OPTIMISATION AND COMPUTATIONAL METHODS TO MODEL THE OCULOMOTOR SYSTEM WITH FOCUS ON NYSTAGMUS

Submitted by

Eleftherios Avramidis

to the University of Exeter
as a thesis for the degree of Doctor of Philosophy
in Mathematics

April 2015

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature ……...

Abstract

Infantile nystagmus is a condition that causes involuntary, bilateral and conjugate oscillations of the eyes, which are predominately restricted to the horizontal plane. In order to investigate the cause of nystagmus, computational models and nonlinear dynamics techniques have been used to model and analyse the oculomotor system.

Computational models are important in making predictions and creating a quantitative framework for the analysis of the oculomotor system. Parameter estimation is a critical step in the construction and analysis of these models. A preliminary parameter estimation of a nonlinear dynamics model proposed by Broomhead *et al.* [1] has been shown to be able to simulate both normal rapid eye movements (*i.e.* saccades) and nystagmus oscillations.

The application of nonlinear analysis to experimental jerk nystagmus recordings, has shown that the local dimensions number of the oscillation varies across the phase angle of the nystagmus cycle. It has been hypothesised that this is due to the impact of signal dependent noise (SDN) on the neural commands in the oculomotor system.

The main aims of this study were: (i) to develop parameter estimation methods for the Broomhead *et al.* [1] model in order to explore its predictive capacity by fitting it to experimental recordings of nystagmus waveforms and saccades; (ii) to develop a stochastic oculomotor model and examine the hypothesis that noise on the neural commands could be the cause of the behavioural characteristics measured from experimental nystagmus time series using nonlinear analysis techniques.

In this work, two parameter estimation methods were developed, one for fitting the model to the experimental nystagmus waveforms and one to saccades. By using the former method, we successfully fitted the model to experimental nystagmus waveforms. This fit allowed to find the specific parameter values that set the model to generate these waveforms. The types of the waveforms that we successfully fitted were asymmetric pseudo-cycloid, jerk and jerk with extended

foveation. The fit of other types of nystagmus waveforms were not examined in this work. Moreover, the results showed which waveforms the model can generate almost perfectly and the waveform characteristics of a number of jerk waveforms which it cannot exactly generate. These characteristics were on a specific type of jerk nystagmus waveforms with a very extreme fast phase.

The latter parameter estimation method allowed us to explore whether the model can generate horizontal saccades of different amplitudes with the same behaviour as observed experimentally. The results suggest that the model can generate the experimental saccadic velocity profiles of different saccadic amplitudes. However, the results show that best fittings of the model to the experimental data are when different model parameter values were used for different saccadic amplitude.

Our parameter estimation methods are based on multi-objective genetic algorithms (MOGA), which have the advantage of optimising biological models with a multi-objective, high-dimensional and complex search space. However, the integration of these models, for a wide range of parameter combinations, is very computationally intensive for a single central processing unit (CPU). To overcome this obstacle, we accelerated the parameter estimation method by utilising the parallel capabilities of a graphics processing unit (GPU). Depending of the GPU model, this could provide a speedup of 30 compared to a midrange CPU.

The stochastic model that we developed is based on the Broomhead *et al.* [1] model, with signal dependent noise (SDN) and constant noise (CN) added to the neural commands. We fitted the stochastic model to saccades and jerk nystagmus waveforms. It was found that SDN and CN can cause similar variability to the local dimensions number of the oscillation as found in the experimental jerk nystagmus waveforms and in the case of saccade generation the saccadic variability recorded experimentally. However, there are small differences in the simulated behaviour compared to the nystagmus experimental data. We hypothesise that these could be caused by the inability of the model to simulate exactly key jerk waveform characteristics. Moreover, the differences between the simulations and the experimental nystagmus waveforms indicate that the proposed model requires further expansion, and this could include other oculomotor subsystem(s).

Acknowledgements

First and foremost, I wish to thank my supervisor Dr. Ozgur Akman. It has been an honour working with him. I am very fortunate for the high quality of training I have received from him. Everyone in academia should follow his example.

I am really grateful for the valuable help provided by Dr. Richard Clement on acquiring and analysing the infantile nystagmus experimental data.

Thanks a lot to Marta for all the support in times of despair and her valuable help on improving my writing skills.

I would also like to thank my mentor Dr. Chris Ferro for helping me solve many problems with acquiring key equipment needed for my Ph.D.

Many thanks to Professor Peter Ashwin for his very helpful comments.

I gratefully acknowledge the funding source that made my Ph.D. work possible. I was funded by the Engineering and Physical Sciences Research Council (EPSRC).

Finally, big thanks to all my family for supporting me throughout the course of this PhD.

Table of Contents

List of Abbreviations

~	Approximately
±	Plus-minus
°	Degrees
°/s	Degrees per second
2D	In two dimensions/two-dimensional space
3D	In three dimensions/three-dimensional space
∞	Infinity
AN	Acquired nystagmus
AOS	Accessory optic system
API	Application programming interface
a-priori	Latin: from what comes before
B.L.O.	Base line oscillation
c.f.	Latin: *confer* - compare
CN	Constant noise
CPU	Central processing unit
CUDA	Compute Unified Device Architecture
CV	Coefficient of variation
deg(s)	Degrees
degs/s	Degrees per second
e.g.	Latin: *exempli gratia* – for example
EBNs	Excitatory burst neurons

et al.	Latin: *et alia* – and others
etc.	Latin: *et cetera* – and so on
FP	Fixed point
FRMD7	FERM (F for 4.1 protein, E for ezrin, R for radixin and M for moesin) domain containing 7 gene
GA	Genetic algorithm
GB	Gigabytes
GB/s	Gigabytes per second
GFLOPS	Giga floating operations per second
GPU	Graphics processing unit
HPC	High-performance computing
Hz	Hertz; a unit of frequency
i.e.	Latin: *id est* – that is
IBNs	Inhibitory burst neurons
IN	Infantile nystagmus
LLBNs	Long-lead burst neurons
MATLAB	Matrix laboratory
MB	Megabyte
MOGA	Multi-objective genetic algorithm
ms	Milliseconds
MVN	Medial vestibular nucleus
NI	Neural integrator
NPH	Nucleus prepositus hypoglossi
NSGA-II	Non-dominated sorting genetic algorithm II

OCT	Optical coherence tomography
ODE	Ordinary differential equations
OpenCL	Open Computing Language
OPNs	Omnipause neurons
pGA	Parallel GA
PPRF	Paramedian pontine reticular formation
RAM	Random access memory
s (or sec)	Seconds
SC	Superior colliculus
SD	Standard deviation
SDE	Stochastic differential equation
SDN	Signal-dependent noise
sGA	Sequential genetic algorithm
SIMD	Single instruction multiple data
SVD	Singular value decomposition
TN	Temporal noise
UPO	Unstable periodic orbit

Nomenclature

$<, \leq$	Less than, less than or equal to
$>, \geq$	Greater than, greater than or equal to
$\in$	Element of
$\mathbb{R}$	Real numbers set
$\sum_{i=1}^{n} a_i$	The sum $a_1 + a_2 + \cdots + a_n$
$\|.\|$	Norm of .
g	Horizontal gaze angle (°)
$\boldsymbol{g}$	Two dimensional rotation vectors of the eye orientation (°)
l	Firing rate of left population of neurons (spikes/s)
m	Motor error (°)
n	Neural integrator firing rate (spikes/s)
r	Firing rate of right population of neurons (spikes/s)
t	Time (s)
$\boldsymbol{u}$	Two dimensional rotation vectors of the motor commands
v	Velocity (°/s)
$\boldsymbol{w}$	Delay vector
$D\mathbf{F}$	Jacobian matrix of $\mathbf{F}$
G	Genetic algorithm generation number
$\boldsymbol{I}$	Identity matrix
M	Multi-objective genetic algorithm objectives number

N	Genetic algorithm population size
$\boldsymbol{T}$	Two dimensional vector with torque values
Δt	Time step size (s)
$\boldsymbol{\Omega}$	Three dimensional vector representing the angular velocity of the eye (°/s)

List of Figures

List of Tables

1 Introduction

The information extracted by analysing eye movements is valuable to both clinicians and physical scientists. For physical scientists, it provides a non-invasive way to study the function of the brain (*e.g.* working memory, visual attention), while for clinicians abnormal eye movements can be an indication of specific pathologies [2]–[6]. Moreover, such data can provide direct information about the functioning of the oculomotor system [3], [7], [8]. The study of oculomotor control is an ideal means for improving our understanding of basic motor control functions. Although the oculomotor system is in many ways simpler compared to other neurobiological control systems, its study allows to extract the important properties of these systems while avoiding non-essential details [8], [9].

Initially, computational models of saccade generation were based on control theory [10]–[12]. These models were highly successful in identifying the complex mechanisms and organisation of the oculomotor system [2], [13]. However, the philosophy of these models was based on the identification of the different functional requirements that need to exist for oculomotor control. Therefore, the primary goal of these models was to identify what functions need to exist in the oculomotor system, whereas the implementation and interactions of these functions was secondary [13]. Consequently, the need for additional analysis approaches emerged towards the end of the 1990s, which aimed to understand how the interactions of specific populations of oculomotor neurons can generate both normal and abnormal eye movements. One of the alternative approaches proposed was to use nonlinear dynamics techniques to model and analyse oculomotor control. The use of these techniques led to several outcomes in the study of the oculomotor system. For example, a nonlinear dynamics model of the oculomotor system proposed by Broomhead *et al.* [1] has been shown to be able to simulate both normal saccades and an oscillatory oculomotor instability, known as infantile nystagmus (initially identified as congenital nystagmus). This demonstrated that infantile nystagmus can be treated as a dynamic disease. In such diseases, there is no structural damage to the system (as had been proposed by some models based on control theory), however the problem is

instead that the system's components interact outside of their normal parameter range [14].

1.1 Aims and novelty of the study

This work focussed on the expansion of our knowledge of the oculomotor system by (i) developing methods and tools which would allow the creation of a more biologically plausible, nonlinear dynamics model, and (ii) examination of the role of neuronal noise on the oculomotor system by extending the Broomhead *et al.* model [1] with the addition of signal dependent and constant noise to the neural commands. We constrained the behaviour and tested the predictions of the extended Broomhead *et al.* model [1] by using experimental normal horizontal eye movements and infantile nystagmus recordings.

The methods and tools developed in this work allow the analysis and further development of oculomotor models by providing the means to estimate the parameter values of the Broomhead *et al.* [1] and stochastic extended models. These parameters set the models to generate the same or similar behaviour recorded experimentally from normal participants and nystagmats (*i.e.* individuals suffering from pathological nystagmus). Furthermore, the inability to find such parameter values would identify the specific experimental behaviours which cannot be replicated by the models, thus showing the need for different modelling approaches for the system of interest.

To test the hypothesis that neuronal noise could be a key process in the dynamics of the oculomotor system, we extended the Broomhead *et al.* [1] model by the addition of signal dependent and constant noise to the burst and neural integrator signals. This test, however, involved the combination of dynamic analysis and the parameter optimisation methods developed in this work. These subsequently allowed us to determine the parameters of the stochastic model which would set it to generate nystagmus waveforms as recorded experimentally and to simulate similar behaviour to that found by Akman *et al.* [15] in experimental nystagmus recordings.

As anticipated, the extended model was able to simulate a broad range of normal and abnormal eye movement behaviour, allowing novel testable predictions to be made. Furthermore, the methods and tools developed in this work could allow further analysis and development of oculomotor models by constraining them to

the experimental eye movement recordings as it will be demonstrated in this work.

1.2 Thesis organisation

Chapter 1 is organised as follows. In section 1.1 we will provide a description of normal eye movements. In section 1.2 we will describe the dynamics of rapid eye movements (saccades). In section 1.3 we will describe the human eyeball and the extraocular muscles. In section 1.4 we will briefly describe the brainstem neural circuits which encode the neural signals for the generation of horizontal saccades. In section 1.5 we will describe the role of neuronal noise in the dynamics of saccades. Lastly, in section 1.6 we will describe infantile nystagmus, its hypothesised aetiology and the types of nystagmus waveforms that are observed experimentally.

The organisation of the rest of the thesis is as follows. In Chapter 2 we will describe different modelling approaches of the oculomotor system and infantile nystagmus, the Broomhead *et al.* [1] model and we will provide an overview of previously performed bifurcation analysis of this model. Moreover, we will describe a nonlinear analysis method and its results from its previously application on experimental infantile nystagmus time series. In Chapter 3 we will provide an overview of parameter estimation that is a special form of mathematical optimisation. We will provide the definition of optimisation, the challenges of the optimisation of biological models and the general theory of the specific optimisation methods which were used in this work. In Chapter 4 we will describe our methodology for the parameter estimation and analysis of the Broomhead *et al.* [1] model and the extended stochastic model. This also includes the description of the experimental data and the methods that we have developed in order to accelerate the parameter estimation using graphics processing units. In Chapter 5 we will present the results obtained from using parameter estimation on the Broomhead *et al.* [1] model to experimental data using the methodology presented in Chapter 4. In Chapter 6 we will describe the extended model and present the results of estimating its noise intensity parameters. Lastly, in Chapter 7 we will provide the conclusions of this work and our suggestions for further work.

1.3 Eye movements

The visual acuity of the human eye is highest in the foveal region of the retina and only when the object of interest is held steady on this region. Therefore, the oculomotor system has evolved to perform eye movements which ensure that these conditions are met. These eye movements can be either voluntary or involuntary, and their role is to fixate the fovea to visual stimuli, to track moving stimuli and to compensate for body movements [2]. Depending on the visual stimuli and the viewing conditions, up to five of the following subsystems are involved: the saccadic, smooth pursuit, vestibular, optokinetic and vergence systems.

The **saccadic** system performs the rapid movements of the eyes, called saccades, which reposition the eyes to new stimuli [16], [17]. The **smooth pursuit** system is used when there is a need to track slowly moving targets [18]. In such movements, the velocity of the target in the visual field is matched to the velocity of the eyes. The **vestibular** system is used to stabilise stimuli during brief head rotations [19]. This is achieved by generating eye movements which have equal velocity but opposite direction to that of the head. The **optokinetic** system stabilises the visual field during sustained head rotations [20]. Lastly, the **vergence** system provides stable fixation when there is a change in distance between the eyes and the stimulus [21].

1.4 Saccade dynamic properties

The function of saccades is the rapid reorientation of the eyes that takes place in a variety of eye movements, such as the quick phases of vestibular and optokinetic movements, catch-up saccades during smooth pursuit, reaction saccades, visual search, free scene viewing or reading [16].

The important characteristics of saccades are their latency, duration, peak velocity, amplitude and velocity profile shape [22]. The **saccadic latency,** which is usually 180-220 ms, is the time from the presentation of the stimulus to the onset of the saccade [16]. The **saccade duration** is the time the eye needs to reorient to stimuli. The onset and termination of a saccade can be located by applying a velocity threshold. For example, saccadic onset is commonly considered as the point when the eye velocity first reaches 15 deg/s, and saccadic termination the point when the eye velocity drops below 15 deg/s [23].

Peak saccade velocity is the maximum eye velocity during a saccade, whereas **saccade amplitude** is the distance in degrees the eye translates during a saccade. The **shape of the velocity profile** is described by the ratio of the accelerating and decelerating parts, which sets the value of the shape's skewness.

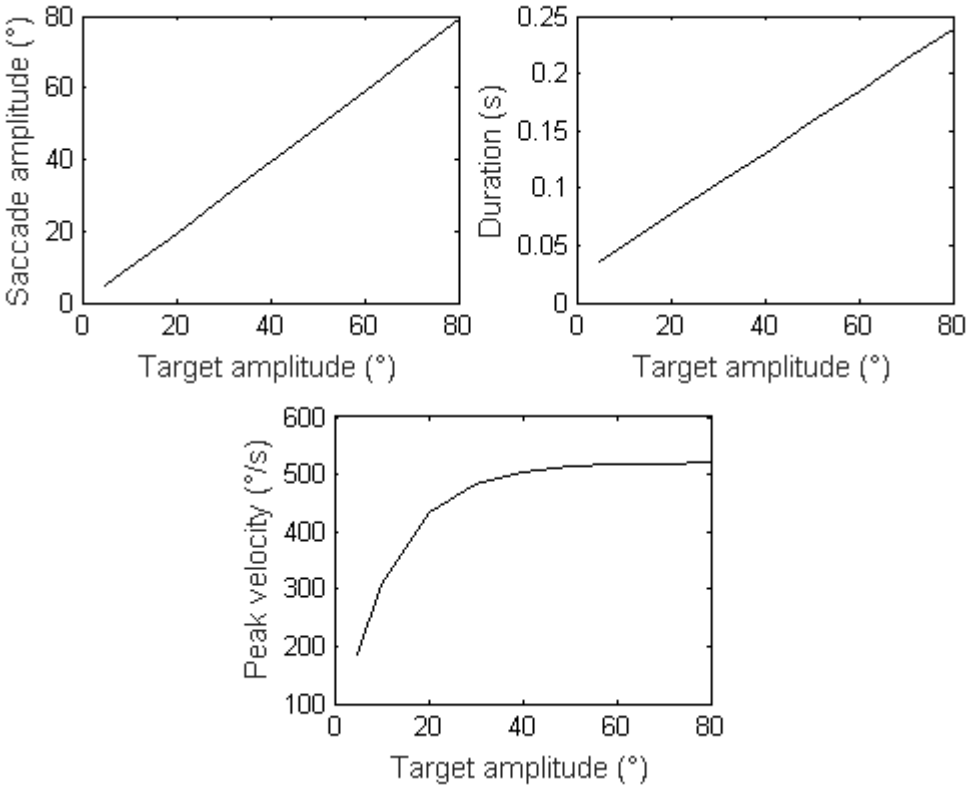

Figure 1.1: Main sequence for horizontal saccade dynamics. Saccadic amplitude (top left), duration (top right) and peak velocity (bottom) for different target amplitudes. The plots were generated using the main sequence equations from [17]. Saccadic amplitude and target amplitude are in degrees (°), duration is in seconds (s), whereas peak velocity is the maximum eye velocity during a saccade measured in degrees per second (°/s).

Saccades exhibit stereotypic behaviour, which can be described using the following relationships between saccadic characteristics. The first two relationships state that the saccade duration increases nearly linearly with amplitude, and peak saccadic velocity increases non-linearly with amplitude at a decreasing rate. Moreover, oblique saccades are virtually straight. These relations are called the **main sequence** (Error! Reference source not found.) [17].

he main sequence is found to hold in a wide range of normal activities, such as reading, viewing pictures, visual search, walking, during optokinetic and vestibular eye movements. This has led to the hypothesis that a common mechanism generates saccades for all of these different tasks [16].

The last relationship describes how the velocity profile shape changes with respect to amplitude [22]. It has been shown that the size of the accelerating part of the velocity profile is crudely persistent for all saccade amplitudes, whereas the size of the decelerating part increases with amplitude size (Figure 1.2). This causes the velocity profile of small saccades to be symmetrical, whereas larger saccades have a more skewed velocity profile [22]. It has also been shown that the main sequence does not hold during all conditions (*e.g.* fatigue), whereas the skewness of the velocity profile against the saccade amplitude does [22].

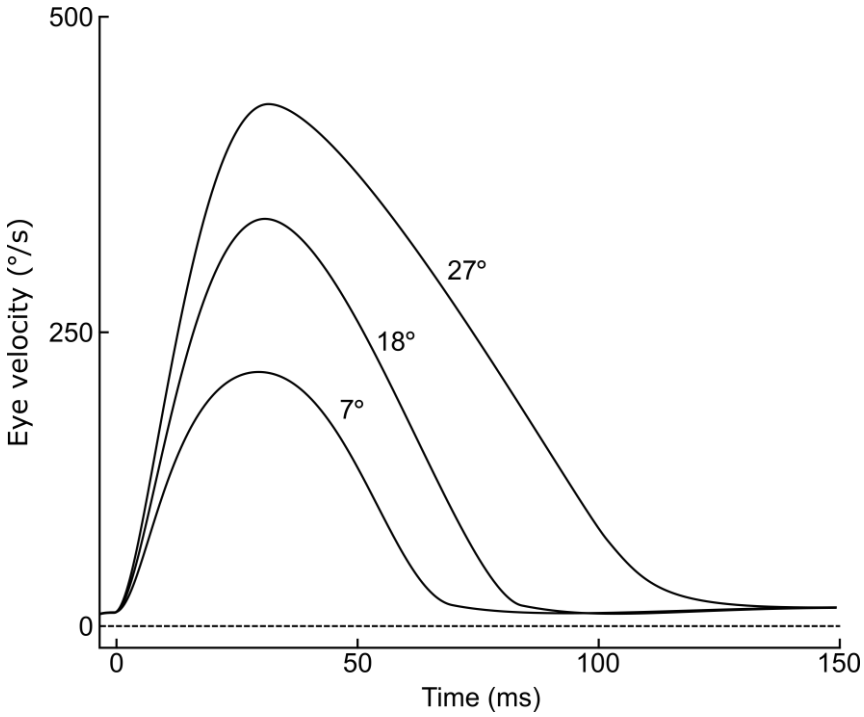

Figure 1.2: Saccadic velocity profiles for three horizontal saccades of different amplitude. Peak velocity occurs at roughly the same time for all saccadic durations, whereas the velocity profile becomes increasingly skewed to the right with increasing saccadic duration. Time is in milliseconds (ms) and eye velocity is in degrees per second (°/s). Adapted from [22].

1.5 The human extraocular muscles

The movements that each eye can perform are generated by six extraocular muscles: lateral rectus, medial rectus, inferior rectus, superior rectus, inferior oblique and superior oblique (Figure 1.3). The muscles work in pairs to rotate the eye along three orthogonal axes. Two axes are for horizontal and vertical movements, whereas the third axis is used to rotate the eye around the retinal axis to correct the torsional effects, which are caused by the rotational of intersecting axes. The lateral and medial rectus muscles are used to rotate the eye around the vertical axis, rotations around the horizontal axis are made by the superior rectus and inferior rectus, and rotations around the retinal axis are made by the superior oblique and inferior oblique [2].

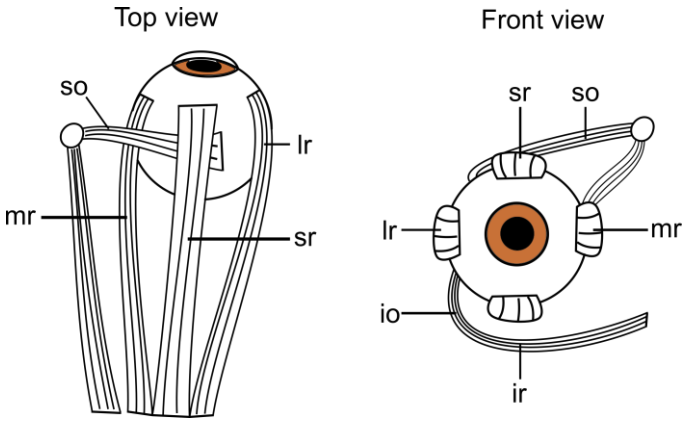

Figure 1.3: Top and front views of the extraocular muscles of the right eye. The extraocular muscles are used to rotate the eye along three orthogonal axes. mr - medial rectus; lr - lateral rectus; sr - superior rectus; ir - inferior rectus; so - superior oblique; io - inferior oblique. Adapted from [24].

1.6 Brainstem control of horizontal saccadic eye movements

The coding of the neural signals that drive the eye muscles to execute the eye movements (described in section 1.3) are generated and modulated by several brain regions located in the cortex, basal ganglia, thalamus, cerebellum and brainstem [16]. In order to generate saccades, which will successfully change the direction of the eyes towards a visual stimulus, a number of information transformations must be performed by the brain. The location of a stimulus in the

visual cortex is represented on a place map where different retinal locations correspond to different locations [2]. By contrast, the oculomotor neurons represent saccades in terms of their temporal discharge (frequency and duration) [2]. This means that the transformations, which must be performed, are from location-encoding on the place map to saccadic commands in terms of temporal discharge of the motor neurons. Moreover, more transformations are required to move the eyes in craniotopic coordinates and from 2D to 3D, because the location of the stimulus is represented in the 2D map, whereas the eye rotates about three axes [2]. The commands for the motor neurons, which generate saccades with specific characteristics issued by higher brain regions, originate in different brainstem regions [9]. As infantile nystagmus is mainly confined to the horizontal plane, in this work we concentrate on the role of specific brainstem regions in generating horizontal saccades, which are believed to be the key areas of the generation of saccades [24].

The motor neurons that directly drive the extraocular muscles are located in cranial nerve nuclei III (oculomotor), IV (trochlear) and VI (abducens) [24]. All these neurons show a typical behaviour of **pulse-slide-step** (Figure 1.4). During horizontal saccades in the direction where the muscle they control is the agonist, motor neurons generate a burst of spikes (**pulse**), with duration approximately equal to the duration of the saccade. By contrast, when saccades are performed in the opposite direction, there is almost no discharge, except near the end of the saccade [25].

During fixations, the neurons fire at a rate that is linearly related to the current position of the eye (Figure 1.5). This behaviour is called the **step,** as there is a difference in the tonic firing rate before and after the saccade (Figure 1.4, bottom panels). Between the pulse and step firing patterns there is an exponential **slide** firing pattern [24] (Figure 1.4, bottom left panel).

