead of running the system from randomly generated initial conditions, we tried to run the simulation with some rules so that the initial conditions are distributed as evenly as possible in the state space. As an ad hoc heuristic, we divided the state space of a single oscillator into only two regions according to the direction of sensor perturbation. Since the direction of the input vector in 2D oscillator space is fixed to $\mathbf{v} = \pm(\delta, \varepsilon)$ (0.013,0.022) and the centre of rotation of the oscillator limit cycle is aligned at the origin, the limit cycle can be divided into two approximately 'symmetric' orbits which tend to receive decelerating input forces when stabilised by neuro-body-environmental interaction. With this in mind, the two regions in the 2nd and 4th quadrants were assigned to each oscillator space for choosing initial conditions, and the initial point for each oscillator was selected randomly within one of the two specified regions (Figure 4.10). Therefore, total $2^8 = 256$ simulations were run to investigate the statistics of stable patterns.

For each oscillator, the initial point was randomly chosen either in region A or region B. Each simulation was indexed by generating initial points from the region A or B according to the binary number converted from the trial number. The ith bit (from the left) of the binary number indicates one of the two regions of the space of oscillator i, that is, 0 for region A and 1 for region B. For example, 179th simulation trial chooses initial conditions from its binary number 10110011, such that the oscillators 1-8 choose their initial points in the regions A-B-A-A-B-B-A-A respectively. In this way, total 256 simulations were run and the stable patterns emerged from the system were categorised.

Figure 4.11 shows the visiting count of each behaviour. We can see that the appearance of the BA pattern is dominant, which indicates that it has the largest basin of attraction in the state space of 4-Fin swimmer. The appearance of other higher performing behaviours were nearly the same, occupying only a small fraction compared to that of BA. Among them, the lowest performing behaviours such as VB were more frequent than the others, that is, the low performing behaviours are more stable. This is because the shape of robot is symmetric, so the symmetric arm movement patterns are more stable, resulting in no forward locomotion. Figure 4.12A shows the distribution of behaviours according to the initial conditions which are ordered by the sequence of trials. When the appearances are grouped as a

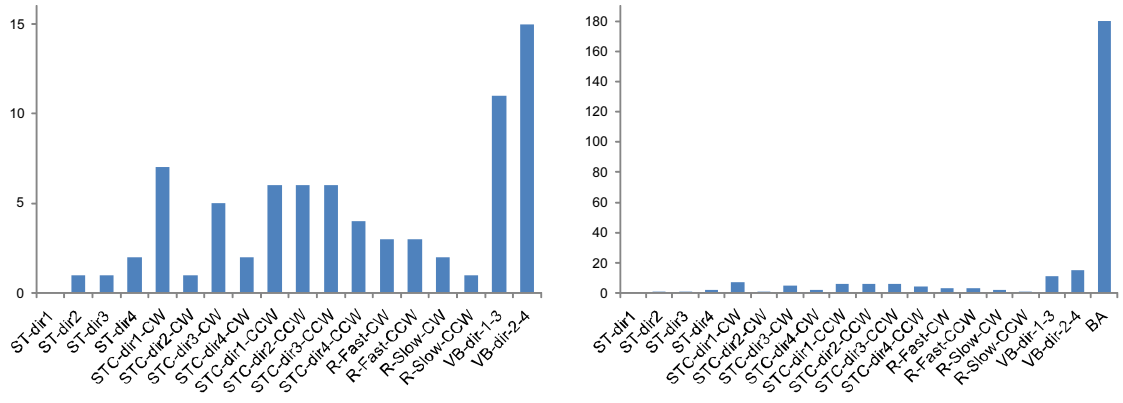

Figure 4.11: Visiting counts of stable behaviours without search (with $g = 1.8$). (Left) Graph excluding BA. (Right) Graph including BA.

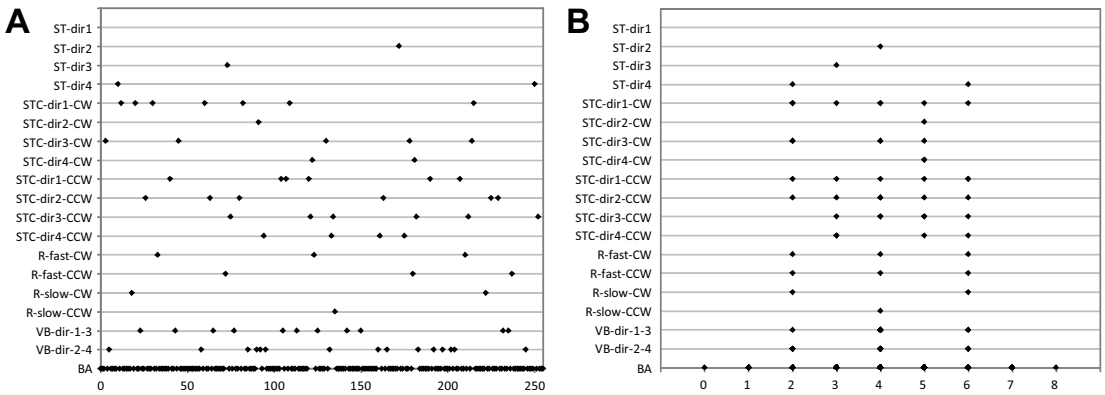

Figure 4.12: Appearance of stable behaviours without search according to the initial conditions. (A) Behaviour appearance vs. trial number. The trial number encodes the initial conditions. (B) Behaviour appearance were shown according to the Hamming distance of initial conditions from 00000000.

function of binary Hamming distances (Figure 4.12B) of initial conditions from zeroth initial point (00000000), the appearance of non-BA behaviours are distributed around the middle, which indicates that the initial conditions whose initial points for eight oscillators are aggregated in similar location (the initial points of most oscillators are in the same region) are almost always entrained to the BA pattern.

4.5 Turning on Chaotic Exploration

We have seen that multiple movement behaviours coexist in the stable regime of the embodied system, while the low-performing behaviours take up most of the state space. Now we will use the chaotic exploration process and see how the system explores and stabilises its state onto good performing behaviours. The chaotic exploration is expected to escape from those undesired behaviours (i.e. Rotation, Vibration, Bound Antiphase) and to be stabilised to one of the swimming-like behaviours.

The exploration process was run by the method described in section 3.4.1. The time constant for measuring locomotion performance in Equation 3.13 was set to $\tau_E = 5T$, where $T = 6.4$ ($\approx 8\tau$, where $\tau = 0.8$ is the time constant of oscillator) is the period of a single oscillator. The time constant (τ_μ) for the global bifurcation parameter μ in Equation 3.14 was set to T, while the previous work of chaotic search (Aida and Davis, 1994) suggests that their experiment has shown the best search performance with τ_μ in the range of $[0.3T < \tau_\mu < 1T]$. The maximum value of the bifurcation parameter was set to $\mu_c = 0.35$ after observing a few simulations not to shut down one of the oscillators in the motor unit. The time scale for the desired performance (τ_d in Equation 3.16) was set to $\tau_d = 25T$.

4.5.1 Exploration Failure (Dead-End) by Weak Chaoticity

The first experiment was done using the sensor gain parameter $g = 1.8$ as in the previous section. Although the previous investigation of the chaoticity of the system in its maximum chaotic regime revealed that the system produced chaotic dynamics with the sensor gain $g = 1.8$, the exploration process did not show the expected result. In most trials, after the system orbit was entrained to the BA pattern, the orbit could not escape from there.

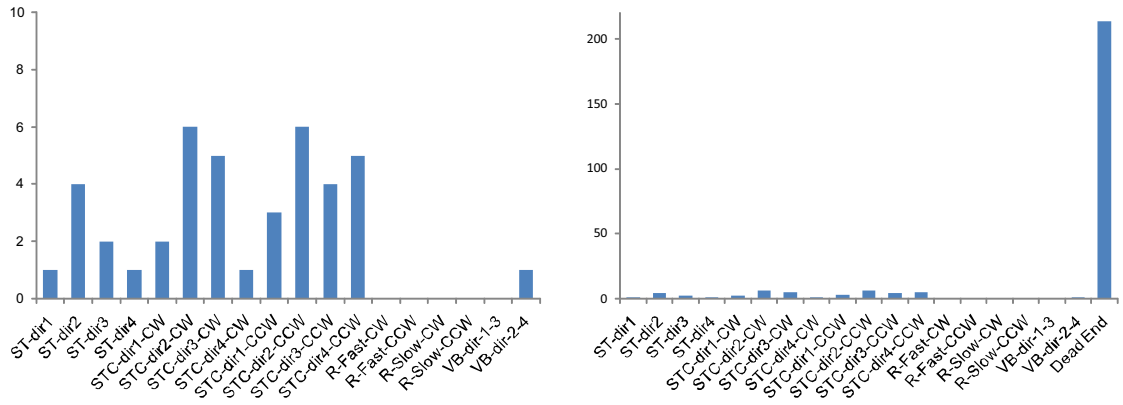

Figure 4.13: Visiting counts of stable behaviours with chaotic search (with $g = 1.8$). (Left) Graph excluding Dead End. (Right) Graph including Dead End.

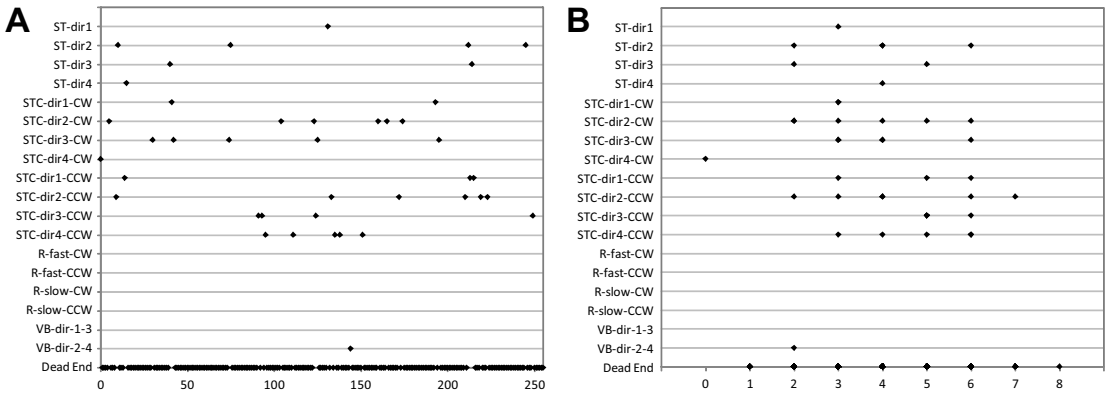

Figure 4.14: Appearance of stable behaviours vs. initial conditions with chaotic search for $g = 1.8$. (A) Behaviours vs. trial number. (B) Behaviours vs. Hamming distance.

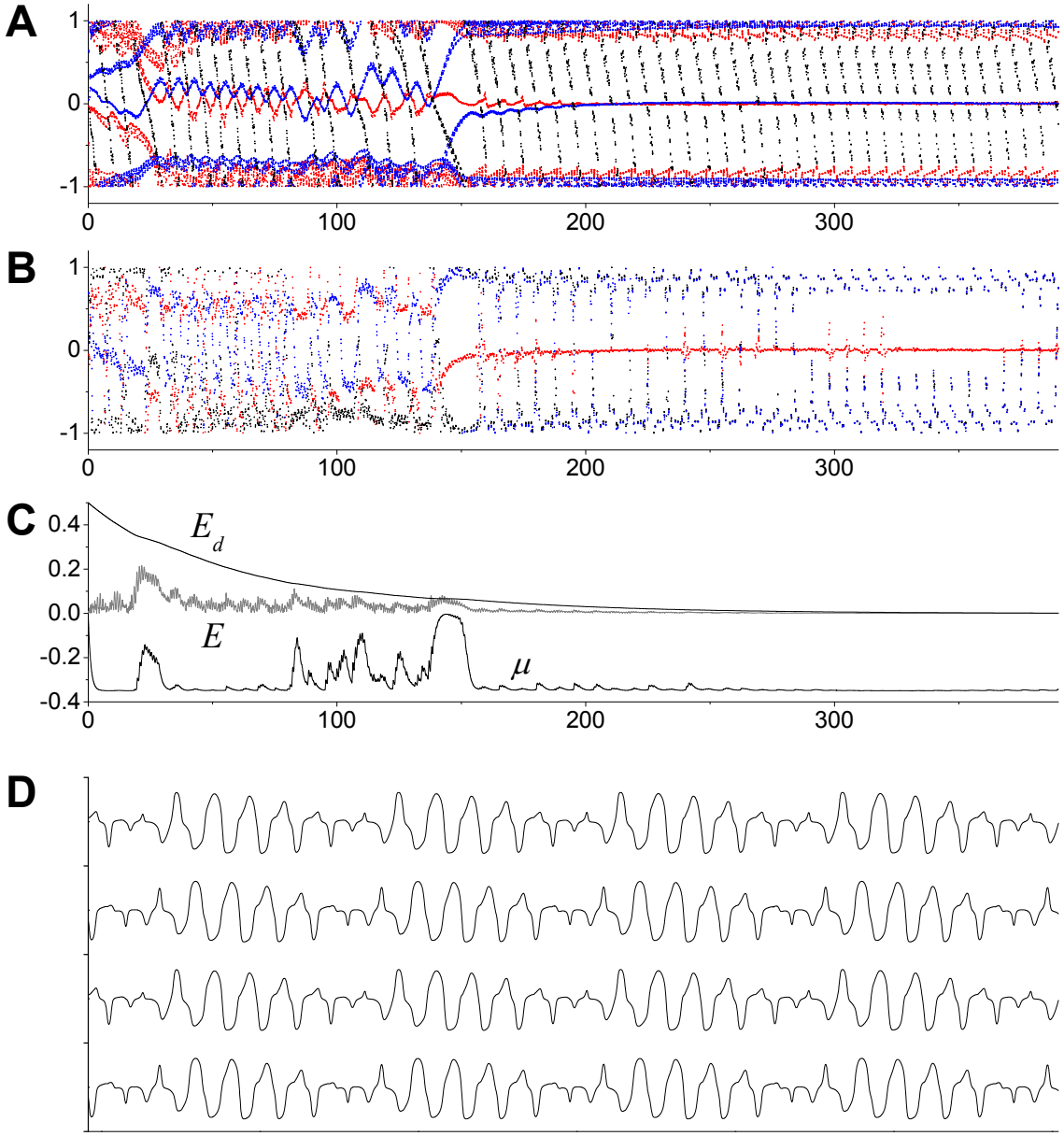

Figure 4.15: Dead end of exploration process with $g = 1.8$. (A) Phase differences between neural outputs. Colours of plots are the same as that of Figure 4.4. (B) Phase relationship of arm movements. (C) The change of actual performance (E), desired performance (E_d), and bifurcation parameter (μ). (D) Arm movements after synchronisation in chaotic regime (Dead End state).

This failure, which we call *dead-end*, is due to the fact that the chaoticity of the destabilised system was simply not strong enough to make the orbit escape from low performing states. The orbit which tries to escape from BA by system destabilisation is quickly re-injected to BA, and the system was virtually stabilised in its chaotic regime by the phase synchronisation between system variables. The re-injection dynamics is repeated regularly, which becomes a part of system synchronisation. Although the time course of phase relationships of neural and robotic parts (Figure 4.15A and B) exhibit the regular pulse-like escaping trials, they are mostly for the phase difference between the agonist and antagonist muscles. This means that only the amplitude of arm movement regularly fluctuates, but there is no significant change in the interlimb phase relationships. The short lasting pulse-like changes of interlimb phase relationships are seen in arm movements (Figure 4.15B), but their occurrences are synchronised with the amplitude fluctuation, that all the phase differences quickly drift by 2π to be returned to their original states. As the result, the actual performance E stays low and the desired performance E_d decays close to E (Figure 4.15C). The final arm movements show the phase relationship similar to BA pattern with the periodic change in their amplitudes as shown in Figure 4.15D.

The statistics of the exploration process (Figure 4.13 and 4.14) reveals that most of the trials yielded to dead-end except that a few number of runs were stabilised to high performing behaviours. Although there are a few seemingly successful results which are stabilised to high performing behaviours, they were not the result of successful escape from low performing behaviour but by being started from 'good' initial conditions which would eventually lead the system to one of the high performing behaviours without encountering low performing behaviours.

4.5.2 Test using Higher Chaoticity

Since the system has failed to have satisfactory explorability when the chaoticity is not strong enough, another test was done using the sensory gain of $g = 2.5$ which has shown the highest chaoticity in section 4.3.2 (green group in Figure 4.5). In this case, most of the trials successfully avoided low-performing patterns and are stabilised to one of the locomotion behaviours. Figure 4.16 shows an example of successful exploration process which stabilises onto STC pattern.

During the search process all variables and control parameters vary continu-

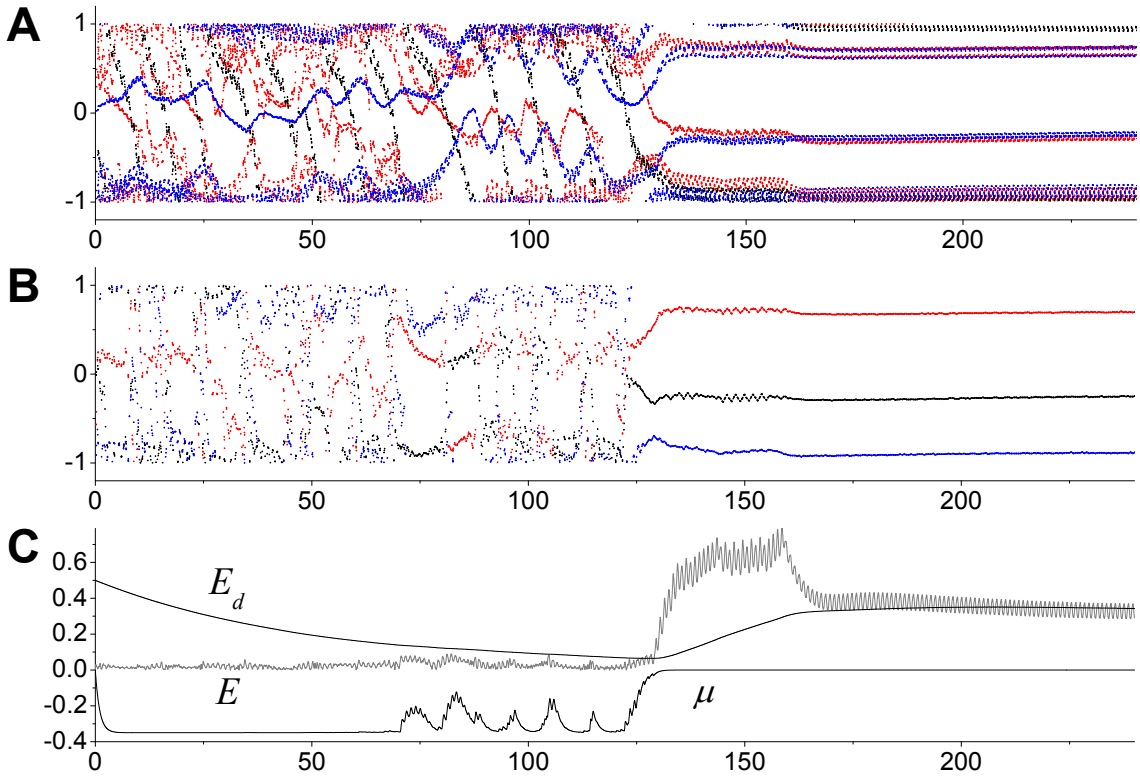

Figure 4.16: Successful exploration of high performing behaviour with $g = 2.5$ (STC-dir4-CCW).

ously as parts of the neuro-body-environment system, and the time evolution plots of phase relationship, performance, and control parameter (Figure 4.16) show that there is no clear distinction between the stabilisation and destabilisation of system. Rather, the change of system behaviour shows a continuous mixture of chaotic and smooth drift of phase relationships until a high performing pattern suddenly emerges within a few tens of oscillation cycles. Since the stable locomotion behaviours have similar performances which are segregated in a certain range of values without any intermediate performing behaviours, the system shows relatively sudden changes between locomotion and non-locomotion behaviours. The system dynamics stays in its chaotic regime usually with some irregular fluctuation around BA pattern (having low E), while initially high desired performance (E_d) decays toward low E (Figure 4.16C). As E_d is decreased to a certain level which is close to E, the bifurc-

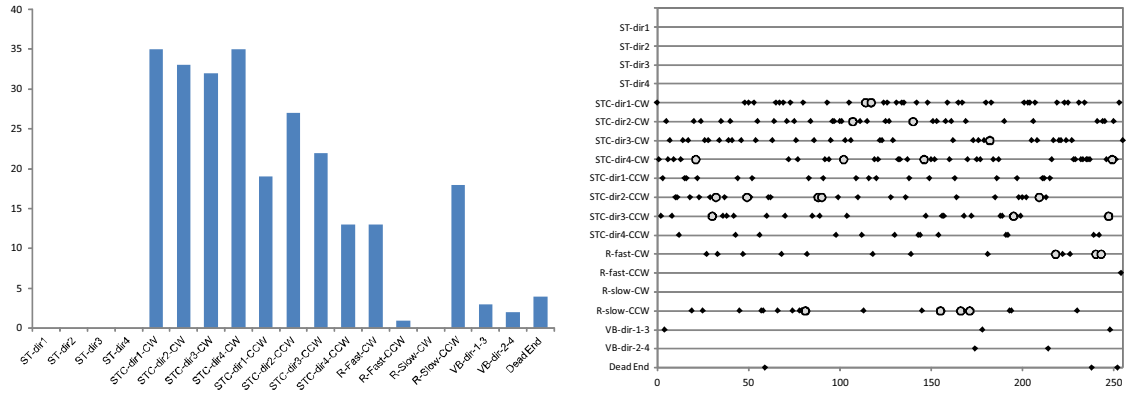

Figure 4.17: Statistics of behaviour appearance for $g = 2.5$. (Left) Visiting counts including Dead End. (Right) Behaviours vs. trial number. Shaded circles indicates the behaviours which are stabilised after Deep-Path.

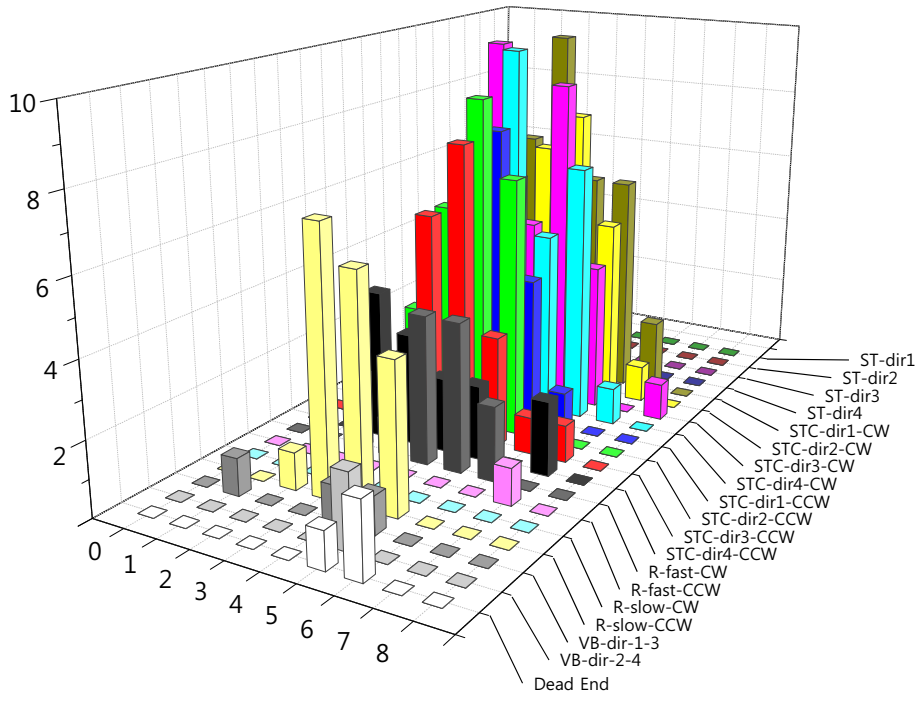

Figure 4.18: The number of appearance of each behaviour vs. 3D visualisation of Hamming distance ($g = 2.5$).

ation parameter μ varies sensitively according to the small changes in locomotion performance, resulting the system to actively try some intermediate patterns in order to be stabilised to a behaviour whose performance is higher than the desired performance ($E \geq E_d$).

The statistics of behaviours (Figure 4.17 and 4.18) shows that the explorability of the system has been improved, and the frequency of the appearance of dead-end was greatly reduced. Note that there was no straight swimming patterns (ST) observed in this experiment. The higher sensor gain strengthened the global coupling between neural elements, and it slightly changed the phase relationship between arm movements as well as the course of exploration process. Most of exploration process in this test exhibited the rotation (R) behaviours as an intermediate stage before being stabilised to the final behaviour, so the asymmetry between passive fin dynamics of contralateral arms was amplified and sustained after the stabilisation of arm movements, resulting only STC motions were stabilised.

4.5.3 Exploration Deficiencies: Bad-Lock and Deep-Path

Although the system exploits chaotic dynamics for the exploration of motor patterns, unwanted synchronisation between system variables in a quasiperiodic regime of the system, resulting in low performance, can arise from some initial conditions. This we call *bad-lock*, where the dynamics of all system variables are locked in a narrow range of phase differences while the precise values of variables vary chaotically (Figure 4.19). The system was locked onto non-locomotion patterns (Rotation and Vibration) where the bifurcation parameter μ does not reach zero but oscillates near zero being phase-locked with other system variables. Although this is undesirable for the purpose of this work, it should be noted that this phenomenon is observed in real biological systems (e.g. in walking and heartbeat rhythms). The bad-lock phenomena occurred more frequently if we set the maximum available bifurcation parameter μ_c below the onset of chaos, indicating that the system has less exploratory 'perturbation force' when using low chaoticity.

Also another deficiency, so called *deep-path* was sometimes observed. It is a similar phenomenon to the exploration dead-end phenomena which was previously seen in section 4.5.1. This involves the orbit becoming entrained in some periodic state for a long time before it eventually reaches the desired state (Figure 4.20).