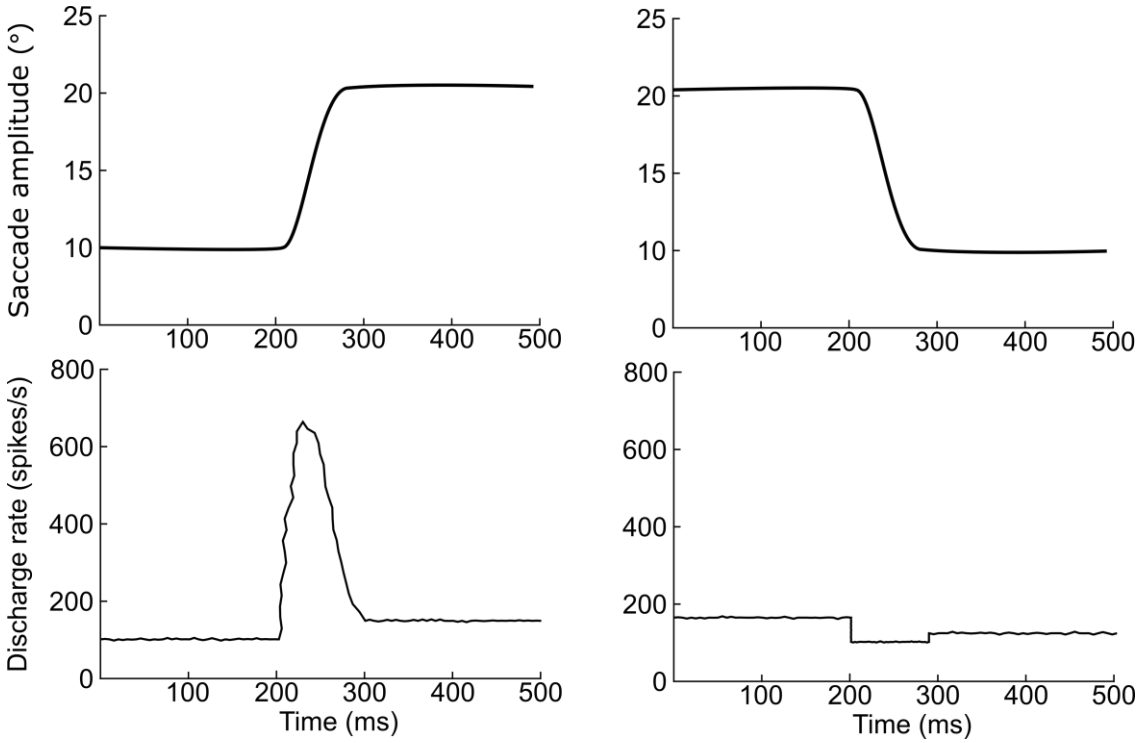

Figure 1.4: Typical firing activity versus time of an abducens motor neuron during rightward and leftward horizontal saccades [24]. The top panels shows the horizontal eye location during a rightward saccade of 10 degrees (top left) and during a leftward saccade of 10 degrees (top right). The bottom left panel shows the neuron's activity during a 10 degrees rightward saccade, whereas the bottom right panel shows its activity during a 10 degrees leftward saccade. Horizontal eye location is in degrees, the motor neuron's firing pattern activity is in spikes per seconds (spikes/s), time is in milliseconds (ms). Adapted from [24].

The input signals for the motor neurons are generated by groups of neurons, which are located in the paramedian pontine reticular formation (PPRF) and in the medulla [24], [26]. The burst and slide signal to the motor neurons is given by the excitatory burst neurons (EBN), which receive signals from the long-lead burst neurons (LLBN) [24], [27]. Both group of neurons produce a high-frequency burst of spikes, but differ in their timing relative to the onset of the saccade. LLBNs fire 40 ms before saccade onset, whereas the EBNs fire 12 ms before saccade onset [2], [24], [27]. The activity of the EBNs is tightly coupled to saccade dynamics. The number of spikes, burst duration and peak firing rate determine the amplitude, duration and peak velocity of each saccade respectively [24].

Furthermore, the EBNs project to inhibitory burst neurons (IBN) which inhibit the firing of the motor neurons controlling the antagonistic muscle [24], [27].

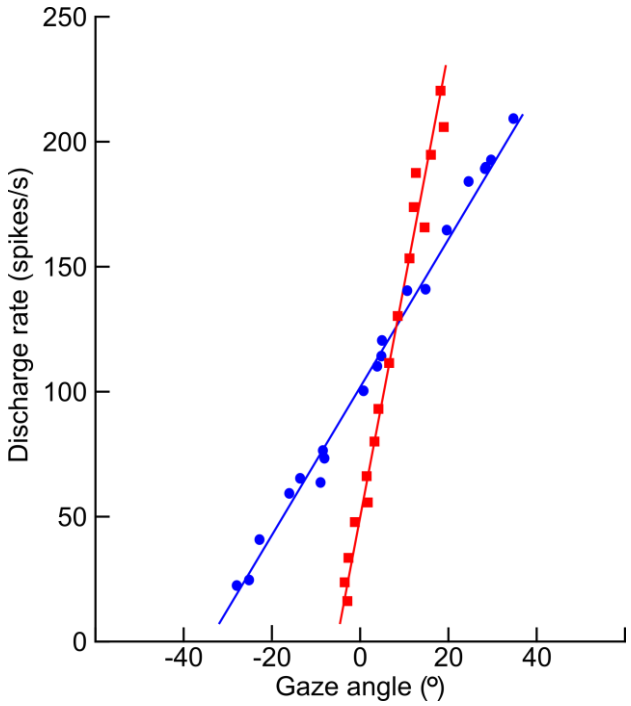

Figure 1.5: Typical discharge rates of two abducens motor neurons versus fixations of different horizontal eye positions. The blue circles show measurements for one neuron; red squares show measurements for the other neuron. The blue and red lines show linear fits to the corresponding measurement points. Eye position is in degrees (°). The motor neuron's discharge rate is in spikes per seconds (spikes/s). Positive gaze angle values represent horizontal eye positions at the right side of the skull, whereas negative at the left side. Adapted from [24].

In order to examine the role of motor neurons in the abducens nucleus and EBNs in the PPRF in the generation of saccades, van Gisbergen *et al.* [25] trained monkeys to follow jumping targets. The recorded neurons were divided into two groups: those which fire maximally during rightward saccades and those which fire maximally for leftward saccades. For a neuron, the direction of maximal firing is called the "**on-direction**", whereas the other direction is called the "**off-direction**". Van Gisbergen *et al.* recorded the action potentials of these neurons in more detail and analysed the relationship between their firing rate, and the

difference between eye position and saccade target (**motor error**). The authors suggested that the firing rate of these neurons is determined by the motor error, quantifying this relationship by averaging the firing curves of individual burst neurons (Figure 1.6). In the on-direction, the mean burst firing curve has the form of an increasing exponential function which saturates at large motor errors, whereas in the off-direction there is a small maximum firing rate close to zero motor error and virtually zero firing rate for larger motor errors (Figure 1.6). This small off-response acts as a **braking signal** that causes the muscles to generate a force directly opposite to the saccade direction, which counteracts the inertia of the eye, thereby preventing overshooting. Moreover, van Gisbergen *et al.* proposed that during saccades, and especially during small saccades (<2 degrees), the motor neurons are driven in push-pull, where both left and right EBNs and IBNs are activated. In this case, van Gisbergen *et al.* suggest that the motor neurons respond to the difference between the EBNs and IBNs activities.

The main input of the LLBNs is believed to be the **superior colliculus** (SC). It is known that stimulation of specific neurons in the deep layers of the SC generates saccades [28]. These specific neurons are topographically organised into a map of saccade vectors in eye-centred coordinates [28]. The neurons near the rostral pole generate small saccades, while neurons in the caudal areas generate large saccades. The high accuracy of the movements is believed to be ensured by an efference copy-based feedback mechanism (using motor error) generated by the LLBNs, which monitors the commands of the firing pattern of EBNs [9]. During saccades, the oculomotor system minimises the motor error, which is the eye displacement in degrees from the target. Therefore, the motor error could be calculated by integrating the firing rate of the EBNs.

The position signal (step), which is sent to the motor neurons, is generated by neurons in the nucleus prepositus hypoglossi (NPH) and the medial vestibular nucleus (MVN), and their firing rate is linearly coupled to eye position [24]. These structures comprise a **neural integrator** which integrates the firing rate of the EBNs in order to produce the step signal [29]. Another important neuron group is the omnipause neurons (OPNs) which act as an inhibitory saccadic gate by inhibiting the EBNs during fixation [24], [27].

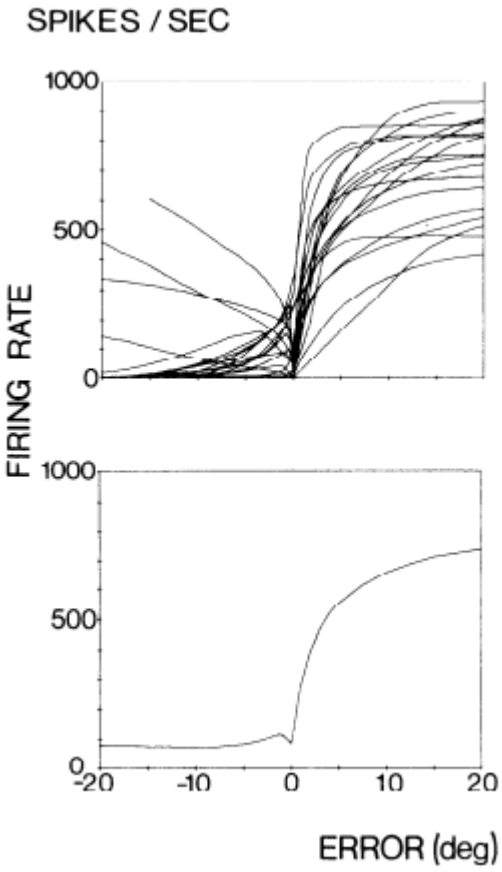

Figure 1.6: Excitatory burst neurons (EBN) firing rate versus different motor error values. The top plot shows smoothed curves of firing rate against motor error recorded experimentally from 18 EBNs. The bottom plot shows the average of the curves derived from the upper plot. Firing rate is in number of spikes per seconds (spikes/sec) and the motor error is in degrees (deg). Figure taken with permission from [25].

1.7 The role of neuronal noise in the saccadic system

Although saccades are highly stereotyped, they exhibit endpoint variability like all body movements [30]–[33]. In the following sections, we will review key studies describing the nature of variability in saccadic eye movements.

1.7.1 Causes of saccadic variability

Harris and Wolpert [33] introduced the idea that **signal-dependent noise** (SDN) on motor commands plays an important role in the dynamics of body movements. According to Harris and Wolpert [33] saccadic eye movements and arm

movements are highly stereotyped due to the strategy employed by the motor system to minimise the variability caused by SDN. They tested their theory by creating two linear models with SDN simulating saccades and arm movements, and a nonlinear model of a two-joint arm. This allowed them to compare the positional variance of simulated saccades and arm movements to empirical data. The authors hypothesised that the trajectories observed in saccades and arm movements can be described as trajectories that minimise the endpoint variance in the presence of SDN in the motor command.

Harris and Wolpert [34] also proposed that not only the endpoint variance of the saccade is important, but also the duration. Due to the lack of visual input during saccades, Harris and Wolpert [34] pointed out that long duration saccades pose an evolutionary disadvantage. Therefore, smaller duration saccades are an advantage, however for these to be possible, the motor signals have to be of large magnitude. This is often associated with extra obstacles, as with the presence of SDN the accuracy is being considerably compromised. Harris and Wolpert [34] suggested that there is an **accuracy-duration trade-off**, which is based on creating the smallest duration saccades possible, given the endpoint variability caused by the SDN. Similarly to their previous study [33], they pointed out that the major factor in endpoint variability is the existence of SDN. They concluded that the trade-off between accuracy and duration is the driving force behind the formation of main sequence saccades.

Harris and Wolpert [33], [34] hypothesised that the only form of noise that could affect the oculomotor system is SDN on motor commands. By contrast, van Beers [30] suggested a model that involved other forms of noise in addition to SDN. He suggested three different sources of noise that could cause the saccadic endpoint variability: uncertainty in the target's location (the perceived location of the target by the visual system), planning of movement (made by higher brain regions that provide input to the brainstem about the type and dynamics of the eye movement *e.g.* saccades to faces or text) and movement execution (the neural signals that are generated during the saccade execution in the brainstem). In his study, van Beers [30] treated the movement planning and execution as one combined source of noise as his approach did not allow him to separate the source of the noise in the motor neurons, apart from the motor commands. He collectively refers to these two types of noise as **motor noise**. The remaining noise source,

referred to as **sensory noise,** is considered to be caused by two factors: (i) the restriction posed by the size and density of the retinal photoreceptors to accurately resolve the target and; (ii) the precision that retinal information is transmitted and processed in the central visual areas. Van Beers [30] did not include the sensory noise in the model but he added its effects to the variability generated by the model which included only the motor noise.

Van Beers' [30] model is based on the estimation of the motor commands which generate the torque causing the eye to generate the experimentally recorded saccades. Van Beers [30] equated the motor commands with the generated torque. The torque equation that van Beers [30] used in his model is the following:

$$\boldsymbol{u} = \boldsymbol{T} = J\frac{d\boldsymbol{\Omega}}{dt} + B\boldsymbol{\Omega} + K\boldsymbol{g} \tag{1.1}$$

where $\boldsymbol{u}$ is a vector with the motor commands, $\boldsymbol{T}$ is a vector with the torque values for the generated saccade, $\frac{d\boldsymbol{\Omega}}{dt}$ is a vector representing the eye's angular acceleration, $\boldsymbol{\Omega}$ is a vector representing the angular velocity of the eye, $\boldsymbol{g}$ is a rotation vectors of the eye orientation, J is the eye's moment of inertia, B is the coefficient of viscosity and K is the stiffness. The motor noise was added to the motor commands at each time step. Van Beers [30] subdivided motor noise into SDN, **constant noise** (CN) and **temporal noise** (TN). The nature of the SDN is based on the proposal of Harris and Wolpert [33], [34], namely a zero-mean, Gaussian white-noise process with variance dependent on the magnitude of the motor command. CN is the noise associated with the activation of muscle pairs. It is a consequence of the fact that in each saccadic movement the muscles which move each eye activate in pairs, and although the torques of the muscles cancel out, the variances do add up. In the case of temporal noise, van Beers [30] pointed out that this kind of noise does not affect the endpoint variability, but causes variability in the temporal aspects of the saccade.

1.7.2 Quantifying the noise

In order to quantify the contribution of each noise type, van Beers [30] performed two experiments which measured the spatiotemporal variability in saccade trajectories as a function of movement direction and amplitude. Then, he fitted the noise parameters of his model by comparing the endpoint variability of the

experimental saccades to the ones generated by the model. This allowed him to distinguish between the different sources and magnitude of the noise that contributed to the endpoint variability. In the first experiment, participants performed horizontal saccades with different amplitudes (2, 4, 6, 8, 10, 12, 14, 16 and 18 degrees). The second experiment consisted of saccades of the same amplitude (9 degrees) but in 24 equally spaced directions. The purpose of the first experiment was to determine the effect of saccade amplitude on the endpoint variability, whereas the purpose of the second experiment was to determine the effect of saccade direction on endpoint variability.

Van Beers' [30] analysis showed that target localisation (sensory noise) is the biggest source of variability (60%). He showed that noise in the motor commands is responsible for a slightly smaller fraction of the variability (40%). This result contradicts the proposal of Harris and Wolpert [33], [34] that SDN is the only source of endpoint variability. Moreover, van Beers found that saccades with longer than average duration have smaller than average peak velocity, and that the variability in pure horizontal and vertical saccades is smaller than in oblique movements. Van Beers [30] hypothesised that SDN could partly originate in the motoneurons and in the brainstem regions, where the motor commands are generated. He also suggested that noisy variations in the representation of the saccade vector in the motor map of the superior colliculus could cause variability in saccade endpoints.

To set the magnitude and the nature of the sensory noise, van Beers [30] used the results of psychophysical studies [35]–[40]. The measurement that was used was the localisation precision, which is the location of the target perceived by the visual system against the true location of the target. Van Beers [30] set the coefficient of variation (CV) of the saccadic endpoint variance caused by the level of localisation precision to be proportional to visual acuity.

The van Beers [30] study showed that there are different types of noise affecting saccades, despite the commonly accepted assumption that SDN is the only source of endpoint variability [33], [34]. Moreover, not only did he provide some quantitative information about the different noise types, but he also showed that these noise types affect saccades by different amounts and in distinct ways. However, the van Beers [30] study cannot provide definitive information about the actual source of the motor noise.

1.7.3 Stereotypy of saccades could be caused by SDN and CN

Building on his previous study, van Beers [31] used the same approach as Harris and Wolpert [33], [34] to propose that motor noise is the cause of the main sequence of saccades. However, van Beers [31] proposed that saccades are optimised to minimise the endpoint variance caused by both SDN and CN. The actual trade-off is between the effects of SDN and CN (Figure 1.7), and not between SDN and the duration of the saccade. Van Beers [31] hypothesised that the main sequence relationships defining saccade dynamics are caused by the above trade-off. If the saccade is of longer duration (smaller speed), the motor neurons would fire with lower power and, consequently, the SDN would be smaller. Thus, CN being independent of motor commands should add up. Consequently, this would yield a bigger endpoint variance. In the opposite case of smaller duration saccades, the SDN would dominate and should produce high endpoint variance.

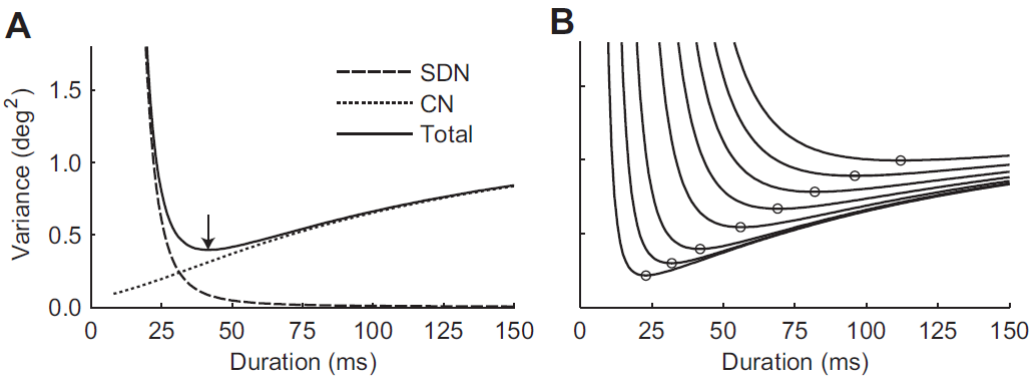

Figure 1.7: Saccadic endpoint variance versus saccadic duration. **A**. The black line represents the total endpoint variance of a 5 degree saccades versus saccadic duration. The dashed lines denotes the endpoint variance caused by the SDN and CN. The arrow indicates the optimal duration that minimises variability caused by the combination of SDN and CN. **B**. Variance curves (as in A) for saccades of amplitude 1, 2.5, 5, 10, 15, 20, 25 and 30 degrees. Duration is in milliseconds (ms). Figure taken with permission from [31].

1.8 Pathological nystagmus

Nystagmus is the term used to describe the involuntary oscillations of the eyes that can be generated physiologically or pathologically [41]. Physiological nystagmus is the optokinetic, vestibular and endpoint oscillations which can be generated by a normally functioning oculomotor system during specific viewing tasks [41]. By contrast, pathological nystagmus is considered an oculomotor disorder which causes the eyes to oscillate involuntarily and almost constantly [42].

Pathological nystagmus can be classified as acquired or infantile. **Acquired nystagmus (AN)** is mainly caused later in life by neurological and vestibular diseases [43]. The onset of AN is usually after 4 months of age and is more frequent in adults [42]. **Infantile nystagmus (IN)** is either present already at birth or develops within 6 months of life [42]–[44]. Infantile nystagmus can be subdivided into a number of categories, with the primary interest of this work being infantile ocular and idiopathic nystagmus. **Ocular infantile nystagmus** is associated with different sensory defects, such as infantile cataract [45], [46], retinal dystrophy, optic nerve hypoplasia, albinism, achromatopsia and strabismus [43], [47], [48]. **Idiopathic infantile nystagmus** so far has not been found to be caused by any specific defects. The prevalence of idiopathic infantile nystagmus was found to be 14 per 10,000, while that of all forms of nystagmus is 24 per 10,000 [49]. The oscillations cause deterioration of visual acuity, as the visual target is no longer held steady in the foveal region [50]. Moreover, the visual acuity could be further reduced by a defective visual system [51]. Pathological nystagmus widely affects the life of patients (*e.g.* physical and social movement restriction, standing out/not fitting in, negativity with regards to the future and relationships) [52], [53]. Therefore, such significant impact on quality of life makes the development of potential treatments for nystagmus patients a high priority.

1.8.1 Aetiology of infantile nystagmus

The aetiology of IN is still the subject of a debate [54]. It is often related to ocular deficits; however how these can lead to nystagmus is not yet known [48], [54]. Some of these ocular defects are congenital cataract, retinal dystrophy and optic nerve hypoplasia [43]. The difficulty in distinguishing the waveforms generated

by idiopathic and ocular nystagmus could mean that the difference between these two types lies in their underlying causes [48]. The wide variety of mechanisms, which have been proposed for the development of nystagmus, are based on: (i) an afferent visual pathway pathology; (ii) ocular deficits that could cause a normal oculomotor system to become misconfigured during development; (iii) abnormal development of the motor central nervous system; and (iv) extraocular muscle pathology.

The events that could lead to IN are probably initiated by specific gene mutations affecting cell physiology, thereby disrupting the development of the connectivity and function of the afferent visual system and/or the oculomotor system [54]. Using optical coherence tomography (OCT), Thomas *et al.* [55] found that mutations in the FRMD7 gene play an important role in retinal formation patterns. They found that the foveal pit of a number of IN patients with the FRMD7 mutations are significantly shallower and show abnormality in the optic nerve's morphology in respect to reduced optic disk area, retinal nerve fibre layer thickness, cup area and depth. These results show that a number of IN patients have abnormal retinas and optic nerves. Moreover, specific gene mutations can affect the structure of the extraocular muscles or the feedback which is generated by their proprioceptors or both [54]. Chen *et al.* [56] proposed that IN could be caused by the existence of a positive feedback system which coexists with the negative feedback. They hypothesised that the positive feedback could be developed by the misrouting of the optic fibers.

1.8.2 Infantile nystagmus waveform characteristics

The key characteristics of the nystagmus waveform are: the oscillation plane, the oscillation period and amplitude, the fast and slow phase, the null zone, the foveation window and the baseline oscillation (Figure 1.8) [51], [57]. The nystagmus most commonly takes place in the horizontal plane, but vertical and torsional movements (and a combination of these two) have also been observed [58]. The amplitude and period of the oscillation can change based on gaze angle and attention level [58].

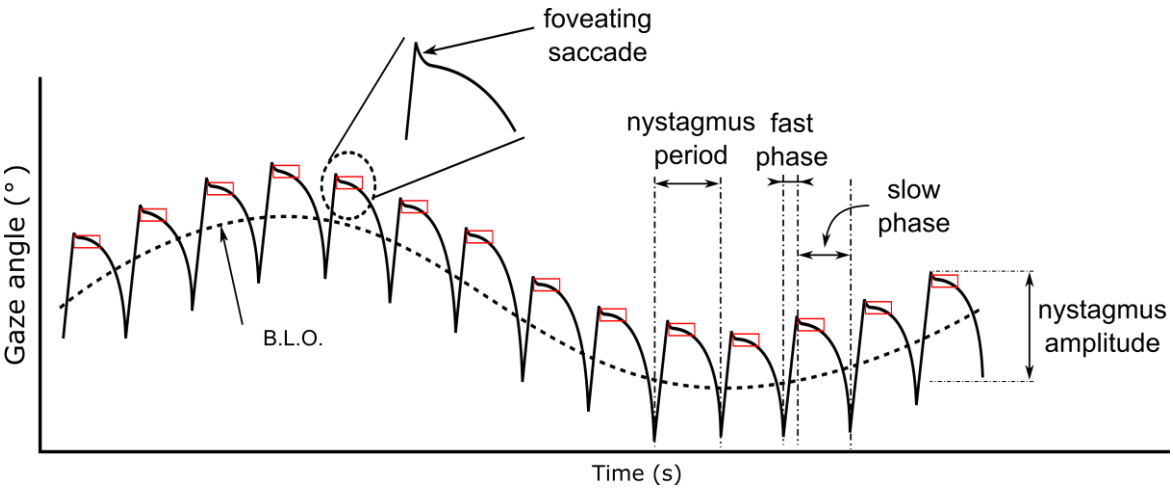

Figure 1.8: Schematic illustration of a unidirectional jerk nystagmus waveform with its most important characteristics highlighted. Time is given in seconds (s). Horizontal eye position (gaze) is in degrees (°). B.L.O.: nystagmus baseline oscillation. For more details please refer to the main text (section 1.8.2). Adapted from [57].

Typically, the **slow phase** takes the eye away from the target, whereas the **fast phase** takes it back to the target. In cases where the target is returned to the fovea by a slow phase, the nystagmus is referred to as **pendular**. The **null zone** is the position of the eye where the intensity of the nystagmus is minimal, and can be observed at left/right/up/down gazes [42]. The **foveation window** is the region where the velocity of the eyes is low enough (< 4 deg/s) for the patient to be able to see clearly. The **baseline oscillation** is a low frequency sinusoidal oscillation, the amplitude which is correlated with the nystagmus amplitude [57]. All nystagmus oscillations are influenced by fixation attempt, gaze angle, fatigue, stress, attention and arousal [43].

Waveform shape is an important characteristic of nystagmus oscillations. There have been at least 12 different idiopathic infantile and ocular nystagmus waveforms identified in clinical studies [59]. These waveforms can be divided into three groups: pendular, jerk and dual. Figure 1.9 shows schematics of these nystagmus waveform groups, whereas Figure 1.10 shows a pure pendular waveform with a shift in the position of the oscillation relative to the (stationary) target.

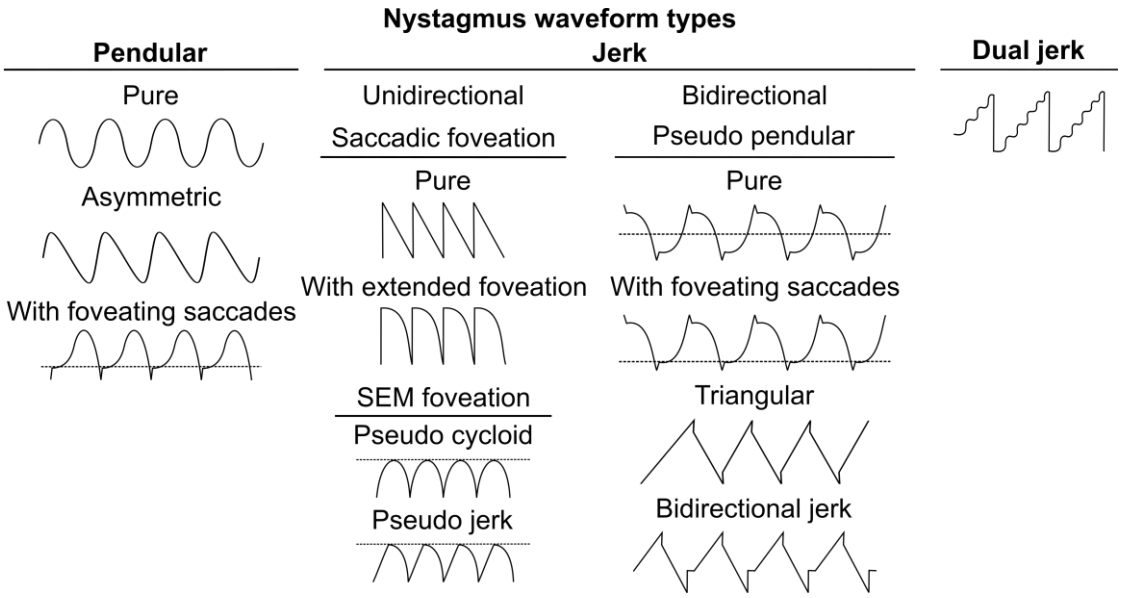

Figure 1.9: Different nystagmus waveform types. The dashed line represents the fixation point. Vertical axis represents horizontal gaze angle in degrees and horizontal axis is time in seconds. SEM: slow eye movement. Adapted from [59].

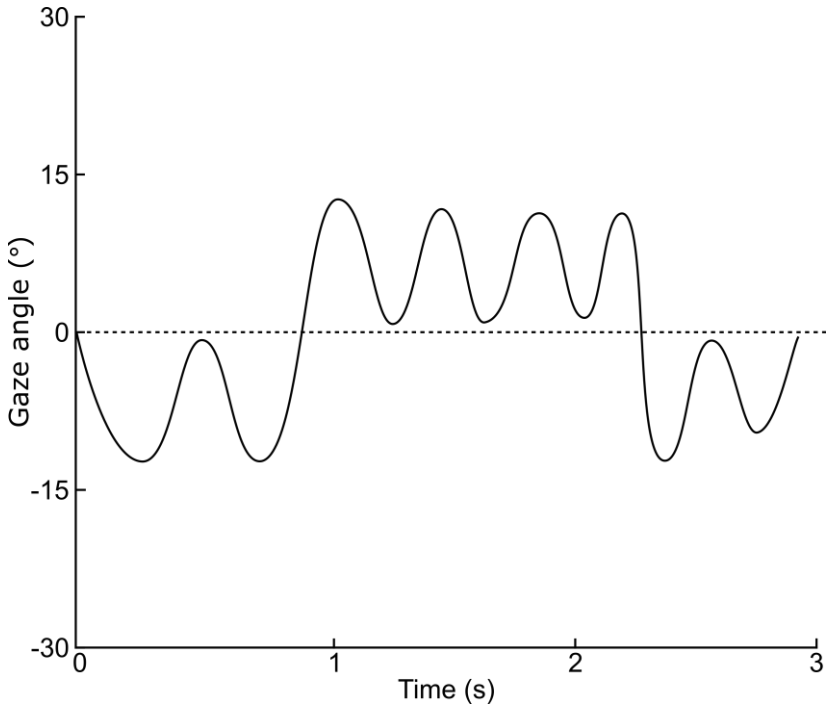

Figure 1.10: Pendular nystagmus with a shift in eye position (the target is at 0 degs). Vertical axis is horizontal gaze angle in degrees (°) and horizontal axis is time in seconds (s). Adapted from [59].

Jerk nystagmus can be divided into two sub-categories: **unidirectional** and **bidirectional**. In unidirectional jerk nystagmus waveforms, there is one saccade in each oscillation period, and this is always towards the target. This type of nystagmus is further subdivided into four categories: **pure jerk**, **jerk with extended foveation** (which the foveation window is larger compared to other jerk waveforms), **pseudo-cycloid** and **pseudo-jerk**. In the first two types, the saccade is of sufficient amplitude to refoveate the target, whereas in the other two it is not.

Bidirectional jerk nystagmus waveforms, are classified into four subtypes: **pure pseudo pendular**, **pseudo pendular with foveating saccades**, **triangular** and **bidirectional jerk** where the direction of the fast phase changes in each cycle.

2 Nonlinear oculomotor modelling and analysis

This chapter will present different approaches used to model the mechanisms of the oculomotor system that could cause infantile nystagmus. We will also describe the nonlinear dynamics of the Broomhead *et al*. [1] model and the nonlinear time series analysis methods which has already been applied by Akman *et al.* [15] to experimental infantile nystagmus recordings.

2.1 Nystagmus modelling approaches

The different approaches to model infantile nystagmus are based on the assumption that a number of malfunctions exists in the oculomotor system. The most straightforward approach is based on the assumption that there is a malfunction of the leaky neural integrator, which is the gaze-holding neural mechanism [60]. The inputs to the neural integrator are eye position and velocity feedback loops. Optican and Zee [60] proposed that in normal subjects the position signal is positive, whereas the velocity signal is negative. According to their work, the velocity signal in patients with infantile nystagmus is reversed and this has been hypothesised to make the neural integrator unstable. Although, the Optican and Zee [60] model is capable of simulating a wide range of infantile nystagmus waveforms by small parameter change values, it cannot generate pure pendular waveforms.

Brodsky and Dell'Osso [61] proposed a model in which nystagmus oscillations can be generated by the activation of the foveate smooth pursuit system and the afoveate full-field optokinetic system. The former could be modulated by the visual cortex and the latter by the **accessory optic system** (AOS). The AOS neurons have large receptive fields, react to slow-moving stimuli and provide visuovestibular abilities in afoveate animals by using input from the retina. The AOS in afoveate animals is used to track moving stimuli. When smooth pursuit is active, the visual scene moves across the retina, thereby providing an optokinetic stimulus to the AOS. Normally in humans, the AOS is active only for the first 2 months of life and is then suppressed. The failure of this suppression, caused by

a number of factors (*e.g.* visual deprivation), can allow the AOS to be active during smooth pursuit and can lead to nystagmus oscillations. The Brodsky and Dell'Osso [61] model is able to generate both normal eye movements and nystagmus.

Harris and Berry [44] proposed a developmental model describing how nystagmus could manifest. Their proposal is based on the hypothesis that abnormal infantile visual experience causes the oculomotor system to develop eye movement strategies that could lead to improved visual experience. It was found that the stimulus speed is an inverse function of the grating spatial frequency [44]. Stimulus motion can improve the contrast sensitivity to a visual grating. They claim that the nystagmus oscillation is a developed strategy for the reduced contrast sensitivity to high-spatial frequencies during the early critical period (<6 months of age), when the oculomotor system is formed. They also proposed that the jerk oscillations are the best compromise between keeping the fovea near the visual target and moving the eye, thus reducing the effects of low contrast sensitivity to high-spatial frequencies.

The nystagmus modelling approach, which we have examined in this work, is based on the assumption that the nystagmus oscillations are generated by an abnormality in the saccadic system. In the next section we will present the Broomhead *et al.* [1] model based on this assumption, and this model is the central focus of this thesis.

2.2 The saccadic model of Broomhead *et al.*

The nonlinear dynamics model of the saccadic system proposed by Broomhead *et al.* [1] is able to simulate both normal saccades and infantile nystagmus oscillations in the horizontal plane. The model has demonstrated that infantile nystagmus can be treated as a dynamical disease. In such diseases, there is no structural damage to the system - as had been proposed by models based on control theory - however, the system's components interact outside of their normal parameter range [14].

The model with parameters shown in Table 2.1 (we call this parameter set **deterministic parameters**) comprises of the following six coupled ordinary differential equations (ODEs):

$$\frac{dg}{dt} = v \tag{2.1}$$

$$\frac{dv}{dt} = -\left(\frac{1}{T_1} + \frac{1}{T_2}\right)v - \frac{1}{T_1 T_2}g + \frac{1}{T_1 T_2}n + \left(\frac{1}{T_1} + \frac{1}{T_2}\right)(r - l) \tag{2.2}$$

$$\frac{dn}{dt} = -\frac{1}{T_N}n + r - l \tag{2.3}$$

$$\frac{dr}{dt} = \frac{1}{\varepsilon}\left(-r - \gamma r l^2 + F_b(m)\right) \tag{2.4}$$

$$\frac{dl}{dt} = \frac{1}{\varepsilon}\left(-l - \gamma l r^2 + F_b(-m)\right) \tag{2.5}$$

$$\frac{dm}{dt} = -(r - l) \tag{2.6}$$

Equations (2.1) and (2.2) model the response of the eye muscles to the saccadic control signal, with g and v representing the horizontal eye position (gaze angle) and eye velocity respectively. These equations (2.1) and (2.2) are also defined as the model's muscle plant. The control signal is a weighted sum of the firing rate $(r - l)$ generated by the EBNs and IBNs (collectively defined as the **burst neurons**) and the position command n (tonic signal), which is obtained from $(r - l)$ through mathematical integration by the neural integrator (NI). The value of $(r - l)$ is highly correlated with eye velocity. The eye muscles are modelled as a second order overdamped linear system with time constants $T_1 = 0.15$ s and $T_2 = 0.012$ s [12], [60]. Equation (2.3) is used to model the leaky neural integrator signal, which is characterised by a time constant $T_N = 25$ s [60]. Equations (2.4) and (2.5) model the activities of the EBNs and IBNs. Following the experimental work of van Gisbergen *et al.* [25] described in section 1.6, these are split into two populations, right (r) and left (l), depending on the direction of motion that elicits maximal firing. The net output from the neurons is the difference between the activities of the two populations, $(r - l)$. The mutual inhibition between the left and right populations of burst neurons is modelled by the terms $\gamma r l^2$ and $\gamma l r^2$ in equations (2.4) and (2.5), with the parameter γ representing the strength of the inhibition. The function $F_b(m)$ is defined by:

$$F_b(m) = \begin{cases} a'\left(1 - e^{-m/\beta'}\right) \; if \; m \geq 0 \\ -\dfrac{\alpha}{\beta} m e^{m/\beta} \qquad if \; m < 0 \end{cases} \tag{2.7}$$

which models the response of the EBNs neurons to the motor error signal m, where m specifies the difference between the current and target eye displacements (motor error). The form of $F_b(m)$ is based on the results of van Gisbergen *et al.* [25] shown in Figure 1.6. A schematic plot of $F_b(m)$ is shown in Figure 2.1.

Table 2.1: Broomhead *et al.* [1] model parameters (deterministic parameters).

Parameter name	Range	Description
α	$0 < \alpha \leq 1000$	Magnitude in firing rate of the off-response
β	$0 < \beta \leq 60$	Value in degrees where the off-response is maximal
ε	$0 < \varepsilon \leq 0.1$	Inverse burst neuron response time
γ	$0 \leq \gamma$	The strength of the mutual inhibition between the left and right populations of EBN neurons
α'	$0 < \alpha' \leq 1000$	Maximum firing rate of the on-response
β'	$0 < \beta' \leq 60$	Value in degrees where the on-response saturates

The four positive parameters α', β', α and β specifying $F_b(m)$ govern the forms of the simulated on- and off-responses for the EBN neurons (see Figure 2.1). The parameters a' and β' in equation (2.7) specify the neurons' mean response to motor error signals generated by motion in the direction of maximal firing (the on-response), whereas the parameters α and β specify the neurons' mean response

to motor error in motion in the opposite direction (off-response). As mentioned in section 1.6, the latter is believed to cause the braking signal and thus slowing the eye towards the end of a saccade in order to prevent overshoot of the target [25]. $F_b(m)$ is strictly increasing when m is positive, with $F_b(m) \rightarrow \alpha'$ as $m \rightarrow \infty$. Thus α' sets the value at which the on-response of the burst neurons saturates as $m \rightarrow \infty$. Moreover, $F_b(\beta') = \alpha'(1-e^{-1})$. Therefore, β' determines how quickly the saturation occurs. When m is negative, there is a single global maximum of $F_b(m)$ at $(-\beta, \alpha/e)$. $F_b(m)$ converges monotonically to 0 as $m \rightarrow -\infty$ from $-\beta$. The magnitude of the off-response is determined by the value of α, while the value of β determines the off-response range. Larger values of α thus simulate a stronger braking signal. The parameter ε in equations (2.4) and (2.5) determines how quickly the burst neurons respond to the motor error. As the value of ε increases, the response time decreases.

The form of equation (2.6) is based on the assumption of a resettable integrator – separate from the NI – that integrates the signal from the burst neurons. This would provide an estimate of the eye's current displacement, from which the motor error m is computed. The initial value of m denotes the initial motor error, which designates the required displacement of the eye to foveate the target. In a normal oculomotor system, equations (2.4)-(2.6) act collectively to drive the motor error to 0, corresponding to the execution of an accurate saccade [1], [62].

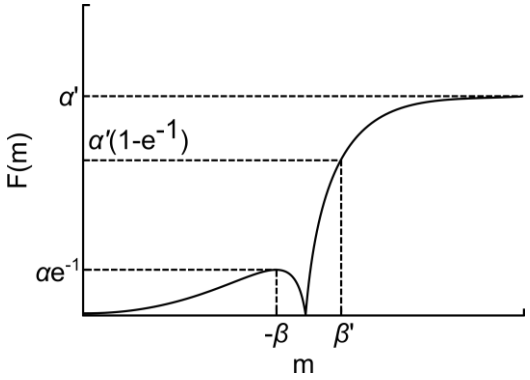

Figure 2.1: The function $F(m)$ modelling EBN firing rate (spikes s^{-1}) versus motor error m (degs). The parameters α' and β' determine the on-response while α and β determine the off-response.

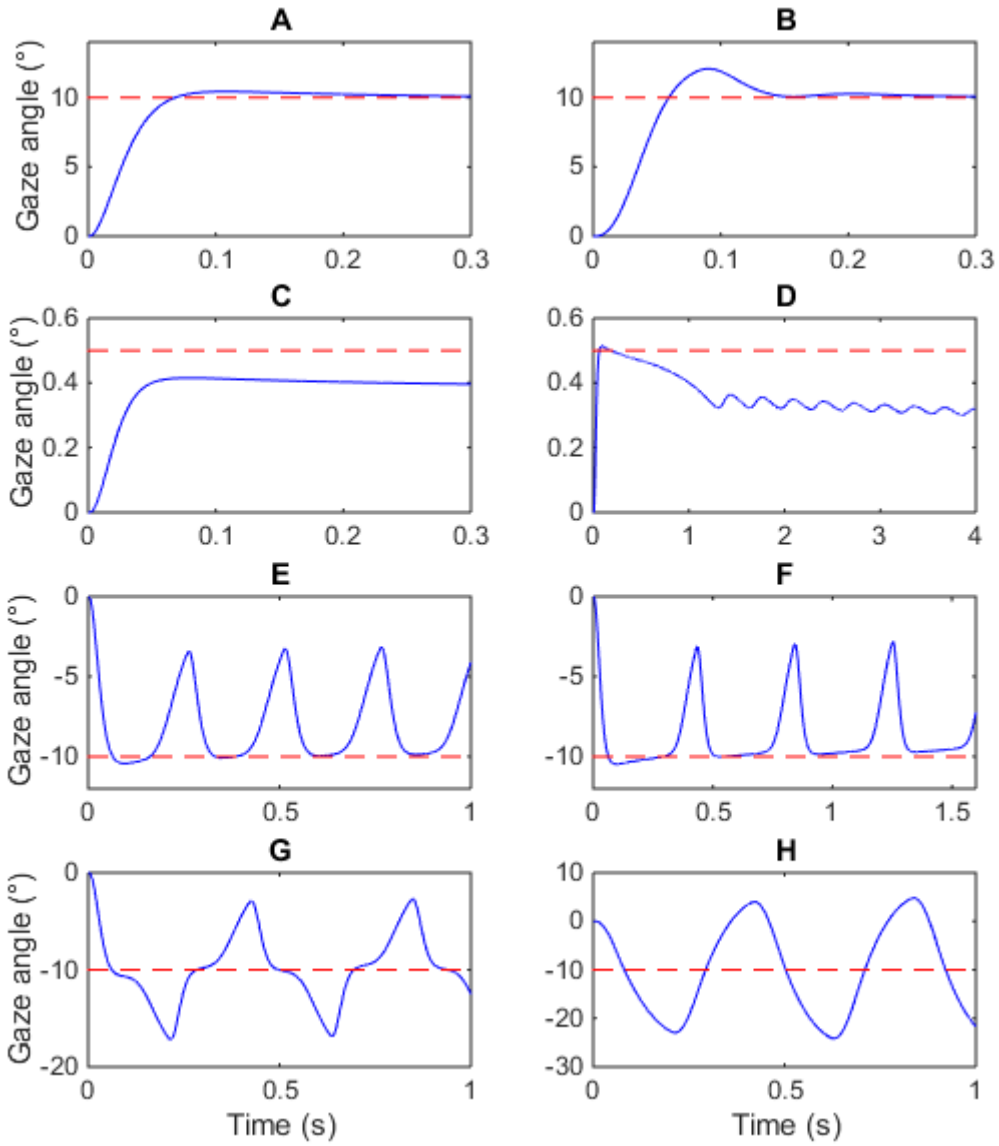

Figure 2.2: Horizontal gaze angle time series versus time generated by the Broomhead *et al.* [1] model using different parameter choices for α, β, and ε, and initial conditions $g_0 = v_0 = n_0 = 0$ and $m_0 = \Delta g$, simulating a saccade of Δg degs executed from rest at the primary position (0 degs). The values of parameters γ, α' and β' were fixed at 0.05, 600 and 9, respectively. **A**: Normometric saccade. $\alpha = 20$, $\beta = 3$, $\varepsilon = 0.001$, $\Delta g = 10$. **B**: Dynamic overshoot. $\alpha = 20, \beta = 3, \varepsilon = 0.015, \Delta g = 10$. **C**: Hypometric saccade. $\alpha = 206$, $\beta = 3$, $\varepsilon = 0.001$, $\Delta g = 0.5$. **D**: Small-amplitude nystagmus. $\alpha = 207.656$, $\beta = 3$, $\varepsilon = 0.006$, $\Delta g = 0.5$. **E**: Jerk IN. $\alpha = 240$, $\beta = 3$, $\varepsilon = 0.004$, $\Delta g = -10$. **F**: Jerk with extended foveation. $\alpha = 240$, $\beta = 3$, $\varepsilon = 0.0048, \Delta g = -10$. **G**: Bidirectional jerk IN. $\alpha = 240, \beta = 3, \varepsilon = 0.006, \Delta g = -10$. **H**: Pendular IN. $\alpha = 240, \beta = 3, \varepsilon = 0.06, \Delta g = -10$. The dashed red

lines represent the saccadic target position. Time is in seconds (s). Gaze angle is in degrees (°). Adapted from [1].

Broomhead *et al*. [1] empirically selected the values for the parameters α', β' and γ to match the main sequence, and then examined the effect of varying the parameters α (off-response magnitude), β (off-response range) and ε (inverse burst neuron response time). This preliminary numerical parameter search of these three parameters revealed that the model is capable of generating accurate and inaccurate saccades, as well as several infantile nystagmus oscillations (refer to Figure 2.2).

2.3 Bifurcation analysis of the Broomhead *et al.* model

In order to quantify the parameter regimes in which different simulated nystagmus waveforms could be generated, Akman *et al.* [62] performed a bifurcation analysis of the model for the parameters α, β and ε, and made a number of predictions on the aetiology of infantile nystagmus. They proposed that: (i) infantile nystagmus could be caused by an abnormally large braking signal at the end of saccades; (ii) the slow-fast form of the jerk and bidirectional jerk oscillation could be caused by the geometry of the saccadic phase space, meaning that the fast phases are not separate corrective saccades; (iii) there is a high likelihood that the oscillation is jerk, bidirectional jerk or pendular in the oscillatory regime (which is consistent with clinical eye movement studies); (iv) saccadic dynamic overshoots can be caused by saccadic control signal reversals; (v) by increasing the value of ε, the model generates a transition from jerk to bidirectional jerk and then pendular oscillations. The value of parameter ε sets the response speed of the burst neurons to the motor error signal, which can be interpreted as the attention level of the participant. This is consistent with the experimental findings that attention levels change the oscillation shape [43].

2.4 Nonlinear time series analysis of jerk infantile nystagmus

Akman *et al.* [15] performed a nonlinear time series analysis of experimental infantile jerk nystagmus time series and found that there is an increase in the local dimensions of the oscillation during the fast phase compared to the fixed point (where the participant tries to fixate). They also showed a reduction at the

slow phase compared to the fixed point. Moreover, Akman *et al.* [15] found a refixation mechanism that is more complex than a homoclinic reinjection, indicating that the jerk nystagmus oscillations are not periodic. Based on their results, they made two hypotheses regarding the origin of this nonperiodicity and the variation in local dimensions: (i) additional oculomotor systems are activated at the start of the fast phase and; (ii) the presence of signal dependent noise in the saccadic system.

As described in section 1.5, Harris and Wolpert [33] hypothesised that SDN on motor commands plays an important role in the form of body movements. Based on their assumptions, Akman *et al.* [15] proposed that SDN in these neurons could cause the observed increase in local dimensions due to the enhanced firing of the burst neurons just before the fast phase. This could also cause the aperiodicity of the waveform and the stochastic refixation. In this work, we examined the hypothesis that a model with SDN on the burst neuron signal could generate the above behaviour.

The methodology used by Akman *et al.* [15] to measure the local dimensions at different points of a nystagmus time series was based on the delay embedding approach, which had been previously applied to experimental nystagmus time series quantifying the stability of the system's fixed points and the geometry of the system's attractor. In the delay embedding approach, also known as the method of delays, the dynamics of the system is first reconstructed by moving a sliding window of d samples through the data, generating a sequence of d-dimensional vectors. Given a time series of gaze position $\{g_1, g_2, \dots g_N\}$ this yields delay vectors $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots \boldsymbol{w}_{N-d+1}\}$, where:

$$\boldsymbol{w}_k = (g_k, g_{k+1}, \dots g_{k+d-1})^T.$$

Because the dynamics of the underlying and embedded system are closely related, the delay vectors can be used to quantify the dynamic behaviour of the oculomotor system [15], [63]. In order to choose the value of d, we apply singular value decomposition (SVD) method to the global trajectory matrix X of the time series. X is defined as the $(N - d + 1) \times d$ matrix which k row is $w_k^T/\sqrt{N}$. SVD identifies the linear subspace of the delay space, in which the reconstructed dynamics exhibit the most variation [64]. Using the singular value spectrum $\{\sigma_1, \sigma_2, \dots, \sigma_d; \sigma_{k+1} \geq \sigma_k\}$ of X, we chose the value of d for which the first $d - 1$

singular vectors $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{d-1}\}$ of X contains a significant proportion of the total variance (>90%) and project the delay vectors onto the subspace spanned by $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_d\}$. By applying this to the experimental data recorded from two subjects with jerk IN, Akman *et al.* [15] chose $d = 7$. The left panel of Figure 2.3 shows a part of one of the experimental time series used in the analysis of Akman *et al.* [15], while the right panel shows the projection of the corresponding delay vectors $\boldsymbol{w}_k$ onto the first two singular vectors $\{\boldsymbol{c}_1, \boldsymbol{c}_2\}$ of X.

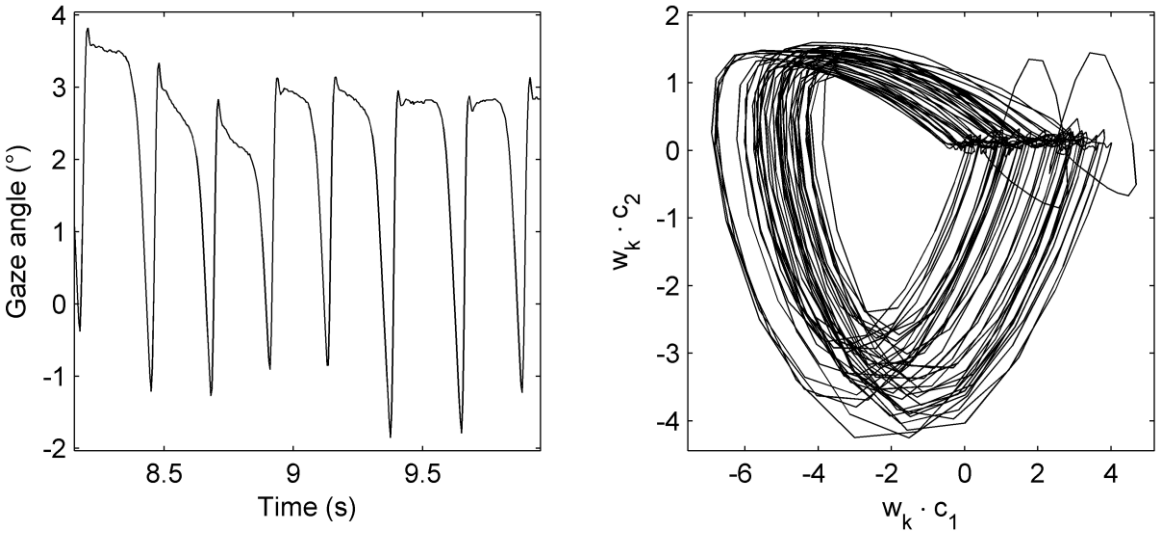

Figure 2.3: Part of a nystagmus time series and its reconstructed global dynamics. **Left panel**: Part of the nystagmus time series used by [15]. Time is in seconds (s). Vertical axis shows the horizontal gaze angle in degrees. **Right panel**: Global dynamics reconstructed from the time series using delay embedding projected onto the first two singular vectors of the trajectory matrix. Adapted from [15].