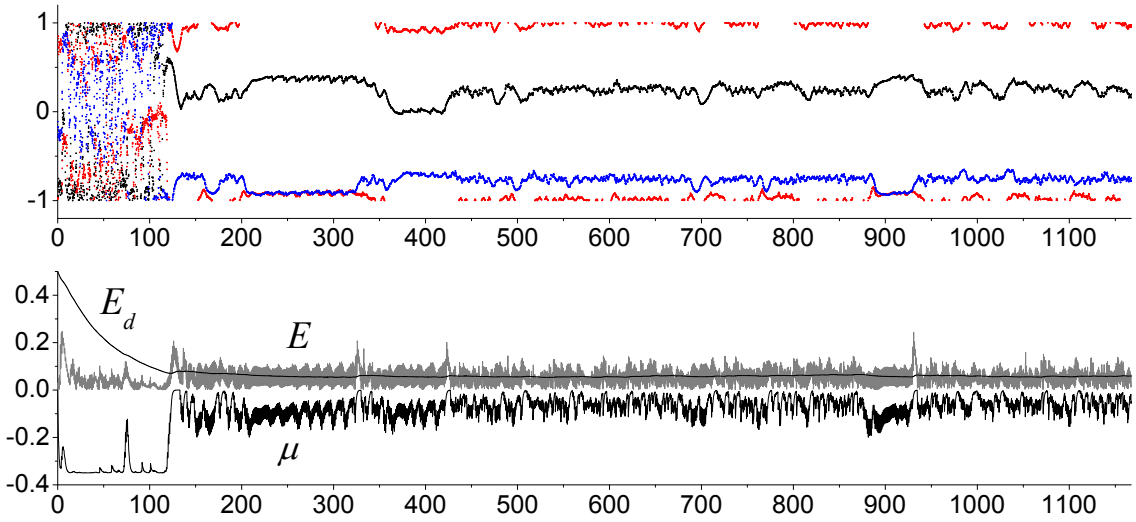

Figure 4.19: Bad-lock.

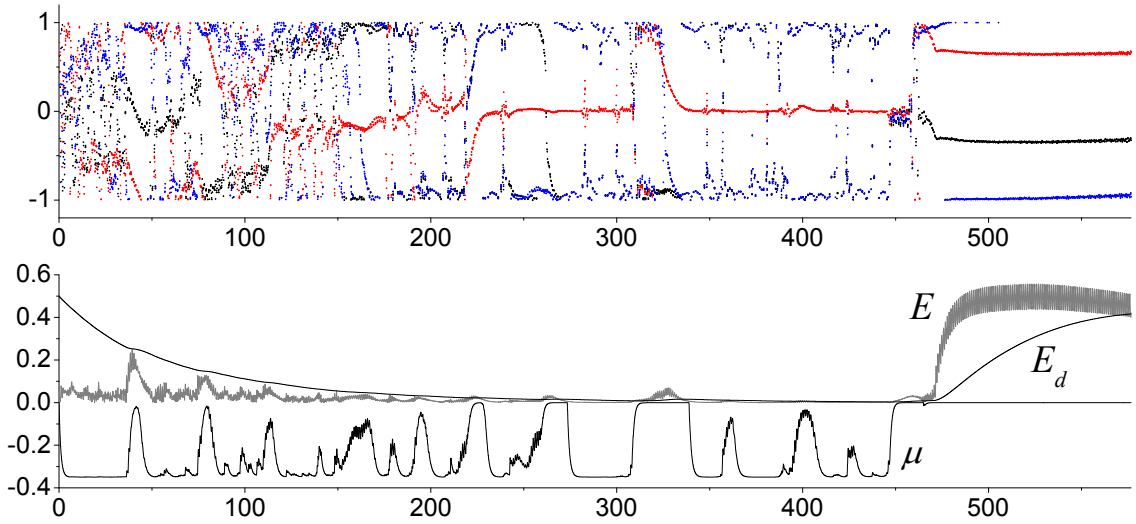

Figure 4.20: Deep-path.

Similar to dead-end, due to the strong attraction of the BA pattern, the orbit which tries to escape from BA by system destabilisation is re-injected to BA and stays there for a long time. The chaoticity of orbit in the chaotic regime of the system is stronger than the previous case, so the system does not stall to dead-end but maintains weak irregular fluctuation of variables. Prolonged stay of the actual performance E in low value makes the desired performance E_d decay close to E so that the sensitivity of μ increases, resulting in increased number of stabilising trials. However, those stabilising trials re-encounter BA pattern repeatedly for a long time before escape. The escape orbit is often stabilised to STC patterns, which indicates that the STC patterns are located in the vicinity of BA in the phase space.

As shown in Figure 4.17 (right), there were a few number of deep-path dynamics which might have become the dead-end of exploration in the previous system using $g = 1.8$. Even so, the total number of dead-end and deep-path occupies much smaller fraction of entire emergent behaviours than previous experiments. The possibilities of bad-lock and deep-path always exist because the system is fully deterministic without stochastic sources, but it should be possible to reduce them by introducing additional adaptive factors as will be presented later. It could be possible to alleviate these defieciencies by introducing a stochastic perturbation such as the random noise when the system stagnation is perceived. However, distinguishing those deficiencies from normal exploration process in an automatic and systematic way is difficult, because there is no clear criteria to identify the occurrence of deficiencies in realtime by tracking the dynamics of system variables. Another option would be to present stochastic perturbation indifferently throughout the exploration process, but it will enigmatise the contribution of chaotic dynamics to the system, which will go against the aim of this work as well.

4.6 Summary

This chapter has presented the initial implementation of the idea of the exploration-stabilisation framework for dynamically searching useful locomotor behaviours among the existing stable behaviours that emerged from the neuro-body-environment system. As a body-environment system, a simple 4-Fin Swimmer model moving in a 2D hydrodynamic environment was combined with the neural part, which consists

of a series of identical motor units, via local sensory feedback. The sensor information was fed to the neural parts by properly designed input equations to form the structure of indirect coupling between CPGs similar to that of the excitatory-to-all connection architecture of coupled BVP oscillators as shown in Figure 3.5. The importance of the global sensor gain was highlighted as it determines the overall strength of embodimental coupling between neural parts, and we investigated how the system behaves using different gains.

Although all CPGs are physically disconnected, the information from one part was transferred through physical embodiment to the others, resulting in multiple synchronised patterns that emerged in the stable regime of system, some of which exhibited biologically plausible locomotor behaviours. The exploration process showed that the system orbit could chaotically wander and be stabilised to one of the high-performing stable patterns. While the emergence of patterns in the stable regime was reliable over a wide range of sensor gain, the system lost exploration capability when the gain was not high enough to generate enough chaoticity. Even if the system showed well-looking exploration process by giving enough chaoticity, there were other unwanted system deficiencies according to some initial conditions, which it seems inevitably arise from the deterministic nature of system. The initial framework will be extended and improved in the next chapter in order to improve performance and alleviate the deficiencies highlighted in the previous section.

Chapter 5

Exploration with Sensor Adaptation and Oscillator Learning

5.1 Introduction

In this chapter we improve the basic framework of chaotic exploration in order to achieve more adaptive system for various kinds of physical embodiment. Since the main direction of our work is to build an integrated neuro-body-environmental system which is essentially a single high dimensional continuous dynamical system, the improvement of the system should be considered within these conceptual boundaries. The basic exploration system is extended by employing two fully dynamic processes which are smoothly integrated with the basic system to form an expanded continuous dynamical system. The first method is inspired by biological processes, and it deals with the preprocessing of the sensor inputs coming from an arbitrary robotic system, so that it is able to maintain an appropriate functional coupling between neural elements in a realtime and continuous manner. The second improvement is to combine a fully dynamic learning method in order to automatically wire initially disconnected neural elements, so that the system can dynamically capture and memorise an emergent locomotor pattern discovered by the exploration process. The following sections will describe each method and their application to the basic system, and the performance of the integrated system will be investigated

by presenting some experimental results.[1]

5.2 Incorporating Sensory Homeostasis

A growing body of literature suggests that the activity levels of biological neurons and networks are homeostatically regulated in order to maintain physiologically proper and stable network performance in the face of growth, channel turnover, and modification of either synaptic or intrinsic parameters or both (Turrigiano et al., 1994; Liu et al., 1998; Turrigiano and Nelson, 2000, 2004; Turrigiano, 2007; Zhang and Golowasch, 2007; Turrigiano, 2008; Wilhelm and Wenner, 2008; Wilhelm et al., 2009; Zhang et al., 2009; Gunay and Prinz, 2010). In biological neural networks, there exist the saturation effects such as hyperexcitation and quiescence. They are not easily observed in living neural systems, because the homeostatic plasticity serves to regulate neural activity. The concrete mechanism of the regulation of neural activity is not known but it is clear that neural activity tends towards a constant level in the long term. Various mechanisms have been revealed about how this homeostasis is accomplished, amongst which are mechanisms affecting the strength of synaptic connections (Turrigiano, 1999; Abbott and Nelson, 2000) and the intrinsic excitability of individual neurons (Desai et al., 1999; Zhang and Linden, 2003).

These findings have greatly inspired both computational biologists and AI researchers serving as a conceptual basis for a simple and powerful adaptive mechanisms. Particularly in the field of neuro-robotics, they have been adapted for use in Continuous-Time Recurrent Neural Networks (CTRNNs) (Beer, 1995b, 2003) and have been shown to make CTRNNs more sensitive and give rise to more interesting behaviours in the control of autonomous agents (Williams, 2004). Homeostatic plasticity has been incorporated into the artificial evolution of CTRNNs for a wheeled robot navigation task while trying to maintain a certain range for the firing rates of neurons (Di Paolo, 2000). The evolved robots were able to adapt to nontrivial

[1]The flash streaming (FLV) as well as downloadable AVI files of the movies in this work are available at: `http://www.informatics.sussex.ac.uk/research/groups/ccnr/movies/yssmovie.html` (Shim, 2012). Where available, the indices of movies (ex. Video 1,2...) in above URL will be indicated using footnotes throughout the text.

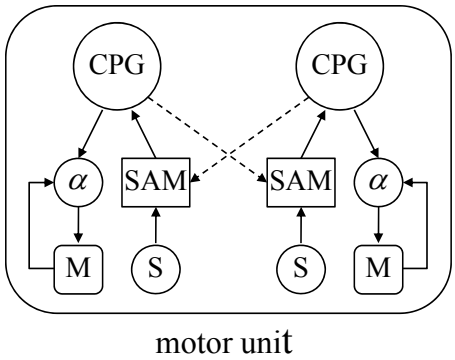

Figure 5.1: Modified motor unit with homeostatic sensor adaptation.

sensorimotor disruptions such as sensor inversions for which they have not been evolved.

As investigated in the previous chapter, the global coupling strength of indirectly coupled neuro-body-environment system significantly influences overall system behaviour. Self-structured coupling of embodied system requires to maintain proper strength in order to provide a range of dynamics from stable to chaotic patterns. Since our goal is to build an exploration system for an arbitrary body-environmental configuration, the sensor signal should be adjusted adaptively to operate in an appropriate range. Using the measurable properties of the CPG output signals of motor units as reference for the homeostatic adaptation of sensor signals, the incoming signals for indirectly coupled CPGs can be regulated to have similar waveforms as in the case of electrically coupled CPG network. Based on the equivalence concept of two coupled BVP models, the sensor input to one CPG is regarded as the incoming signal from the antagonistic CPG via coupling between them. Therefore the input signal is adjusted to have similar amplitude and offset with that of the antagonistic CPG output.

5.2.1 Homeostatic Regulation of Sensor Signals

In order to implement sensor adjustment, a sensor adaptation module (SAM) is added to the motor unit (Figure 5.1). The sensor adaptation module performs homeostatic adaptation for sensor input by calibrating the raw sensor signal using a linear transformation, which continuously adjusts the amplitude and offset of

the periodic sensor signal in order to closely match its waveform to that of an antagonistic oscillator output. The sensory signal (in most cases mechanosensory information from haptic sensors or muscle afferents) may vary according to the choice of sensors and the different body-environment interaction conditions. If the incoming signal is too large the chaoticity of the system will be lost, if too small the neural signals will be uncorrelated. The regulation of sensory activation ensures that the oscillator pair in a motor unit maintains a certain level of information exchange close to that of a weakly coupled oscillator pair so that the network dynamics are regulated within an appropriate range to generate flexible yet correlated activities. This also ensures that the chaoticity of a motor unit is controlled in a systematic and collective way by the feedback signal regardless of the physical properties of the robotic system and the type of sensors.

As the properties of the reference signal for which the sensory adjustment should follow, the energy and average of the periodic signal were used. Both values can be measured in a realtime and dynamical way and they effectively represent the amplitude and offset of the reference neural output to be compared to the actual sensor signal. Let us recall the equations for a pair of CPGs in a motor unit (from Equations 3.9-3.12) in a reduced form,

$$\tau \dot{x}_{l,r} = c(x_{l,r} - \frac{x_{l,r}^3}{3} - y_{l,r} + z_{1,2}) + \delta F_x(s_{l,r}, x_{l,r}) \tag{5.1}$$

$$\tau \dot{y}_{l,r} = \frac{1}{c}(x_{l,r} - by_{l,r} + a) + \varepsilon F_y(s_{l,r}, x_{l,r}) \tag{5.2}$$

where $F_x(s_{l,r}, x_{l,r})$ and $F_y(s_{l,r}, x_{l,r})$ are sensor input functions in order to form a proper coupling topology based on the equivalence concept of two coupled BVP models (as described in section 3.3). With sensor adaptation these functions are denoted by manipulating raw sensor input as; $F_x(H_{l,r}(s_{l,r}, t), x_{l,r})$ and $F_y(H_{l,r}(s_{l,r}, t), x_{l,r})$, where $H_{l,r}(\cdot)$ are the sensor adaptation functions for each CPG in a motor unit which varies in time. The function H takes the form of a time varying linear transformation such as $H(s,t) = P(t)s + Q(t)$, so at each time instance, every individual CPG is subject to its own different sensor transformation.

Given a raw sensor signal s and its reference neural signal n whose amplitude and offset should be followed by $H(s,t)$ (usually an antagonistic oscillator output in the corresponding motor unit), the common form of sensor adaptation function

is written as:

$$H(s,t) = (s - \overline{s})e^{A(t)} + (\overline{s} + B(t)) \tag{5.3}$$

$$\tau_h \frac{dA(t)}{dt} = \sqrt{\overline{(n - \overline{n})^2}} - \sqrt{\overline{(H(s,t) - \overline{H(s,t)})^2}} \tag{5.4}$$

$$\tau_h \frac{dB(t)}{dt} = \overline{n} - \overline{H(s,t)} \tag{5.5}$$

where $\overline{x}$ represents the continuous running average of x as calculated from $\tau_h d\overline{x}/dt = -\overline{x} + x$. The raw sensor signal s is linearly transformed by a multiplicative factor $e^{A(t)}$ and an additive factor $B(t)$. The function $A(t)$ is updated by comparing the difference of the root mean square of the temporal average of the squares of the antagonistic neural output n and the transformed incoming signal $H(s,t)$, which is analogous to the signal energy that reflects the strength or amplitude. $B(t)$ is used as part of the scheme to remove the offset bias: each signal is subtracted by its average offset ($\overline{n}$ and $\overline{H(s,t)}$) before calculating the energy difference. $B(t)$ is updated by the offset difference between two signals. The time scale of adaptation should be set longer than that of the oscillator, and we used τ_h as the time scale of performance evaluation (τ_E) throughout this work, as described in the next section.

5.2.2 Movement Patterns in the Stable Regime

By applying sensory homeostasis to the previously introduced 4-Fin Swimmer robot, another set of experiments was performed to investigate the newly extended system. The fin bending angle ϕ was used as the raw sensor signal for a corresponding motor unit as before, but the raw sensor variables are now put through the adaptation function without using the fixed gain just by setting $s_l = \phi$ and $s_l = -\phi$. Then the full equations describing a motor unit for the 4-Fin Swimmer are written from the previous equations (Equations 4.3-4.6) as:

$$\tau \dot{x}_l = c(x_l - \frac{x_l^3}{3} - y_l + z - \mu) + \delta H(s_l, t) \tag{5.6}$$

$$\tau \dot{y}_l = \frac{1}{c}(x_l - by_l + a) + \varepsilon(H(s_l, t) + x_l) \tag{5.7}$$

$$\tau \dot{x}_r = c(x_r - \frac{x_r^3}{3} - y_r + z) + \delta H(s_r, t) \tag{5.8}$$

$$\tau \dot{y}_r = \frac{1}{c}(x_r - by_r + a) + \varepsilon(H(s_r, t) + x_r). \tag{5.9}$$

Table 5.1: Categorised emergent behaviours of 4-Fin Swimmer with sensory homeostasis existing in the stable regime.

Pattern	# of variations	Avrg E
1. straight (ST)	4 (each dir)	0.7
2. circular (STC)	8 (4×(CW,CCW))	0.6
3. rotate (R)	2 (CW,CCW)	0.06
4. peg-leg (PL)	4 (each arm)	0.04
5. vibration (VB)	2 (dir 1-3 & 2-4)	0.03
6. bound antiphase (BA)	1	0.0

Since the fin sensor reflects the output difference of the oscillator pair in the corresponding motor unit as previously defined (i.e. $s_{l,r} = f(x_{r,l} - x_{l,r})$), the reference neural signals for sensory adaptation in Equations 5.4 and 5.5 should also be changed to $n_l = -n_r = x_r - x_l$. All other parameters were same as previous experiment. Since $H(s,t)$ is operated for each CPG, a motor unit includes a pair of $H(s,t)$ each of which has its own adaptation variables $A(t)$ and $B(t)$ for agonist and antagonist CPGs. Because the 4-Fin Swimmer has only a single sensor for a motor unit as previously described, there is a relation; $A_l(t) = A_r(t)$ and $B_l(t) = -B_r(t)$.

Again, we fixed the bifurcation control parameter to the stable regime ($\mu = 0$) and ran the simulation to see the existing behaviours emerged from various initial states. More than 1000 simulations were run in order to observe and categorise the behaviours. Basic movement behaviours of the Swimmer were categorised into motion in the four directions of swimmer. The result was categorised into six different behaviours which is almost the same as the previous system without using sensory adaptation. The representative behaviours existing in the stable regime were classified according to the locomotion performance and are listed in Table 5.1. The time courses of the phase relationships and the sensor adaptation variables for each pattern were depicted in Figure 5.2.

The list of emergent behaviours was similar to that of non-adaptive case, that the straight movements (ST), moving in circles (STC), Rotation (R), Vibration (VB), and Bound Antiphase (BA) exhibited similar or almost same movement patterns to

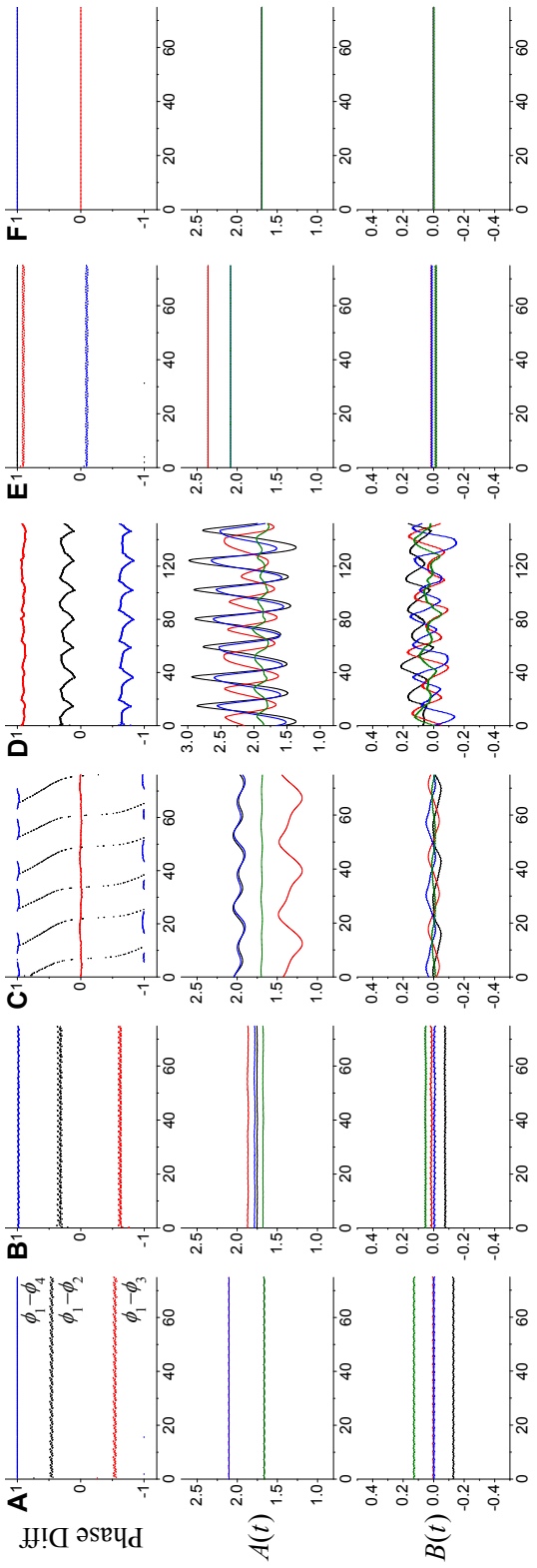

Figure 5.2: Existing behaviours of the 4-Fin Swimmer selected from Table 5.1. Time (oscillator cycle) courses of instantaneous phase differences ($\times\pi$rad) and sensor adaptation variables ($A(t)$ and $B(t)$ from Equations 5.4 and 5.5). (A) ST-dir2, (B) STC-dir2-CCW, (C) PL-arm2, (D) Rotate-CCW (period was doubled for clarity), (E) VB-dir-1-3, and (F) BA. Each colour in the graphs for phase relationships indicates the phase difference between arm1-arm2 (black), arm1-arm3 (red), and arm1-arm4 (blue). Coulors for $A(t)$ and $B(t)$ represent those of corresponding fin sensors which are: fin1 (black), fin2 (red), fin3 (blue), and fin4 (green).

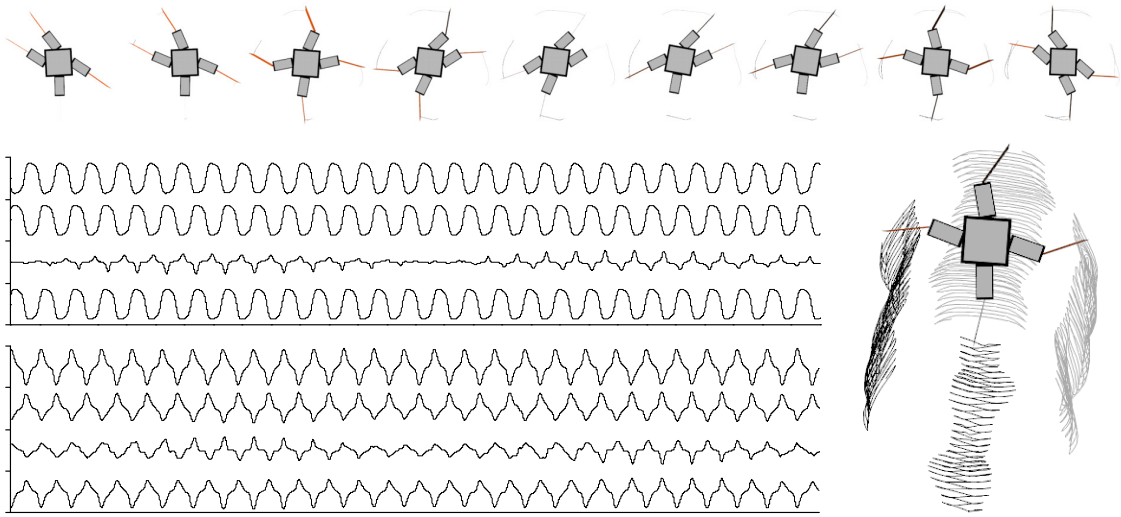

Figure 5.3: (Upper) Snapshots of the Peg-leg behaviour (PL-arm3) behaviour of 4-Fin Swimmer. Images were taken every 1/10 gait cycle. The tip trajectories of the fore (black) and rear (grey) fins are shown. The graphs show joint angles (upper) and fin bending angles (lower). Each segment along the vertical axis indicates the range [-1,1] rad.

those of previous case.[2] While the identical categorisation and naming were drawn for the behaviours of both cases, the precise motions show slight differences in that the sensor adaptation actively participates in the overall dynamics, resulting in the final interlimb coordination which converges to a certain state whose dynamics is dynamically balanced with that of sensor adaptation. The interlimb coordination can fluctuate (Figure 5.2D) or even show continuous shift (Figure 5.2C), however their qualitative behaviours are maintained by the body-environmental coupling.

Among these stable behaviours, the forward locomotion involves straight movements (ST), moving in circles (STC), and peg-leg (PL) motions. PL motions involve one of the arms moving with a small amplitude while the other three arms all use the same large amplitude. The phase relationship of the PL pattern is essentially similar to that of BA except that the amplitude of one arm is smaller than the others and its phase continuously shifts (with an irregular fluctuation) compared to the others,

[2]Video 1-5 at Shim (2012).