The next step in the analysis is to find the fixed point of the reconstructed dynamics. This point is assumed to be where the individual tries to fixate [15], [63]. If a fixed point exists, it will be located on the principal diagonal of the delay space and in a region where the flow velocity is low. A rough approximation is therefore provided by the delay vector of minimum velocity lying within a small distance of the diagonal [15], [63] (*e.g.* see the right panel of Figure 2.3).

Next, to calculate the local dimensions in the neighbourhood of a given point $\bar{\boldsymbol{w}}$, we calculated the SVD of the local trajectory matrix $X_\varepsilon(\bar{\boldsymbol{w}})$ at $\bar{\boldsymbol{w}}$. This is defined in terms of the ε-neighbourhood of $\bar{\boldsymbol{w}}$, $B_\varepsilon(\bar{\boldsymbol{w}}) = \{\mathbf{w}_\mathrm{k} : \|\mathbf{w}_\mathrm{k} - \bar{\boldsymbol{w}}\| < \varepsilon\}$, as the $|B_\varepsilon(\bar{\boldsymbol{w}})| \times d$ matrix which rows consists of the vectors $\left\{\frac{(\boldsymbol{w}_k - \bar{\boldsymbol{w}})^T}{\sqrt{|B_\varepsilon(\bar{\boldsymbol{w}})|}} : \boldsymbol{w}_k \in B_\varepsilon(\bar{\boldsymbol{w}})\right\}$. To estimate the local dimensions, we observe the scaling behaviour of the singular values of $X_\varepsilon(\bar{\boldsymbol{w}})$ as the neighbourhood size ε varies. The local dimensions is the number of singular values which are either independent of ε or scale linearly with ε that are above a noise floor [64]. The noise floor is the ratio of the measurement device resolution (*i.e.* eye tracker) to the standard deviation of the normalised time series recording [15]. Figure 2.4 shows the reconstructed dynamics obtained by the method of delays for a 10 s extract from an experimental nystagmus time series, whereas Figure 2.5 shows the singular value scaling for each of the points shown in Figure 2.4.

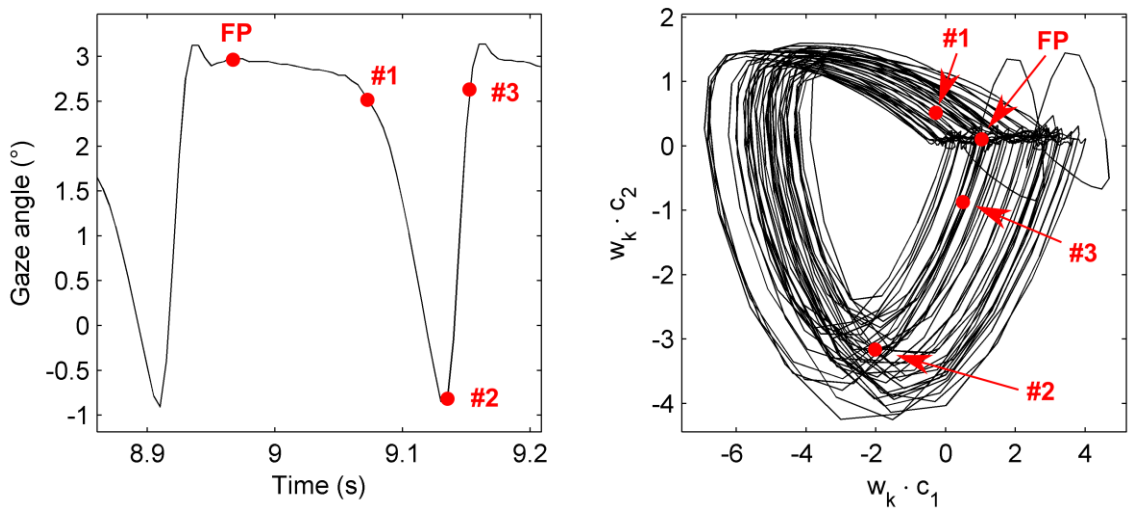

Figure 2.4: Reconstructed global dynamics from a 10 s portion of a jerk nystagmus time series and the locations of specific points on the waveform. **Right panel**: Projection of the reconstructed attractor onto the two singular vectors of the trajectory matrix. FP denotes the estimated fixed point, while #1, #2 and #3 denote points on the attractor corresponding to the start of the slow phase, the beginning of the fast phase and the end of the fast phase, respectively. **Left panel**: Section of the time series indicating the phases of the oscillations corresponding to points FP, #1, #2 and #3. Adapted from [15].

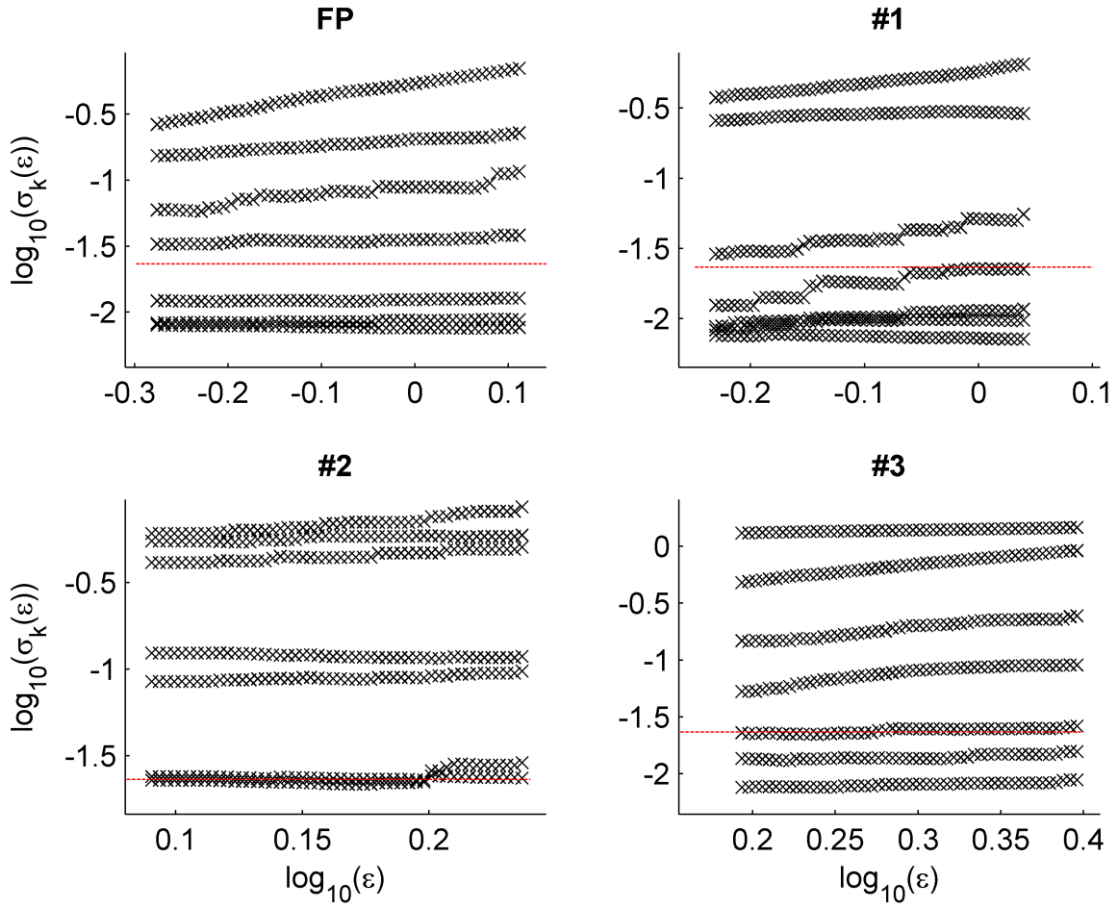

Figure 2.5: Scaling of local singular values at the points shown in Figure 2.4. FP denotes the estimated fixed point, while #1, #2 and #3 denote points on the attractor corresponding to the start of the slow phase, the beginning of the fast phase and the end of the fast phase, respectively. In each plot, the red horizontal dashed line indicates the estimated noise floor. The scaling of singular values with ball size ε is indicated by the X points. The local dimensions pattern, which was measured by Akman *et al.* [15], is shown. The local dimensions at the fixed point (FP) is four, at the start of the slow phase (#1) it reduces to three, at the start of the fast phase (#2) it increases to five, which is the largest local dimensions on the attractor, and at the end of the fast phase (#3) the local dimensions reduces back to four.

The local dimensions at the fixed point (Figure 2.5 FP) is four. At the start of the slow phase (Figure 2.5 #1) it reduces to three, at the start of the fast phase (Figure 2.5 #2) it increases to five, which is the largest local dimensions on the attractor, whereas at the end of the fast phase (Figure 2.5 #3) the number of local

dimensions reduces back to four. Interestingly, Akman *et al.* [15] found this pattern in two jerk time series originated from different nystagmats.

3 Optimisation of computational biological models

The main objective of any optimisation process is to find the best solution to a specific problem [65]. One can appreciate the strength of the optimisation process by taking a closer look at the evolution of life on Earth. Evolution through natural selection is an optimisation process in which individuals within the population of a species adapt to their environment in order to increase their chances of reproduction. The most successfully adapted individuals are more likely to pass their genes onto the next generation. Iteratively, this process will yield a population comprised of individuals with characteristics adapted to the environment allowing them to reproduce.

In this Chapter we will describe the methods that we used to optimise the Broomhead *et al.* [1] and stochastic models that we developed in this work. This Chapter is organised as follows. In section 3.1, we will provide a definition of the mathematical optimisation. In section 3.2, we will describe the challenges in optimising biological models. We will describe genetics algorithms in section 3.3 and multi-objectives genetic algorithms in section 3.4. In section 3.6, we will describe the graphics processing units for scientific applications and in section 3.7 how graphics processing units can be used to accelerate the execution of genetic algorithms.

3.1 Mathematical optimisation

Mathematical optimisation or **optimization problem** can be defined in the form of:

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{x}) \\ \text{subject to} \quad & f_i(\boldsymbol{x}) \leq b_i, \qquad i = 1, \dots, m \\ & f_i(\boldsymbol{x}) = b_i, \qquad i = m+1, \dots, k \end{aligned} \tag{3.1}$$

Where x is a vector $\boldsymbol{x} = x_1, \dots, x_n$ which is the optimisation variable (or **decision variable**) of the optimisation problem, $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the **objective function**,

$f_i: \mathbb{R}^n \to \mathbb{R}, i = 1, \dots, m, m+1, \dots, k$ are the **constraint** functions, and $b_1, \dots, b_m, b_{m+1}, \dots, b_k$ are the constraint values [66].

This process finds the specific variable value (s) that minimises the output of a given objective function [65], [66]. The decision variable values are varied during the optimisation process. The objective functions quantify the quality of the solution for each decision variable value, whereas the constraint is the requirement which must be met by the optimisation process.

In systems biology, optimisation has been used in the formulation of mathematical models for biological systems [67] and to test whether different modelling assumptions are correct [68]. The most commonly used optimisation in biological computational modelling is **parameter estimation**. The aim of parameter estimation is to find the model's parameter values that set it to exhibit the same or similar behaviour (qualitative or quantitative) as that measured from the real system which is being modelled. Parameter estimation is also often called parameter **fitting**. This process is a critical step in the construction and analysis of biological models [69] and by finding the optimal parameter values for a model it can be assessed whether the assumptions made in forming the model were likely to be correct. For consistency in this work, we will consider **parameter estimation (or fitting) as the default optimisation process of biological models**.

Different optimisation methods have been developed which can be applied to complex computational biological models. These methods search the space of decision variables (parameters) in order to find the minimum of the objective function. In the next section, we will introduce some of the challenges in the optimisation of computational biological models.

3.2 Challenges in optimising computational biological models

The optimisation of computational models of biological systems contains a number of difficulties and these are based on: (i) the existence of a single global minimum or a number of local minima; (ii) how feasible it is to find the exact solution to the problem, rather than an approximation; (ii) the smoothness or continuity of the objective function surface; and (iv) the number of objectives the optimisation problem includes. To address these obstacles, different optimisation methods have been developed. These can be classified based on the following

characteristics: deterministic or stochastic in nature, suitable for global or local optimisation, use in linear or nonlinear problems, and the use of the objective function's derivative [70].

3.2.1 Existence of global and local minima

The existence of multiple minima in the optimisation problem depends on the objective function's output surface. The surface is formed by the objective function's values for different decision variable (parameter) combinations, $\{(x, f(x))|x \in A\}$. In the case where the surface contains one global minimum, the optimisation problem is designated as **unimodal**, whereas in the case of multiple minima it is designated as **multimodal**.

Based on the ability of the optimisation method to find the global minimum in unimodal or multimodal objective function surfaces, we can classify two types of optimisation methods: local and global [70]. Local optimisation methods are able to find local minima. The main advantage of these methods is that they converge fast to the local optimum values, whereas their main weakness is that they can be used only for problems with a unique minimum. This means that if there are multiple local minima, the method cannot find the global minimum if it has converged to a local minimum. One of the widely used local search methods is **gradient descent,** which uses the derivative of the objective function to find the minimum. In this method, the search for the minimum starts from a point that is usually chosen randomly. The search process is iterative: in each iteration of the loop, a step is taken proportional to the gradient of the objective function. Usually, the stopping criterion is that there is a small change to the function value compared to the previous iteration. If there is no previous knowledge about the existence of multiple local minima of the optimisation problem, this method would be the first choice. To test whether a local optimisation method is sufficient, multiple runs are performed using different initial search points in each run. If the solutions found from all the runs are the same minima, there is strong indication that the objective function has only one minimum.

In the case where there are multiple minima of the objective function, a global optimisation method has to be used. Global optimisation methods are classified as **deterministic** or **stochastic** [71]. Deterministic methods have the advantage of ensuring that a global optimum of the problem will be found, but in the case of

nonlinear problems, there is a high computational cost to do so. On the other hand, stochastic methods use a heuristic based approach to find an approximation of the global minimum. However, these methods do not ensure finding the global minimum, but rather offer the advantage of fewer computational resources, when compared with the deterministic methods.

3.2.2 Exact or approximate solutions

Based on the precision of the results, optimisation methods can be classified into two major categories: **exact** and **approximate** [70]. Exact methods guarantee finding the exact optima of the problem. However, their main disadvantage is that they cannot provide the optima in a practical time when the optimisation problem is large and complex. In the case of approximate methods, there is no guarantee that the exact optimal solution will be found, but they provide an alternative solution, or a number of solutions, which are close to the optimal solution, and in a substantially shorter time compared to the exact methods [70].

The high complexity of biological systems does not often allow the creation of models which can be optimised using the exact method in reasonable time scales. The most widely used approximate methods are **metaheuristic**. They are usually stochastic methods which combine heuristic procedures in order to search the parameter space of the objective function. Here, we quote the definition of Romanycia and Pelletier [70]. Although their definition is widely used in the field of Artificial Intelligence (AI), we believe that it can be applied to the field of optimisation as well:

> *"a heuristic in AI is any device, be it a program, rule, piece of knowledge, etc., which one is not entirely confident will be useful in providing a practical solution, but which one has reason to believe will be useful, and which is added to a problem-solving system in expectation that on average the performance will improve".*

Metaheuristic methods have been used in systems biology with considerable success in parameter estimation [72], such as the optimisation of artificial neural network architecture [73].

3.2.3 Smoothness and continuity of objective function

Non-smooth optimisation problems are the ones for which the objective function's surface has sharp changes. This could arise, for example from a nonlinear model with a rich bifurcation structure, thereby creating extreme changes to the objective function's surface in certain parameter directions. This can hinder the performance of any optimisation method and it is more problematic for methods that use the derivative of the objective function, as they are not able to determine the direction of the optimal value. A non-continuous objective function can cause similar problem, where the objective function does not provide output for a number of parameter values.

3.2.4 Multi-objective optimisation

In the case where multiple objective functions need to be optimised, the optimisation is defined as multi-objective. This case is often used when the objectives that need to be optimised are conflicting. One approach would be to use the weighted-sum method [74]. In this method, the weighted sum of the objectives is minimised, thus converting the problem into a standard single-objective optimisation. However, this method requires the weight values to be specified and in most cases there is no *a-priori* knowledge of the relative importance of the objectives.

In the case where there is no *a-priori* knowledge of the weights, multi-objective optimisation algorithms are often used which are based on the concept of *Pareto* dominance [75]. These algorithms return a set of estimated *Pareto* optimal solutions, or estimated ***Pareto* optimal set**, that represent the optimal trade-off between the individual objectives. The *Pareto* set consists of solutions for which none of the objectives can be improved without compromising one of the others. The image of this set in objective space is referred to as the ***Pareto* front**. The solutions are categorised on their ***Pareto* dominance**, which states that solution A dominates solution B if and only if the following two conditions hold: first, solution A is no worse than solution B on all objectives and second, solution A is better for at least one objective [76]. Using this rule on a set of solutions, levels of dominance can be created such that the first level's solutions are not dominated by any other solution (**non-domination levels**). An effective multi-objective optimisation algorithm returns estimated *Pareto* optimal solutions,

which lie close to the *Pareto* optimal set, and for which the corresponding points on the *Pareto* front are well spread in objective space (Figure 3.1). A good spread of solutions facilitates the selection of a final solution from the *Pareto* set, and the method used to do this is based on the problem of interest. The final selection of the best individual often uses the fitness value of all the objectives of the individuals which are part of the *Pareto* optimal set [77].

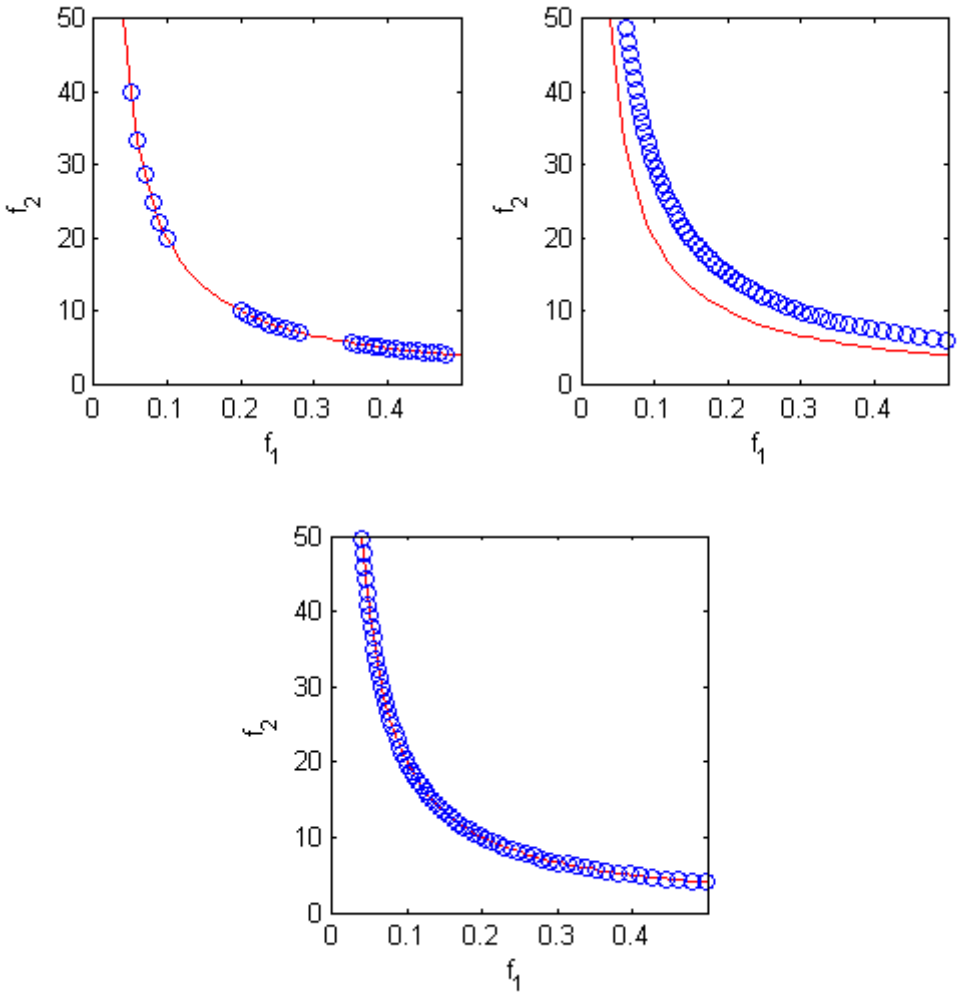

Figure 3.1: Different approximations of the *Pareto* optimal front of a hypothetical two-objective optimisation problem (minimisation). The red line represents the *Pareto* optimal front and the blue circles represent solutions found by the optimisation method. The axes f_1 and f_2 represent the objective values of the optimisation problem. **Top left**: good convergence of the solutions to the *Pareto* optimal front but with poor spread. **Top right**: good spread but poor convergence of the found solutions. **Bottom**: Good convergence and spread of the solutions to the *Pareto* optimal front. Adapted from [78].

3.3 Genetic algorithms

The Genetic Algorithm (GA) is a widely used global, stochastic and metaheuristic optimisation method. It is based on mimicking natural selection, in which an evolving population of candidate problem solutions (*e.g.* parameter sets) is used to approximate the optimal solution (the parameter set yielding the highest fitness) [70]. The advantage of genetic algorithms over other optimisation methods (*e.g.* gradient descent) is that they do not need any assumptions about the shape of the objective function's output (*i.e.* continuous, non-smooth).

3.3.1 Genetic algorithm - general procedure

Initially, the GA creates a population of individuals, each of which represents a candidate solution for the problem by encoding a particular combination of the model's parameter values (in the case of model parameter estimation). The individuals can be represented (encoded) using a number of different coding schemes, such us binary string, real-valued numbers [79]. During the GA execution, the evolution of the individuals is an iterative process, in which at each iteration (called a generation) a value is set to each individual, which represents how successful it is in solving the problem (*i.e.* the goodness-of-fit). Using these values, the GA applies a number of stochastic genetic operators to the individuals in order to create a new population. The individuals with the best fitness values are more likely to be used in order to guide the search towards the optimal solution [79]. The termination condition of the GA is often based on the number of generations reached, or a fitness value threshold reached [70]. Figure 3.2 represents a flowchart for the GA described in this section.

The development and analysis of selection and genetic operators have been an active subject in the genetic algorithm research area [80]–[83]. In the next sections, we will present a number of the most commonly used selection and genetic operators.

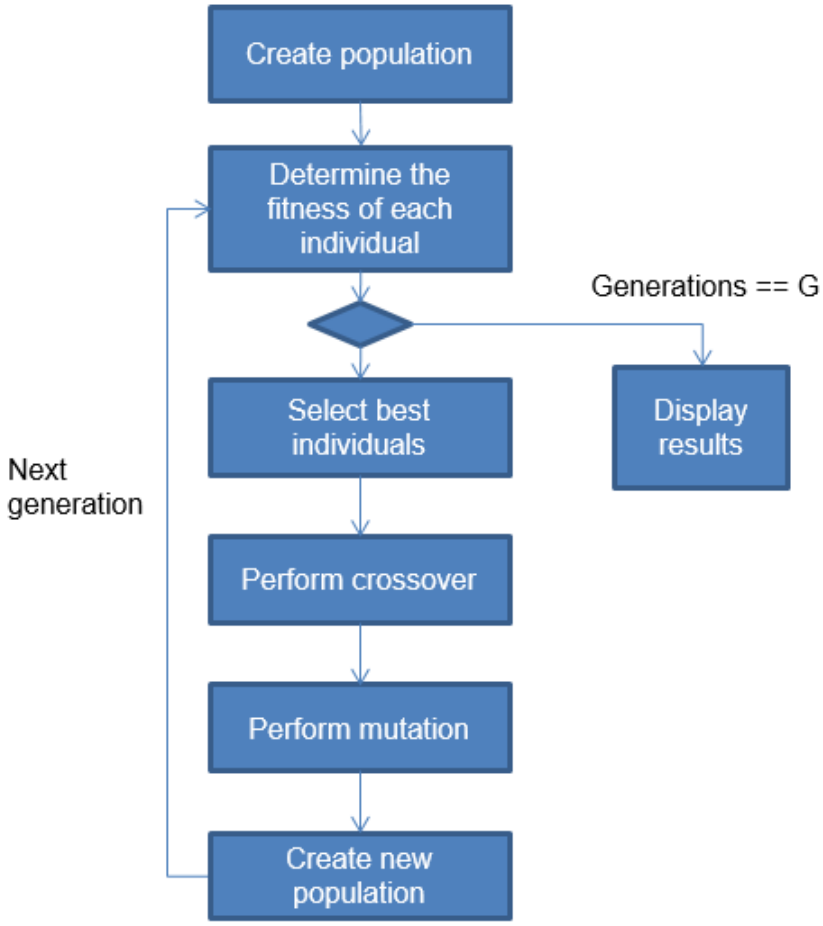

Figure 3.2: Flowchart for the genetic algorithm described in this section. G indicates the number of generations at which the genetic algorithm will terminate. For more details please refer to the main text (section 3.3.1).