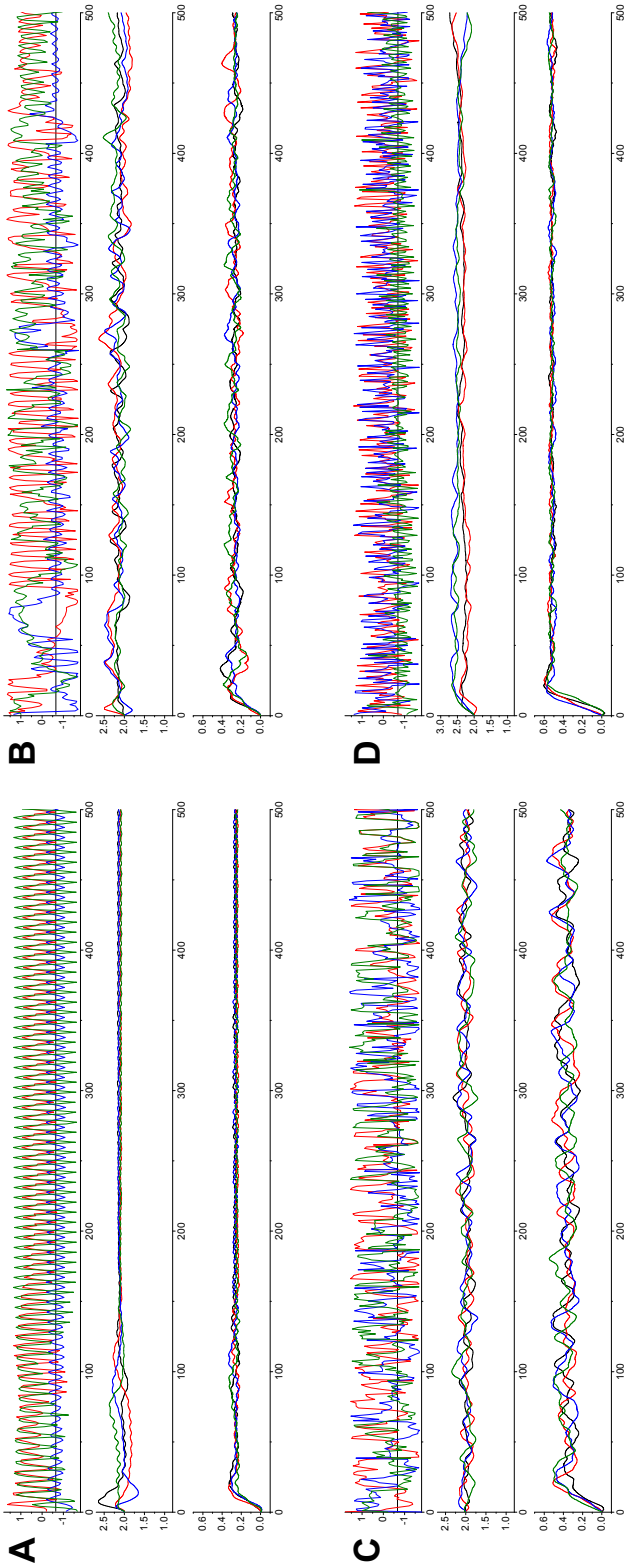

Figure 5.4: Time plot of the neural output and sensor adaptation variables ($A(t)$ and $B(t)$) of the system with different μ fixed to (A) 0.32, (B) 0.33, (C) 0.35 and (D) 0.37. The agonist signals of three motor units (${}^2x_l$, ${}^3x_l$, ${}^4x_l$) are plotted when ${}^1x_l$ crosses the central axis (vertical axis located at the unstable equilibrium point) of the periodic orbit. Each colour in the plots represents same data as described in Figure 4.2, and the colours for $A(t)$ and $B(t)$ are same as Figure 5.2.

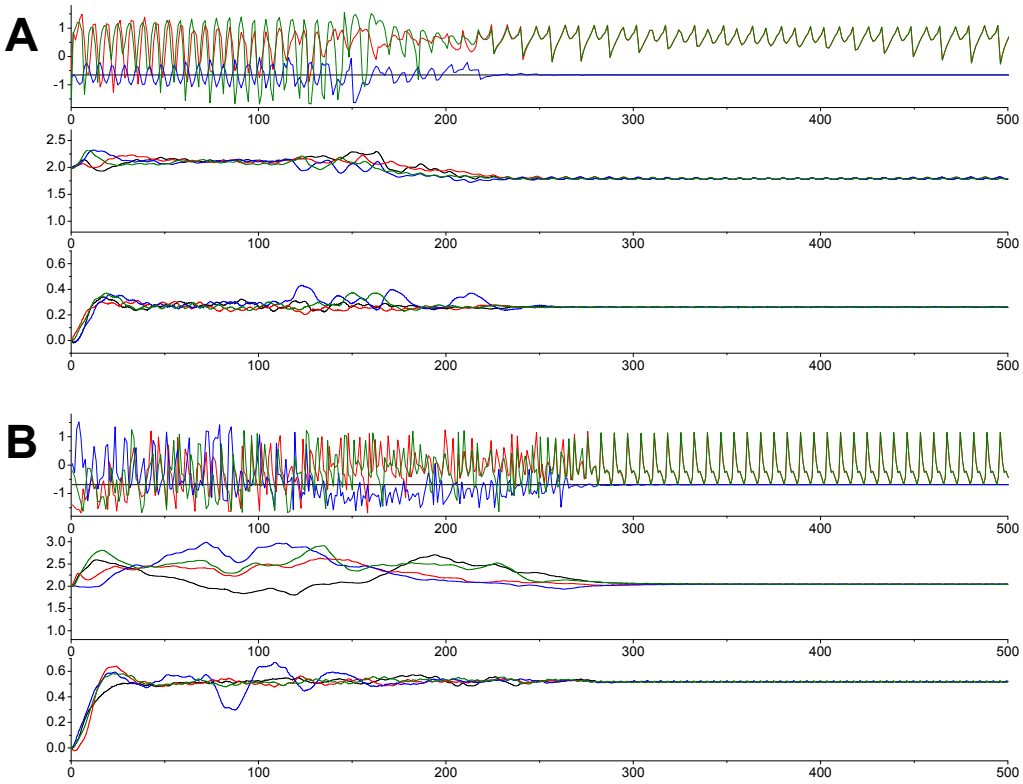

Figure 5.5: Synchronised behaviours emerged from the same μ as those of Figure 5.4B and D, but run from different initial conditions. (A) $\mu = 0.33$, (B) $\mu = 0.37$.

which achieves a slow forward locomotion by asymmetric propulsion forces. The sensor adaptation makes the lame arm synchronise with the corresponding motor unit with a small amplitude, resulting in the partial loss of the phase correlation with the other arms. This results in the motion of the other three arms being coordinated in such a way that the inertial/hydrodynamic forces are transferred less strongly to the remaining arm through physical embodiment.

The movements of arms in the rotation motion are out of phase with each other and fluctuate irregularly with relatively large amplitude. The fluctuation and shifting of phase relationship suggests that an emergent behaviour does not necessarily exhibit concrete phase locking between subsystems in the neuro-body-environment system when it is incorporated by dynamic sensor adaptation. Note that the PL patterns only appear when sensory homeostasis is present. Therefore, the homeostatic

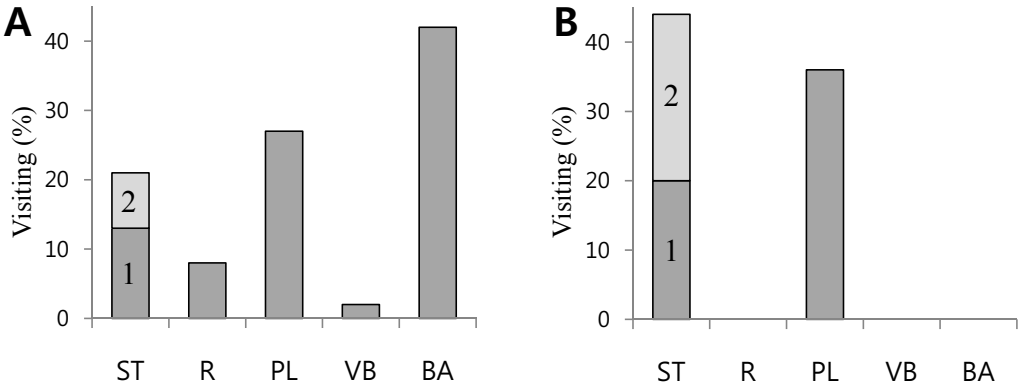

Figure 5.6: Visiting ratio of each patterns. Appearance of behaviours with (A) no control ($\mu = 0$) and (B) chaotic search. The ST (number 1) and STC (number 2) movements are depicted as a single bar. Although PL movements have low performances, they were frequently searched because of their high stability.

regulation of sensory signal results in an opposite effect which is the diversification of limb motion, that is, the multiple combinations of the amplitudes and offsets of limb motions can be explored and stabilised by sending the standardised sensory input signals to the neural controller. In turn, different limb-wise oscillations may cause different interlimb coordination as well.

5.2.3 Exploration of Stable Locomotor Patterns

Observation of the system behaviour at different fixed values of μ revealed that the stable dynamics of the system begin to fluctuate as the global bifurcation parameter μ increases, exhibiting a series of transient dynamics from quasiperiodicity to chaos (Figure 5.4). In the higher chaotic regime complex transitory dynamics similar to chaotic itinerancy occurs which drives the system to briskly explore the phase space.

The behaviour of the system in the chaotic regime with sensor adaptation allowed a slightly wider range of μ, such that the system did not stall at higher values of μ up to $\mu_{max}\approx 0.45$, where some of the system CPGs without sensor adaptation lost active oscillation at $\mu_{max}\approx 0.4$. Different from the non-adaptive case, multiple isolated 'attractors' were observed in some of the chaotic regime of the system according to different initial conditions, that the chaotic (Figure 5.4B,D) and non-chaotic (Figure

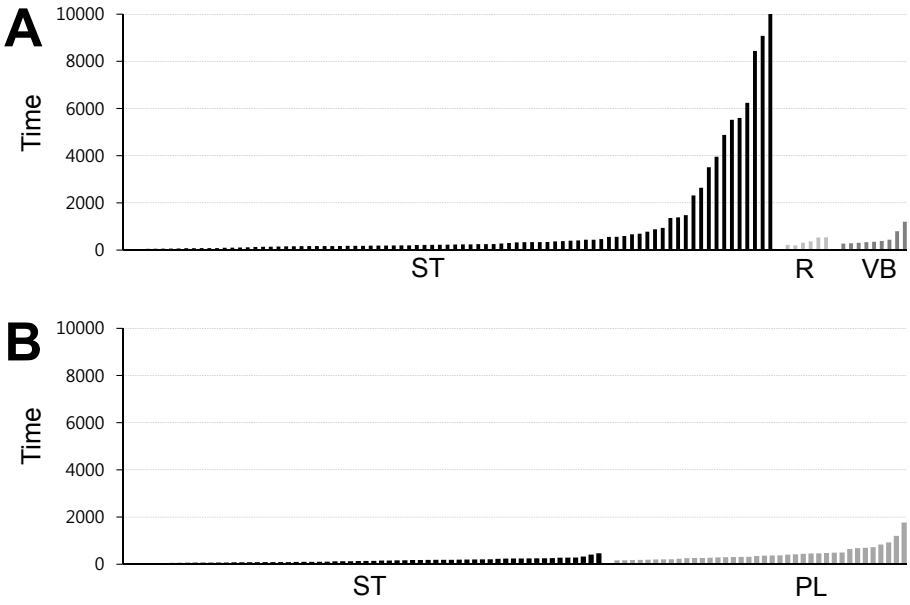

Figure 5.7: Final behaviours and their exploration time (oscillation cycles) until stabilisation without (A) and with (B) sensory adaptation. For each case 100 simulations were run and each behaviour was ordered by the time taken to stabilise.

5.5) behaviours coexist.

Like the preceding experiment, the distributions of visits to each of the behaviours identified in Table 5.1 were investigated under the presence and absence of chaotic search to see the effect. 100 simulations were performed for each case and the visiting counts of six major behaviours were recorded by observation. Figure 5.6 shows a clear difference between the visiting ratios of the two cases, suggesting the effectiveness of chaotic search which tended to settle on dynamically stable locomotion. During the search process all variables and control parameters vary continuously as parts of the neuro-body-environment system, and the time evolution plots of phase differences, performances, and bifurcation parameter (Figure 5.8A,B) show that the stabilisation and destabilisation of the system occurs repeatedly in a trial-and-error manner until it settles on an effective form of locomotion. The sensor parameters (Figure 5.8C,D) also change continuously and settle to different values via homeostatic adaptation.

It has been revealed that the use of sensory adaptation has helped to signific-

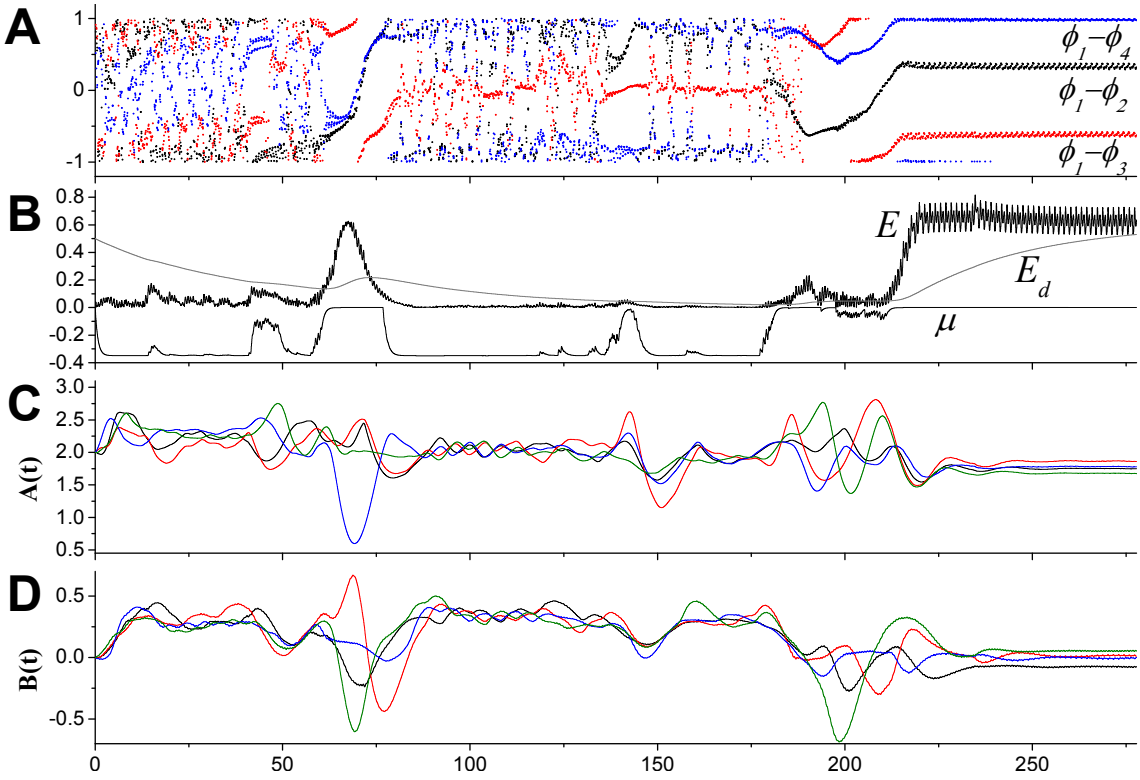

Figure 5.8: An example of exploration and stabilisation on to straight (ST2) locomotion. (A) Time plot of instantaneous phase differences between the arms 1-2, 1-3 and 1-4. (B) Performance (E), desired performance (E_d), and control parameter (μ). (C,D) Sensory adaptation parameters $A(t)$ and $B(t)$ in equation 7.

antly alleviate the issue of exploration deficiencies such as dead-end, bad-lock, and deep-path that the non-adaptive system have encountered in previous experiments. Although the deep-path motion appeared in a few tests, no appearance of dead-end and bad-lock was observed with the use of sensory adaptation. Figure 5.7 shows an example of the exploration time taken for stabilisation of the systems with and without adaptation. The fixed sensor gain of the non-adaptive system was chosen to produce a similar behaviour category to the adaptive case. While the adaptive system was stabilised within 1000 cycles in general, a number of runs of the non-adaptive system showed it could take up to ten times as long to stabilise compared to the adaptive system. The escape orbit is often stabilised to the PL patterns,

which indicates that the PL patterns are located in the vicinity of BA in the phase space.

5.3 Merging with a Dynamic Learning System

As the final building block for the system to achieve a continuous exploration-learning procedure, a fully dynamic oscillator learning algorithm is employed and integrated with the exploration process. Often, there are high performing locomotion patterns which are not completely stable and only appear for a while during the exploration process. These transient target behaviours can be captured and memorised by the oscillator learning process.

Figure 5.9 illustrates the final version of the integrated exploration and learning scheme, which has been progressively improved from Figures 3.1 and 5.1. The overall procedure of the exploration-learning process is explained as follows. As previously described, each robot joint has a dedicated motor unit (right side of figure) comprising oscillator-based central pattern generator neurons (CPG) with sensory input (S) and motor output (M). Sensory input undergoes homeostatic adaptation as it passes through a sensor adaptation module (SAM). This enhances the synchronicity between the neural and physical system thus allowing the neural system to cope with an arbitrary robotic system. An evaluation feedback signal controls a global bifurcation parameter that alters the chaoticity of the CPGs. Connections between the oscillators (expressed by broken lines) are initially inactive by setting their weights to zero. The chaotic dynamics of the neural-body-environment system drives a search process that finds motor patterns that perform well according to the evaluation criteria. As the system stabilises on a high performing pattern the bifurcation parameter reduces to zero and the connections between the oscillators become active, their weights being set by a learning procedure that is smoothly linked to the chaotic exploration process. In this way the learning process further stabilises, captures and memorises the motor patterns.

5.3.1 Learning by Adaptive Synchronisation

Since the wiring of oscillator coupling has to be dynamically modified in realtime in parallel with the exploration and discovery of useful patterns, we adapted a fast

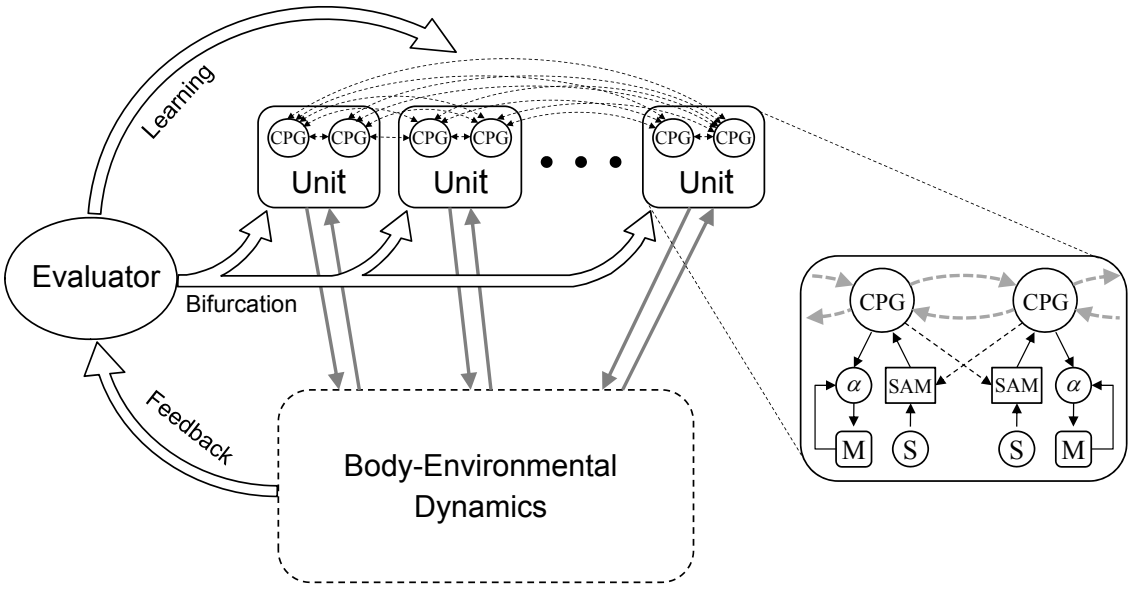

Figure 5.9: An overview of the integrated exploration and learning scheme.

and fully dynamic learning model developed by Doya and Yoshizawa (1992). The learning algorithm decomposes the problem of weight learning between oscillators into a collection of cellular-wise processes by adjusting the input connection weights (also called the phase-lock matrix) of individual neurons to maintain a given phase relationship between the cellular activity and incoming signals. This is available only when the phase relationship between the neuronal activity and input signals are presented in advance, which provides a suitable interface for our exploration system.

The learning algorithm for a network of identical oscillator was derived from the original work of Doya and Yoshizawa (1992). Let us consider n oscillators which are fully connected to each other. We denote the state j of the oscillator i as x_i^j and write a compact expression with the coupling term C which is the total input from other oscillator variables,

$$\tau \frac{dx_i^j}{dt} = f^j(\mathbf{x}_i, t) + C_i^j, \quad j = 1, 2, ..., m \tag{5.10}$$

where $\mathbf{x}_i = (x_i^1, x_i^2, ..., x_i^m)^T$ is the state vector. Assuming that the oscillators produce sinusoidal waveforms which can be approximated by sine waves (ex. $A\sin(\omega t +$

ϕ)), the phase-locked solution of the state vector of oscillator i and those of the other oscillators can be expressed as a linear relationship,

$$\mathbf{x}_i = \sum_{k=1}^{n} \mathbf{P}_{ik}\mathbf{x}_k, \quad i \neq k \tag{5.11}$$

where $\mathbf{P}_{ik}$ is the $m \times m$ phase-lock matrix for the oscillators i and k. Suppose we already have a certain phase relationship between $\mathbf{x}_i$ and other oscillators, then we can drive $\mathbf{x}_i$ in order to satisfy the equality in Equation 5.11 by using a simple error feedback to the oscillator using the gradient of an objective function E_i such that:

$$E_i = \frac{1}{2}\|\mathbf{x}_i - \sum_{k=1}^{n} \mathbf{P}_{ik}\mathbf{x}_k\|^2 = \frac{1}{2}\sum_{j=1}^{m}\{x_i^j - \sum_{k=1}^{n}\sum_{l=1}^{m} p_{ik}^{jl}x_k^l\}^2 \tag{5.12}$$

$$C_i^j = -g\frac{\partial E_i}{\partial x_i^j} = g(\sum_{k=1}^{n}\sum_{l=1}^{m} p_{ik}^{jl}x_k^l - x_i^j) \tag{5.13}$$

p_{ik}^{jl} represents the (j,l)th element of the matrix and g is a feedback gain which should be set small enough so that the ongoing oscillation is not distorted. Thus we can rewrite the above state Equation 5.10 by neglecting the small decay term gx_i^j in Equation 5.13 as

$$\tau\frac{dx_i^j}{dt} = f^j(\mathbf{x}_i, t) + g\sum_{k=1}^{n}\sum_{l=1}^{m} p_{ik}^{jl}x_k^l \quad i \neq k \tag{5.14}$$

We can see that the feedback term actually represents the coupling term from other oscillators in that gp_{ik}^{jl} is the coupling connection strength from x_k^l to x_i^j. The coupling matrix $\mathbf{P}_{ik}$ can be obtained using the same gradient descent learning with regard to p_{ik}^{jl}. In order to eliminate any bias effect, the deviation of signal from its temporal average ($\tau_v\frac{d\bar{x}}{dt} = -\bar{x} + x$) was used for learning,

$$\frac{dp_{ik}^{jl}}{dt} = -\gamma\frac{\partial E_i}{\partial p_{ik}^{jl}} = \gamma\{(x_i^j - \overline{x_i^j}) - \sum_{r=1}^{n}\sum_{s=1}^{m} p_{ir}^{js}(x_r^s - \overline{x_r^s})\}(x_k^l - \overline{x_k^l}) \tag{5.15}$$

where γ is the learning rate.

5.3.2 Integrated Exploration-Learning System

As the exploration process stabilises the system by discovering a high performing locomotor behaviour, the synaptic connections between oscillators are dynamically

wired using an adaptive synchronisation learning scheme. The coupling strengths are continually adjusted to follow the emergent patterns in parallel with the exploration process until the system is stabilised by discovering a desired pattern.

Since the learning is performed by adjusting the connection weights between individual CPGs, let us consider the equations for a single CPG i in compact form with the coupling term,

$$\tau \frac{dx_i}{dt} = f^x(x_i, y_i, t) + C_i^x \tag{5.16}$$

$$\tau \frac{dy_i}{dt} = f^y(x_i, y_i, t) + C_i^y \tag{5.17}$$

$$f^x(x_i, y_i, t) = c(x_i - \frac{x_i^3}{3} - y_i + z_b) + \delta F_x(H(s_i, t), x_i) \tag{5.18}$$

$$f^y(x_i, y_i, t) = \frac{1}{c}(x_i - by_i + a) + \varepsilon F_y(H(s_i, t), x_i) \tag{5.19}$$

The subscription for z_b becomes $b = ((i + 1) \bmod 2) + 1$. The sensory input term was regarded as a part of the oscillator dynamics to promote sensory influence in the global coordination of the learnt oscillator network. In order to improve the readability of the following equations, let us denote the states x and y of oscillator i as x_i^1 and x_i^2 and rewrite above equations in more reduced expression,

$$\tau \frac{dx_i^j}{dt} = f^j(\mathbf{x}_i, t) + C_i^j, \quad j = 1, 2 \tag{5.20}$$

where $\mathbf{x}_i = (x_i^1, x_i^2)^T$ is the state vector for the CPG i. Considering M (=$2N$, where N is the number of degrees of freedom of the robot) fully connected oscillators, the coupling term C for state j of oscillator i (x_i^j) can be written as:

$$C_i^j = g \sum_{k=1}^{M} \sum_{l=1}^{2} p_{ik}^{jl} x_k^l \quad i \neq k \tag{5.21}$$

$$\frac{dp_{ik}^{jl}}{dt} = \gamma\{(x_i^j - \overline{x_i^j}) - \sum_{r=1}^{M} \sum_{s=1}^{2} p_{ir}^{js}(x_r^s - \overline{x_r^s})\}(x_k^l - \overline{x_k^l}). \tag{5.22}$$

where g is a small feedback gain term and gp_{ik}^{jl} represents the adaptive connection strength coupling from x_k^l to x_i^j, which forms a covariance-like learning rule. $\overline{x}$ is the continuous running average of x calculated by $\tau_E \frac{d\bar{x}}{dt} = -\bar{x} + x$, where the time constant was set as same as that of the performance evaluation.