3.3.2 Selection operators

A selection operator selects individuals who are combined with each other in order to create the new population. Here we present the most well-known methods [84].

Elitism copies the best individuals of the current population to the next one, ensuring that the best individuals will be in the new population.

Roulette-wheel is the selection operator in which the selection of the individual is proportional to its fitness evaluation. The probability of an individual being selected is given by the following equation:

$$p(i) = \frac{f(i)}{\sum_{j=1}^{n} f(j)}$$

where i is an individual from the population, n is the population number and $f(i) > 0$ is the fitness value of the i individual.

Rank selection is a form of roulette-wheel selection, where the probability of an individual to be selected is based on its rank. In a population of size n, the individual with the best fitness value will get the rank value n and the worst individual will get a rank value of 1. The probability of an individual to be selected is given by the following equation:

$$p(i) = \frac{rank(i)}{n(n-1)}$$

where i is an individual of the population and $rank(i)$ is the rank value of the ith individual.

Tournament selection, where n tournaments with K individuals are created. The best individuals of each tournament are selected based on their fitness values.

3.3.3 Crossover operators

The crossover operators specify how two individuals are combined to create a new individual who will be a part of the new population [85]. Here we present a number of different crossover operators.

Single-point crossover occurs when a single point in the parental genome is randomly chosen to split their genomes into two parts. These parts are combined to form two new individuals where each child will have one part of the genome from each parent. **Two-point** crossover is when two positions are chosen in the genome of the parents to create two new individuals.

Intermediate crossover, which is defined for real encoding, occurs when the children are created by the weighted average of the parents. The weight of the average is taken as a uniform random number between the values 0 and 1. The child genome using intermediate crossover is given by the following equation:

$$child = parent1 + random \times (parent2 - parent1)$$

Heuristic crossover, which is defined for real encoding, occurs when the child is created on the line connecting the genotype of the two parents, but on a distance

away from the parent with the better fitness and at the direction away from the parent with the worst fitness. The child genome using heuristic crossover is given by the following equation:

$$child = parent2 + Ratio \times (parent1 - parent2)$$

where $Ratio$ is the constant that weights the distance of the child away from the parent with the highest fitness. The fitness of $parent1$ is higher than the other parent's. The child's genome is in a distance away from the parent with the best fit.

3.3.4 Mutation operators

Mutation in the population provides an extra variability to the explored solutions and ensures that the population will not be fixed permanently in a single area of parameter space [79]. Depending on the population encoding, different mutation operators can be used, such as binary, uniform and Gaussian.

Binary mutation is applied to populations with bit string encoding. In this case a randomly chosen bit from the individual's genome is flipped.

Uniform mutation can be applied to integer and float encoded populations. In this case, the value in the individual's genome is replaced with a uniformly chosen value within the lower and upper constraint bounds (if constraints exist).

Gaussian mutation also can be applied to integer and float population encoding. Here however, the gene value is replaced with a new value chosen from a Gaussian distribution. If the new value is outside the genome's constraints, the value is resampled to the nearest constraint value.

Adaptive feasible mutation mutates the individuals towards locations of parameter space which, based on previous generations, are more likely to produce individuals with better fitness.

3.3.5 Selection of GA parameters

The performance of each genetic operator depends on the fitness function, the population encoding, population size and other GA-specific parameters, such as crossover type. However, there is no definite proof regarding which combinations of genetic operators are the best [79]. Moreover, when forming the fitness function for a specific optimisation problem, each version of the fitness function

should be tested using the GA, thus resulting in a complex and time consuming process [67]. This might lead to an iterative process, where each candidate fitness function and GA parameter combination (*e.g.* population size, mutation rate/type, crossover type, selection type) should be tested. These tests involve multiple runs for each fitness function and GA parameter combination, due to the stochastic nature of the GA. These multiple runs will allow the estimation of the distributions of key benchmark measures, such as the distribution of the returned fitness values and the execution time. Usually the goal is to find the GA parameter combination with the smallest mean and standard deviation fitness values. Low mean and standard deviation values show that a specific GA parameter combination provides consistently small fitness values (good fit) for a large number of GA runs [70]. This could indicate that the GA has converged to specific solutions that provide good fits to the problem. The process of forming the appropriate fitness function and choosing the most suitable GA parameter values can be time consuming, and requires considerable computational resources [79].

3.4 Non-dominated sorting genetic algorithm II

Multi-objective genetic algorithms (MOGAs) are variants of standard GAs which incorporate procedures specialised in optimising multi-objective problems. One of the most widely used and successful MOGA is the Non-dominated Sorting Genetic Algorithm II (NSGA-II) [86]. The main characteristics of NSGA-II is the presence of elitism, a fast non-dominated sorting approach and a selection operator that ensures better spread of the solutions and better convergence near the *Pareto* optimal front [86]. In this section we describe the procedure of the NSGA-II used in this work. It should be noted that this procedure may not apply to all MOGAs.

The procedure of NSGA-II is as follows (see Figure 3.3): initially, a random population P_0 of size N is created. Each individual is assigned with a rank designating its non-domination level and all the individuals are sorted based on that rank. Subsequently, the operators of binary tournament selection, crossover and mutation are applied to the population. These operators create a new population Q_0 of the same size as the initial population.

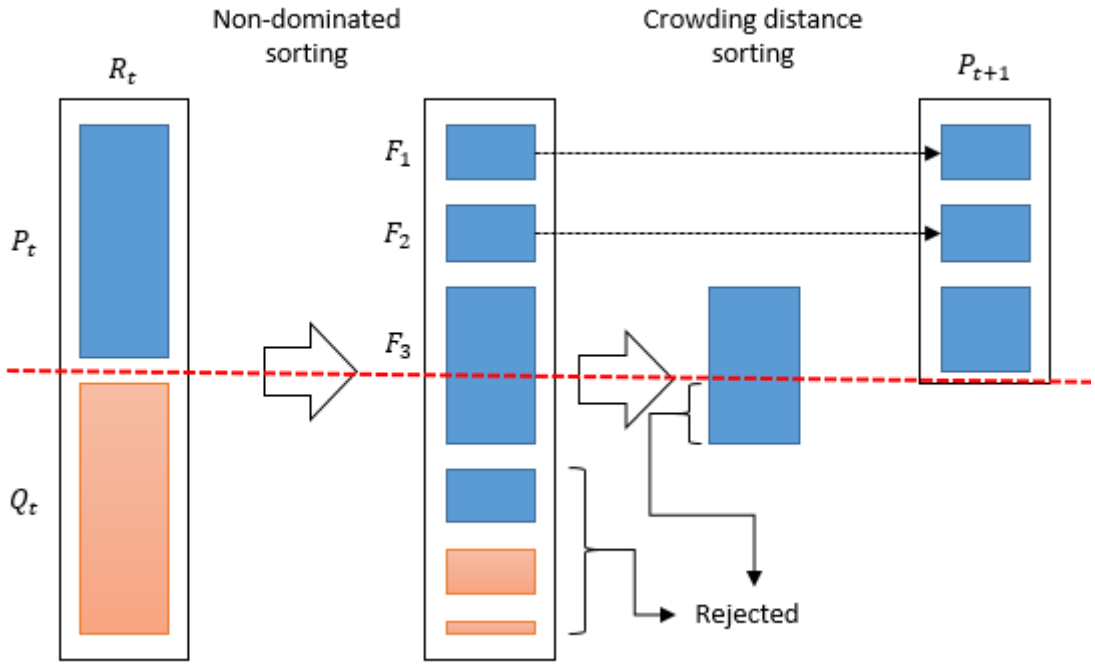

Figure 3.3: NSGA-II procedure. See main text for more details. Figure adapted from [86].

In the following iterations of the algorithm, the populations P_t and Q_t are merged and form a new population R_t, which has the size of the two combined populations. The population R_t is sorted by non-domination rank. Individuals in the best non-dominated front are in the F_1 set, those in the second non-dominated front are in the F_2 set, and so on. The individuals in the F_1 set are the best individuals of the P_t and Q_t populations. This ensures elitism, namely the best individuals always survive in the new population. If the number of members of set F_1 is smaller than the population size *N*, then all the members of the set are chosen for the new population and the next individuals to fill the new population are chosen from the other non-dominated levels. If the last set that is chosen to fill the new population with *N* individuals is bigger than the number of individuals required, then the members of that set are sorted by the crowded-comparison operator ($\prec$) and the best of these are chosen to complete the population. The crowded-comparison operator is based on an estimate of the density of individuals around another individual. The operator sorts the individuals according to the following rules:

$$i \prec j \text{ if } (i_{\text{ rank}} < j_{\text{ rank}}) \text{ or } ((i_{\text{ rank}} = j_{\text{ rank}}) \text{ and } (i_{\text{ distance}} > j_{\text{ distance}}))$$

Where i is better than j, $i_{\text{ rank}}$ is the non-domination rank and $i_{\text{ distance}}$ is the crowding distance of individual i. In other words, the individual with lower rank is chosen, or, if the individuals are a part of the same front (same rank), the individual that is in the less crowded region (larger crowding distance) is chosen. This selection provides better spread of individuals that form the population P_{t+1}. After this procedure, the selection and genetic operators are applied again to create a new child population Q_{t+1}. When the algorithm terminates, a population of individuals is returned which represents the best trade-off of the set of criteria being evaluated, and at this point the best individual of the returned population is chosen.

The crowding distance of the individuals in a set is calculated using the following steps [87]:

1. For all the individuals $i = 1, 2, \dots, N$ of the set we assign the initial crowding distance value of $d_i = 0$.
2. For each objective function $m = 1, 2, \dots, M$ the individuals are sorted from worse to best values and create the sorted set D^m.
3. For all the objectives $m = 1, 2, \dots, M$ we assigned infinity as the distance value to the boundary solutions. For the rest of the individuals of the set the crowding distance is calculated using the following:

$$d_{D_j^m} = d_{D_j^m} + \frac{f_m^{\left(D_{j+1}^m\right)} - f_m^{\left(D_{j-1}^m\right)}}{f_m^{(max)} - f_m^{(min)}}$$

where D_j^m is the index of the j-th member of the m objective in the sorted list. $f_m^{(max)}$ and $f_m^{(min)}$ are the maximum and minimum fitness values for each objective designated by m, respectively.

3.5 Parallel Genetic Algorithms

Effective search of the parameter space can require a large population size (*i.e.* a large number of candidate solutions). Moreover, fitting the parameters of computational biological models to multiple experimental datasets often gives rise to complex fitness functions and high-dimensional search spaces; therefore, GAs can be very computationally demanding. Furthermore, the application of the

GA genetic operators to big populations can also lead to high computational needs. This can be further amplified by two factors (refer to section 3.3.5). Firstly, the GA itself has a number of parameters (population size, crossover type, mutation type, *etc.*), which can be tuned. Secondly, the stochastic nature of the GA means that multiple runs are necessary to check whether the solutions have stably converged with a given set of GA parameters [70]. Hence, it is necessary to systematically test different combinations of GA parameters to ensure that the solution obtained is a global optimum, rather than a local optimum.

For these reasons, considerable efforts have been made to reduce the execution time of GAs. Different genetic operators have been developed with the aim of decreasing the requirement for computational resources [85], [88]. Other improvements have focussed on adding parallelism to GAs in order be able to use parallel computer architectures [70]. For a given set of GA parameters, the same fitness function and genetic operators are applied to the solutions in the population. This makes GAs very suitable for parallelisation and many parallel GA models have been developed which can speed up their execution [70].

The GA, depicted in the flowchart of Figure 3.2, is a **Sequential GA** (sGA), and was not developed to be used in parallel computer architectures. The **Parallel GA** (pGA) was formed in order to take advantage of parallel organisations and parallel hardware computer architectures. The pGA can provide results for optimisation problems which need a very large population in order to approximate the global optima of the fitness function. This large population often requires considerable computational resources in the form of CPU time or computer memory [89]. All modern computer CPUs have the capability to run a number of parallel processes (most mid-range CPUs have 4 or more computing cores). A GA implementation must be able to be executed on a parallel computer architecture, otherwise a lot of computational resources will be wasted.

In our work, we will examine two GA parallel models which can be executed on computer hardware with high parallel capabilities: the **independent runs parallel model** and the **master-slave model**. The independent runs parallel model describes the execution of multiple sGAs which run on a pool of processors [70]. For example, if one has at his disposal 10 computers, he can use them to run one sGA on each one. This will allow to quickly acquire data for the statistical analysis about the found solutions by the GA (*e.g.* convergence) [70]. In the

master-slave model the evaluation of the fitness function is distributed to a number of different processors (called “Slaves”) and the genetic algorithm operations are performed by a single processor (called “Master”).

3.6 General-purpose computing on graphics processing units (GPGPU)

Generally, the high computational demands of GAs means that high-performance computing (HPC) clusters (*i.e.* multiple connected computers) are required to obtain results within a reasonable time frame. As described section 3.5, the intrinsic parallel architecture of GAs allows them to be executed over multiple CPU cores within an HPC cluster. However, although HPC clusters provide significant processing power, they have a number of disadvantages. Among these is their purchase price, as well as the high costs associated with maintenance and energy consumption.

An alternative approach to HPC clusters is to use consumer level graphics processing units (GPUs). These cards possess parallel computing power that can be compared to that of all the CPUs in a small HPC cluster. Furthermore, GPUs have a considerably lower cost and require substantially less maintenance.

Current GPU models that are commonly used for gaming in personal computers, can also be utilised for scientific computing. They include a large number of relatively slow processing units compared to CPUs, they are optimised for single instruction multiple data (SIMD) processing and allow massive parallelism [90]. The collective processing power of the processing units in a GPU is much higher than that of a single CPU. Therefore, GPUs are considerably faster in parallel computations but slower in serial computations compared to CPUs. The GPU features that are important for scientific applications are the following: (i) the peak single floating point and double floating point precision processing performance measured in giga floating-point operations per second (GFLOPS); (ii) the memory bandwidth size measured in gigabytes per seconds (GB/s); (iii) the memory size measured in gigabytes (GB), and (iv) the programming framework that they can be programmed with.

In most scientific applications, high precision arithmetic (double floating point precision) is required and so hardware providing this is crucial. Nowadays, high double precision (more than 1 teraflops of double precision) processing

performance is only supported by some high-end GPU models (such as the AMD Radeon R9 280X). To use a GPU, the data to be manipulated have to be copied to the GPU memory and, after the manipulation, results have to be copied back to the main systems memory. This procedure creates a bottleneck; however, this is not significant when the manipulation of the data takes considerably more time than the transfer of the data itself. High memory bandwidth allows quick access to the data and plays a key role in the reduction of running time, which includes the transfer and accessing of the data. Memory size may be an important factor in using GPUs for scientific applications, as the performance could be considerably reduced when manipulating large amount of data.

Programs that are executed on a GPU have to be written using a specific programming framework. The two most widely used frameworks are the NVIDIA Compute Unified Device Architecture (CUDA) and Open Computing Language (OpenCL) [91]. The CUDA framework is a proprietary architecture specifically designed to be run only on NVIDIA GPUs. Alternatively, OpenCL which is managed by the Khronos Group[1], is an open framework which can be used to program almost any parallel architecture, such us multicore CPUs, GPUs, a number of connected GPUs and/or CPUs. Code written in OpenCL by other researchers can be easily used without the need for special hardware, unlike code written in CUDA. This is important, because current and future hardware will be most probably based on parallel architectures. The proprietary nature of CUDA and its requirement to be executed only on NVIDIA hardware is a major drawback for scientists. Therefore, we present here only the OpenCL framework.

OpenCL includes a C99 based language for writing functions (kernels), which are executed on OpenCL devices (*e.g.* CPUs, GPUs) and an application programming interface (API). The API is used to define and control the OpenCL devices and execute the kernels on them. A simple OpenCL program involves the steps of finding the computing device, compiling the code that will be run on the device, copying the data to the device, performing the computation and copying back the results of the computation.

[1] https://www.khronos.org/opencl/

Table 3.1: Different GPU models with their most important characteristics used in scientific computing applications. MB: megabytes; GB/S: gigabytes per second; GFLOPS: gigaflops operations per second; API: application programming interface. The prices shown are for February 2014.

Model	Memory size (MB)	Memory Bandwidth (GB/S)	Single precision/Double precision (GFLOPS)	API support	Price (British Pounds)
AMD Radeon HD 7970 GHz Edition (R9 280X)	3072	288	4096/1024	OpenCL 2.0	200
AMD Radeon R9 295X2	2×4096	2×320	11466/1433	OpenCL 2.0	600
AMD FirePro W8100	8192	320	4219/2109	OpenCL 2.0	850
AMD FirePro W9100	16384	320	5238/2619	OpenCL 2.0	2550
NVIDIA GeForce GTX Titan	6144	288.4	4500/1500	OpenCL 1.2 CUDA	710
NVIDIA GeForce GTX Titan Black	6144	336.4	5121/1707	OpenCL 1.2 CUDA	750
NVIDIA GeForce GTX Titan Z	2×6144	2×336.4	8122/2707	OpenCL 1.2 CUDA	2000
Quadro K6000	12288	288	5196/1732	OpenCL 1.2 CUDA	3800

Table 3.1 shows the performance characteristics of current GPU models which could be considered for scientific applications. We selected these models based on their ability to provide high double precision performance and their ability to be installed on personal computers. We also compared the acquisition cost, which may be an important factor in making a decision regarding purchase, especially for scientific groups with limited budgets. For example, for scientific computing with no need for high memory, we suggest the AMD R9 280X. This GPU chip provides 1024 GFLOPS of double precision performance for a considerably lower price. The other options would involve dual chip cards and workstation cards at a considerably higher price. NVIDIA GPUs, which are gaming and workstation models, provide high double precision performance, but at high cost. Although NVIDIA provides GPUs with high double precision performance, they do not support the latest version of OpenCL. NVIDIA supports version 1.2, whereas AMD supports version 2.0.

3.7 GPGPU and EAs

Parallel EAs have been developed in such a way that they use the parallel computational abilities of GPUs, which depending on application and hardware, yields considerable speedups when compared to CPUs [90], [92], [93]. The GPU (NVIDIA GT540M) based genetic algorithm implementation by Cavuoti *et al.* [92] provided an average of 75 times speedup compared to the CPU (Intel Core i7 2630QM) based implementation. They used the GPU to run the most computational intensive computations, which were the generation of the random chromosome and the calculation of the fitness function. Kalonias *et al.* [94] used a GPU accelerated GA to solve the examination timetabling problem. They found that the GA running on a GPU (NVIDIA GTX580) is, depending on the problem size, up to 26 times faster than running the same GA using one single thread of a CPU (Intel i7 at 3.40 GHz). Ben-Shalom *et al.* [95] showed that a GPU can accelerate a GA used to constrain parameters of Markov ion channel models. They showed that their GPU (NVIDIA GTX480) implementation is roughly 180 times faster than the NEURON-MPI implementation executed on an 80 CPU cluster.

Although the speedup is depended on the hardware and application, GPUs do provide considerable speedup. As discussed in section 3.5, in a GA the same operations are applied to the individuals of a population and the fitness function can be readily parallelised, due to its application to all the individuals comprising the population. The calculation of the fitness function, genetic operators and selection operators can be implemented in a format that can be run on GPUs.

Moreover, a GPU can provide further speedup in the case where the population representation and the calculation of the genetic operators can be done using single precision floats. The reason for that is that most of the GPUs provide better performance using single precision floats compared to double precision. As shown in Table 3.1, single precision arithmetic performance of high end GPUs is between 4000-12000 GFLOPS. Current low-end GPU models, which are integrated on laptops, provide single precision performance of 276-2304 GFLOPS, and this is sufficient to allow the speedup of the GA considerably.

In the case of optimising computational models comprising systems of ordinary differential equations (ODEs), the most computationally intensive task is the integration of the model for all its parameter combinations represented by the GA population. In this case, the GPU can compute the integration of the model for all parameter combinations, while the fitness evaluation and genetic operators can be implemented using the CPU. The advantage of this approach is that only the GPU implementation of the ODE solver has to be programmed. Any GA implementation that can be run on the CPU can be used via an interface, thus allowing access to the GPU results.

4 Methods for analysing oculomotor models

This Chapter will present: (i) the methods for parameter estimation that can be applied to the Broomhead *et al.* [1] model and what we believe can also be used in other oculomotor (and biological) models; (ii) the analysis methods which we used in this work to test the hypothesis that a stochastic version of the Broomhead *et al.* [1] model can generate similar behaviour to that measured in the experimental nystagmus and saccadic data sets.

The parameter estimation methods were developed with the objective of providing robust methods for fitting the Broomhead *et al.* [1] model to the experimental recordings of nystagmus oscillations and saccades. These methods are based on the use of the NSGA-II MOGA algorithm described in section 3.4, which was developed for optimising multi-objective problems for which conflicting objectives are present. Additionally, we present the procedure that we have used to select the NSGA-II parameters (*e.g.* population size, crossover method), which allows the NSGA-II to provide sufficiently near-optimal fits. Moreover, we have developed a method to accelerate the NSGA-II by parallelising its most computationally intensive component (*i.e.* the integration of the oculomotor model) and then executing the NSGA-II on a CPU/GPU combination (refer to section 3.7). We have further developed a method for the execution of multiple NSGA-II instances in parallel on a CPU combined with a GPU. By applying these methods, we show that this approach can alleviate the need for an HPC cluster. We believe that these methods provide fast parameter fitting that can be used for the development of oculomotor models.

As discussed in section 2.4, Akman *et al.* [15] performed a nonlinear time series analysis of experimental recordings from jerk infantile nystagmus and discovered that there is an increase in the local dimensions of the waveform during the fast phase. They proposed that this is due either the activation of additional oculomotor components at the beginning of the fast phase or a signal-dependent noise process in the saccadic system. The second main objective of the methodology presented here was to test the hypothesis that noise on the neural

command signal could cause the saccadic system to exhibit the behaviour observed experimentally. The testing was performed with the following steps. Firstly, we develop a stochastic saccadic model based on the Broomhead *et al.* [1] model, with signal-dependent and constant noise added to the burst and neural integrator signals. Secondly, we fit the model's local dimensions pattern to experimental data by performing a parameter search of the model's noise intensity values. Thirdly, we fitted the noise intensity values which cause the model to exhibit the same endpoint variability as found in saccades made by participants without any pathologies. The comparison of the noise parameter values of the participants with and without nystagmus could show that noise levels may play a role in the development of nystagmus, or that nystagmus could cause different noise levels.

The results acquired by following these methods could show the strengths and weaknesses of the developed stochastic model, and identify any testable predictions. It could also recognise the need for additional experimental data that will allow further extension of the model.

This Chapter is organised as follows: In section 4.1 we will provide an overview of the methodology we used. In section 4.2 we will present the experimental database that was used to fit the Broomhead *et al.* [1] and stochastic models, and the procedure that we used to transform the data so that they could be used by the NSGA-II. In section 4.3 we will provide a detailed description of fitting the Broomhead *et al.* [1] model to experimental recordings of nystagmus oscillations and saccadic velocity profiles. This will include the description of the procedure we utilised to find the specific NSGA-II parameter values which will return sufficiently accurate and converged fittings. Moreover, we will describe how we accelerated the NSGA-II using a GPU, and finally, in section 4.4 we will describe the fitting method of the parameters from the stochastic model and the experimental data that was used in this work.

4.1 Methodology

The stochastic model (described in section 6.1) will include the parameters of the Broomhead *et al.* [1] model (equations 2.1-2.6) and the parameters for setting the noise intensity values. To distinguish the two sets of parameters, we refer to

the parameters of the Broomhead *et al.* [1] model as **deterministic parameters** and the stochastic noise intensity parameters as **stochastic parameters**.

The procedure for estimating the values of the stochastic model parameter set is split into two pipelines (Figure 4.1). The right pipeline attempts to find the parameter values for the case of the experimental nystagmus data (nystagmus pipeline), while the left pipeline attempts to find the values which cause the model to exhibit the same endpoint variability as found in saccades (saccadic pipeline).

Each pipeline is split into two sub-processes (Figure 4.1). The first sub-process of each pipeline finds the deterministic parameters, which sets the model to generate the same behaviour as measured experimentally in nystagmats (waveform shape, period and amplitude), or in the case of normal participants, finds the parameters which cause the model to generate normal saccades (saccadic velocity profile). This also tests whether the model can generate the experimentally measured nystagmus waveforms and saccades. The future use of this method could test any modification or extension of the model. This sub-process for both experimental data sets uses the NSGA-II described in section 3.4 for fitting the deterministic parameters. The second sub-process of each pipeline was developed to use the parameters found in the first sub-process in the stochastic model and attempts to estimate the stochastic parameters that set it to exhibit the same local dimensions pattern found in experimental nystagmus data, or in the case of normal participants, the same saccade endpoint variability. The existence of these combined parameters (deterministic and stochastic) will show whether the model can generate these type of nystagmus waveforms, saccade dynamics and stochastic properties. However, if that is not the case, it will show that more processes have to be included in the model. Additionally, it will indicate whether more experiments are required in order to further develop the model and/or test any predictions of the model. It has to be noted, that the optimisation of the stochastic model for the deterministic and stochastic parameters is split into disjoint procedures. Using the split optimisation method could provide inaccurate results only in the case where the noise processes could affect the nystagmus waveform characteristics (period size, amplitude size and shape) or the saccade dynamics. However, in Chapter 6 (sections 6.2.5 and 6.2.7) we will show that in the case of the specific stochastic model we developed here, this is not the case.

Figure 4.1: Parameter fitting methodology. The methods are grouped into two pipelines: one for fitting the model to nystagmus data (right column) and the other for fitting the model to saccadic data (left column). Each pipeline is divided into two sub-processes: one for estimating the deterministic parameters to experimental data and another one for estimating the stochastic parameters using characteristic statistics of experimental data.

The input for each sub-process is experimental data from nystagmus participants or participants without any pathology. In the following sections, we will describe the experimental data, the procedure that we used to change the form of the data so that they can be used by the NSGA-II, the procedure for fitting the stochastic model and the analysis of the results.

4.2 Experimental data

The experimental data that we used for the model fitting is a collection of experimental nystagmus recordings, mean saccadic velocity profiles and endpoint saccadic variance measures for different saccade amplitudes.

4.2.1 The experimental nystagmus database

The experimental nystagmus database, from which the data used in this study, consists of horizontal eye movement recordings from 48 nystagmats enrolled in a drug trial [96]. From this database, data from 20 participants was chosen based on the idiopathic pathology of their nystagmus. Ocular and extra-ocular motility examinations, electroretinogram recordings and visual evoked potentials were used to make the diagnosis. The mean age of the idiopathic group was 40.05±8.29. The group consisted of 13 males and 7 females. The remaining subjects (28) in the drug trial had additional conditions and therefore were not used in the study. The eye movement recordings used here were the initial ones before any drugs were administered to the participants. Binocular eye movements were recorded with the head stabilised in the primary position by a chin rest. Eye movements were recorded in 3 positions of gaze: straight ahead, 15 degrees right and 15 degrees left. Five of the subjects in the idiopathic group had a latent component to their nystagmus. The sample rate of the recording was 250 Hz with a resolution of 0.005 degrees and with maximum noise level of 0.01 degrees root mean square. Full details of the investigations of the subjects and equipment are given in [96].

In the nystagmus pipeline (see Figure 4.1, section 4.1), the NSGA-II fits the output of the Broomhead *et al.* [1] model to the experimental waveform characteristics of the nystagmus waveforms. This process requires one period of the waveform to be used as a target for the optimisation process. The identification of this waveform may be challenging due to the non-deterministic nature of real nystagmus oscillations, namely each oscillation cycle has a slightly different

period, amplitude and shape (Figure 4.2). This raises the question of which of these cycles should be chosen for the fitting procedure. In the next section we will describe the method that we used to extract the single period segment of the time series to be fitted into the model.

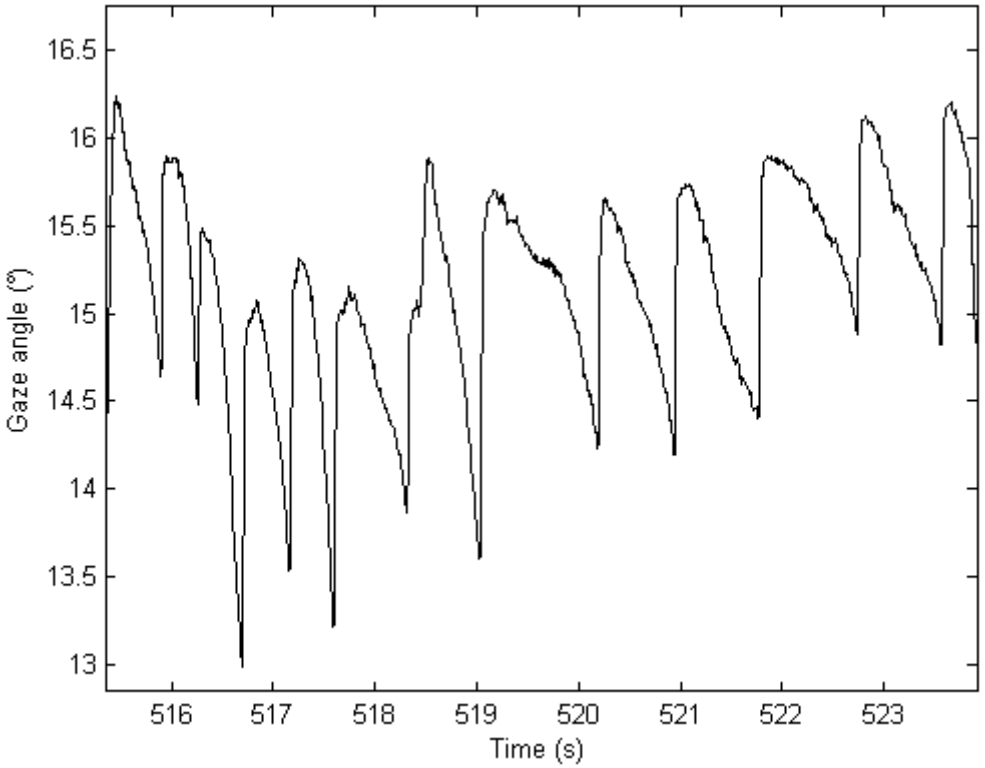

Figure 4.2: Part of a horizontal eye-tracking recording derived from a participant with nystagmus. On the vertical axis, positive values represent rightward eye positions in degrees (°). The horizontal axis shows time in seconds (s).

4.2.2 Single period waveform extraction for experimental nystagmus recordings

The method proposed by So *et al.* [97], [98] allows one to estimate either unstable periodic orbits (UPOs) generated by a deterministic model or stable periodic orbits destabilised by a noise process [97]–[101]. Their method is based on the assumption that there is a significant deterministic component in the system and that the UPOs can be used to characterise the system's dynamics [99], [100].

This method transforms the experimental time series in such a way that the data of the transformed time series which are closer to the periodic orbits are concentrated on the periodic orbits. The location of the UPOs in the experimental

time series is identified by sharp peaks on a histogram plot of the transformed experimental data. Shortly, this transformation is a dynamical lens that enhances the probability of the data points which are close to the periodic orbits of the system [99].

So *et al.* method has been applied to nystagmus recordings by Clement *et al.* [100] and Theodorou and Clement [101]. In order to extract the UPO target waveform from the nystagmus time series, we used this method as it was applied by Theodorou and Clement [101]. This method was applied not to the actual gaze angle time series, but to the intervals $\{\tau_1, \tau_2, \dots, \tau_L\}$ between individual nystagmus cycles, identified by thresholding the velocity time series (first derivative of the experimental time series). The threshold value was chosen to be roughly in the middle of the fast phase of each nystagmus waveform. The resulting duration data was then transformed using the So *et al.* method so the interval data were concentrated on the duration of the nystagmus unstable periodic orbit. The interval points which were of close value to the unstable periodic orbit duration, were moved closer to that value. The histogram of the transformed data have a sharp peak at the duration of the nystagmus unstable periodic orbit.

To apply the So *et al.* transformation to the experimental nystagmus data, first we applied the method of delays [102] to the intervals. In the method of delays, a sliding window of *d* samples is moved through the data, generating a sequence of d-dimensional vectors. Given our interval time series $\{\tau_1, \tau_2, \dots, \tau_L\}$, the delay vectors $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_{N-d+1}\}$ are given by $\boldsymbol{w}_k = (\tau_k, \tau_{k+1}, \dots, \tau_{k+d-1})^T$. The evolution of the delay vectors is assumed to be governed by a nonlinear map $\mathbf{F}$, $\boldsymbol{w}_{k+1} = \mathbf{F}(\boldsymbol{w}_k)$ [98]. Due to the close relationship between the dynamics of the oculomotor system and the dynamics of the delay vectors, the delay vectors can be used to quantify the dynamic behaviour of the oculomotor system. In particular, fixed points of $\mathbf{F}$, which are vectors in the delay space $\boldsymbol{w}_*$ for which $\mathbf{F}(\boldsymbol{w}_*) = \boldsymbol{w}_*$, correspond to periodic orbits of the oculomotor system, and for these points $\tau_1 = \tau_2 = \dots = \tau_L$. Such points lie on the long diagonal of the delay space (the set of *d*-dimensional vectors with equal entries). Candidate fixed points can therefore be identified by looking for delay vectors that lie close to this diagonal. The method of So *et al.* [97], [98] facilitates this identification step by applying a data transformation z to the delay vectors and concentrates them onto the fixed points. This data transformation is given by:

$$z(\boldsymbol{w}_n) = \left(\boldsymbol{I} - \boldsymbol{s}(\boldsymbol{w}_n)\right)^{-1}(\boldsymbol{w}_n - \boldsymbol{s}(\boldsymbol{w}_n)\boldsymbol{w}_{n-1}), \tag{4.1}$$

where $\boldsymbol{I}$ is the $d \times d$ identity matrix and:

$$\boldsymbol{s}(\boldsymbol{w}_n) = D\mathbf{F}(\boldsymbol{w}_n) + (\boldsymbol{w}_n - \boldsymbol{w}_{n-1}) \tag{4.2}$$

where $D\mathbf{F}(\boldsymbol{w}_n)$ is the $d \times d$ Jacobian matrix of $\mathbf{F}$ evaluated at $\boldsymbol{w}_n$. The transformed data lying, within a small cross sectional tube along the diagonal, is then projected onto the diagonal itself. Fixed points correspond to sharp peaks in the histogram obtained by binning the 1-dimensional projected data [98]. In its original form, the method of So *et al.* [98] also includes the addition of a tensor of random numbers to the function (4.1), which eliminate spurious fixed points that could be caused by data artefacts. In the case of nystagmus time series, this is not necessary because the time series is examined for any artefacts, such as blinks, fixations, attention changes or other abnormalities and those are removed before the application of the method [101].

To extract UPOs from the nystagmus recordings using this method, we chose a d value of 2 and a histogram bin size of 0.025 s [100]. In order to further refine the solutions, we rejected any transformed points that lay 0.5 units from the diagonal. The estimation of $D\mathbf{F}(\boldsymbol{w}_n)$ can be obtained by using finite differences [101]. The middle left plot in Figure 4.3 shows the interval frequency polygon obtained from an experimental recording of a jerk nystagmus oscillation. This plot has a clear peak indicating the presence of a UPO with a period close to the corresponding interval length. All the candidate periodic orbits have maximum length difference to the peak location of the histogram of 0.0125 seconds (Figure 4.3, middle right). From this collection of periodic waveforms we chose the one with the smallest difference between its first and the last point (Figure 4.3, bottom left). Because it is more convenient to use waveform targets that start at the beginning of the fast phase, we concatenated three copies of the UPO and chose the part between two successive beginnings of the fast phase. This waveform will be subsequently used as target for the fitting process (Figure 4.3, bottom right).

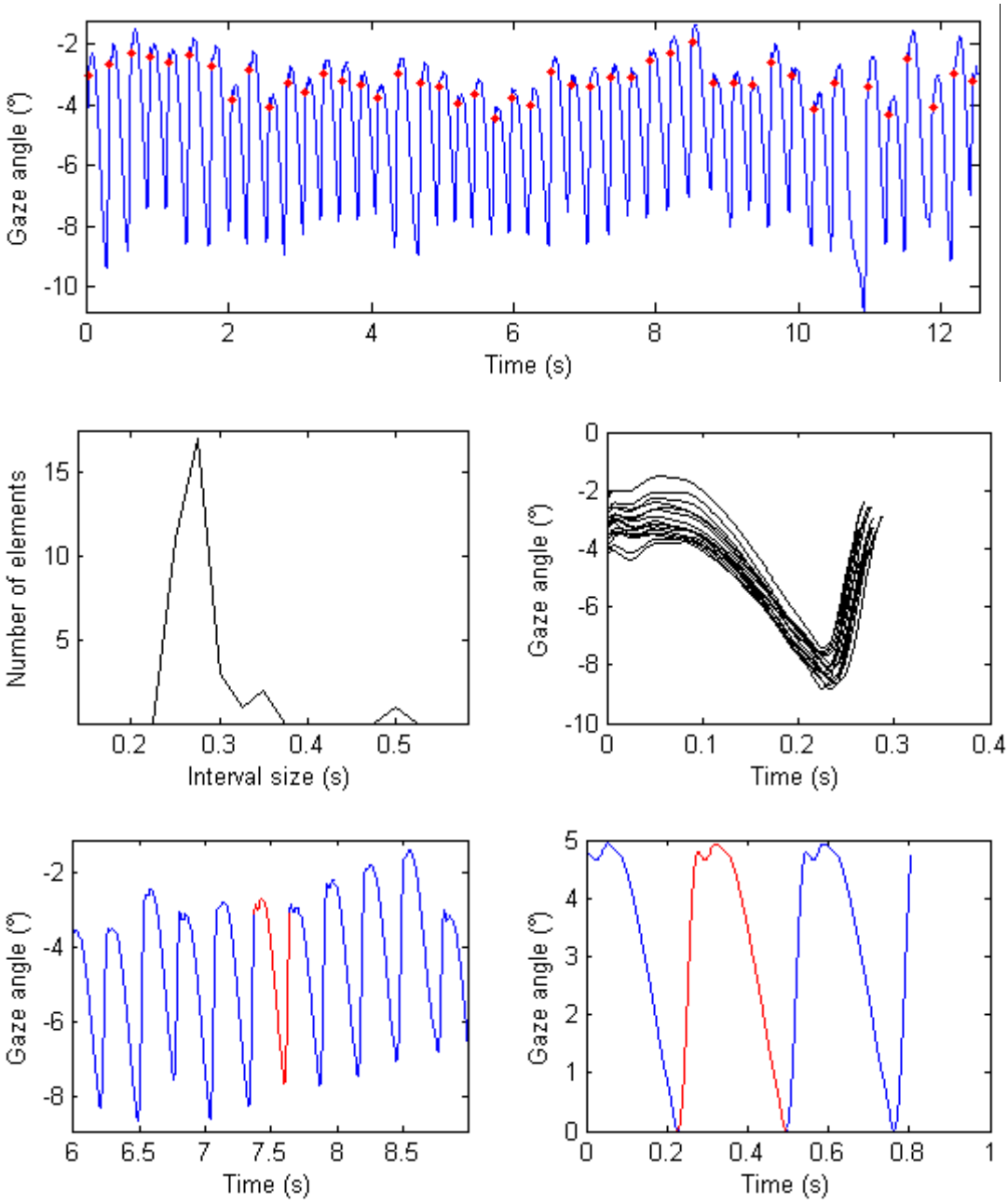

Figure 4.3: Plots showing key stages of the unstable periodic orbit (UPO) extraction method. **Top:** Part of an experimental time series of a right-beating jerk nystagmus with the red points (threshold of 70 deg/s) showing the different intervals. **Middle left:** Frequency polygon which shows the distribution of the interval size of the transformed interval data for the time series shown in the top plot. There is a clear peak at 0.28 s. **Middle right:** Candidate waveform solutions extracted from the time series based on their period. **Bottom left:** Part of an experimental recording (blue), where the extracted UPO is shown in red. **Bottom right:** Three concatenated copies of the extracted UPO. The final selected UPO, which starts at the beginning of the fast phase, is shown in red. Gaze angle is in degrees (°); time is given in seconds (s).

4.2.3 The experimental saccadic database

The experimental data of the saccadic velocity profiles was taken from [23]. The data we used in this work are the typical velocity profiles observed for horizontal saccades of 5, 10 and 20 degrees (Figure 4.4). These saccade amplitudes were chosen because they are in the interval of the most frequently made saccades [103], [104].

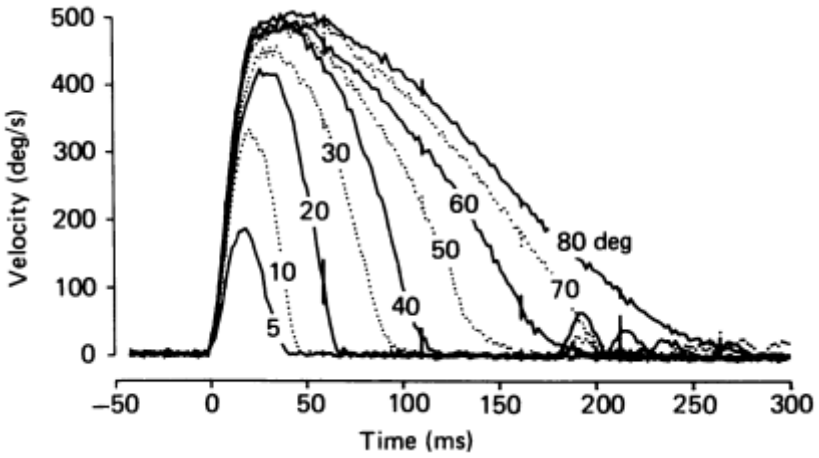

Figure 4.4: Typical horizontal saccadic velocity profiles for amplitudes ranging from 5 to 80 degrees. Velocity is in degrees per second (deg/s); time is in milliseconds (ms). Figure reproduced with permission from [23].

4.2.4 Endpoint saccadic variability data

The experimental data for saccadic endpoint variability was taken from [30]. The data represent the mean coefficient of variation for horizontal saccades of 2, 4, 6, 8, 10, 12, 14, 16 and 18 degrees from five subjects (Figure 4.5). Full details of the investigations on the subjects and equipment are given in detail in [30].

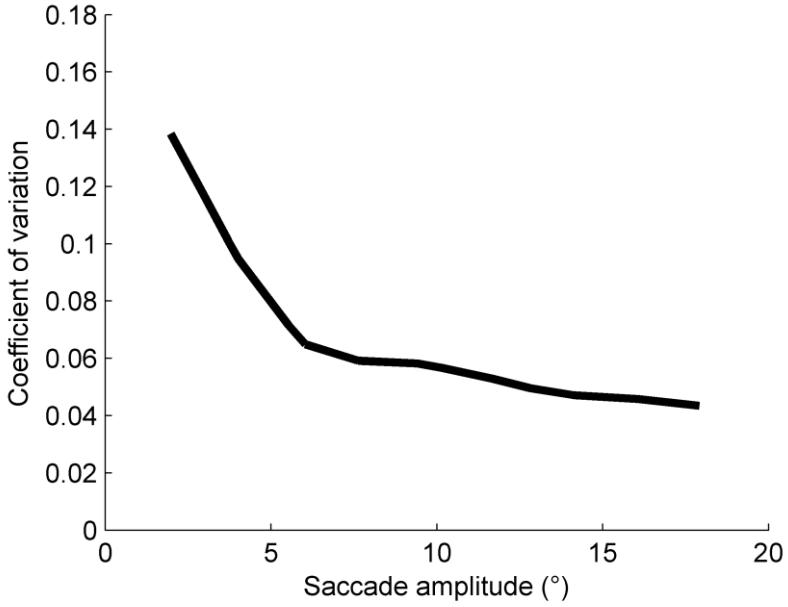

Figure 4.5: Mean coefficient of variation (CV) of saccadic endpoints for different saccade amplitudes from five participants of the study by van Beers [30]. The black line denotes the mean CV for horizontal saccades. Saccade amplitude is in degrees (°). Adapted from [30].

4.2.5 Experimental data for nystagmus local dimensions pattern

The data we used for fitting the local dimensions pattern has been generated by applying the method of Akman *et al.* [15] on the experimental time series section which was also used in the UPO extraction method (refer to section 4.2.2). This analysis generates data consisting of the local dimensions at each data point of the time series (discussed in section 2.4).

4.3 Parameter fitting of the Broomhead *et al.* model

The goal of the optimisation method presented here, was to find the values of the six parameters of the saccadic model $\{\alpha, \beta, \varepsilon, \gamma, \alpha', \beta'\}$ (deterministic parameters, see Table 2.1), which would set it to exhibit behaviour that is the same as, or similar to, the experimental eye movement recordings. In the case of velocity profiles recorded from normal subjects, the values of the model parameters, which would set the model to generate the profiles found experimentally for each saccade amplitude, have to be found. In the case of nystagmus data, the model fitting will involve finding the parameter values which cause the model to generate

nystagmus waveforms with the same amplitude, period and shape as in the experimental data. In both cases, different characteristics require fitting, thus making the parameter fitting a multi-objective optimisation problem, where these conflicting objectives need to be optimised together.

One approach would be to use the weighted-sum method [74], as discussed in section 3.2.4. However, this approach requires the weight values to be specified, for which there is usually no *a-priori* knowledge. Indeed, in fitting the three characteristics of the nystagmus waveforms and in the case of the saccadic velocity profiles for all saccade amplitudes, we do not know the weight values which will maximally exploit the information provided by the different objectives.

The multi-objective optimisation method of our choice is the Non-dominated sorting genetic algorithm II (NSGA-II) multi-objective genetic algorithm (MOGA). MOGAs (and the NSGA-II) compared to other optimisation methods (*e.g.* gradient descent) can overcome the problems caused by the complexity of the Broomhead *et al.* [1] model's bifurcation structure. Akman *et al.* [62] found that the Broomhead *et al.* [1] model has a rich bifurcation structure, which includes Hopf, homoclinic, saddlenode, pitchfork and gluing bifurcations. The bifurcations create sharp changes in waveform type under modest parameter variations. This causes difficulties in optimisation of algorithm in crossing bifurcation surfaces by reducing its ability to efficiently explore the parameter space. However, the genetic algorithm is able to overcome this problem (discussed in section 3.3). In order to successfully fit the Broomhead *et al.* [1] model using the NSGA-II, a fitness function has to be formed. This fitness function would evaluate the similarity of the model's behaviour to the experimental recordings with different parameter values. Moreover, the NSGA-II has parameters of its own (*e.g.* population size or crossover method), which need to be set for the specific optimisation in order to succeed. In the next section, we will describe the method that we used to find the specific NSGA-II parameter values which approximated the optimal *Pareto* set.

4.3.1 Choosing the NSGA-II parameters and forming the fitness functions

We used the MATLAB NSGA-II implementation which is a part of the MATLAB Global Optimisation Toolbox [105]. Each individual is encoded in floating point values, a particular combination of the model parameters $\alpha, \beta, \varepsilon, \gamma, \alpha'$ and β'. The

initial NSGA-II population was uniformly distributed in the following parameter space:

$$1 \leq \alpha \leq 1000;\ 0.1 \leq \beta \leq 60;\ 0.000001 \leq \varepsilon \leq 0.1;\ 0.00001 \leq \gamma \leq 12;\ 50 \leq \alpha' \leq 1000;\ 0.1 \leq \beta' \leq 60.$$

The bounds on α, β, α' and β' where chosen based on the findings of [25], whereas the bounds on ε were based on the previous analysis of [62]. The bounds on parameter γ were set using multiple NSGA-II runs that were shown to return solutions yielding biologically feasible waveforms. The initial conditions for each individual were set at:

$$g_0 = v_0 = n_0 = r_0 = l_0 = 0 \text{ and } m_0 = \Delta g,$$

where for the nystagmus pipeline (Figure 4.1, right column), Δg was set to 1.5 simulating a rightward saccade of 1.5 degrees, and for the saccadic pipeline (Figure 4.1, left column), Δg was set to each of the different saccade amplitudes 5, 10 and 20 degrees in turn. These saccade amplitudes were chosen as they are in the interval of the most frequently made saccades [103], [104]. The termination criterion was set to 100 generations in order to keep the running time short. The selection type was binary tournament (see section 3.3.2): this selects each parent by randomly drawing two individuals from the population and then chooses the one with the highest fitness for crossover (see section 3.3.3). The mutation type was set to adaptive feasible: this chooses a direction and step length which satisfy the model's parameter constraints (see section 3.3.4).

The NSGA-II parameters that we have explored for better fitting results were the population size, crossover type and distance measure type. The population sizes that we explored were 500, 1000, 2000, 4000 and 8000, whereas the crossover types that we examined were intermediate and heuristic. The distance measure types we explored were based on objective value (phenotype) and on model parameter values (genotype). The reason we chose not to explore all the NSGA-II parameters is that it would considerably increase the computation time, making the process impractical. Therefore, we executed a number of exploratory optimisation runs with combinations of different population size, crossover function and distance measure type. The waveform targets used in these exploratory optimisation runs were generated by the oculomotor model (synthetic data) and experimental waveforms. Further details of the four synthetic

waveforms are shown in Chapter 5. The results of the exploratory runs allowed us to find the NSGA-II parameters, which were most influential on the returned solutions, namely the population size and crossover function. It was expected that the population size would be an important parameter for the NSGA-II accuracy and convergence. If the population size is too small, the NSGA-II may be unable to find close estimates of the global optima of the problem. Increasing the population size usually allows the NSGA-II to find close estimates of the global optima with increasing accuracy and precision, but requires greater computational resources [106]. Hence, when choosing the population size, a balance needs to be found between accuracy and convergence, and the resulting computational load. In the case of the crossover method, we found that heuristic crossover with ratio 0.8 allowed the NSGA-II to return more accurate and precise results, with a considerably smaller population size compared with the other crossover methods (such as intermediate crossover). The crossover ratio is the fraction of the population, which are not part of the elite children, and which will be created by crossover.

The formulation of the fitness function for each data fitting protocol was “a trial and error” method. Each time we created a new fitness function, or made some changes, we again tested it on the synthetic and experimental waveform targets. In the case of the nystagmus data, we initially used the squared difference between the simulated and target waveform. However, we found that a multi-objective fitness function based on the period and shape of the nystagmus waveform gave the best results. In the case of the saccadic data, we initially chose a multi-objective function which was based on the characteristics of the saccadic main sequence (saccade amplitude, duration and peak velocity). However, we found that it did not fit the model very accurately to the experimentally measured saccadic velocity profiles. Therefore, we formulated a multi-objective fitness function based on the sum of the squared difference between the velocity profiles of the experimental and simulated data for different saccade amplitudes. In Chapter 5 we will describe the fitness functions and the procedure in more detail.

The **hypervolume indicator** [107] was used to measure and compare the convergence of the solutions found by the NSGA-II for different NSGA-II parameter combinations and fitness functions. Due to the stochastic nature of

NSGA-II, we used 16 runs of the NSGA-II to calculate the mean and variance of the hypervolume indicator. Here we set the reference point for the hypervolume calculation as the objective-vector dominated by all non-dominated sets produced by the 16 runs, with the minimal values on each objective to satisfy this (see Figure 4.6 for an illustration of this with two objectives and two non-dominated sets generated by runs).