During the exploration process, the feedback gain g and the weight learning rate γ are adaptively adjusted according to the global control parameter μ so that the couplings between oscillators are gradually activated around the onset of system stabilisation. g and γ are controlled according to:

$$g = \alpha(1 - \eta), \quad \gamma = \beta\eta \tag{5.23}$$

$$\tau_\eta \frac{d\eta}{dt} = -\eta + D(\mu - \epsilon) \tag{5.24}$$

where α and β are constants, and $D(x)$ is the heaviside function with very small $\epsilon = 0.0001$. As the incoming weights are learned in order to match the sum of afferent signals close to the oscillator's signal, it is sufficient to use $\alpha = \sqrt{\delta^2 + \varepsilon^2}$ (input weights for sensor), which has similar intensity to the sensory input. $\beta = 1/\tau_E$ was set for Equation 5.22 to have the same time scale as the evaluator. η is the smooth activation signal which controls both the learning rate of connection weights and feedback gain according to the value of μ. It acts as a switching signal determined by the global bifurcation parameter. When η is triggered around the onset of system stabilisation, the decrease of the learning rate of the phase-lock matrix and the activation of oscillator couplings simultaneously takes effect.

The learning signal gradually activates the functionally connected network rather than suddenly switching it on, thus preventing the destruction of stable patterns while allowing unstable ones to be filtered out. Since the coupling is not strong and is activated gradually, highly unstable patterns which show short-lived high performance are naturally filtered by the instability of the pattern itself during the activation period (the system destabilises and returns to the exploration phase).

The learning rules are set up such that during the exploration phase the couplings effectively remain functionally inactive. As dictated by Equations 5.23 and 5.24 the coupling gain g is turned on when the bifurcation parameter μ goes to zero which means learning is activated when the system is stabilised to some discovered pattern. Otherwise ($\mu \neq 0$), the system is in an exploration phase and g is set to zero which turns off the learning. Thus, exploration and learning are merged as a continuous dynamical process such that the desired locomotion pattern is spontaneously explored, discovered, and memorised in a coherent way.

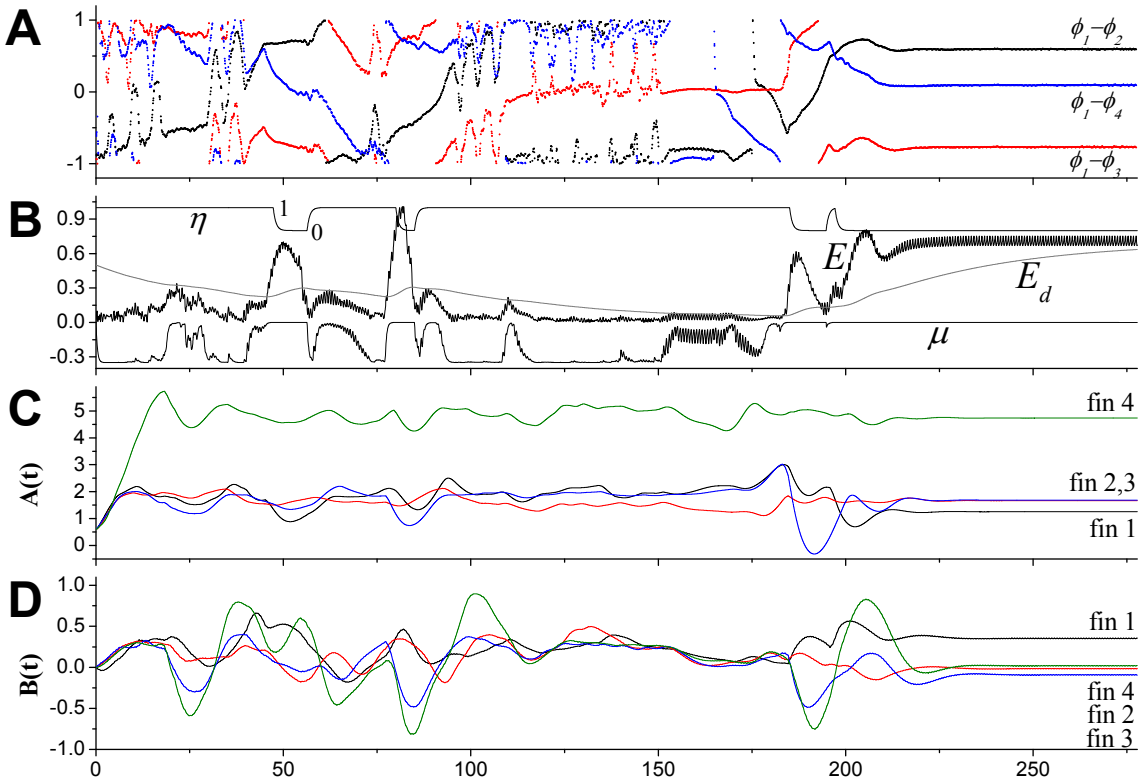

Figure 5.10: Exploration and capture of transient locomotor behaviour in damaged-fin swimmer by oscillator learning. The oscillator wiring is triggered (η in Eq.(13), 1:off, 0:on) at each stabilising trial and attempts to sustain the performance of the discovered pattern. The sensor parameters $A(t)$ and $B(t)$ cooperate with the exploration-learning for a given physical embodiment (see text).

5.3.3 Stabilising Transient Patterns by Oscillator Learning

First, the integrated exploration-learning system was tested using a 'damaged' version of the Swimmer robot by reducing the length of one of its fins (damaged fin) or removing one of its arms (3-armed), such that there are few or no stable patterns in the stable regime but there exist a series of useful transient patterns.

Figure 5.10 shows the exploration and learning of the robot with a damaged fin where the length of the fin on arm 4 was reduced by 90%. It had only one stable pattern whose phase relationship is the same as that of the BA pattern in the undamaged robot, which has almost zero performance. With learning, it captured

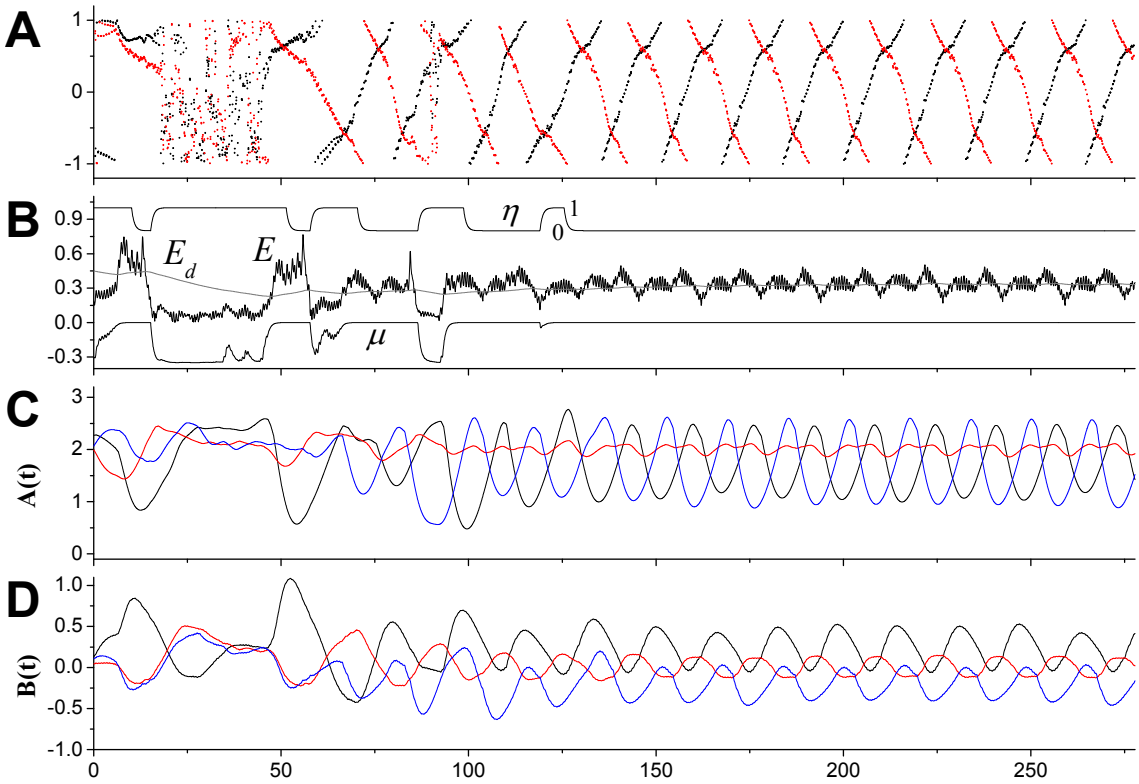

Figure 5.11: Alternating behaviour of 3-Arm Swimmer.

one of the high performing transient patterns after a few trials. The approximate direction of locomotion is toward dir-3. Figure 5.10C,D show that the sensor gain ($A(t)$) of the damaged fin (fin 4) was increased to amplify its signal, and the fact that fin 1 has the smallest gain tells us that arm 1 is the main source of propulsion. The salient deviation of the offset ($B(t)$) of the fin 1 sensor (opposite side of fin 2 and 3) indicates that the discovered transient pattern involves the oscillation of fin 1 in a tilted position, granted by its mechanical compliance; consequently it compensated the asymmetric hydrodynamic forces and achieved forward locomotion. The homeostatic sensory regulation participates in the exploration process as the slow variables diversify the course of transient patterns during search and slows them down at the onset of discovery, which is beneficial to the realtime pattern capture by oscillator learning.

Figure 5.11 shows a particular case of an alternative 3-armed robot (formed by

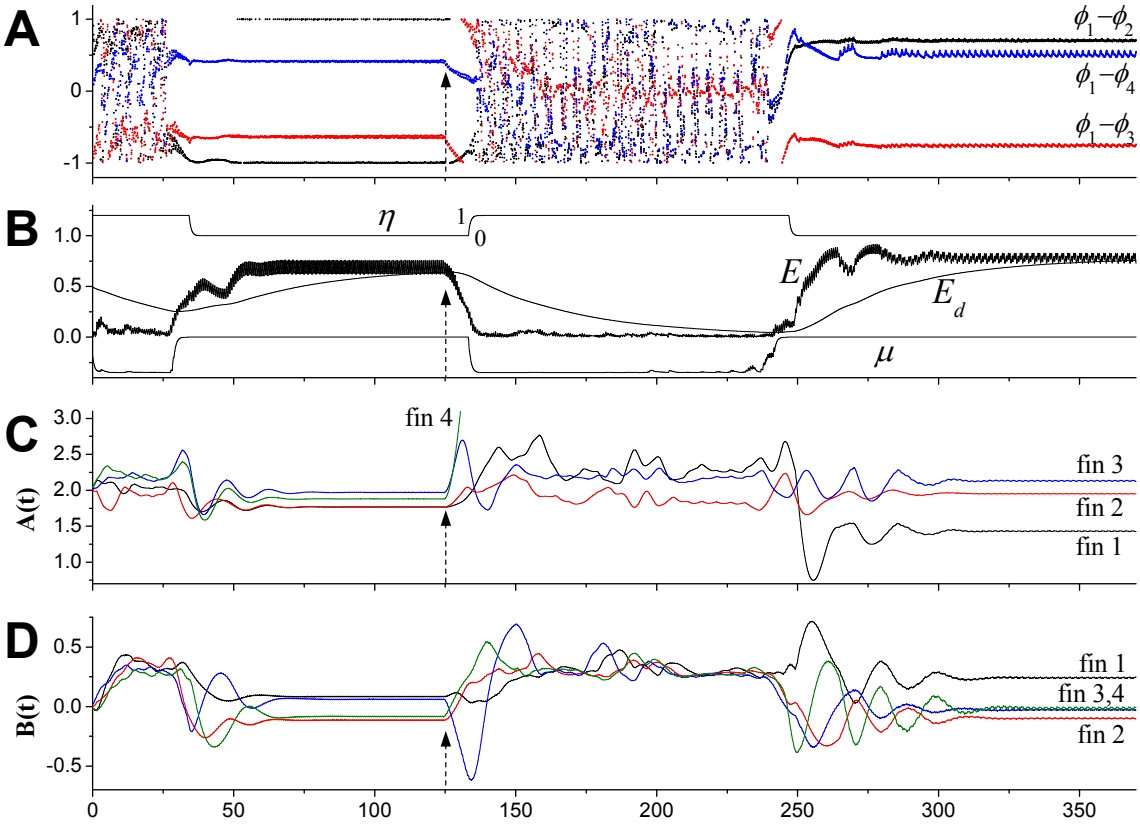

Figure 5.12: Realtime recovery after a radical change to the body (damage). Dashed lines and arrows indicates the time of damage, when the length of fin 4 is decreased to 1/10 of its original length. The sensor gain of (damaged) fin 4 ($A(t) \approx 5.0$) in (C) was truncated for a better view of the other gain plots.

removing arm 4) where two different locomotion patterns are periodically exchanged while not losing the overall stability of the whole behaviour.[3] The robot alternates its moving direction between dir-3 and dir-4 by exchanging two unstable undulating motions. The periodicity of this conjoined behaviour also exhibits a small degree of irregular fluctuation as in the case of loosely coordinated behaviours previously shown in Figure 5.2. However, being captured by oscillator coupling, it is sustained by global coordination between subsystems which include adaptive sensor dynamics.

Since the oscillator learning process is automatically regulated by a control

[3]Video 6 at Shim (2012).

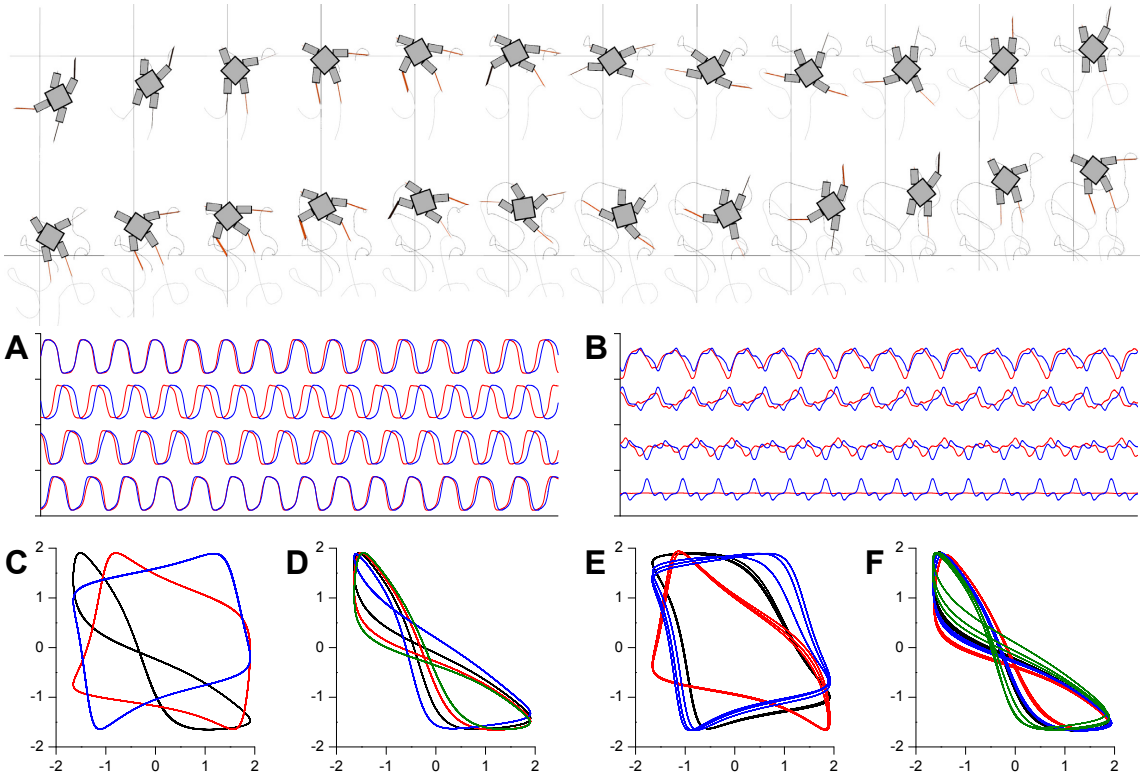

Figure 5.13: Uppermost images show the snapshots of the movement of 4-Fin Swimmer after re-adaptation (the movement before damage is virtually same as ST motion), which is continued from Figure 5.12. 24 images were taken every 1/10 gait cycle. The tip trajectories of the fore (black) and rear (grey) fins are shown. (A) and (B) show the joint angles and the fin angles respectively, where the undamaged motion (blue) and the readapted motion (red) are superposed. The fiducial point for the superposed plots was set to the starting point of arm angle 1 in (A). (C,D,E,F) illustrate the 2D plots of neural outputs before (C,D) and after (E,F) re-adaptation. (C) and (E) depict the relationships between agonist CPGs, where x-axis shows ${}^{1}x_l$ and y-axis is for ${}^{2}x_l$ (black), ${}^{3}x_l$ (red), and ${}^{4}x_l$ (blue). (D) and (F) show each output pair of the antagonistic CPGs (${}^{i}x_r$ vs. ${}^{i}x_l$ for each motor unit): arm1:black, arm2:red, arm3:blue, and arm4:green.

parameter (η) it is possible to operate the exploration-learning system continually without reset. Figure 5.12 shows a typical successful example of the realtime recovery of locomotion behaviour after body damage of an unknown variety, i.e. with no a priori knowledge.[4] During an initially learnt stable behaviour (similar to STC-dir3-CCW), the same damage as in Figure 5.10 was sustained. The performance of the robot immediately dropped below E_d and the system entered into the search phase. After a few hundreds of cycles the system found a new locomotion behaviour for the changed body (undulating movement similar to Figure 5.10). The superposed graphs of two behaviours (Figure 5.12E,F) show a slight frequency increase in arm movements after recovery due to the change of mechanical impedance of the robot.

5.4 Experiment with Terrestrial Walker

In order to investigate the generality of the approach, the system was applied to another body-environment system totally different from the previous swimmer model, which is a legged robot walking in a 3D terrestrial environment. For the convenient configuration of various body morphologies for simulation, the legs of walking robots were modelled by assembling multiple capped cylinders all of which have identical dimensions. In the following sections, the experiments using a quadruped robot is presented as a representative walking machine. More results of the quadruped and other terrestrial robots are presented in Appendix B (with corresponding movies in above URL).

5.4.1 Preparing Quadruped Robot

The quadruped body was configured as bilaterally symmetric as shown in Figure 5.14, and the Coulomb friction model with a coefficient of 1.0 was used. It has 8 motor units (16 oscillators) which is as many as the number of degrees of freedom. The sprawl posture of the legs and the proper setting of joint ranges ensures that the robot will not overturn during exploration.

Inspired by the vertebrate muscle proprioceptors in biology, the stretching force experienced by a torsional muscle was used as the sensory signal and fed to the

[4] Video 7 at Shim (2012).

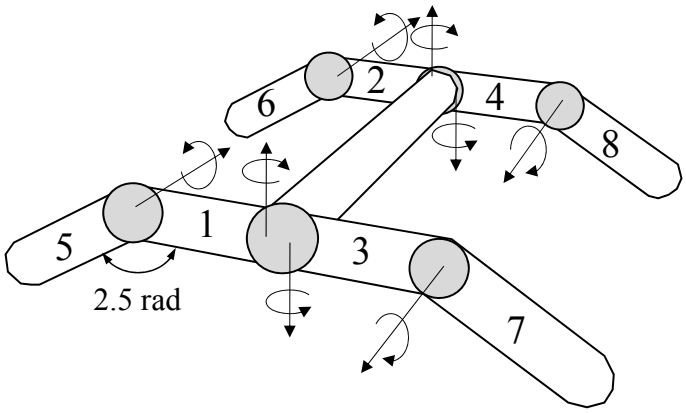

Figure 5.14: Quadruped model in 3D terrestrial environment which has 8 degrees of freedom. Arrows represent the axis of each joint, and the indices of limbs are marked as numbers. The index of each joint is assigned as the same as its child limb. The lower limbs (limb 5-8) are tilted down by a same angle, and all the joints in above posture are set to the neutral position (zero angle).

CPG in the relevant motor unit. The stretching force that a muscle experiences (the collective effect not only of the antagonistic muscle but also the whole-body motion) was chosen as the appropriate sensory signal which contains well-blended neural-body-environmental information. Neurobiological studies on the spinal or decerebrate cat (Conway et al., 1987; Pearson et al., 1992; Pearson and Collins, 1993; Pearson, 2008) suggest that a major component of sensory signals from the ankle extensor muscle receptors, which influence the central rhythm generating network, is related to the force in the muscles, which primarily arises from the Golgi tendon organs (Group Ib afferent), while the Group Ia afferent from muscle spindles has a lesser effect.

The muscle stretching force can be considered equivalently as the pulling torque of torsional muscle model. At any given time instance, if a torsional muscle is 'stretched' from its resting angle, the pulling torque that the muscle receives is proportional to the product of the muscle activation level (which linearly controls the muscle spring constant) and the stretched angle. Since the actual sensor value is processed by the sensory adaptation before it is fed to the CPG, it is sufficient to use a simple formulation which implements this relationship. Given the muscle

activation level σ and the joint angle θ, the pulling torque can be written in a processed form as

$$\sigma_l = \tanh(x_l + k_m s_{m,l}) \tag{5.25}$$
$$\sigma_r = \tanh(x_r + k_m s_{m,r}) \tag{5.26}$$
$$s_l = (\sigma_l + 1)(1 - \frac{\theta}{\Phi}) - 2 \tag{5.27}$$
$$s_r = (\sigma_r + 1)(1 + \frac{\theta}{\Phi}) - 2 \tag{5.28}$$

where $k_m = 0.1$ and s_m are the stretch reflex from the muscle spindle which is described in Equation 3.18-3.20. The equation was centered around zero to have a range [-2,2] and fed to the relevant SAM. The denominator Φ (as in Equation 3.19) normalises the angle and the angular velocity of torsional muscle by the unit of its resting angle, and it is set to the maximum available joint angle (refer to the joint range in Table A.1).

Because there are two separate sensor signals for a motor unit coming from each muscle, the sensor design matches the forms of original equations for a pair of coupled CPGs. The coupling functions $F(s, x)$ in Equations 3.9-3.12 becomes;

$$F_x(s_{l,r}, x_{l,r}) = H(s_{l,r}, t) - x_{l,r} \tag{5.29}$$
$$F_y(s_{l,r}, x_{l,r}) = H(s_{l,r}, t). \tag{5.30}$$

Therefore, the CPG equations of a motor unit for quadruped are written as:

$$\tau \dot{x}_l = c(x_l - \frac{x_l^3}{3} - y_l + z_1) + \delta(H(s_l, t) - x_l) \tag{5.31}$$
$$\tau \dot{y}_l = \frac{1}{c}(x_l - b y_l + a) + \varepsilon H(s_l, t) \tag{5.32}$$
$$\tau \dot{x}_r = c(x_r - \frac{x_r^3}{3} - y_r + z_2) + \delta(H(s_r, t) - x_r) \tag{5.33}$$
$$\tau \dot{y}_r = \frac{1}{c}(x_r - b y_r + a) + \varepsilon H(s_r, t) \tag{5.34}$$

where the time constant was set to $\tau = 0.4$ and the maximum available value of the global bifurcation parameter $\mu_c = 0.32$ was used for the quadruped. All other parameters were the same as the previous experiment.

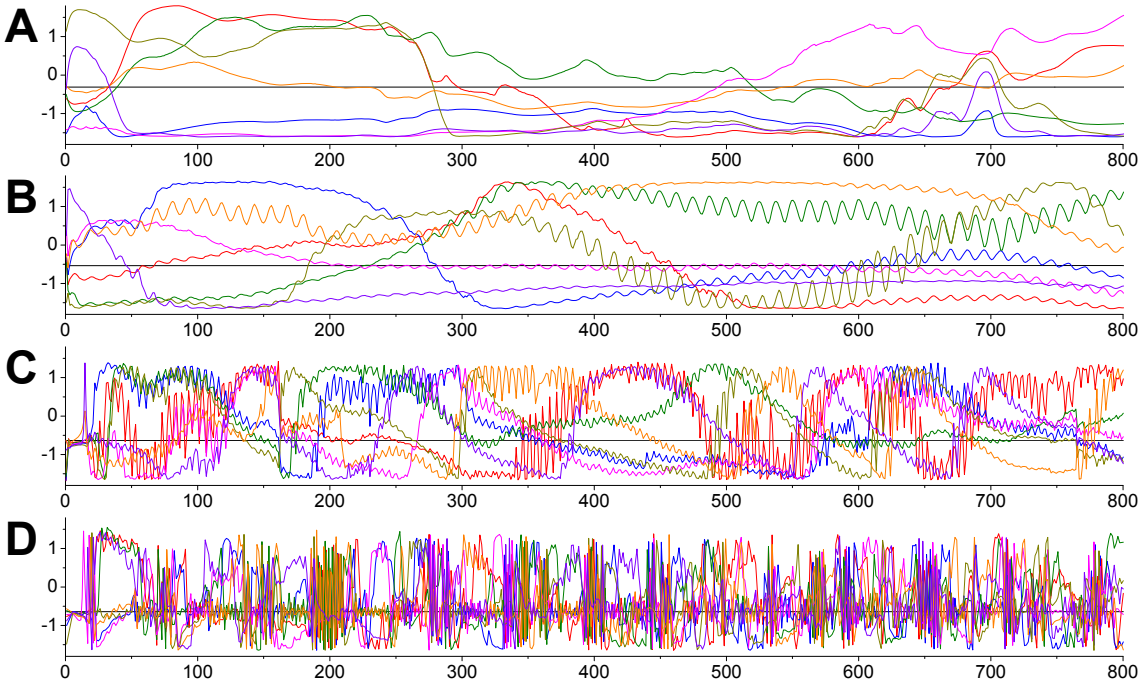

Figure 5.15: Time plot of the neural output of the quadruped system with different μ fixed to: (A) 0.0, (B) 0.2, (C) 0.3 and (D) 0.32. The agonist signals of seven motor units (from ${}^{2}x_l$ to ${}^{8}x_l$) are plotted when ${}^{1}x_l$ crosses its singular point of the periodic orbit. Colour codes are ordered by the index of motor units: 1-black, 2-red, 3-blue, 4-green, 5-magenta, 6-dark yellow, 7-orange, and 8-purple.