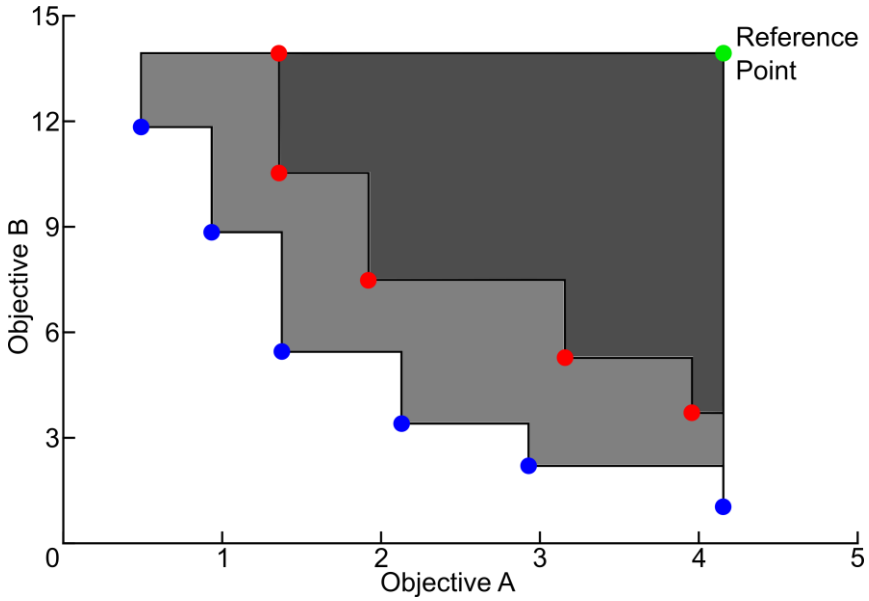

Figure 4.6: Two hypervolume areas for two estimated *Pareto* optimal fronts. The points of *Pareto* front 1 are shown as blue dots and those of *Pareto* front 2 are indicated as red dots. The hypervolume of *Pareto* front 2 is shown in dark grey, whereas the hypervolume of *Pareto* front 1 is the union of the light grey and dark grey surfaces. The reference point is chosen to be dominated by all the points in all the *Pareto* fronts which will be computed for all NSGA-II parameter configurations.

The reference point is chosen to be dominated by all the points in all the estimated *Pareto* optimal fronts and it will be computed for all the NSGA-II parameter configurations. Different reference point values could give similar results. Bigger hypervolume values indicate better convergence. The particular hypervolume indicator that we used in this work is the ratio of the estimated *Pareto* front hypervolume to the hypervolume defined by the points of the reference point, the origin of the objective axes and the points defined by the value of the reference

point on each objective axis (Figure 4.6). Furthermore, in the case of nystagmus waveform fitting, we additionally calculated the mean of the shape fitness of the final selected individuals (refer to section 5.2.1 for selection method) using the synthetic targets. Also, the precision was calculated using the coefficient of variation of the shape fitness values. In the case of the saccadic velocity profile fitting, we calculated the mean and coefficient of variation for all three objectives, using the synthetic velocity profiles as targets.

4.3.2 Acceleration of the parameter fitting method using GPGPU

As the Broomhead *et al.* [1] model is a stiff system, a small time step is required for accurate integration. Moreover, the NSGA-II needs to integrate the model a large number of times over many different parameter combinations. Initially, we tried to develop and test our NSGA-II configuration using an HPC cluster, however some difficulties emerged. As the cluster is used extensively by other users, job queues were long, and each job (a single NSGA-II run) could not be completed in less than 2 days. Therefore, we explored the use of GPUs due to their ability to accelerate the execution of the NSGA-II (described in section 3.7).

In our method, we chose to parallelise only the model integration because this is the most computationally intensive task. This was achieved by integrating the model for all the individuals in the NSGA-II population using the GPU. The remaining NSGA-II processes were performed by the CPU. This parallel implementation of the NSGA-II is based on the master-slave parallel model discussed in section 3.5. This model includes a single NSGA-II, where the most computationally intensive task (*i.e.* integration of the oculomotor model for each individual) is distributed to multiple processors, whereas a master processor is used to control the parallelisation and to run the genetic operations of the NSGA-II [70]. In our model, the master processor is the CPU and the multiple processors are the computing units of the GPU. In the following section, we will describe the program that integrates the oculomotor model for different parameter combinations in parallel. Furthermore, we will compare its speedup when it is run on a GPU versus a CPU.

4.3.3 Oculomotor model OpenCL numerical solver

Programs that are executed on a GPU, have to be written in the Open Computing Language (OpenCL). Our OpenCL program includes a kernel with the functions

which numerically solve the Broomhead *et al.* [1] model (2.1)-(2.6) for different parameter combinations (see Table 2.1). Moreover, our program includes code written in C++ which, using the OpenCL API, performs the actions necessary to run the kernel on the GPU. The numerical method that we used in the kernel was Implicit Euler, with time step $\Delta t = 10^{-5}$ s. The input to the program is a text file in which each line represents a particular parameter combination. The output of the executable is a binary file with the gaze angle g for each parameter combination which was used as an input.

Figure 4.7 shows the logical flow of the GPU executable program. Initially, the program reads the file with the parameter combinations which need to be integrated. Subsequently, these parameters and initial conditions for the model are transferred to the GPU's Random Access Memory (RAM). Consequently, the OpenCL kernel is called to integrate the model for 40 (2500 Hz sampling rate) steps of $\Delta t = 10^{-5}$ s. The 2500 Hz sampling rate means that the output of the integration is saved to the disk every 40 steps of 10^{-5} s. In parallel with the integration (current kernel call), the result of the integration of the last time step (previous kernel call) is written into the binary file on the disk. Once the integration for the timespan required for all parameter combinations is completed, the program quits.

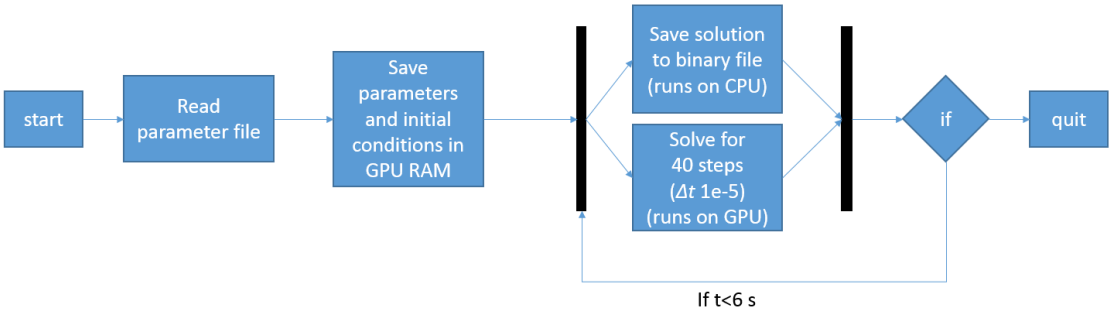

Figure 4.7: Flow diagram of the OpenCL executable used to run the model simulations. After the start of the program, a text file is read containing the oculomotor parameter set. This parameter set and initial conditions are copied to the GPU's RAM. Next, the integration of the model is performed in parallel and the results are written to a binary file. The program quits after 6 s of simulation. See main text for more details.

The sampling frequency of the results plays an important role in the speedup provided by the GPU compared to the CPU. This is based on the bottleneck caused by the memory transfer from and to the GPU memory, and the time required for storing the data on the hard disk or RAM for analysis. A low sampling frequency does not save a lot of data on the disk (250-1000 Hz). For example, 20000 of 6 s model integrations produces a file of size 1.34 GB using 2500 Hz sampling rate, whereas if 1000 Hz is used, the file size falls to 0.54 GB. In the extreme case of a sampling frequency of 100000 Hz, the data file is of 53.6 GB in size, which is 40 times bigger than the one when using a 2500 Hz sampling rate. The size is very important because it sets the number of model integration results, which can be saved on the computer's RAM. Moreover, a larger number of returned data points could increase the analysis time, whereas a low number may not be enough for an accurate analysis. Therefore, we used the following method in order to choose the sampling frequency.

The evaluation of each nystagmus oscillation involves calculating the period and amplitude of the target waveform. Interestingly, by considering fits to the synthetic waveforms, we found that the sampling frequency of the data points is very important. A low sampling frequency did not allow the extraction of an accurate period and amplitude measure, whilst a high frequency might have an impact on the computational time required for model integration and solution analysis (fitness evaluation). In order to choose the lowest value of sampling frequency, allowing enough data for the solution analysis, we used the following method. First, we integrated the model for the four different waveforms, with $\Delta t = 10^{-5}$ s and with sampling frequency 10^5 Hz (all the points were used). These target waveforms were chosen because they represent a wide distribution of amplitudes and periods. Next, we subsampled the waveforms by using the sampling frequencies of 250, 500, 1000, 1250, 2500 and 100000 Hz to obtain a set of candidate waveforms. Subsequently, we calculated the sum of squared errors for amplitude and period between the target and candidate waveforms for each sampling frequency (see Figure 4.8). The final sampling frequency was chosen to be 2500 Hz, as the use of a higher frequency caused no significant drop in the errors.

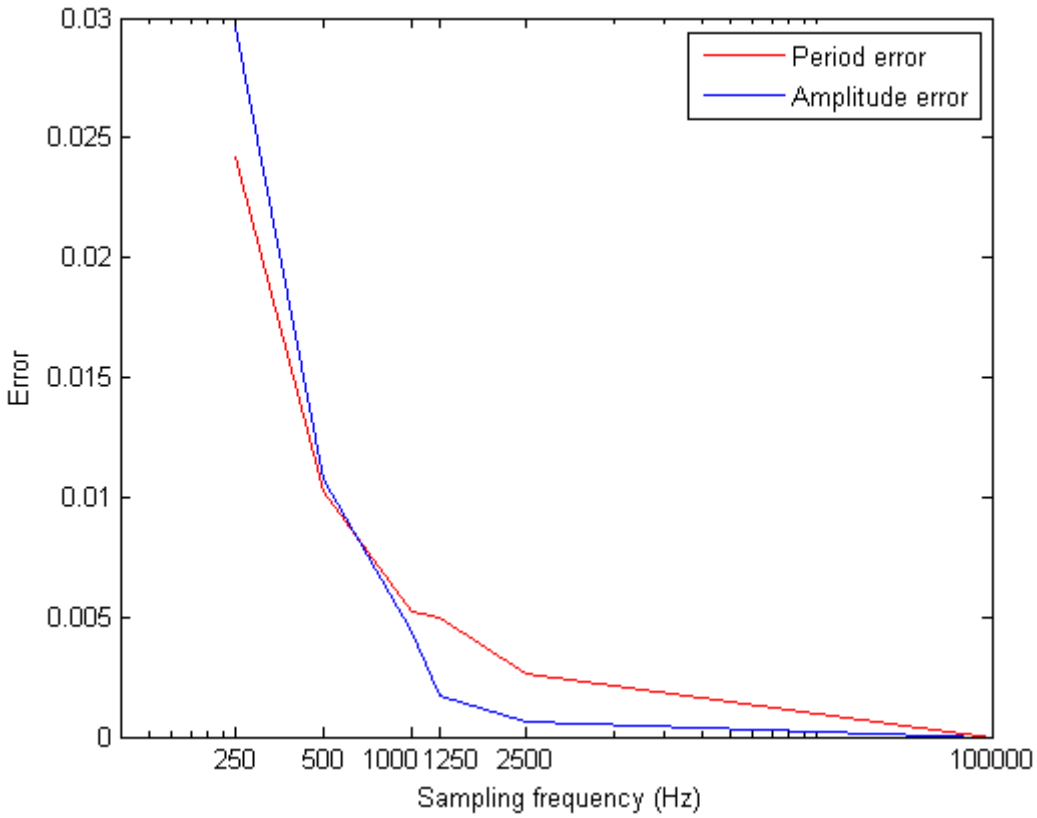

Figure 4.8: Error in estimating the correct period and amplitude from four synthetic nystagmus waveforms versus the sampling frequency of the solutions obtained using the numerical solver described in section 4.3.3. The target waveform in each case was obtained from the solver using a time step of 10^{-5} s, equivalent to a sampling frequency of 100000 Hz. The error is the sum of squared differences between the amplitude and period estimated from the nystagmus waveforms of different sampling frequency. The red line denotes the period error, whereas the blue line denotes the amplitude error. Sampling frequency is in Hertz (Hz).

Returning results using the lower frequency of 2500 Hz, allowed us to minimise the effect of the GPU call bottleneck, the amount of memory used and the computational time required to analyse the results. Most importantly, it did not hinder the accuracy required to calculate the period and amplitude of each individual waveform.

4.3.4 OpenCL numerical solver speedup using a GPU

We tested our executable on an AMD Radeon HD 7970 Graphics GHz Edition GPU and on an Intel Core i5-3550 Ivy Bridge CPU (full hardware and software specification can be found in the Appendix: see Table A.1 and Table A.2). Figure 4.9 shows the execution time of the OpenCL program in seconds, executed with either the CPU or the GPU. For large numbers of simulations, the GPU is ~30 times faster than the CPU (Figure 4.10).

Figure 4.10 shows that in order to fully exploit the capabilities of the particular GPU that we used, the number of model simulations has to be more than 30000. Usually, a population this large is not required for the NSGA-II. Hence, we have explored the possibility of executing multiple NSGA-IIs on a single GPU card. In the next section, we will describe the method that we have developed to run in parallel multiple independent NSGA-IIs.

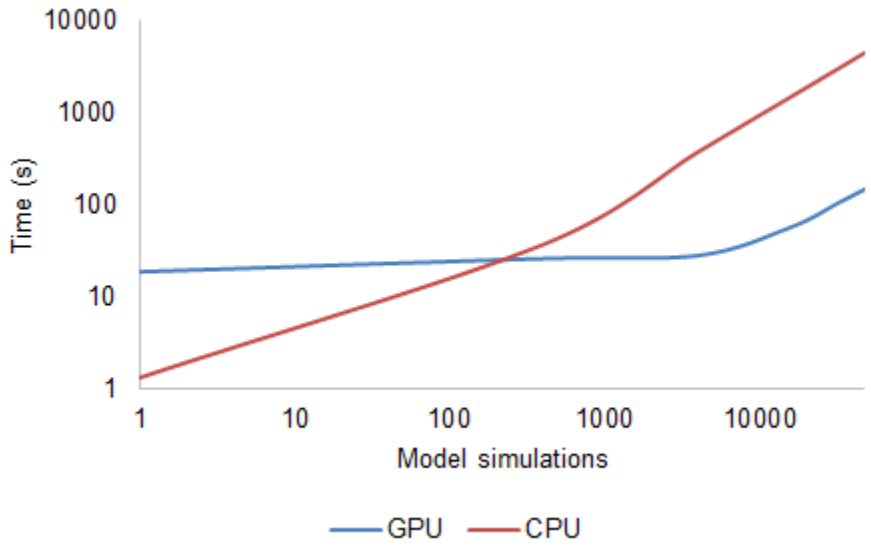

Figure 4.9: CPU and GPU execution time in seconds (s) vs the number of model simulations. The numerical solver was the Implicit Euler method (time step 10^{-5} s) and the total integration time was 6 s. The oculomotor model parameters were set to uniformly distributed parameters in the following ranges: $\alpha = [1,1000]$, $\beta = [0.001,10]$, $\varepsilon = [0.00001,0.1]$, $\gamma = [0.001,6]$, $\alpha' = [1,2000]$, $\beta' = [0.001,60]$, and the initial motor error was set to 2 degrees.

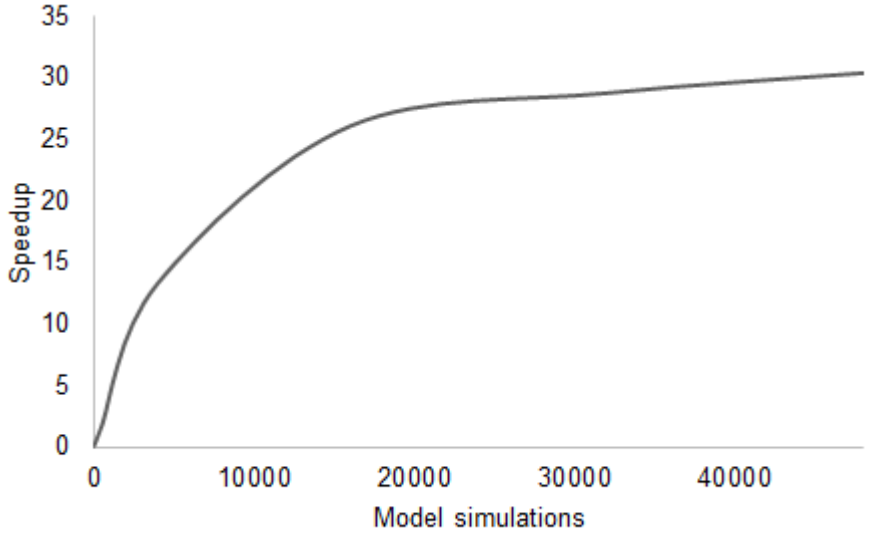

Figure 4.10: CPU/GPU execution time ratio (speedup) vs the number of model simulations performed. The numerical solver was the Implicit Euler method (time step 10^{-5} s) and the total integration time was 6 s. The oculomotor model parameters were set to uniformly distributed parameters in the following ranges: $\alpha = [1,1000]$, $\beta = [0.001,10]$, $\varepsilon = [0.00001.0.1]$, $\gamma = [0.001,6]$, $\alpha' = [1,2000]$, $\beta' = [0.001,60]$. The initial motor error was set to 2 degrees.

4.3.5 Method for running multiple NSGA-IIs in parallel

In order to run multiple instances of the NSGA-II embarrassingly parallel (*i.e.* there is no dependency between the parallel instances), we used the program organisation shown in Figure 4.11. The general idea of the method is as follows: an NSGA-II manager program (written in MATLAB) creates the NSGA-II MATLAB instances that run in parallel. Each of these NSGA-II instances sends the parameter combinations that it has to evaluate (individuals) to the GPU server program (another MATLAB instance), which writes these combinations into an input file. This file is read by the GPU executable, which calculates the oculomotor model solutions (gaze angle time series). The results of the GPU executable are written as a binary file, which is split by the GPU server program into multiple binary files containing the solutions corresponding to each NSGA-II instance. Finally, the NSGA-II instances read the binary files containing the results which correspond to their population, they calculate the fitness and apply the genetic

operations in order to create the new population of individuals. This method is based on the independent runs parallel model discussed in section 3.5, in which multiple sequential NSGA-IIs are executed on a pool of processors [70]. For example, if one has at his disposal 10 computers, he can use them to run one sequential NSGA-II on each computer. This allows to obtain statistics about the solutions the NSGA-II is returning (*e.g.* convergence measures) [70]. In our model, multiple master-slave NSGA-IIs run independently on the CPU/GPU combination and there is no communication between them. Furthermore, the only part of each NSGA-II which runs in the GPU is the integration of the oculomotor model.

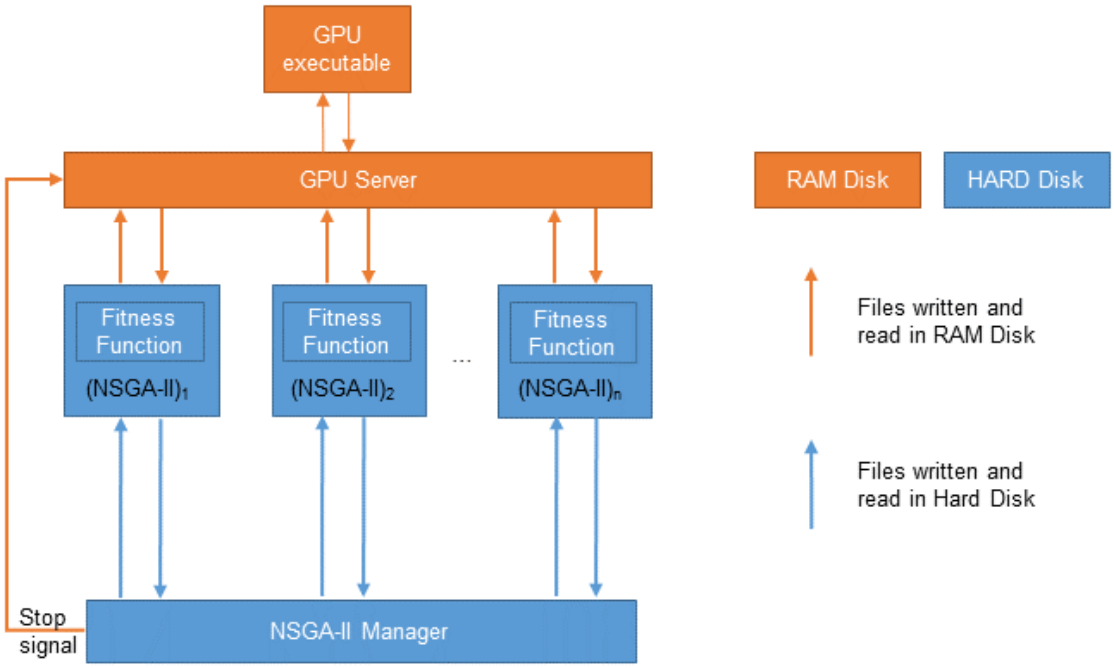

Figure 4.11: Flow diagram of the method of running multiple NSGA-II instances in parallel. {$(NSGA\text{-}II)_1$, $(NSGA\text{-}II)_2$,…, $(NSGA\text{-}II)_n$} are individual MATLAB NSGA-II instances that run in parallel and which are created by the NSGA-II Manager. The GPU server is the interface between the NSGA-II instances and the GPU executable that performs the integration of the oculomotor model for each parameter set sent by the NSGA-II instances. The arrows indicate the communication between the programs using files. Orange arrows indicates that the program or a file is saved on the RAM disk, whereas blue arrows indicate saving to the hard disk. See main text for more details.

Running multiple NSGA-II instances causes two main problems, due to the very large files which are generated (*e.g.* 32000 simulations will create a 3.58 GB file). First, writing and reading to/from the hard disk takes up a considerable amount of time compared to writing to the RAM, thereby causing a speed bottleneck. Second, writing a file of such size multiple times could have a significant impact on the life expectancy of the hard disk. To address these two problems, we decreased the size of the output file by writing the output of only the last 3.2 s from the integration. This was done only for nystagmus waveform fitting, because only the last 3.2 s of the integration are required for the fitness evaluation (see section 5.1.1). This decreased the written data almost by half. In the case of saccadic velocity profile fitting, the integration time was considerably smaller (<1 s). Secondly, we used a RAM drive (a block of the system's RAM that the software uses as a disk drive) which proves to be considerably faster than the hard disk. For example, writing a binary file with size of 3.58 GB on the RAM DISK with MATLAB takes 1.38 seconds, whereas writing the same file on the hard drive takes 19 seconds (for the specific RAM DISK software and hard disk model we used in this study see Table A.2 in the Appendix).

4.3.6 Speedup obtained using the multiple NSGA-II parallel method

Figure 4.12 shows that running multiple NSGA-II instances in parallel on a single GPU yields a considerable speedup. In order to compare the execution time of running different numbers of parallel NSGA-II instances in both serial and parallel with the parallel model, we used the following methodology. The number of instances that we chose was dependent on the population size of each NSGA-II instance. Next, we chose a fixed number of parallel simulations to run on the GPU. We choose 32000 parallel simulations, which is the number of simulations at which the CPU/GPU execution time ratio starts to saturate (Figure 4.10). Therefore, the equation which gives the number of parallel instances i is $i = 32000/ps$, where ps is the population size of each NSGA-II instance. The different population sizes tested were: 500, 2000, 4000 and 8000. In the case of the serial execution, we run the NSGA-II instances serially. In the parallel case, we run all the NSGA-II instances in parallel. Figure 4.12 shows that the speedup of the parallel model decreases with increasing population size. This was as anticipated, because larger population sizes use more of the GPU capabilities when a single NSGA-II instance is run. It should also be noted that the population

size and number of parallel instances that can be used are constrained by both the specifications of the GPU card and the system's RAM. Our test system had 32 GB of RAM, of which 6 GB was apportioned to the RAM disk.

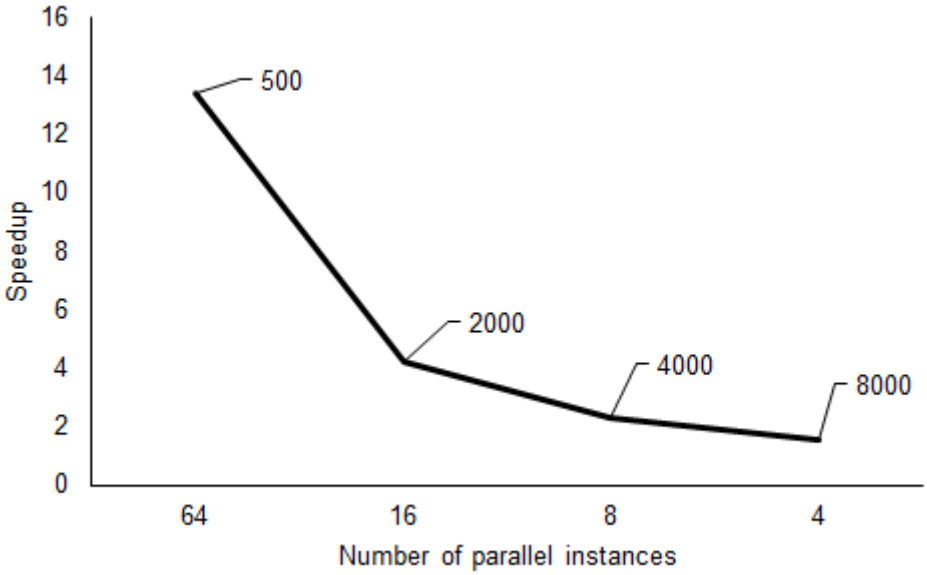

Figure 4.12: Speedup of the parallel NSGA-II method versus the number of parallel instances. Speedup is defined as the ratio of parallel execution time to serial execution time for the same number of NSGA-II instances. The indexed numbers indicate the population size for each NSGA-II instance: in each case, the sum of the populations of all NSGA-IIs running in parallel is 32000, which is the number of parallel integrations that fully use the GPU resources.