5.4.2 Experiment Result

Under conditions where static stability against gravitational force is guaranteed both in the 2D swimmer (balancing against gravity does not need to be considered) and 3D quadruped, the walking machine has fewer behavioural constraints for producing forward locomotion since the resistance force is not always present in the 3D terrestrial environment (for instance, there is no friction on a leg as it moves through the air during a swing phase). The neural-body-environmental phase space of the quadruped can be envisaged as an undulating landscape of rolling hills, while the 2D Swimmer case has a few deep basins of attraction. While this increased the number of candidate patterns for forward locomotion in the quadruped, there existed latent instabilities such as slipping due to dynamic friction or the spontaneous occurrence

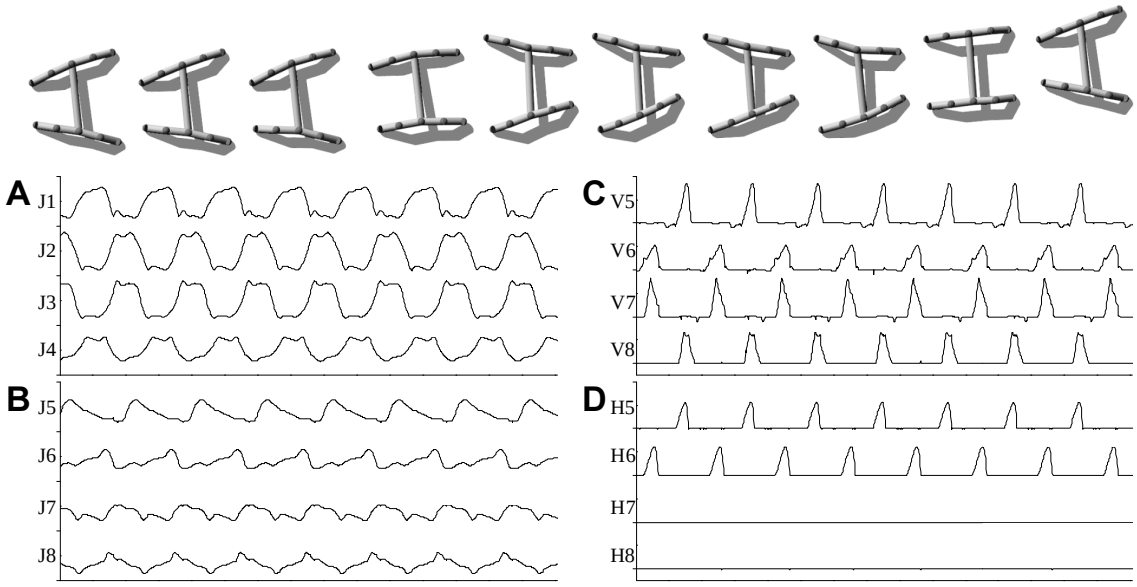

Figure 5.16: An example of a movement which is similar to the quadruped walking gait captured by the exploration-learning process. Snapshots were taken every 1/10 gait cycle. (A,B) The joint angles of limbs. (C) The horizontal speeds of each foot (the tips of limbs 5-8) in the direction of locomotion. (D) The height of each foot from the ground. The two rear feet (V7,V8,H7,H8) show stick-and-slip movements on the ground under Coulomb friction. The range of each plot is as follows; J1-J8: [-1.0,1.0]rad, V5-V8: [0.0,2.0]m/s, H5-H8: [0.0,0.08]m.

of sharp-amplitude, high-frequency perturbation stemming from the ground contact, all of which caused slow degeneration of the ongoing locomotor pattern.

In practice, the movement patterns of the quadruped which are observed in the stable regime of the system exhibited no ultimately permanently sustained behaviour (Figure 5.15A), this was also true in tests on other walking robots. However, the transient patterns which were observed in the stable and quasiperiodic regimes exhibited several working locomotor behaviours. Among them, interestingly, locomotor patterns which are similar to the quadruped walking gait frequently emerged during exploration (Figure 5.16). Other kinds of as-it-could-be gait patterns and their variations which exploit given body compliance also emerged, which are difficult to categorise qualitatively.

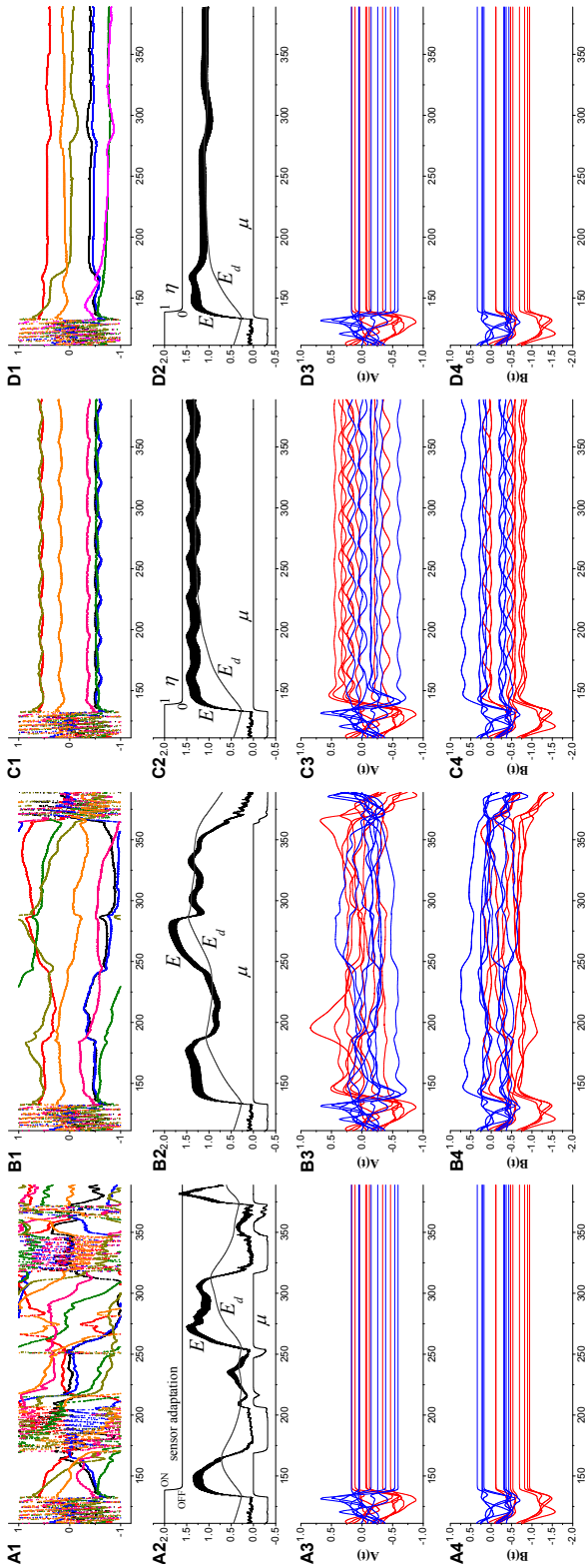

Figure 5.17: Four experiments (A,B,C,D) of quadruped exploration with different system settings starting from the same initial condition. (A1-D1) depict the phase differences between the limbs 1-2:black, 1-3:red, 1-4:blue, 1-5:green, 1-6:magenta, 1-7:dark yellow, 1-8:orange. The system behaves differently after the discovery of the first pattern ($t \approx 140$). The blue and red lines in (A3-D3) and (A4-D4) represent the parameters ($A(t)$, $B(t)$) of the sensor functions $H_l(s_l)$ and $H_r(s_r)$ respectively. (A1-A4) No sensor adaptation nor oscillator learning was enabled. The sensor adaptation was turned off at the time of the first stabilisation and their values were maintained throughout the running period. The patterns degenerate quickly and the system is hard to stabilise. (B1-B4) Sensor adaptation was turned on. The initial pattern changes much more slowly and its performance stays at higher values for a long time until the system re-enters into the search phase following performance loss. (C1-C4) The pattern was captured by oscillator learning and completely stabilised. (D1-D4) Same as C but the sensory adaptation was turned off as in A. Initially captured pattern was slightly shifted after fixing sensor parameters and showed slow degeneration over a long time.

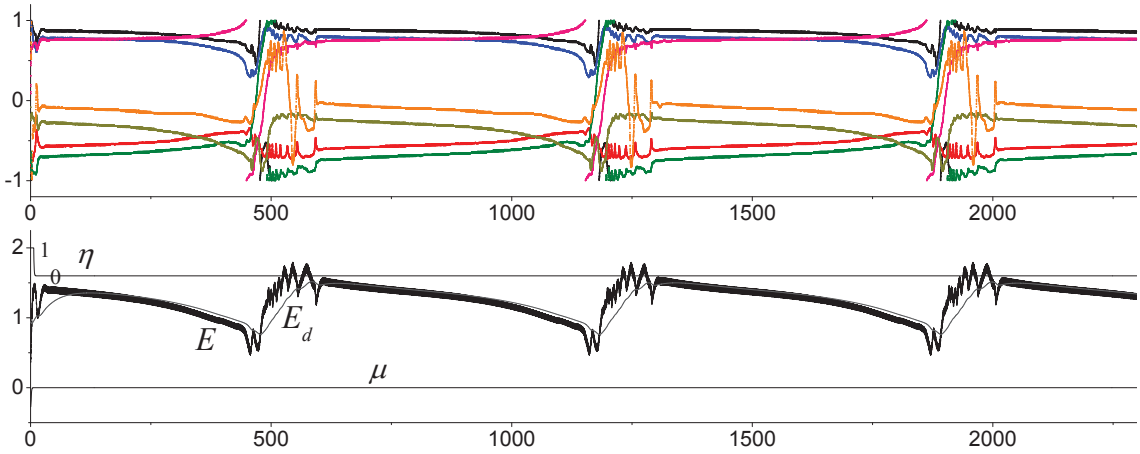

Figure 5.18: Long-term periodicity of quadruped behaviour observed in a small number of cases. The coordinated pattern (straight walking) slowly shifts over a period of hundreds of cycles, followed by a short catastrophic transition (circling motion by asymmetric gait) and then re-entrained to the same walking pattern.

The degeneracy of locomotor behaviour could be greatly improved by using homeostatic sensory adaptation and then completely stabilised by oscillator learning. Figure 5.17A and 5.17B show the system behaviours for quadrupeds with and without sensor adaptation. All experiments were started from the same initial condition. In Figure 5.17A the sensor adaptation was turned off when the system was stabilised to the first discovered pattern. The performance of the emergent pattern in the adaptive system (Figure 5.17B) degenerate much slower than in the non-adaptive case. Sensor adaptation prevented abrupt changes in phase relationships by buffering sudden changes of incoming sensor signals, so the initial movement pattern slowly changed, giving it a greater probability of being maintained. In the case of terrestrial walking on flat ground, the sharp noise-like signals arises from the discrete nature of the ground friction environment. In simulation this can occur due to numerical inaccuracy in ground contact calculations, and in the real world it may come from small irregularities in surface friction or the slight ruggedness of a flat surface, or both. These sharp signals are regarded as undesirable for the purpose of stabilising and learning the emergent behaviours since it slowly destroys them.

The patterns could be completely stabilised by introducing oscillator learning

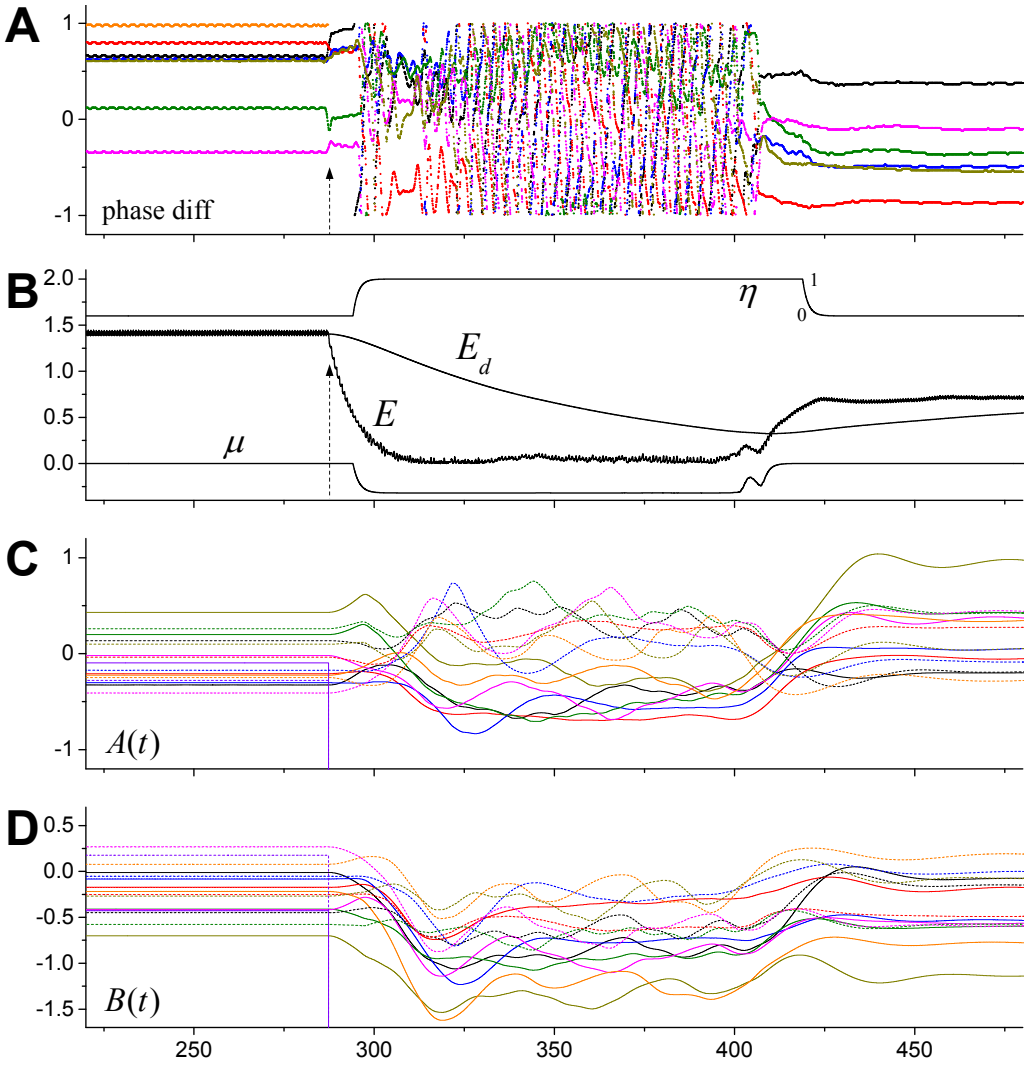

Figure 5.19: Realtime recovery after a radical change to the body (damage). Dashed arrows indicates the time of damage, that the limb 8 is removed. (A) depict the phase differences between the limbs 1-2:black, 1-3:red, 1-4:blue, 1-5:green, 1-6:magenta, 1-7:dark yellow, 1-8:orange. (B) Performance plot. (C,D) The solid and dashed lines in (C) and (D) represent the parameters ($A(t)$, $B(t)$) of the sensor functions $H_l(s_l)$ and $H_r(s_r)$ respectively. Colours are for joint: 1:black, 2:red, 3:blue, 4:green, 5:magenta, 6:dark yellow, 7:orange, and 8:purple.

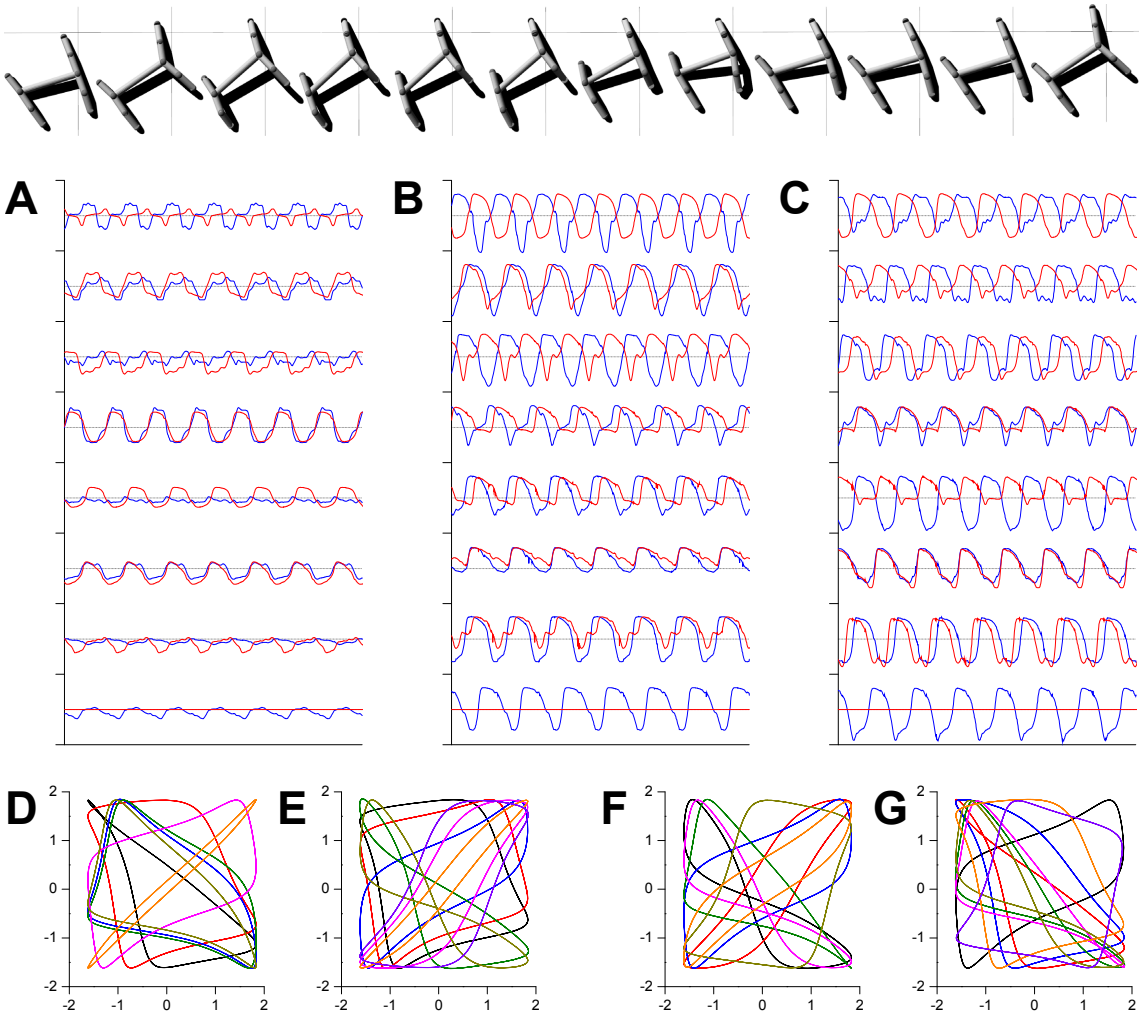

Figure 5.20: Uppermost images show the snapshots of the movement of Quadruped after re-adaptation (the movement before damage is virtually same as quadruped walking gait), which is continued from Figure 5.19. 12 images were taken every 1/10 gait cycle. Joint angles (A) and the raw sensor values of agonist (B) and antagonist (C) muscles are depicted, where the undamaged motion (blue) and the readapted motion (red) are superposed. The fiducial point for the superposed plots was set to the starting point of joint angle 4 in (A). (D,E,F,G) illustrate the 2D plots of neural outputs before (D,E) and after (F,G) re-adaptation, which represent the output of CPGs as same as that of Figure 5.12. Colours for (D) and (F) are: ${}^{2}x_l$:black, ${}^{3}x_l$:red, ${}^{4}x_l$:blue, ${}^{5}x_l$:green, ${}^{6}x_l$:magenta, ${}^{7}x_l$:dark yellow, and ${}^{8}x_l$:orange. Colours for (E) and (G) are for joint: 1:black, 2:red, 3:blue, 4:green, 5:magenta, 6:dark yellow, 7:orange, and 8:purple.

(Figure 5.17C). However, if oscillator learning was presented without sensory adaptation (Figure 5.17D) the pattern could not be sustained completely because the oscillator coupling was not strong enough to maintain the coordinated pattern against the degeneracy. In a few cases, the speed of degeneracy under the control of the learnt oscillators after adaptation is so slow that the locomotor behaviour, which appears stable, is eventually destroyed after a very long period of simulation, which triggers a period of re-adaptation. This appears in the form of a long-term behavioural periodicity (Figure 5.18), which is analogous to a slow version of alternating pattern observed in the 4-Fin Swimmer experiment as shown in Figure 5.11. The realtime recovery of locomotion behaviour after body change was also tested in Figure 5.19 and 5.20.[5] During an initially learnt stable behaviour (quadruped walking gait), limb 8 was suddenly detached from the rest of body. The system found and stabilised to a new locomotion behaviour for the changed body within 200 cycles. The new locomotion was deviated from the human expectation (for example, limped version of quadruped walking gait), in that it crawled sideways by reversing the phase relationships between ipsilateral legs.

5.5 Summary

This chapter has presented two dynamic methods for improving the basic exploration system which are the homeostatic sensory adaptation and the oscillator learning. Each method was implemented as a realtime dynamical process, and they were integrated with the exploration system to be upgraded into a larger system. The sensor adaptation dynamically adjusted incoming sensor signals in order that each CPG receive a regulated input of a proper intensity and offset to deal with an arbitrary robotic system. The oscillator learning was able to capture and sustain useful transient patterns by rewiring the CPGs dynamically using adaptive synchronisation process. To investigate the generality of the improved system, the system was tested using a quadruped robot which is a new body-environment system totally different from 4-Fin Swimmer. Both robotic systems were able to simultaneously explore, capture, and maintain a useful transient pattern. The sensor adaptation actively participated the exploration process as a part of the whole dynamics and contributed

[5] Video 7,8 at Shim (2012).

to the emergence of new stable patterns as well as slowing the speed of degeneration of ongoing transient patterns. The learning was cooperated by sensor adaptation to make the capturing of transient pattern easier by increasing the sustainability of the pattern.

Comparative observation of the two different robotic system revealed that the role of homeostatic sensory adaptation becomes more prominent in the case of terrestrial behaviours. The experiments with the 2D swimmer have shown little variance of sensor parameters after convergence, and pattern degeneracy was hardly observed, which indicates that the transient patterns of the swimmer are strongly attracted to a small number of stable patterns. This is because the swimmer robot is fully surrounded by hydraulic resistance and continually influenced by counteracting forces, what we call a 'densely structured' physical environment, which results in the phase differences between the arms of a symmetric shape yielding only a few highly symmetric configurations. While the case of the Swimmer robot has shown a relatively limited variety of patterns due to its strong embodied coupling resulting from the densely structured physical environment it inhabits, the environmental forces in the case of terrestrial robots are relatively discrete and intermittently exerted, which allowed more diverse coordinated limb motions. It was seen that the adaptation of the sensor parameters of the quadruped yielded more diverse values, where the offset parameter ($B(t)$) of lower leg muscles (leg 5-8) typically showed notable deviation under the effect of constant body weight.

More results of the quadruped and other terrestrial robots are presented in Appendix B with the URL for their movies. All the cylinder-shaped limbs used in the robots have the same dimensions and weights. The parameters of non-cylindrical body parts were separately specified with images. All robots used the same muscle parameters for all joints.

Chapter 6

Discussion

6.1 Summary

We have presented an integrated system which can explore and learn the emergent behaviours of a neuro-body-environment system coupled through physical embodiment by applying a chaotic search method. The overall neural architecture generalised and extended that presented in (Kuniyoshi and Sangawa, 2006). Each joint in an articulated robot was connected to a motor unit comprising a pair of central pattern generator (CPG) neurons which receive sensory input. The neurons produced motor outputs for an antagonistic muscle pair that control the movement of the joint. The CPG neurons were all connected to each other but these connections were initially made inactive. The CPG neurons received sensory signals that integrate information from the body-environment interaction dynamics experienced by the system (e.g. from force and position/angle sensors). Hence, while the direct connections are inactive, any coupling between the oscillators are indirect via bodily and environmental interactions. The network of oscillators, coupled through physical embodiment, had multiple synchronised states (modes) that reflect the body schema and its interactions with the environment, each of which can be regarded as a potential candidate for meaningful motor behavior. The whole system was treated as a single high dimensional dynamical system containing intrinsic chaotic dynamics as a necessary driving force for the exploration of its own emergent patterns.

The exploration process, powered by adaptive bifurcation through the feedback evaluation signal, allowed the system to be continuously driven between stable and

unstable regime, until a pattern if found that is sufficiently stable and high performing for the bifurcation parameter to reduce to zero and the system to fully stabilise. The evaluation signal was determined by the ratio of the actual performance (e.g. forward speed) to the desired performance. Since the method is intended for use in the most general case, where the robotic system is arbitrary, we do not have prior knowledge of what level of performance it can achieve. Thus the dynamics of the desired performance were modelled as a temporal average of the actual performance, such that the expectation of a desired goal is influenced by the history of the actual performance experienced, resulting in a simple adaptive evaluation signal. The search process was completely deterministic, and was able to selectively entrain the system orbit to one of the patterns by imposing goal directedness toward a desired behaviour.

The adaptive calibration of incoming sensor signals was established by using homeostatic sensory regulation. The sensor signal fed to a CPG neuron underwent homeostatic adaptation as it passes through a sensor adaptation module (SAM) before reaching the neuron. The SAMs were introduced because by adjusting the waveforms of input signals to be close to those of the neural activities, the synchronicity between the neural and physical system was enhanced thus allowing the neural system to cope with an arbitrary robotic system. This regulation also resulted in a diversification of output behaviors, increasing the scale of the search process, in that the same neuro-sensory coordination could be achieved by different limb movements, accomplishing multi-scale exploration.

The discovered rhythmic pattern is memorised and sustained by wiring initially disconnected oscillators using an adaptive synchronisation method. As the system stabilises, the connections between oscillators were dynamically activated using an adaptive synchronisation learning scheme. The learning rule was controlled by the bifurcation parameter and was set up such that the connections between the oscillators are effectively zero (inactive) during the exploration process but gradually adjusted (become active) as the system nears stability. The learning system was able to capture and maintain the high performing locomotor pattern by using the pattern itself as a supervising signal. Thus, exploration and learning has been merged as a continuous dynamical process such that the desired motor behavior is spontaneously explored, discovered, and memorised in a coherent way.

The system was evaluated by using it to control a range of realistically simulated

articulated robots that were required to locomote in an effective way. The generality of system was demonstrated by a series of experiments with a swimming robot and a variety of walking robots of differing morphologies. In each case a range of stable locomotion behaviors were discovered and learnt. It was also shown that the robots can readily readapt after radical body change.

6.2 Chaos as Source of Creativity?

Chaos is often referred to as a source of creativity which provides a steady and controllable source of noise (Skarda and Freeman, 1987, 1990). The overall process from the perspective of creating a new behaviour can be briefly sketched as follows. The mutual entrainment between neural system and physical embodiment initially creates a phase space that contains several stable and transient patterns. If the current entrained state is not satisfactory, the system bifurcates to a chaotic state in order to escape from that state and restabilises when a system performance meets the criteria. However, the phase space of the restabilised system differs from the previous one because some of the system parameters (sensor parameters: let us consider them as parameters, not variables, although both interpretations are possible) have also been changed by the chaotic drive. If we define the onset of stabilisation (at the time μ becomes 0) as the time of returning, whenever the state orbit returns to the target space, it never experiences exactly the same phase space as before. This process is what we call multi-scale exploration, and its eventual behaviour after the onset of stabilisation varies over the different physical embodiments.

The final dynamics of sensor adaptation after returning involves each sensor parameter being locked around a particular value (potentially different for each parameter) with small oscillations. This diversity of sensor parameter convergence can be regarded as the 'neutral stability' of the system since different motor movements can cause the same sensory input to CPGs. Conversely, it can be thought that the same coordinated pattern of CPGs can result in different limb motions. However, for the case of the 2D swimmer, which has a small number of strong basins of attraction, the sensor parameters tend to converge to one of the previous distributions although their precise values may differ. The neutrality in the convergence of sensor parameters has a wider range in the case of the quadruped, hence

more diverse stabilised behaviours are exhibited. Even in the case where the sensor parameters eventually converge to the same set of distributions, the intermediate trajectories before convergence can take various routes, which can be captured by oscillator learning, resulting in the 'creation' of a new behaviour. Therefore, this process differs from a simple action selection mechanism where predetermined stable patterns are selected by a chaotic jump. Rather, it creates various streams of transient patterns by driving both the state orbit and the system parameters using chaotic dynamics.

6.3 Biological Relevance

As Kuniyoshi and Sangawa (2006) have stated, completely decoupled CPGs are an extreme model which might deviate from biological reality. However, a few studies point out the biological evidence for the functional decoupling of neural system and its importance in the emergence of new behaviour. It has been hypothesised that decoupling of locomotor CPGs serves as a potential mechanism for the evolution of novel behaviors (Dubbeldam, 2001). Motion analysis of Great Siren (*Siren lacertina*), an eel-like amphibian (Azizi and Horton, 2004) has found strong evidence that the axial and appendicular CPGs are decoupled during 'aquatic walking' (the pattern somewhere between aquatic and terrestrial locomotion), which supports the hypothesis that the decoupling of CPGs has led to the evolution of this novel locomotor behavior. In a broader perspective, Rosslenbroich (2009) pointed out that the locomotor neural processes of more evolved vertebrates are uncoupled from one another so that these parts can act in more differentiated and partly independent ways, which may contribute to the increase in organismic autonomy necessary for evolutionary innovation.

These emergent patterns may be refined and selected at the supraspinal level by reward based reinforcement, which is thought to be one of the primary functions of the Basal Ganglia (BG) (Redgrave et al., 1999; Schultz, 2006; Chakravarthy et al., 2010). Recent modelling studies on BG (Sridharan et al., 2006; Magdoom et al., 2011) hypothesise that the indirect striato-pallidal pathway through the subthalamic nucleus subserves exploratory behaviour for goal-directed learning, gated by the dopamine signal from the substantia nigra which serves as the global learning signal

for reward prediction. We hypothesize that goal-directed chaotic exploration may possibly take a role in such mechanisms in connection with self-organised behaviours. In this context, it might be possible to our system to draw some implications about the previously mentioned optimal parameters in relation to the metalearning and neuromodulation centered around the BG (Doya, 2002).

6.4 Future Directions

6.4.1 Predefined System Parameters

Predetermined parameters primarily lie in the control part of the system. The CPG model and its parameters in this work were adopted from previous work (Asai et al., 2003a) whose parameter setting was rigorously studied in terms of exhibiting various dynamical regime from stable to chaos which meets the purpose of our work, and they were also verified by later studies using robotic platform as in (Kuniyoshi and Sangawa, 2006; Kinjo et al., 2008). Although our system has demonstrated a good degree of generality and an ability to automatically adapt to unknown bodies and environments, further analysis is necessary in order to determine the optimum values of fixed parameters used in the search process. One of the future directions would be to determine the effect of fixed parameters, such as time scale of each adaptive sub-processes.

The precise determination of the optimal parameters for evaluation and feedback bifurcation requires a massive number of simulation experiments together with formulating a proper definition of the system performance related to those parameters. For instance, the problem of measuring the 'performance' of overall system dynamics in terms of explorability is vague. One possible solution would be to investigate how frequently the system orbit visits each patterns in the stable regime of the system during the exploration phase. In the previous studies on chaotic search (Nara and Davis, 1992; Aida and Davis, 1994), they have provided a partial solution to this problem by observing the system orbit wandering through the basin of attractors which are identified by roughly dividing the state space of the system. This is a very complicated task but was possible because the target system implemented the discrete state memory patterns using binary neural elements, and the degrees of freedom of the system was exactly known in advance. However, our system is expressed

as a continuous dynamical system, and the complete set of variables is not known. Moreover, different from previous work, our system is not supposed to search a specific pattern which has a certain evaluation value, but to deal with the open-ended search in a way that the system tries to maximise the performance, which makes it more difficult to quantify the goodness of system over different robotic systems.

However, a few qualitative insights about the role of some parameters can be sketched. For example, the time scales of slow dynamics such as evaluation (τ_E), goal seeking (τ_d), and feedback bifurcation (τ_μ) impact the search dynamics. Preliminary results of investigating the effect of different time scales revealed that the ratio between the time scales for evaluation, goal seeking, and feedback bifurcation determines the balance between the 'memorising' and 'forgetting' of patterns during the search process (Aida and Davis, 1994), implying there might be an optimal ratio which allows the system to stay in the chaotic regime for an optimal duration (just enough to be uncorrelated with the previously visited pattern) enabling fast search with a very small probability of being trapped in a bad state for a long time.

The time scale of the sensor adaptation (determined by τ_h) can influence the landscape of phase space as well as the neutrality of convergence. For example, a test using the 2D swimmer showed that when τ_h was decreased by 1/2 (double the speed of adaptation), STC and VB pattern disappeared in the stable regime of system. When decreased to 1/4, a new stable pattern appeared where the two diagonal arms moved with large amplitudes whereas the movements of the other two were small and fluctuating, and the phase differences between neighbouring arms drifted continually (reminiscent of the PL motion with two fluctuating arms). If the time constant was doubled (half the adaptation speed), PL motion disappeared. Generally, too slow adaptation caused the long-term fluctuation of seemingly stable behaviours by eventually collapsing them toward more stable behaviours, too fast time scale caused large fluctuations in sensor parameters, being synchronised with the fast state dynamics, which tended to diminish the diversity of behaviours.

While this work has concentrated mainly on the interlimb coordination between the identical neuromuscular modules, the frequency of CPGs is an important factor. The future direction would be to investigate the effect of different CPG frequencies. Since we use an arbitrary robotic system, one of the interesting options would be to adopt the concept of the adaptive frequency oscillator whose frequency can be tuned to the resonant frequency of the mechanical system (Buchli et al., 2006; Righetti

et al., 2006).

6.4.2 Influence of Physical System

Another factor which influences the system is the amount of bandwidth resulting from the design of body-environment interactions. In the case of the 4-fin swimmer presented here, the functional coupling strength between motor units varies with the ratio of body inertia to environmental forces (similar concept to the Reynolds number in fluid mechanics). Increased body mass will result in an increased moment of inertia which causes less transmission of the hydraulic force from one leg to the others, and vice versa. A similar effect will be caused by decreasing the density of the surrounding fluid or by increasing fin joint stiffness.

One possible future direction would definitely be to investigate the effect of increased body compliance by e.g. using flexible limbs. For legged movement, the flexible limbs will soften the sharp discreteness of ground reaction forces in the surface normal direction at the time of swing-stance transition. On the other hand, the elastic limb may cause unexpected snap out in the tangential direction. It could also be possible to get richer behaviours by using flexible limbs for the planar swimmer. However, the current model provides the sensory information available in every compliant body part. Therefore, without incorporating proper sensors for the newly added body compliance, whether there would be considerable effect is questionable in some cases. For instance, the flexibility of the limb of the 4-Fin swimmer may function as merely another physical property of passive fin bending. Introducing additional sensors raises a new problem of sensor fusion, where the local sensorimotor circuit needs to deal with multiple sensor signals being fed to a single CPG.

6.4.3 Experiment with Real Robot

Transfer to robotic hardware is one of the prerequisites for the validation of system performance. The most influential factor in dealing with a real robot would be the presence of noise. In the chaotic regime of the system, the noise will prevent the system from being trapped in the undesirable states such as bad-lock and deep-path. However, it will disturb the system in stabilising the desired locomotion

behaviours. The disturbance in stabilisation could be overridden by the oscillator learning process, which strengthen the asymptotic stability of ongoing motor pattern by wiring neural elements. However, the actual formation of the desired locomotor pattern usually begins before the triggering of the oscillator learning process, which means that the emerged pattern has to be sustained until the learning takes place. The noise might increase the possibility of the destruction of the locomotor pattern in its early stage of formation, but the counter-possibility also exist that the noise can push the system orbit to better patterns. Although the noise can be viewed as a participant of system, at least it is necessary that a mechanism exists to maintain the ongoing patterns during the intermediate stage of the exploration process. Also, the wiring of the neural controller by oscillator learning should be strong enough to maintain the ongoing locomotion behaviour under the presence of noise.

Another minor issue which arises from the practical implementation (although it is termed minor, it becomes a major problem when the robot is built using limited resource) include the measurement of locomotion performance. Self-awareness of the position and speed of a legged robot is a nontrivial problem. One of the most popular methods for self positioning is the global positioning system (GPS) using an external module such as camera or satellite. Apart from the problems of resource and implementation, GPS systems usually have low resolution and sparse sampling of measurement which will impact on the overall exploration process. Embedded device such as an inertial measurement unit (IMU) is another option, but even the state-of-the-art IMU system cannot satisfactorily cope with the drift error by sensor noise unless incorporated by GPS support. Since the noisiness and inconsistency of real world system will make the exploration process discrete and non-deterministic, future work should deal with not only the practical implementation but also the theoretical aspect of the system in this regard.

6.5 Closing Remarks

The neuro-robotic system presented in this work has been shown to be sufficiently general and effective. The seamless interaction between the exploration and learning processes results in a system that can be thought of as continually self-monitoring in order to maintain an appropriate level of motor function. As well as being an

effective means of developing robotic controllers, the method has more general implications for truly autonomous artificial systems which must maintain their integrity on several levels, including behavioral. This work provides a few possible biological inspirations and thus it is expected that it can also serve as an indication of the kinds of processes that may be operating in natural systems.

Although the movement patterns produced by our work are by no means poor, they may deviate from perfectly optimised patterns for highly adaptive locomotion. However, we believe our system can make an important contribution as both a stand-alone control architecture, and as a powerful and general exploratory-learning element in a more complex robotic system which may involve further adaptive mechanisms to refine and switch between locomotion patterns.

Appendix A

Robot Parameters

Table A.1: Robot simulation parameters.

4-Fin Swimmer	
torso dimension (m)	0.2×0.2×0.2
arm dimension (m)	0.075×0.075 ×0.15
torso weight (Kg)	1.6
arm weight (Kg)	0.34 (×4)
joint range (rad)	$\pm 0.25\pi$
fin dimension (m)	0.2×0.2
fin weight (Kg)	0.375
fin stiffness (Nm)	0.1
fin damping (Nms)	0.045
fluid density (Kg/m^3)	1000.0
muscle param	
α (Nm)	1.076
β (Nm)	0.108
γ (Nm)	20.0
δ (Nms)	0.152

Quadruped	
torso dimension (m)	R: 0.05, L:0.9
leg dimension (m)	R: 0.05, L:0.3
torso weight (Kg)	7.6
leg weight (Kg)	1.44 (×8)
joint range (rad)	upper: $\pm 0.15\pi$
	lower: $\pm 0.1\pi$
friction coefficient	1.0
muscle param	
α (Nm)	7.935
β (Nm)	1.684
γ (Nm)	20.0
δ (Nms)	1.156

Appendix B

Extra Walking Robots

Some selected results using other types of terrestrial robots are shown. Their movie clips can be found in the URL:
http://www.informatics.sussex.ac.uk/research/groups/ccnr/movies/yssmovie.html.

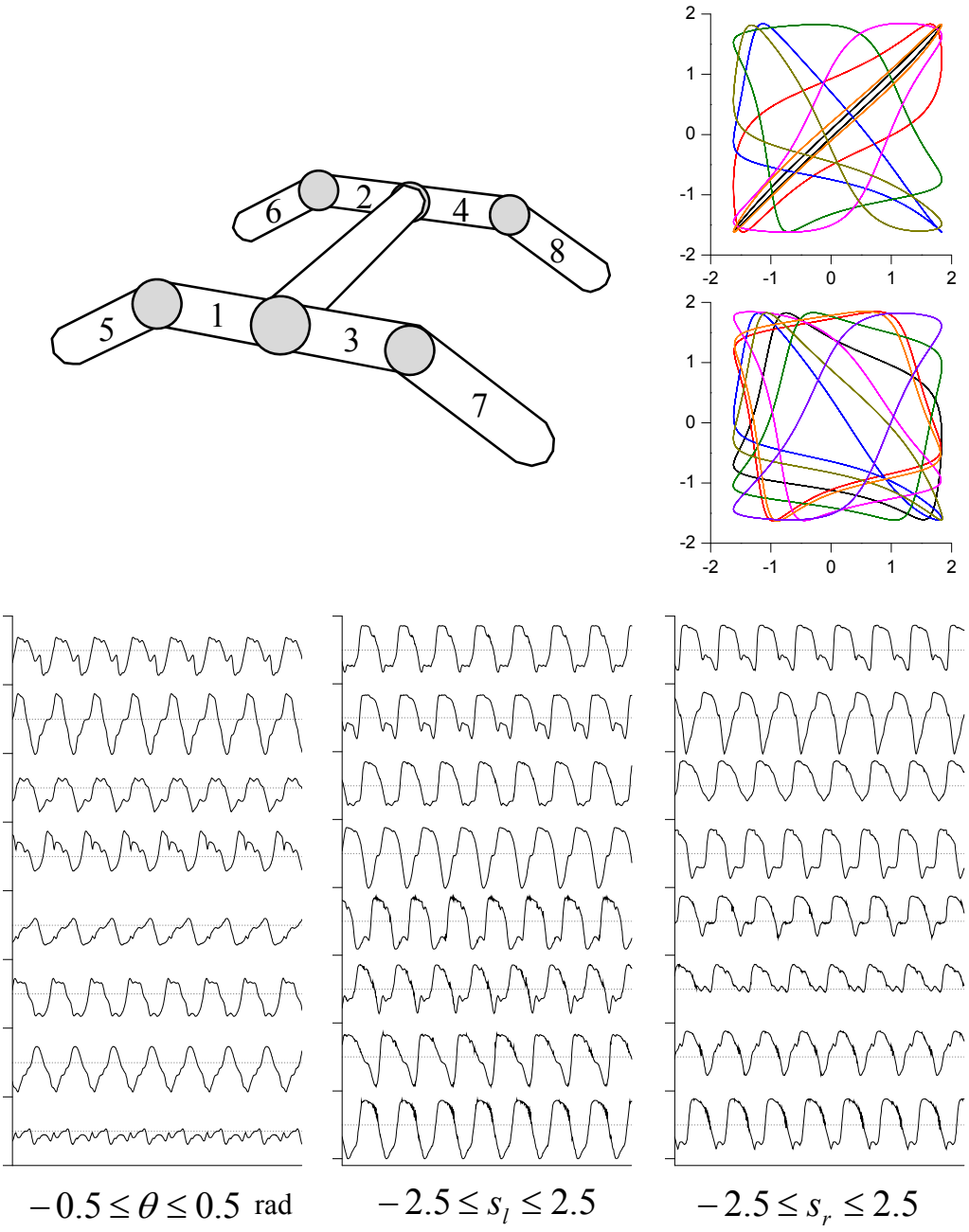

Figure B.1: Quadruped Type-I (the same as Figure 5.14) bounding-like gait. Upper two images are neural outputs, and the lower graphs are joint angles (θ) and raw sensor signals (s_l and s_r). The range of each segment in the graphs are shown respectively.

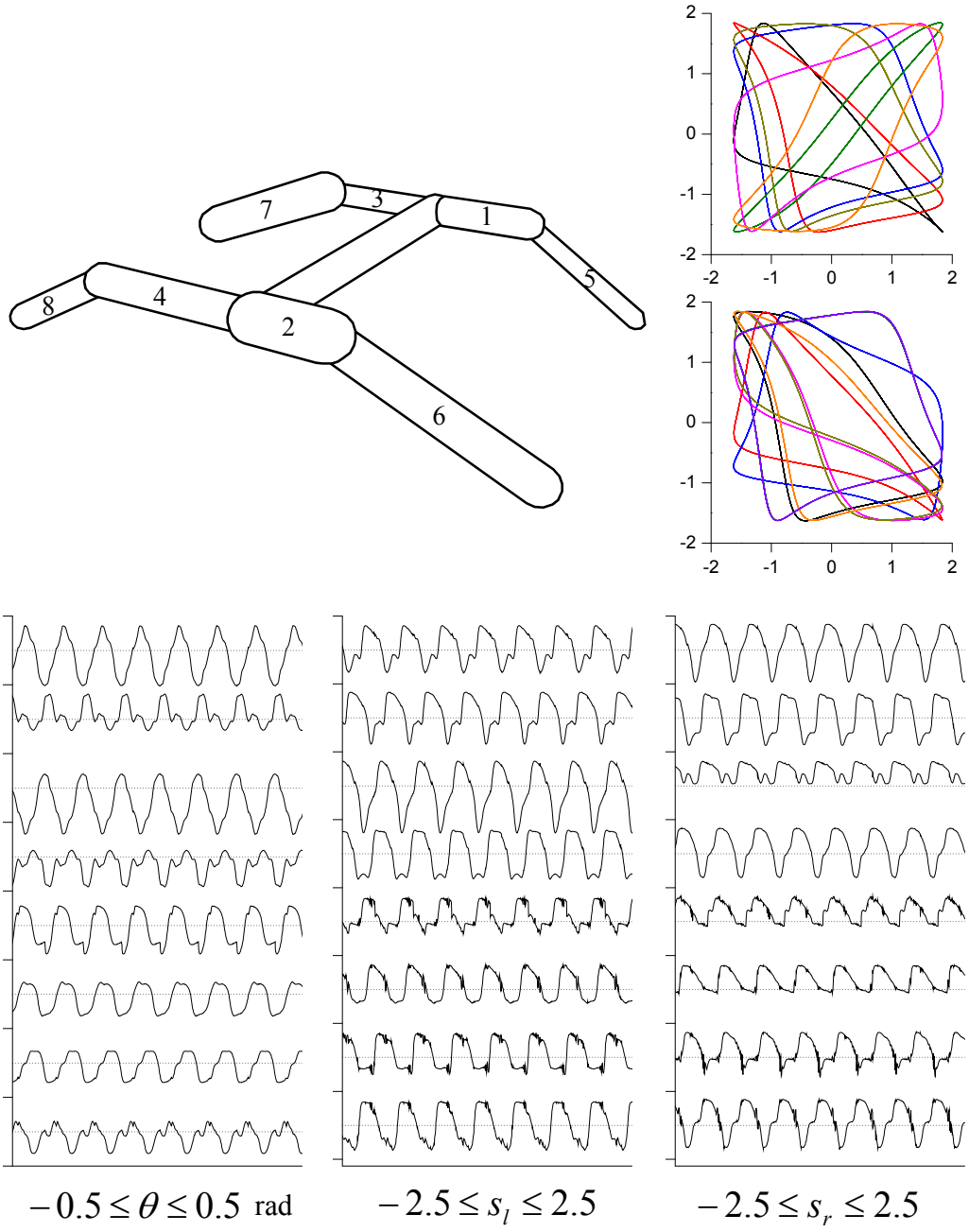

Figure B.2: Quadruped Type-I with different limb dimensions. The radius (m), length (m), and weight (kg) of limbs are: limb-1 (0.05, 0.3, 0.288), limb-2 (0.07 ,0.2, 0.451), limb-3 (0.04, 0.4, 0.228), limb-4 (0.05, 0.5, 0.445), limb-5 (0.03, 0.43, 0.133), limb-6 (0.05, 0.6, 0.523), limb-7 (0.07, 0.5, 0.913), and limb-8 (0.04, 0.4, 0.228). Other settings are the same as the original quadruped.

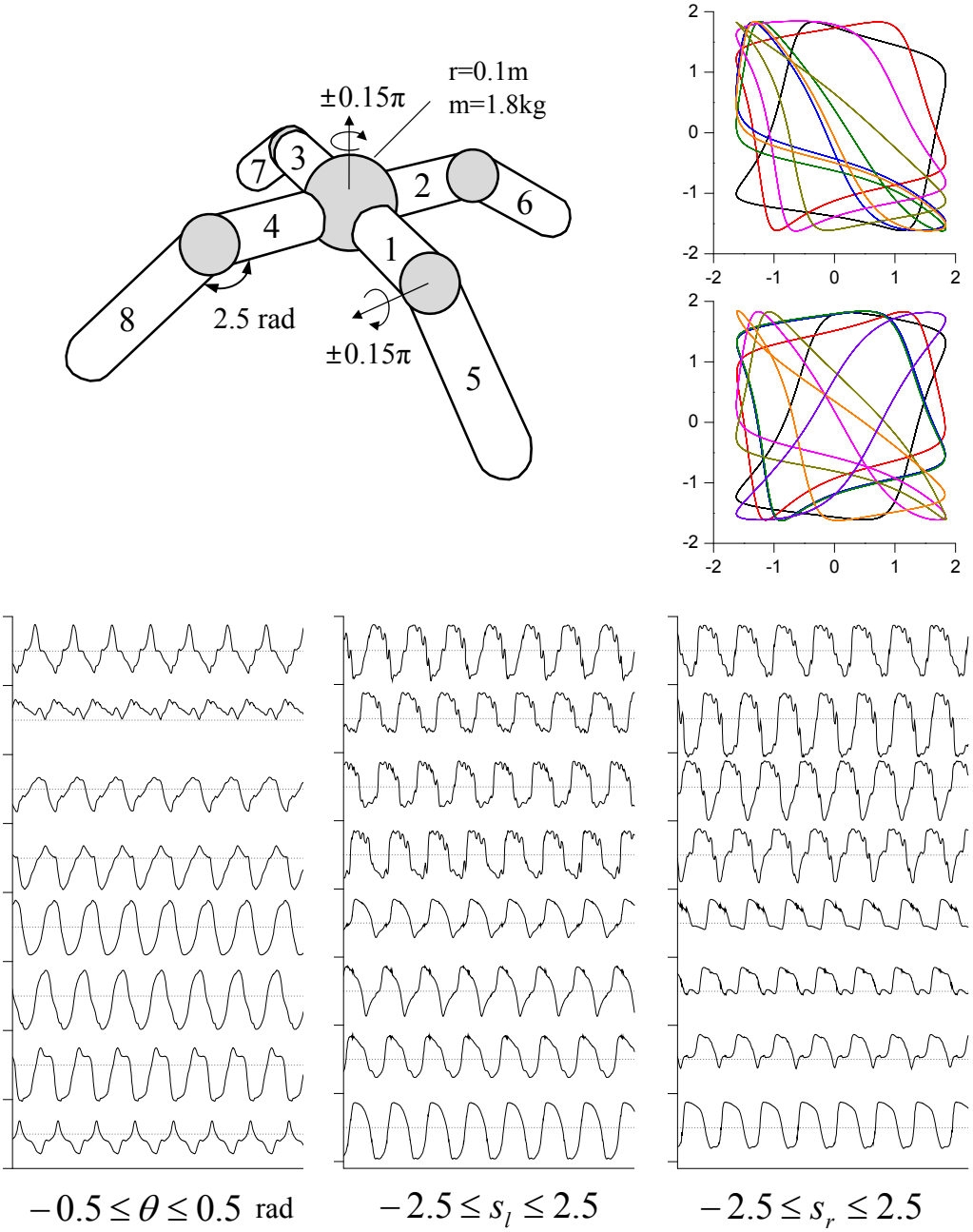

Figure B.3: An example of the locomotion of Quadruped Type-II.

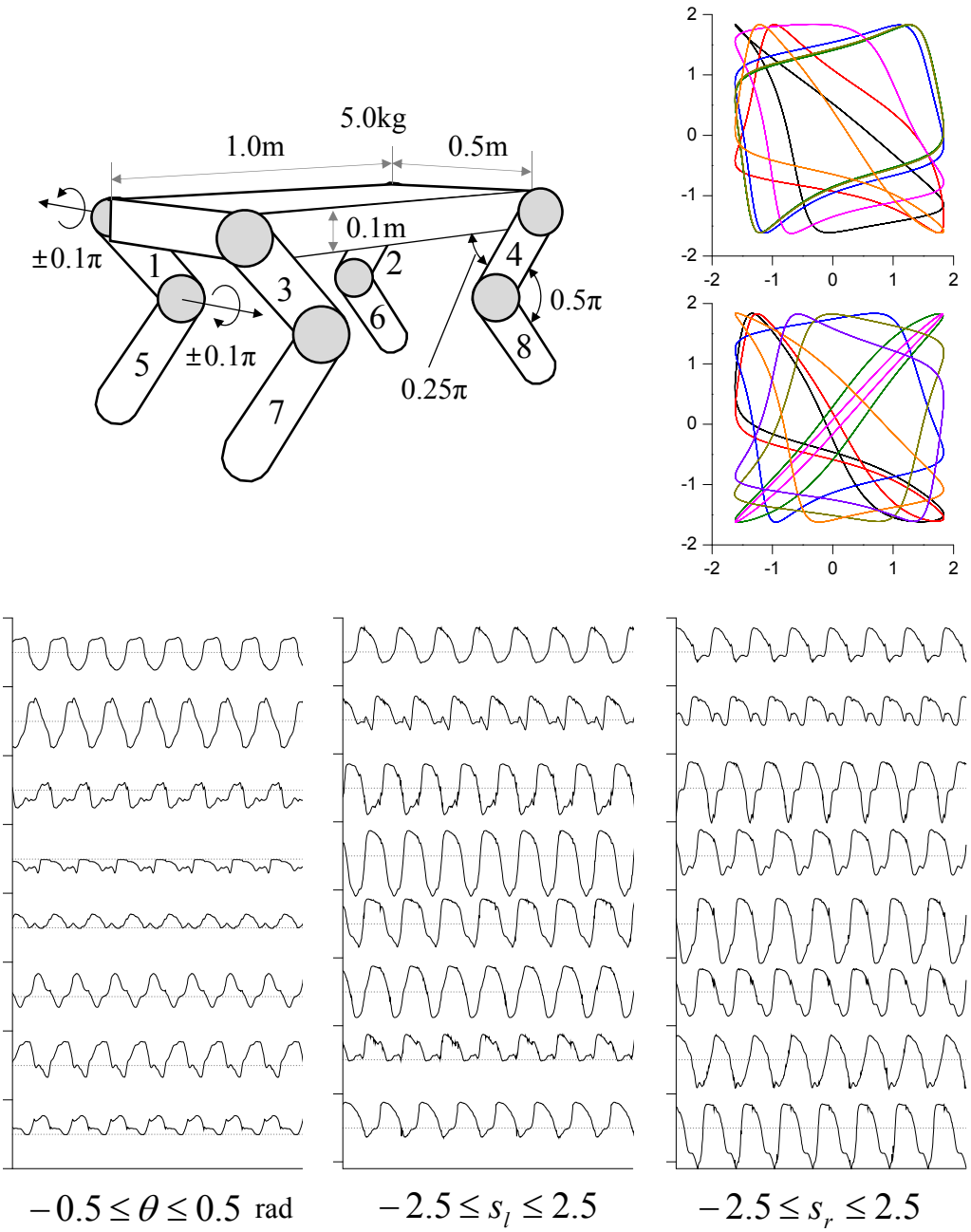

Figure B.4: An example of the locomotion of Quadruped Type-III.

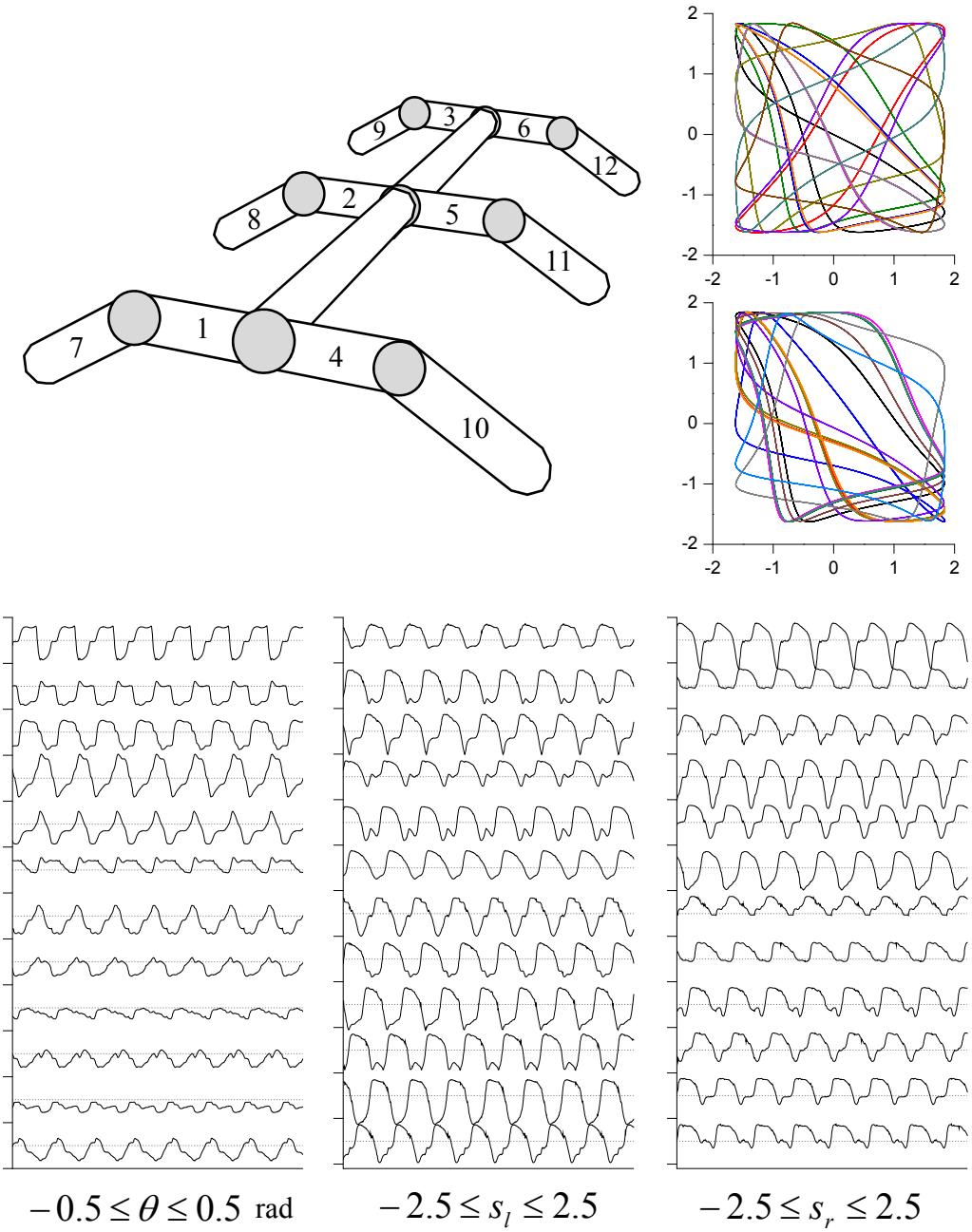

Figure B.5: An example of the locomotion of Hexapod. It is a simple extension of Quadruped Type-I to a six-legged body.

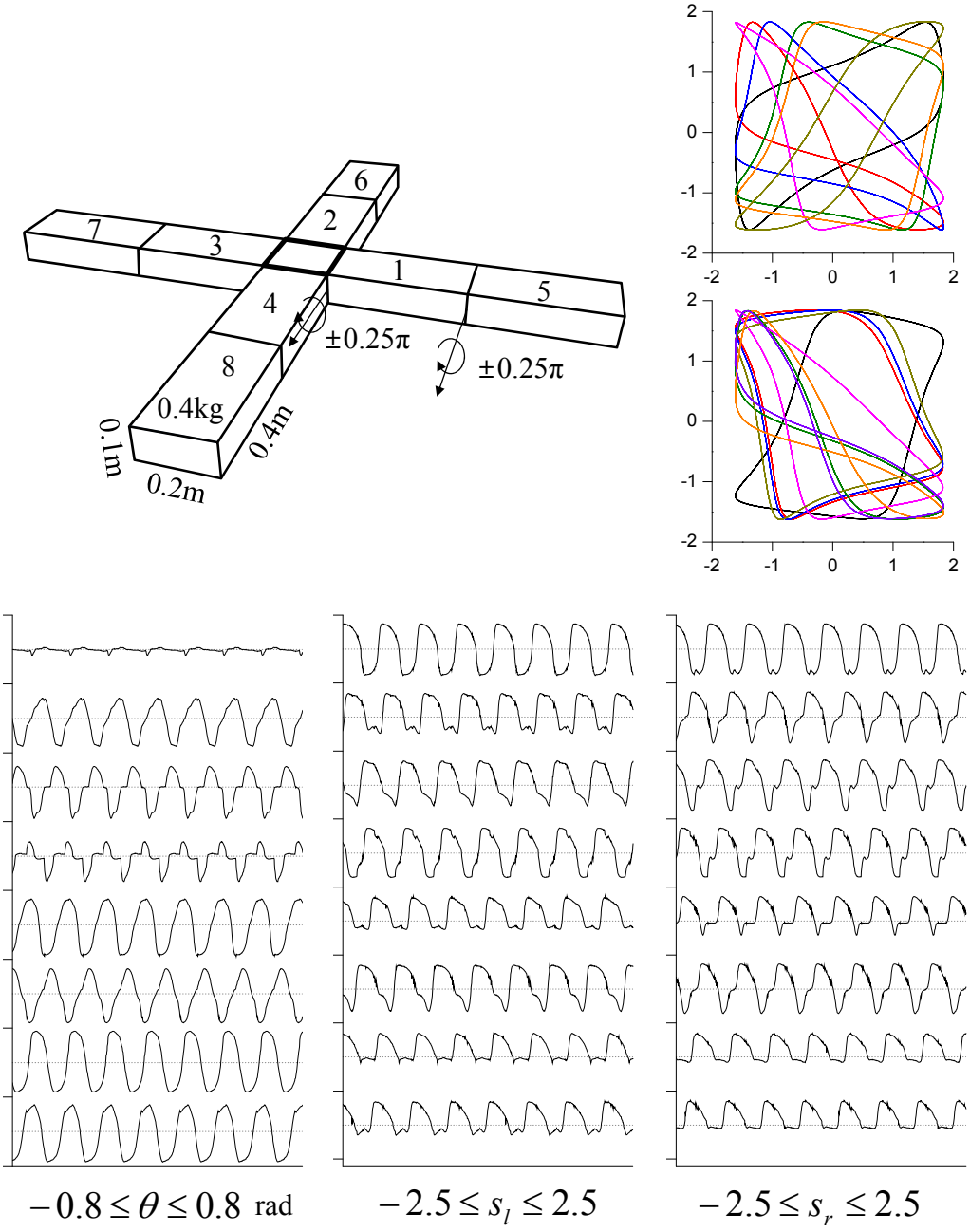

Figure B.6: An example of the locomotion of Cross-shaped Articulation.

Bibliography

Abbott, L. and Nelson, S. (2000). Synaptic plasticity: Taming the beast. *Nature Neuroscience*, 3:1178–1183.

(Referenced on page 99.)

Abe, K., Asai, Y., Matsuo, Y., Nomura, T., Sato, S., Inoue, S., Mizukura, I., and Sakoda, S. (2002). Classifying lower limb dynamics in Parkinson's disease. *Brain Research Bulletin*, 61:219–226.

(Referenced on page 46.)

Adachi, M. and Aihara, K. (1997). Associative dynamics in a chaotic neural network. *Neural Networks*, 10:83–98.

(Referenced on page 16.)

Aida, T. and Davis, P. (1994). Oscillation mode selection using bifurcation of chaotic mode transitions in a nonlinear ring resonator. *IEEE Transactions on Quantum Electronics*, 30(12):2986–2997.

(Referenced on pages 20, 54, 88, 135, and 136.)

Aihara, K. (2003). Chaos in neural systems. In Arbib, M., editor, *The Handbook of Brain Theory and Neural Networks*, pages 208–212, Cambridge, MA. MIT Press.

(Referenced on pages 15 and 16.)

Aihara, K., Takabe, T., and Toyoda, M. (1990). Chaotic neural networks. *Physics Letters A*, 144(6-7):333–340.

(Referenced on page 16.)

Anderson, J. D. (2004). *Introduction to Flight*. McGraw-Hill, 5th edition.

(Referenced on page 64.)

Andras, P. and Lycett, S. (2007). An advantage of chaotic neural dynamics. In *Proceedings of International Joint Conference on Neural Networks*, pages 1417–1422. IEEE Press.

(Referenced on page 16.)

Anitescu, M. and Potra, F. A. (1997). Formulating rigid multibodydynamics with contact and friction as solvable linear complementarity problems. *Nonlinear Dynamics*, 14:231–247.

(Referenced on page 57.)

Asai, Y., Nomura, T., Abe, K., and Sato, S. (2003a). Classification of dynamics of a model of motor coordination and comparison with Parkinson's disease data. *Biosystems*, 71:11–21.

(Referenced on pages 12, 44, 45, 47, and 135.)

Asai, Y., Nomura, T., and Sato, S. (2000). Emergence of oscillations in a model of weakly coupled two Bonhoeffer-van der Pol equations. *Biosystems*, 58:239–247.

(Referenced on pages 12, 44, and 46.)

Asai, Y., Nomura, T., Sato, S., Tamaki, A., Matsuo, Y., Mizukura, I., and Abe, K. (2003b). A coupled oscillator model of disordered interlimb coordination in patients with Parkinson's disease. *Biological Cybernetics*, 88:152–162.

(Referenced on pages 12, 44, and 46.)

Ashby, W. R. (1952). *Design for a Brain*. Chapman and Hall, London.

(Referenced on page 22.)

Azizi, E. and Horton, J. M. (2004). Patterns of axial and appendicular movements during aquatic walking in the salamander *Siren lacertina*. *Zoology*, 107:111–120.

(Referenced on page 134.)

Babloyantz, A. and Destexhe, A. (1986). Low-dimensional chaos in an instance of epilepsy. *Proceedings of the National Academy of Sciences of the United States of America*, 83:3513–3517.

(Referenced on page 15.)

Babloyantz, A. and Salazar, J. M. (1985). Evolution of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*, 111:152–156.
(Referenced on page 15.)

Barbeau, H. and Rossignol, S. (1994). Spinal cord injury: Enhancement of locomotor recovery. *Current Opinion in Neurology*, 7:517–524.
(Referenced on page 10.)

Barlas, Y. and Yasarcan, H. (2006). Goal setting, evaluation, learning and revision: A dynamic modeling approach. *Evaluation and Program Planning*, 29(1):79–87.
(Referenced on page 55.)

Barto, A. G., Sutton, R. S., and Watkins, C. J. C. H. (1990). Learning and sequential decision making. In Gabriel, M. and Moore, J., editors, *Learning and Computational Neuroscience*, pages 539–602, Cambridge, MA. MIT Press.
(Referenced on page 14.)

Beer, R. D. (1995a). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72:173–215.
(Referenced on page 25.)

Beer, R. D. (1995b). On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, 3:469–509.
(Referenced on pages 13 and 99.)

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3):91–99.
(Referenced on page 25.)

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4):209–243.
(Referenced on pages 13 and 99.)

Beer, R. D. (2008). The dynamics of brain-body-environment systems: A status report. In Calvo, P. and Gomila, A., editors, *Handbook of Cognitive Science: An Embodied Approach*, pages 99–120. Elsevier.
(Referenced on page 25.)

Beer, R. D. (2011). Dynamical systems and embedded cognition. In Frankish, K. and Ramsey, W., editors, *The Cambridge Handbook of Artificial Intelligence*. Cambridge University Press.

(Referenced on pages 24 and 25.)

Blankenship, A. G. and Feller, M. B. (2010). Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nature Reviews Neuroscience*, 11:18–29.

(Referenced on page 23.)

Bongard, J. (2010). The utility of evolving simulated robot morphology increases with task complexity for object manipulation. *Artificial Life*, 16(3):201–223.

(Referenced on page 13.)

Bongard, J. (2011). Morphological change in machines accelerates the evolution of robust behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 108(4):1234–1239.

(Referenced on page 13.)

Bongard, J., Zykov, V., and Lipson, H. (2006). Resilient machines through continuous selfmodeling. *Science*, 314:1118–1121.

(Referenced on pages 2 and 23.)

Bonhoeffer, K. F. (1948). Activation of passive iron as a model for the excitation of nerve. *The Journal of General Physiology*, 32:69–91.

(Referenced on pages 12 and 44.)

Bracewell, R. (1999). The Hilbert transform. In *The Fourier Transform and Its Applications*, pages 267–272, New York, NY. McGraw-Hill.

(Referenced on page 29.)

Brooks, R. A. (1999). *Cambrian Intelligence: The Early History of the New AI*. The MIT Press, Cambridge, MA.

(Referenced on page 2.)

Brown, T. G. (1914). On the nature of the fundamental activity of the nervous centres; together with an analysis of the conditioning of rhythmic activity in

progression, and a theory of the evolution of function in the nervous system. *Journal of Physiology, London*, 48:18–46.

(Referenced on page 11.)

Buchli, J., Righetti, L., and Ijspeert, A. J. (2006). Engineering entrainment and adaptation in limit cycle systems. *Biological Cybernetics*, 95(6):645–664.

(Referenced on pages 16 and 136.)

Buono, P. L. (2001). Models of central pattern generators for quadruped locomotion II. secondary gaits. *Journal of Mathematical Biology*, 42(4):327–346.

(Referenced on page 12.)

Caponetto, R., Fortuna, L., Fazzino, S., and Xibilia, M. G. (2003). Chaotic sequences to improve the performance of evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 7(3):289–304.

(Referenced on page 20.)

Chakravarthy, V. S., Joseph, D., and Bapi, R. S. (2010). What do the basal ganglia do? a modeling perspective. *Biological Cybernetics*, 103(3):237–253.

(Referenced on page 134.)

Cliff, D., Harvey, I., and Husbands, P. (1993). Explorations in evolutionary robotics. *Adaptive Behavior*, 2(1):71–108.

(Referenced on pages 12 and 13.)

Cohen, A. H., Rossignol, S., and Grillner, S. (1988). *Neural Control of Rhythmic Movements in Vertebrates*. Wiley, New York.

(Referenced on pages 1 and 10.)

Collins, J. and Richmond, S. (1994). Hard-wired central pattern generators for quadrupedal locomotion. *Biological Cybernetics*, 71:375–385.

(Referenced on page 12.)

Conway, B. A., Hultborn, H., and Kiehn, O. (1987). Proprioceptive input resets central locomotor rhythm in the spinal cat. *Experimental Brain Research*, 68:643–656.

(Referenced on page 121.)

Cordier, P., Dietrich, G., and Pailhous, J. (1996). Harmonic analysis of a complex motor behavior. *Human Movement Science*, 15(6):789–807.

(Referenced on page 15.)

Crair, M. C. (1999). Neuronal activity during development: permissive or instructive? *Current Opinion in Neurobiology*, 9:88–93.

(Referenced on pages 2 and 16.)

Davis, P. (1990). Application of optical chaos to temporal pattem search in a nonlinear optical resonator. *Japanese Journal of Applied Physics*, 29:L1238–L1240.

(Referenced on pages 3 and 20.)

Davis, P. (1998). Adaptive mode selection using on-off switching of chaos. *International Journal of Bifurcation and Chaos*, 8(8):1671–1674.

(Referenced on pages 3 and 20.)

Desai, N., Rutherford, L., and Turrigiano, G. G. (1999). Plasticity in the intrinsic excitability of neocortical pyramidal neurons. *Nature Neuroscience*, 2:515–520.

(Referenced on page 99.)

Determan, J. and Foster, J. A. (1999). Using chaos in genetic algorithm. In *Proceedings of the 1999 Congress on Evolutionary Computation*, volume 3, pages 2361–2364, Piscataway, NJ. IEEE Press.

(Referenced on page 20.)

Di Paolo, E. (2000). Homeostatic adaptation to inversion in the visual field and other sensorimotor disruptions. In Meyer, J., Berthoz, A., Floreano, D., Roitblat, H., and Wilson, S., editors, *From Animals to Animats 6: Proceedings of the 6th International Conference on Simulation of Adaptive Behavior*, pages 440–449, Cambridge, MA. MIT Press.

(Referenced on page 99.)

Dietz, V., Colombo, G., Jensen, L., and Baumgartner, L. (1995). Locomotor capacity of spinal cord in paraplegic patients. *Annals of Neurology*, 37:574–582.

(Referenced on page 10.)

Doya, K. (1996). Temporal difference learning in continuous time and space. In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., editors, *Advances in Neural*

Information Processing Systems 8, pages 1073–1079, Cambridge, MA. MIT Press.
(Referenced on page 14.)

Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Computation*, 12:219–245.
(Referenced on page 14.)

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks*, 15:495–506.
(Referenced on page 135.)

Doya, K. and Yoshizawa, S. (1992). Adaptive synchronization of neural and physical oscillators. *Advances in Neural Information Processing Systems*, 4:109–116.
(Referenced on page 112.)

Dubbeldam, J. L. (2001). Evolution of playlike behaviour and the uncoupling of neural locomotor mechanisms. *Netherlands Journal of Zoology*, 51:335–345.
(Referenced on page 134.)

Eckmann, J. P. and Ruelle, D. (1992). Fundamental limtations for estimating dimensions and lyapunov exponents in dynamical systems. *Physica D*, 56:185–187.
(Referenced on page 77.)

Ekeberg, O. (1993). A combined neuronal and mechanical model of fish swimming. *Biological Cybernetics*, 69:363–374.
(Referenced on pages 1 and 57.)

Endo, G., Moromoto, J., Matsubara, T., Nakanishi, J., and Cheng, G. (2008). Learning CPG-based biped locomotion with a policy gradient method: Application to a humanoid robot. *The International Journal of Robotics Research*, 27(2):213–228.
(Referenced on pages 2 and 14.)

Feudel, U. and Grebogi, C. (1997). Multistability and the control of complexity. *Chaos*, 7:597–603.
(Referenced on page 20.)

Feudel, U., Grebogi, C., and Yorke, J. A. (1996). Map with more than 100 coexisting low-period periodic attractors. *Physical Review E*, 54:71–81.
(Referenced on page 20.)

Fitzhugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1:445–466.

(Referenced on pages 12 and 43.)

Floreano, D., Husbands, P., and Nolfi, S. (2008). Evolutionary robotics. In Siciliano, B. and Khatib, O., editors, *Springer Handbook of Robotics*, pages 1423–1451. Springer.

(Referenced on pages 1 and 13.)

Fraser, A. M. and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33:1134–1140.

(Referenced on page 34.)

Freeman, W. J., Kozma, R., and Werbos, P. J. (2001). Biocomplexity: Adaptive behavior in complex stochastic dynamical systems. *Biosystems*, 59:109–123.

(Referenced on page 15.)

Freeman, W. J. and Skarda, C. (1960). Chaotic dynamics versus representationalism. *Behavioral and Brain Sciences*, 13(1):167–168.

(Referenced on page 15.)

Freeman, W. J. and Viana Di Prisco, G. (1986). EEG spatial pattern differences with discriminated odors manifest chaotic and limit cycle attractors in olfactory bulb of rabbits. In Palm, G. and Aertsen, A., editors, *Brain Theory*, pages 97–119, London. Springer-Verlag.

(Referenced on pages 3 and 15.)

Gallagher, J. C., Beer, R. D., Espenschied, K. S., and Quinn, R. D. (1996). Application of evolved locomotion controllers to a hexapod robot. *Robotics and Autonomous Systems*, 19:95–103.

(Referenced on pages 1 and 13.)

Goldstein, S. (1942). Approximate two-dimensional aerofoil theory: Part 1: Velocity distributions for symmetrical aerofoils. Technical Report ARC/CP-68, British A.R.C. Research Reports.

(Referenced on page 63.)

Golubitsky, M., Stewart, I., Buono, P. L., and Collins, J. J. (1999). Symmetry in locomotor central pattern generators and animal gaits. *Nature*, 401:693–695.

(Referenced on page 12.)

Goulding, M. (2004). A matter of balance. *Nature*, 429:515–517.

(Referenced on page 11.)

Goulding, M. (2009). Circuits controlling vertebrate locomotion: moving in a new direction. *Nature Reviews Neuroscience*, 10(7):507–518.

(Referenced on page 9.)

Granmo, M., Petersson, P., and Schouenborg, J. (2008). Action-based body maps in the spinal cord emerge from a transitory floating organization. *The Journal of Neuroscience*, 28(21):5494–5503.

(Referenced on page 23.)

Grillner, S. (2003). The motor infrastructure: from ion channels to neuronal networks. *Nature Reviews Neuroscience*, 4(7):573–586.

(Referenced on pages 8 and 9.)

Grillner, S., Cangiano, L., Hu, G., Thompson, R., Hill, R., and Wall, P. (2000). The intrinsic function of a motor system from ion channels to networks and behavior. *Brain Research*, 886(1-2):225–236.

(Referenced on page 8.)

Grillner, S., Hellgren, J., Menard, A., Saitoh, K., and Wikstrom, M. A. (2005). Mechanisms for selection of basic motor programs: roles for the striatum and pallidum. *Trends in Neuroscience*, 28:364–370.

(Referenced on page 8.)

Grillner, S. and Wallen, P. (1985). Central pattern generators for locomotion control, with special reference to vertebrates. *Annual Review of Neuroscience*, 8:233–261.

(Referenced on pages 9, 10, and 59.)

Grillner, S., Wallen, P., Saitoh, K., Kozlov, A., and Robertson, B. (2008). Neural bases of goal-directed locomotion in vertebrates. an overview. *Brain Research Review*, 57:2–12.

(Referenced on page 9.)

Guevara, M. R., Glass, L., Mackey, M. C., and Shrier, A. (1983). Chaos in neurobiology. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13:790–798.
(Referenced on page 15.)

Gunay, C. and Prinz, A. A. (2010). Model calcium sensors for network homeostasis: Sensor and readout parameter analysis from a database of model neuronal networks. *The Journal of Neuroscience*, 30:1686–1698.
(Referenced on page 99.)

Hansel, D. and Sompolinsky, H. (1992). Synchronization and computation in a chaotic neural network. *Physical Review Letters*, 68:718–721.
(Referenced on page 16.)

Hanson, M. G. and Landmesser, L. T. (2004). Normal patterns of spontaneous activity are required for correct motor axon guidance and the expression of specific guidance molecules. *Neuron*, 43:687–701.
(Referenced on page 11.)

Hiebert, G. W. and Pearson, K. G. (1999). Contribution of sensory feedback to the generation of extensor activity during walking in the decerebrate cat. *Journal of Neurophysiology*, 81(2):758–770.
(Referenced on page 59.)

Hindmarsh, J. L. and Rose, R. M. (1981). A model of neuronal bursting using three coupled first order differential equations. *Proceedings of the Royal Society B*, 221:87–102.
(Referenced on page 12.)

Hirsch, M. W., Smale, S., and Devaney, R. L. (2012). *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier, 3rd edition.
(Referenced on page 25.)

Hodgkin, A. and Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117:500–544.
(Referenced on pages 11 and 43.)

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A*, 454:903–995.

(Referenced on page 30.)

Huang, X. and Cao, J. (2006). Generalized synchronization for delayed chaotic neural networks: A novel coupling scheme. *Nonlinearity*, 19:2797.

(Referenced on page 16.)

Husbands, P. and Harvey, I. (1992). Evolution versus design: Controlling autonomous robots. In *Proceedings of 3rd Annual Conference on Artificial Intelligence, Simulation and Planning*, pages 139–146, Piscataway, New Jersey. IEEE Press.

(Referenced on page 12.)

Husbands, P. and Meyer, J. A., editors (1998). *Evolutionary Robotics: First European Workshop, EvoRobot'98*. Springer, Berlin.

(Referenced on page 12.)

Iida, F., Gomez, G. J., and Pfeifer, R. (2005). Exploiting body dynamics for controlling a running quadruped robot. In *Proceedings of the 12th International Conference on Advanced Robotics*, pages 229–235.

(Referenced on page 17.)

Ijspeert, A. J. (2001). A connectionist central pattern generator for the aquatic and terrestrial gaits of a simulated salamander. *Biological Cybernetics*, 84(5):331–348.

(Referenced on page 1.)

Ijspeert, A. J. (2008). Central pattern generators for locomotion control in animals and robots: A review. *Neural Networks*, 21:642–653.

(Referenced on pages 10, 11, and 12.)

Ijspeert, A. J., Crespi, A., Ryczko, D., and Cabelguen, J. M. (2007). From swimming to walking with a salamander robot driven by a spinal cord model. *Science*, 315(5817):1416–1420.

(Referenced on page 1.)

Itoh, Y., Taki, K., Kato, S., and Itoh, H. (2004). A stochastic optimization method of CPG-based motion control for humanoid locomotion. In *IEEE Conference on Robotics, Automation and Mechatronics*, pages 347–351.

(Referenced on page 1.)

Jean-Xavier, C., Mentis, G. Z., O'Donovan, M. J., Cattaert, D., and Vinay, L. (2007). Dual personality of GABA/glycine-mediated depolarizations in immature spinal cord. *Proceedings of the National Academy of Sciences of the United States of America*, 104:11477–11482.

(Referenced on page 10.)

Johnson, M. H. (2005). *Developmental Cognitive Neuroscience*. Blackwell, 2nd edition.

(Referenced on pages 2 and 16.)

Kamimura, A., Kurokawa, H., Toshida, E., Tomita, K., Murata, S., and Kokaji, S. (2003). Automatic locomotion pattern generation for modular robots. In *IEEE International Conference on Robotics and Automation*, pages 714–720.

(Referenced on page 1.)

Kaneko, K. (1994). Relevance of dynamic clustering to biological networks. *Physica D*, 75(3):55–73.

(Referenced on page 21.)

Kaneko, K. (2003). Chaotic itinerancy. *Chaos*, 13(3):926–936.

(Referenced on page 20.)

Kantz, H. and Schreiber, T. (1997). *Nonlinear time series analysis*. Cambridge University Press, Cambridge.

(Referenced on pages 35 and 38.)

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. MIT Press.

(Referenced on page 3.)

Kennel, M. B., Brown, R., and Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys-*

ical Review A, 45(6):3403–3411.
(Referenced on pages 36 and 37.)

Kimura, H., Akiyama, S., and Sakurama, K. (1999). Realization of dynamic walking and running of the quadruped using neural oscillator. *Autonomous Robots*, 7:247–258.
(Referenced on page 1.)

King, F. W. (2009). *Hilbert Transforms*. Cambridge University Press, Cambridge, 2nd edition.
(Referenced on page 29.)

Kinjo, K., Nabeshima, C., Sangawa, S., and Kuniyoshi, Y. (2008). A neural model for exploration and learning of embodied movement patterns. *Journal of Robotics and Mechatronics*, 20(3):358–366.
(Referenced on pages 16, 19, 45, and 135.)

Korn, H. and Faure, P. (2003). Is there chaos in the brain? II. Experimental evidence and related models. *Comptes Rendus Biologies*, 326:787–840.
(Referenced on page 3.)

Kuniyoshi, Y. and Sangawa, S. (2006). Early motor development from partially ordered neural-body dynamics: Experiments with a cortico-spinal-musculoskeletal model. *Biological Cybernetics*, 95:589–605.
(Referenced on pages 3, 18, 19, 24, 42, 43, 45, 131, 134, and 135.)

Kuniyoshi, Y. and Suzuki, S. (2004). Dynamic emergence and adaptation of behavior through embodiment as coupled chaotic field. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pages 2042–2049.
(Referenced on pages 3, 16, 17, and 24.)

Kuniyoshi, Y., Yorozu, Y., Suzuki, S., Sangawa, S., Ohmura, Y., Terada, K., and Nagakubo, A. (2007). Emergence and development of embodied cognition: A constructivist approach using robots. *Progress in Brain Research*, 164:425–445.
(Referenced on page 16.)

Liu, Y. and Davis, P. (1998). Adaptive mode selection based on chaotic search in a Fabry-Perot laser diode. *International Journal of Bifurcation and Chaos*,

8(8):1685–1691.
(Referenced on page 20.)

Liu, Z., Golowasch, J., Marder, E., and Abbott, L. F. (1998). A model neuron with activity-dependent conductances regulated by multiple calcium sensors. *The Journal of Neuroscience*, 18:2309–2320.
(Referenced on page 99.)

Lungarella, M. and Berthouze, L. (2002). On the interplay between morphological, neural, and environmental dynamics: a robotic case-study. *Adaptive Behavior*, 10(3-4):223–241.
(Referenced on page 3.)

Magdoom, K. N., Subramanian, D., Chakravarthy, V. S., Ravindran, B., Amari, S., and Meenakshisundaram, N. (2011). Modeling basal ganglia for understanding Parkinsonian reaching movements. *Neural Computation*, 23(2):477–516.
(Referenced on page 134.)

Mataric, M. and Cliff, D. (1996). Challenges in evolving controllers for physical robots. *Robotics and Autonomous Systems*, 19(1):67–83.
(Referenced on page 13.)

Matsubara, T., Morimoto, J., Nakanishi, J., Sato, M., and Doya, K. (2006). Learning CPG-based biped locomotion with a policy gradient method. *Robotics and Autonomous Systems*, 54:911–920.
(Referenced on pages 2 and 14.)

Morihiro, K., Isokawa, T., Matsui, N., and Nishimura, H. (2005). Reinforcement learning by chaotic exploration generator in target capturing task. *Lecture Notes in Computer Science*, 3681:1248–1254.
(Referenced on page 20.)

Morihiro, K., Isokawa, T., Matsui, N., and Nishimura, H. (2008). Effects of chaotic exploration on reinforcement learning in target capturing task. *International Journal of Knowledge-based and Intelligent Engineering Systems*, 12(5):369–377.
(Referenced on page 20.)

Morris, C. and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal*, 35(1):193–213.

(Referenced on page 12.)

Mpitsos, G. J., Burton, R. M., Creech, H. C., and Soinila, S. O. (1988). Evidence for chaos in spike trains of neurons that generate rhythmic motor patterns. *Brain Research Bulletin*, 21:529–538.

(Referenced on page 3.)

Nagumo, J., Arimoto, S., and Yoshizawa, S. (1962). An active pulse transmission line simulating nerve axon. In *Proceedings of the IRE 50*, pages 2061–2071.

(Referenced on pages 12 and 43.)

Nakamura, Y., Mori, T., Sato, M., and Ishii, S. (2007). Reinforcement learning for a biped robot based on a CPG-actor-critic method. *Neural Networks*, 20(6):723–735.

(Referenced on pages 2 and 14.)

Nara, S. and Davis, P. (1992). Chaotic wandering and search in a cycle-memory neural network. *Progress of Theoretical Physics*, 88(5):845–855.

(Referenced on pages 20 and 135.)

Nishimura, H., Iizuka, M., Ozaki, S., and Kudo, N. (1996). Spontaneous motoneuronal activity mediated by glycine and GABA in the spinal cord of rat fetuses in vitro. *The Journal of Physiology*, 497:131–143.

(Referenced on page 11.)

Nolfi, S. and Floreano, D. (2001). *Evolutionary robotics. The biology, intelligence, and technology of self-organizing machines.* MIT Press, Cambridge, MA, 2 edition.

(Referenced on page 12.)

Nomura, T., Sato, S., Doi, S., Segundo, J. P., and Stiber, M. D. (1993). A Bonhoeffer-van der Pol oscillator model of locked and non-locked behaviors of living pacemaker neurons. *Biological Cybernetics*, 69:429–437.

(Referenced on page 44.)

Ohgane, K., Ei, S. I., and Mahara, H. (2009). Neuron phase shift adaptive to time delay in locomotor control. *Applied Mathematical Modelling*, 33(2):797–811.
(Referenced on pages 12 and 44.)

Ohgi, S., Morita, S., Loo, K., and Mizuike, C. (2008). Time series analysis of spontaneous upper-extremity movements of premature infants with brain injuries. *Physical Therapy*, 88(9):1022–1033.
(Referenced on page 15.)

Olusola, O. I., Vincent, U. E., and Njah, A. N. (2010). Synchronization, multistability and basin crisis in coupled pendula. *Journal of Sound and Vibration*, 329(4):443–456.
(Referenced on page 21.)

Orlovsky, G. N., Deliagina, T., and Grillner, S. (1999). *Neuronal Control of Locomotion. From Mollusc to Man*. Oxford University Press.
(Referenced on page 9.)

Ott, E., Sauer, T., and Yorke, J. A. (1994). *Coping with chaos: Analysis of chaotic data and the exploitation of chaotic systems*. John Wiley & Sons, Inc.
(Referenced on page 20.)

Parker, T. S. and Chua, L. O. (1989). *Practical numerical algorithms for chaotic systems*. Springer-Verlag.
(Referenced on page 20.)

Pearson, K. G. (2008). Role of sensory feedback in the control of stance duration in walking cats. *Brain Research Review*, 3139:1–26.
(Referenced on page 121.)

Pearson, K. G. and Collins, D. F. (1993). Reversal of the influence of group Ib afferents from plantaris on activity in medial gastrocnemius muscle during locomotor activity. *Journal of Neurophysiology*, 70(11):1009–1017.
(Referenced on page 121.)

Pearson, K. G., Ekeberg, O., and B"uschges, A. (2006). Assessing sensory function in locomotor systems using neuro-mechanical simulations. *Trends in Neurosciences*,

29(11):625–631.
(Referenced on page 59.)

Pearson, K. G., Ramirez, J. M., and Jiang, W. (1992). Entrainment of the locomotor rhythm by group Ib afferents from ankle extensor muscles in spinal cats. *Experimental Brain Research*, 90:557–566.
(Referenced on page 121.)

Pfeifer, R. and Bongard, J. (2007). *How the Body Shapes the Way We Think: A New View of Intelligence*. Bradford Books.
(Referenced on pages 2, 3, and 16.)

Pfeifer, R. and Iida, F. (2004). Embodied artificial intelligence: Trends and challenges. *Embodied Artificial Intelligence*, 3139:1–26.
(Referenced on page 16.)

Pfeifer, R., Lungarella, M., and Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science*, 314:1088–1093.
(Referenced on pages 3 and 16.)

Pinto, C. M. A. and Santos, A. P. (2011). Modelling gait transition in two-legged animals. *Communications in Nonlinear Science and Numerical Simulations*, 16(12):4625–4631.
(Referenced on page 12.)

Pitti, A., Lungarella, M., and Kuniyoshi, Y. (2005). Quantification of emergent behaviors induced by feedback resonance of chaos. *Advances in Natural Computation*, 15(3):199–213.
(Referenced on pages 3, 16, and 17.)

Pitti, A., Lungarella, M., and Kuniyoshi, Y. (2009). Generating spatiotemporal joint torque patterns from dynamical synchronization of distributed pattern generators. *Frontiers in NeuroRobotics*, 3(2):1–14.
(Referenced on page 18.)

Pitti, A., Niiyama, R., and Kuniyoshi, Y. (2010). Creating and modulating rhythms by controlling the physics of the body. *Autonomous Robots*, 28(3):317–329.
(Referenced on pages 3, 16, and 18.)

Prandtl, L. (1918a). Tragflügeltheorie. I Mitteilung. In *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, page 151.

(Referenced on page 63.)

Prandtl, L. (1918b). Tragflügeltheorie. II Mitteilung. In *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, page 107.

(Referenced on page 63.)

Prochazka, A. (1999). Quantifying proprioception. *Progress in Brain Research*, 123:133–142.

(Referenced on page 59.)

Raftery, A., Cusumano, J., and Sternad, D. (2008). Chaotic frequency scaling in a coupled oscillator model for free rhythmic actions. *Neural Computation*, 20(1):205–226.

(Referenced on page 16.)

Rakic, P. (1988). Specification of cerebral cortical areas. *Science*, 241:170–176.

(Referenced on pages 2 and 16.)

Rapp, P., Zimmerman, I., Albano, A., Deguzman, G., and Greenbaun, N. (1985). Dynamics of spontaneous neural activity in the simian motor cortex: The dimension of chaotic neurons. *Physics Letters A*, 110(6):335–338.

(Referenced on pages 3 and 15.)

Redgrave, P., Prescott, T., and Gurney, K. N. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, 89:1009–1023.

(Referenced on page 134.)

Rescorla, R. A. and Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, pages 64–69.

(Referenced on page 55.)

Righetti, L., Buchli, J., and Ijspeert, A. J. (2006). Dynamic hebbian learning in adaptive frequency oscillators. *Physica D*, 216(2):269–281.

(Referenced on pages 16 and 136.)

Riley, M. and Turvey, M. (2002). Variability and determinism in motor behaviour. *Journal of Motor Behavior*, 34(2):99–125.

(Referenced on page 15.)

Rosenstein, M. T., Collins, J. J., and DeLuca, C. J. (1993). A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena*, 65:117–134.

(Referenced on pages 38 and 77.)

Rossignol, S. (2000). Locomotion and its recovery after spinal injury. *Current Opinion in Neurobiology*, 10(6):708–716.

(Referenced on page 10.)

Rossignol, S., Dubuc, R., and Gossard, J. P. (2006). Dynamic sensorimotor interactions in locomotion. *Physiological Review*, 86:89–154.

(Referenced on pages 9 and 59.)

Rosslenbroich, B. (2009). The theory of increasing autonomy in evolution: A proposal for understanding macroevolutionary innovations. *Biology and Philosophy*, 24(5):623–644.

(Referenced on page 134.)

Schouenberg, J. (2004). Learning in sensorimotor cicuits. *Current Opinion in Neurobiology*, 14:693–697.

(Referenced on page 11.)

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Physiology*, 57:87–115.

(Referenced on page 134.)

Shik, M. L., Severin, F. V., and Orlovsky, G. N. (1966). Control of walking by means of electrical stimulation of the mid-brain. *Biophysics*, 11:756–765.

(Referenced on pages 9 and 10.)

Shim, Y. S. (2012). Webpage for supplementary videos. http://www.informatics.sussex.ac.uk/research/groups/ccnr/movies/yssmovie.html.

(Referenced on pages 99, 105, 118, 120, and 129.)

Shim, Y. S. and Husbands, P. (2007a). Feathered flyer: Integrating morphological computation and sensory reflexes into a physically simulated flapping-wing robot for robust flight manoeuvre. In Almeida e Costa, F., Rocha, L. M., Costa, E., Harvey, I., and Coutinho, A., editors, *Advances in Artificial Life, 9th European Conference on Artificial Life (ECAL 2007)*, pages 756–765. Springer.

(Referenced on pages 7 and 13.)

Shim, Y. S. and Husbands, P. (2007b). A flapping-wing flying machine with an evolved neurocontroller. In *Proceedings of International Symposium on FLYING INSECTS & ROBOTS*, pages 109–110.

(Referenced on pages 7 and 13.)

Shim, Y. S. and Husbands, P. (2010). Chaotic search of emergent locomotion patterns for a bodily coupled robotic system. In *Proceedings of 12th International Conference on the Synthesis and Simulation of Living Systems (ALIFE XII)*, pages 757–764.

(Referenced on pages 6 and 22.)

Shim, Y. S. and Husbands, P. (2012). Chaotic exploration and learning of locomotion behaviours. *Neural Computation*, 24(8):2185–2222.

(Referenced on page 7.)

Shim, Y. S. and Kim, C. H. (2003). Generating flying creatures using body-brain co-evolution. In *Proceedings of ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, pages 276–285, San Diego, CA.

(Referenced on page 7.)

Shim, Y. S. and Kim, C. H. (2006). Evolving physically simulated flying creatures for efficient cruising. *Artificial Life*, 12(4):561–591.

(Referenced on pages 7, 13, and 64.)

Shim, Y. S., Kim, S. J., and Kim, C. H. (2004a). Evolving flying creatures with path following behaviour. In *The 9th International Conference on the Simulation*

and Synthesis of Living Systems (ALIFE IX), pages 125–132, Boston, MA.
(Referenced on page 7.)

Shim, Y. S., Shin, S. Y., and Kim, C. H. (2004b). Two-step evolution process for path-following virtual creatures. In *Proceedings of Computer Animations and Social Agents (CASA)*, pages 85–93, Geneva, Switzerland.
(Referenced on page 7.)

Sims, K. (1994). Evolving 3d morphology and behavior by competition. In Brooks, R. and Maes, P., editors, *Proceedings of Artificial Life IV*, pages 28–39. MIT Press.
(Referenced on page 13.)

Skarda, C. and Freeman, W. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10:161–195.
(Referenced on pages 15 and 133.)

Skarda, C. and Freeman, W. (1990). Chaos and the new science of the brain. *Concepts in Neuroscience*, 1(2):275–285.
(Referenced on page 133.)

Smith, R. (1998). Open Dynamics Engine (ODE). http://ode.org/.
(Referenced on page 57.)

Sproewitz, A. and Berthouze, L. (2005). Robust robot bouncing: Passive compliance and flexible phase locking. In *Proceedings of 3rd International Symposium on Adaptive Motion in Animals and Machines*, page 59.
(Referenced on pages 12 and 44.)

Sproewitz, A., Moeckel, R., Maye, J., and Ijspeert, A. J. (2008). Learning to move in modular robots using central pattern generators and online optimization. *The International Journal of Robotics Research*, 27(3):423–443.
(Referenced on page 2.)

Sridharan, D., Prashanth, P. S., and Chakravarthy, V. S. (2006). The role of the basal ganglia in exploration in a neural model based on reinforcement learning. *International Journal of Neural Systems*, 16:111–124.
(Referenced on page 134.)

Stein, P. S. G., Grillner, S., Selverston, A., and Stuart, D. G., editors (1997). *Neurons, Networks and Motor Behavior*. MIT Press.

(Referenced on pages 1 and 10.)

Steingrube, S., Timme, M., Worgotter, F., and Manoonpong, P. (2010). Self-organized adaptation of a simple neural circuit enables complex robot behaviour. *Nature Physics*, 6:224–230.

(Referenced on page 3.)

Stewart, D. E. and Trinkle, J. C. (1996). An implicit timestepping scheme for rigid-body dynamics with inelastic collisions and Coulomb friction. *International Journal of Numerical Methods in Engineering*, 39:2673–2691.

(Referenced on page 57.)

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Bradford Books, MIT Press, Cambridge, MA.

(Referenced on page 14.)

Taga, G., Takaya, R., and Konishi, Y. (1999). Analysis of general movements of infants towards understanding of developmental principle for motor control. In *Proceedings of 1999 IEEE International Conference on Systems, Man, and Cybernetics*, pages 678–683.

(Referenced on page 3.)

Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. and Young, L., editors, *Dynamical Systems and Turbulence. Lecture Notes in Mathematics vol 898*, pages 366–381, Berlin. Springer-Verlag.

(Referenced on page 34.)

Terman, D. and Rubin, J. (2007). Neuronal dynamics and the basal ganglia. *SIAM News*, 4(2):http://siam.tekdevelopment.com/old–issues/2007/march–2007/.

(Referenced on pages 3 and 15.)

Thelen, E. and Smith, L. B. (1994). *A dynamic systems approach to development of cognition and action*. MIT Press.

(Referenced on pages 2 and 16.)

Tolwinski, S. (2007). The Hilbert transform and empirical mode decomposition as tools for data analysis. Technical Report Technical Report, University of Arizona Program in Applied Mathematics.

(Referenced on pages 31 and 32.)

Turrigiano, G. G. (1999). Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. *Trends in Neuroscience*, 22:221–228.

(Referenced on page 99.)

Turrigiano, G. G. (2007). Homeostatic signaling: The positive side of negative feedback. *Current Opinion in Neurobiology*, 17:318–324.

(Referenced on page 99.)

Turrigiano, G. G. (2008). The self-tuning neuron: Synaptic scaling of excitatory synapses. *Cell*, 135:422–435.

(Referenced on page 99.)

Turrigiano, G. G., Abbott, L. F., and Marder, E. (1994). Activity-dependent changes in the intrinsic properties of cultured neurons. *Science*, 264:974–977.

(Referenced on page 99.)

Turrigiano, G. G. and Nelson, S. B. (2000). Hebb and homeostasis in neuronal plasticity. *Current Opinion in Neurobiology*, 10:358–364.

(Referenced on page 99.)

Turrigiano, G. G. and Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, 5:97–107.

(Referenced on page 99.)

Vadivasova, T. E., Sosnovtseva, O. V., Balanov, A. G., and Astakhov, V. V. (1999). Phase multistability of synchronous chaotic oscillations. *Discrete Dynamics in Nature and Society*, 4:231–243.

(Referenced on page 20.)

Venaille, A., Varona, P., and Rabinovich, M. (2005). Synchronization and coordination of sequences in two neural ensembles. *Physical Review E*, 71(6):061909.

(Referenced on page 15.)

Wadden, T. and Ekeberg, O. (1998). A neuro-mechanical model of legged locomotion: Single leg control. *Biological Cybernetics*, 79:161–173.

(Referenced on page 57.)

Weis-Fogh, T. and Jensen, M. (1956). Biology and physics of locust flight i: Basic principles in insect flight. a critical review. *Philosophical Transactions of the Royal Society of London B*, 239:415–458.

(Referenced on page 62.)

Wheeler, M. (2005). *Reconstructing the Cognitive World: The Next Step*. MIT Press, Cambridge, MA.

(Referenced on pages 2 and 16.)

Whelan, P. (1996). Control of locomotion in the decerebrate cat. *Progress in Neurobiology*, 49:481–515.

(Referenced on page 10.)

Wilhelm, J. C., Rich, M. M., and Wenner, P. (2009). Compensatory changes in cellular excitability, not synaptic scaling, contribute to homeostatic recovery of embryonic network activity. *Proceedings of the National Academy of Sciences of the Unites States of America*, 106:6760–6765.

(Referenced on page 99.)

Wilhelm, J. C. and Wenner, P. (2008). GABAA transmission is a critical step in the process of triggering homeostatic increases in quantal amplitude. *Proceedings of the National Academy of Sciences of the Unites States of America*, 105:11412–11417.

(Referenced on page 99.)

Williams, H. (2004). Homeostatic plasticity in recurrent neural networks. In Schaal, S., Ijspeert, A., Billard, A., Vijayakumar, S., Hallam, J., and Meyer, J. A., editors, *From Animals to Animats 8: Proceedings of the 8th International Conference on Simulation of Adaptive Behavior*, pages 344–353, Cambridge, MA. MIT Press.

(Referenced on page 99.)

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3):285–

317.

(Referenced on page 37.)

Wright, J. and Liley, D. (1996). Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behavioral and Brain Sciences*, 19:285–320.

(Referenced on page 15.)

Yakovenko, S., Gritsenko, V., and Prochazka, A. (2004). Contribution of stretch reflexes to locomotor control: A modeling study. *Biological Cybernetics*, 90(2):146–155.

(Referenced on page 59.)

Yamada, Y., Nishikawa, S., Shida, K., and Kuniyoshi, Y. (2011a). Emergent locomotion patterns from a quadruped pneumatic musculoskeletal robot with spinobulbar model. In *International Workshop on Bio-Inspired Robots, Poster*, page 48.

(Referenced on page 24.)

Yamada, Y., Nishikawa, S., Shida, K. Niiyama, R., and Kuniyoshi, Y. (2011b). Neural-body coupling for emergent locomotion: A musculoskeletal quadruped robot with spinobulbar model. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011)*, pages 1499–1506.

(Referenced on page 24.)

Zhang, C. K. and Shao, H. H. (2001). A hybrid strategy: Real-coded genetic algorithm and chaotic search. In *IEEE International Conference on Systems, Man, and Cybernetics 4*, pages 2361–2364.

(Referenced on page 20.)

Zhang, W. and Linden, D. (2003). The other side of the engram: Experience-driven changes in neuronal intrinsic excitability. *Nature Reviews Neuroscience*, 4:885–900.

(Referenced on page 99.)

Zhang, Y. and Golowasch, J. (2007). Modeling recovery of rhythmic activity: Hypothesis for the role of a calcium pump. *Neurocomputing*, 70:1657–1662.

(Referenced on page 99.)

Zhang, Y., Khorkova, O., Rodriguez, R., and Golowasch, J. (2009). Activity and neuromodulatory input contribute to the recovery of rhythmic output after decentralization in a central pattern generator. *Journal of Neurophysiology*, 101:372–386.

(Referenced on page 99.)