4.3.7 The need for more parallel NSGA-II processes

In our application, we chose to run only the model integrations on the GPU, as we found this to be the most computationally intensive task. We measured the time required for running an NSGA-II instance for different population sizes and compared the model integration time to that required to perform the NSGA-II operations and to evaluate the fitness function (Figure 4.13). The analysis confirmed that the GPU executable is more time consuming than the rest of the NSGA-II processes.

Future applications of our method will require the researcher to compare the model execution time to that of the other NSGA-II process and decide whether

more parallelisation has to be performed. Moreover, by running the MATLAB profiler, which measures the execution time of each MATLAB function, one can get further information regarding which particular function could be accelerated. In order to get the ratio of GPU execution time to the time required to evaluate the fitness function and to perform the other NSGA-II operations, the algorithm was run for a small number of generations. Figure 4.13 shows the mean execution time for 8 NSGA-II executions for 5 generations. The functions that are part of the NSGA-II operations increased considerably for population sizes greater than 10000. This is as expected, because the NSGA-II is of $O(MN^2)$ computational complexity [86], where M is the number of objectives and N is the population size. The profiler output shows that the MATLAB function which causes this increase, is nonDominatedRank and this ranks the individuals in the population. If there is a requirement for more reduction of the NSGA-II execution time, an accelerated version of the nonDominatedRank function for a GPU could be developed.

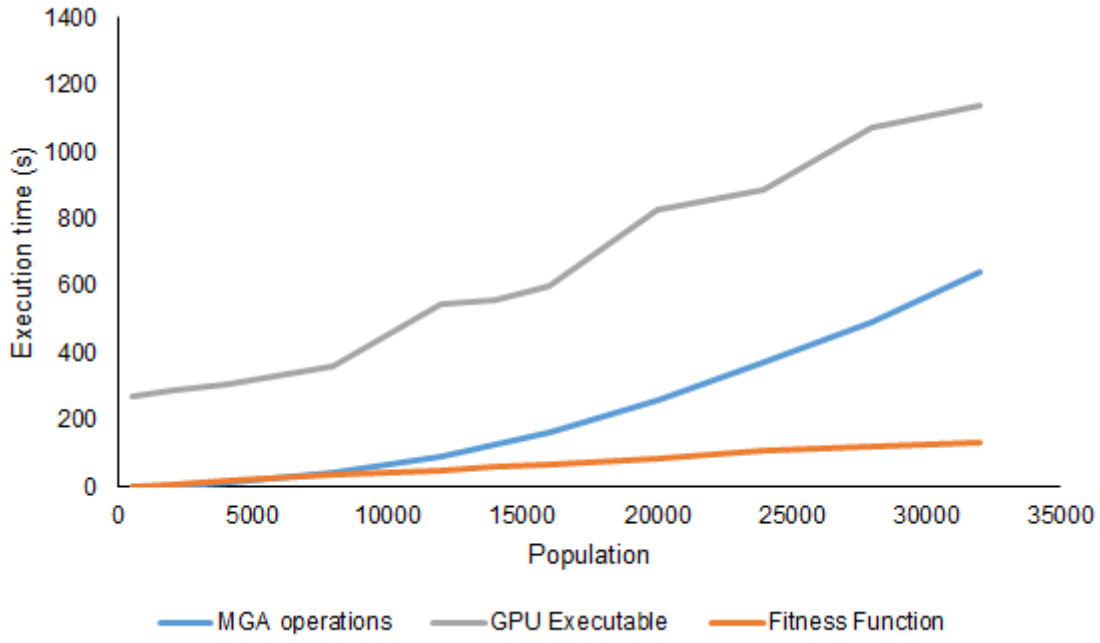

Figure 4.13: Mean execution time of three key computational tasks in an NSGA-II run versus the population size. Mean values were taken from 8 runs of the NSGA-II for 5 generations. Execution time increases with population size for each task. The three computational tasks are the processing time needed for the NSGA-II operations, GPU executable (model integrations) and fitness function evaluation.

4.4 Fitting the noise intensity parameters of the stochastic model

Having found the deterministic parameter sets (parameters of the Broomhead *et al.* [1] model, see Table 2.1), which enable the oculomotor model to generate both saccades with normal velocity profiles and nystagmus waveforms, we fitted the noise intensity parameters (stochastic parameters) of the stochastic model (described in section 6.1) to match the experimental saccadic endpoint variance data, and in the case of nystagmus, the experimental local dimensions pattern.

In both cases (saccadic endpoint variability and nystagmus local dimensions pattern) an iterative uniform grid search of the parameter values of the noise intensity parameters was performed. In each iteration of the method a grid is formed with dimension size equal to three which is the number of the stochastic model parameters (refer to section 6.1). Each grid dimension has a set of values which are uniformly spaced in a value range. The number of values that we chose was five. In each iteration the parameter values in the grid were chosen so they generate the best fit compared to the other points of the grid. The next step of the grid search method involves the creation of a new grid centred to the found point bound with smaller value ranges. After a number of iterations the found point will be chosen to be the optimum in the optimisation process.

4.4.1 Saccadic endpoint variance fitting

The fitness function used in this study is based on the difference simulated and experimental coefficient of variation (CV), which is the standard deviation divided by the mean of the saccadic endpoint for saccades of amplitude 5, 10 and 15 degrees. The following equation evaluates the fitness of the simulated data generated by the Broomhead *et al.* [1] model.

$$cvd = \sum_{i=5,10,15} \left(CV_{sim}(i) - CV_{exp}(i) \right)^2 \tag{4.3}$$

where $CV_{sim}(i)$ is the CV of the saccadic endpoints generated by the Broomhead *et al.* [1] model for a specific saccadic amplitude i , whereas the $CV_{exp}(i)$ is the experimental data measured experimentally by [30] for a specific saccadic amplitude i. $CV_{sim}(i)$ was calculated from 640 simulated saccades generated by the Broomhead *et al.* [1] model.

4.4.2 Fitting the nystagmus local dimensions pattern

In order to fit the stochastic model to the nystagmus data (see section 4.2.5), we used the Broomhead *et al.* [1] model parameters (deterministic parameters), which were found by fitting experimental UPO nystagmus waveforms, and we explored the noise intensity parameters using grid search. The evaluation of the fit was analysed visually, as it was difficult to quantify. We aimed to reproduce the qualitative pattern and to find the local dimensions to be highest during the fast phase of the waveform, lowest during the slow phase and intermediate at the fixed point as shown in [15]. Grid search was used due to a relatively small number of parameters (3 parameters compared to 6 of the Broomhead *et al.* [1] model) that had to be explored. Moreover, in contrast to a NSGA-II approach, we found that using grid search requires less computational resources. Furthermore, we found that changing the stochastic parameter values of the stochastic model did not cause the behaviour to change dramatically as it happens with the Broomhead *et al.* [1] model.

Moreover, we used a different approach to set the noise floor of the time series. Akman *et al.* [15] defined the noise floor as the ratio of the eye-tracker resolution to the standard deviation of the un-normalised time series. We found that this cannot work in different data sets (simulated and/or experimental) where the noise could be caused by numerical error, experimental noise and/or signal quantisation noise. In order to have the same procedure to set the noise for both experimental and simulated data, we used a different method to estimate the level we consider to be the noise floor.

We consider the noise level to be above the concentrated local singular values at the fixed point where their locations are the lowest of all singular values. To choose the local singular values that are below this level, we used the k-means clustering method [108] to group the singular values of the fixed point into two groups. We consider the noise floor to be the mean of the distance between the two found groups.

In the next Chapter, we will present in detail the parameter estimation method that we have developed to fit the Broomhead *et al.* [1] model to experimental nystagmus and saccadic velocity profiles data sets.

5 Using parameter estimation to explore the predictive capacity of the Broomhead *et al.* model

In this chapter, we will present the results of the first subsections from the saccadic and nystagmus pipelines (refer to Figure 4.1) described in Chapter 4. The results include: (i) in section 5.1 we will describe the fitness functions for fitting the parameters of the Broomhead *et al.* [1] model (deterministic parameters) to synthetic and experimental nystagmus waveforms, and synthetic and experimental saccadic velocity profiles; (ii) in section 5.2 we will describe the procedures used for selecting the final solution from the found approximation of the *Pareto* fronts; (iii) in section 5.3.1 we will show the NSGA-II parameter values which provide a good convergence to the *Pareto* fronts based on the hypervolume indicator; and (iv) in sections 5.3.2-5.3.5 we will present the results of fitting the deterministic parameters to synthetic and experimental nystagmus waveforms, and synthetic and experimental saccadic velocity profiles. As described in Chapter 4, the synthetic dataset is used to test whether the NSGA-II is able to accurately recover the model parameter values used to generate them.

As mention in section 3.1, we optimised the Broomhead *et al.* [1] model by minimising the multi-objective function values which means that smaller values show better or higher fitness of the model to the experimental data.

5.1 NSGA-II fitness functions

In order to measure the fitness of each candidate solution in the population of the NSGA-II, two fitness functions were developed: one for saccadic velocity profiles and one for nystagmus waveforms. These functions were constructed using a “trial and error” method and were based on fitting key characteristics derived from the experimental data (described in Chapter 4). In this section, we will present the final versions of these fitness functions.

5.1.1 Fitness function for nystagmus waveforms

The fitness function of fitting the model to nystagmus waveforms is based on the period and shape of the oscillation. The multi-objective fitness function evaluates each individual to extract two values: (i) the difference in shape between the target and the scaled-to-target-period simulated waveforms; and (ii) the difference in period between the waveforms.

An amplitude comparison was not included, as exploratory NSGA-II runs using experimental and synthetic waveform targets indicated that it was redundant; the shape comparison was sufficient for the NSGA-II to converge to the solution with the correct amplitude size.

The input to the fitness function is a single period of the target waveform and a single period of the simulated time series generated from the model for a particular combination of parameter values.

5.1.1.1 Extracting a single period of a synthetic (simulated) waveform

In order to calculate the two fitness values, a single period of both the simulated and target waveforms was used. The procedure for extracting a single period from experimental time series is detailed in section 4.2.2. For synthetic waveforms, the period extraction procedure is as follows.

First, the model was integrated for 6 s. This time interval was chosen as it enables one period of the waveform to be reliably extracted. The first 2.4 s of the time series were discarded, as this is transient dynamics corresponding to the saccade that precedes the onset of oscillations. Next, the simulated gaze time series was examined to determine whether it is an oscillation. This was assessed by calculating the variance of the waveform in the sections lying between 2.8-3.2 s and 3.2-3.6 s. If both of these variances lie below a threshold value (0.001) which was estimated to give the best results using a trial-and-error method, then the input is classified as non-oscillatory and the process is terminated, with the individual arbitrarily assigned a high value of 10^{60} for each objective. The variance is calculated for two time series segments rather than just one, as some nystagmus oscillations generated by the model have parts that are almost linear. In this case, the calculated variance of only one part of the waveform can be very

small (<0.00001). Additionally, in non-oscillatory eye movements, there is a post-saccadic drift (Figure 5.1) which is generated by the model, and it has also been measured experimentally in human eye movements [60]. Measuring the variance of the sections lying between 2.8-3.2 s and 3.2-3.6 s provides a distinction between oscillatory and drift movement.

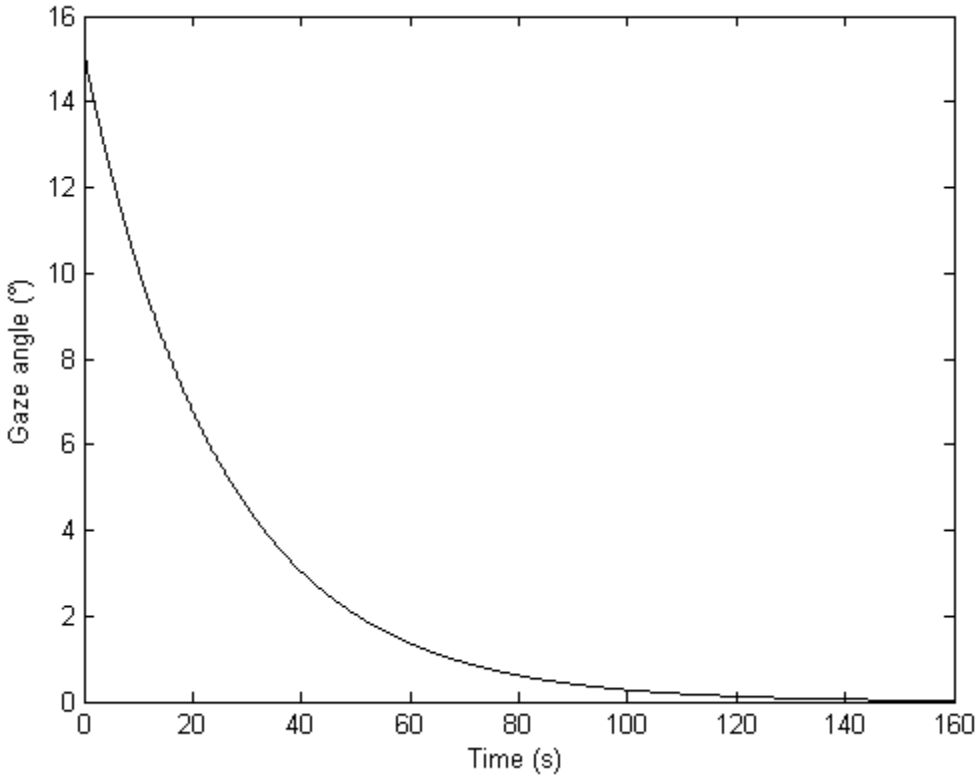

Figure 5.1: Post-saccadic drift of a 15 degree horizontal saccade simulated by the Broomhead *et al.* [1] model. The drift is the effect of the leaky integrator of equation (2.3) (see section 2.2 for more details). Post-saccadic drift is also observed in experimental conditions, when the eyes of the participants are deprived of visual stimuli (*e.g.* during darkness). Time is in seconds (s) and gaze angle is in degrees (°).

If the variances lie above the threshold, then there is an oscillation and one period of the oscillation is extracted by locating the global minima of each period of the time series as follows. With the input time series as $\{x_1, \ldots, x_K\}$, where x_i denotes the horizontal gaze angle at the time point indexed by i, local minima are first found by choosing the points x_P for which:

$$x_{P-1} > x_P \text{ and } x_P < x_{P+1}.$$

In order to find the global minimum, the locations of these 6 local minima are rotated clockwise by 5 degrees in the (P, x_P) plane and the x_P values are sorted in decreasing order (Figure 5.2). The first two sorted local minima are identified as two consecutive global minima and all points between these two minima are extracted as one period of the oscillation. It should be noted that even though the majority of the waveforms have one local minimum per period, there are other cases for which there are multiple local minima, thus necessitating the method we have developed for locating global minima.

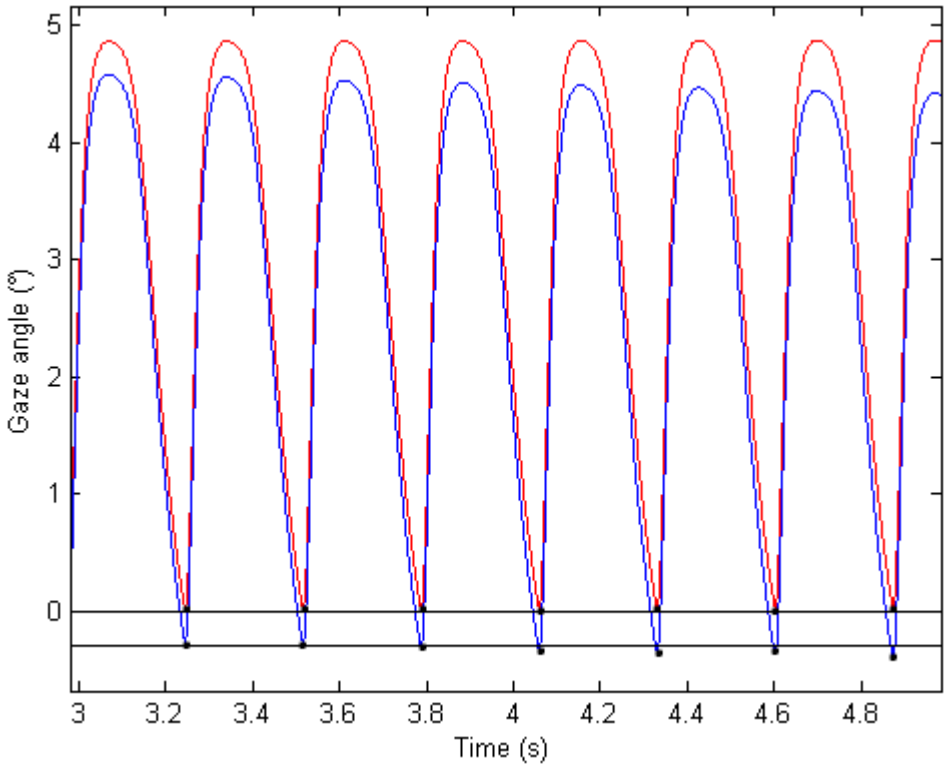

Figure 5.2: Rotation of local minima for the extraction of a single period waveform. The rotation is performed on the local minima only, but the original and rotated time series are also shown for clarification. A part of the input time series to the fitness function is shown in red. The rotation of the time series by 5 degrees clockwise is shown in blue. Black dots represent the locations of the local minima in each case. The horizontal black lines allow better identification of the changes in position of the local minima under rotation. In the case of the rotated local minima, their horizontal position decreases with time.

MATLAB provides the *findpeaks* function for locating local optima; however we found that our version of the method (*findpeaks2*) is faster in processing time by ~50%. The MATLAB version of the function does have extra capabilities; however these were not applicable to our work.

5.1.1.2 Calculating the shape and period difference

In order to calculate the shape difference, the extracted waveform is scaled to the target waveform on the time axis, so that they have the same period. Subsequently, cubic interpolation is applied to the scaled waveform, so that it has the same time mesh as the target. The difference in shape, sd, is then calculated as a sum-of-squares:

$$sd = \sum_{i=1}^{N} (S_i - T_i)^2,$$

where S_i and T_i denote the values of the scaled and target waveforms, respectively, at time t_i, where $t_{i+1} - t_i$ is the (fixed) data sampling interval.

The period difference, *pd*, is given by:

$$pd = (\tau_E - \tau_T)^2,$$

where τ_E is the period of the extracted unscaled waveform and τ_T is the period of the target waveform.

5.1.2 Fitness function for saccadic velocity profiles

The input of the fitness function is the target saccadic velocity profiles for saccades of amplitude 5, 10 and 20 degrees, and the corresponding simulated profiles generated from the model for a particular combination of parameter values. As described in section 4.2.3, we chose these particular amplitudes because they are in the amplitude interval of the most frequently made saccades. The fitness function comprises three main objectives $\{d_1, d_2, d_3\}$; these are the squared sum of the squared difference between the simulated and target data for each saccade amplitude:

$$d_j = \sum_{i=1}^{N} \left(S_{ij} - T_{ij}\right)^2.$$

Here, S_{ij} and T_{ij} denote the values of the simulated and target profiles, respectively, at time t_i, where $t_{i+1} - t_i$ is the (fixed) data sampling interval, and j indexes the different saccade amplitudes. To extract the simulated velocity profile that is compared to the target one, we used the target's velocity profile end point as the end point of the simulated profile, ensuring that both had the same length.

5.2 Selection of the best individual

As described in Chapter 3, when the NSGA-II terminates it returns a set of points which is an approximation to the optimal *Pareto* front. In the next two sections, we will describe the method that we used to choose the individuals from the returned approximated *Pareto* front, for both nystagmus and saccade fitting.

5.2.1 Selection procedure for nystagmus waveforms

In order to select the final individual from the returned approximated *Pareto* front in nystagmus waveform fitting, we chose the individual of the approximated *Pareto* front which had the smallest Euclidean distance to the zero value in both objectives. This procedure allows an individual with a sufficient trade-off between the objectives.

5.2.2 Selection procedure for saccadic velocity profiles

To select the final individual from the returned approximated *Pareto* front in saccade fitting, we applied the same procedure as in the case of nystagmus fitting, namely the individual with the smallest Euclidean distance to the zero value in all objectives was selected. We further explored different final solution selection methods. These selections methods were based on different objective values of the approximated *Pareto* front found by the NSGA-II using the method discussed in 4.3. One alternative method chooses the solution with the lowest value on objective 1 (*i.e.* the best fit to a 5 deg saccade), whereas other two selection methods choose the solution with the lowest value in objective 2 and 3, respectively (*i.e.* the best fit to a 10 and 20 deg saccade, respectively).

5.3 Results

In the following sections, we will present the results of selecting the NSGA-II parameter values, which yielded the best results among all the NSGA-II parameter values we examined.

The results for nystagmus and saccadic velocity profiles are presented as box plots. Each box plot axis is scaled to the constraints values that we used in the optimisation method, except for parameter ε which we set the upper bound of 0.05. Moreover, a heat map is used to present results for fitting experimental saccadic velocity profiles. The box plots show the median and spread of the solutions, whereas the heat maps allow visualisation of all the objective values for all the solutions without the need of data compression or removal [77]. The solutions presented on the heat map were seriated (ordered) by the objective values of the returned non-dominated *Pareto* front [77]. The seriation repositions solutions with similar objective fitness values closer together [109], thus allowing visual identification of trends in the model [77]. The code used for the seriation was developed by [110] [2].

5.3.1 Selection of NSGA-II parameter values

In this section, we will present the results of the NSGA-II parameter selection method. As described in Chapter 4, we explored different combinations of NSGA-II parameters to find the ones that returned the best possible fits of the Broomhead *et al.* [1] model to synthetic and experimental nystagmus, and saccade data. As described in section 4.3.1, we have not explored all the NSGA-II parameters because this would increase considerably the computation time. Therefore, we executed a number of exploratory optimisation runs with different combinations of population size, crossover function (intermediate or heuristic), and distance measure type (phenotype or genotype), whilst keeping the mutation type set to adaptive feasible and the selection type set to binary tournament. As discussed in section 4.3.1, the exploratory runs for both nystagmus and saccadic datasets showed that population size and crossover function had the greatest influence on the returned solutions, and that heuristic crossover allowed the NSGA-II to converge for considerably smaller populations. Of the remaining NSGA-II parameter, best results were obtained with a distance measure based on phenotype. The population sizes that we tested were 500, 1000, 2000, 4000 and 8000. The final size of the population was chosen to provide a sufficient trade-off between accuracy, convergence and NSGA-II running time.

[2] The code was downloaded from https://github.com/fieldsend

The time required for testing the NSGA-II parameter configurations was found to be considerably long (Table 5.1). During the process of forming the fitness function, a small change may alter the fitness function output landscape substantially. Consequently, the NSGA-II parameter optimisation has to be performed again for each putative fitness function. The time required to run 16 optimisations on the four target nystagmus waveforms (shown in Figure 5.3) for the 5 different population sizes was estimated 99 hours (4.2 days) and this was achieved using one GPU card (Table 5.1). The run times that we specify here are derived from the multiple NSGA-II parallel method described in section 4.3.5.

Table 5.1: Time required to run the NSGA-II 16 times for the four synthetic nystagmus waveform targets shown in Figure 5.3 against population size. The run times shown here were obtained using the multiple NSGA-II parallel method described in section 4.3.5.

Population size	Hours	Days
500	3.943	0.16
1000	7.278	0.30
2000	18.17	0.75
4000	38.43	1.60
8000	65.06	2.71

In the following sections, we will present the results obtained from population sizes of 500, 1000, 2000, 4000 and 8000 used to fit the model to nystagmus and saccadic data. For each NSGA-II parameter combination, we ran the NSGA-II 16 times to obtain statistical measurements about the quality (accuracy and convergence) of the solutions.

5.3.2 Synthetic nystagmus oscillations - fitting results

The parameters of the Broomhead *et al.* [1] model, which were used to generate the synthetic nystagmus targets are shown in Table 5.2. These parameters were chosen as they generate a range of the waveforms observed experimentally. The time series generated using these parameters and the extracted single period

waveforms are shown in Figure 5.3 and Figure 5.4, respectively. In each case the extraction of the single period waveform was performed using the technique described in section 5.1.1.1.

The hypervolume indicator, which is we used to measure the convergence of the NSGA-II (refer to section 4.3.1), shows that with population size of 4000, the NSGA-II returns solutions with the best convergence compared to the other population sizes. Figures 5.7-5.10 show the hypervolume size for the different population sizes against generation number for all the four waveform targets, whereas Figures 5.11-5.14 show the hypervolume size of the last generation for the different population sizes for all the four waveform targets. Moreover, Table 5.4 shows how the mean and coefficient of variation of the shape difference objective varies with population size. This confirms that a population of 4000 returns solutions with sufficient convergence.

Table 5.2: Parameter combinations of the Broomhead *et al.* [1] model used to generate the four synthetic nystagmus waveforms shown in Figure 5.3. The waveforms are designated as NSA, NSB, NSC and NSD, and simulate the following nystagmus types: NSA - asymmetric pseudo-cycloid; NSB - pseudo-cycloid; NSC - jerk; NSD - bidirectional jerk.

Target ID	α	β	ε	γ	α'	β'
NSA	270	3.5	0.0035	0.06	600	10
NSB	210	1.5	0.0020	0.03	380	6
NSC	110	1.5	0.0035	0.05	600	9
NSD	110	1.5	0.0065	0.07	550	9

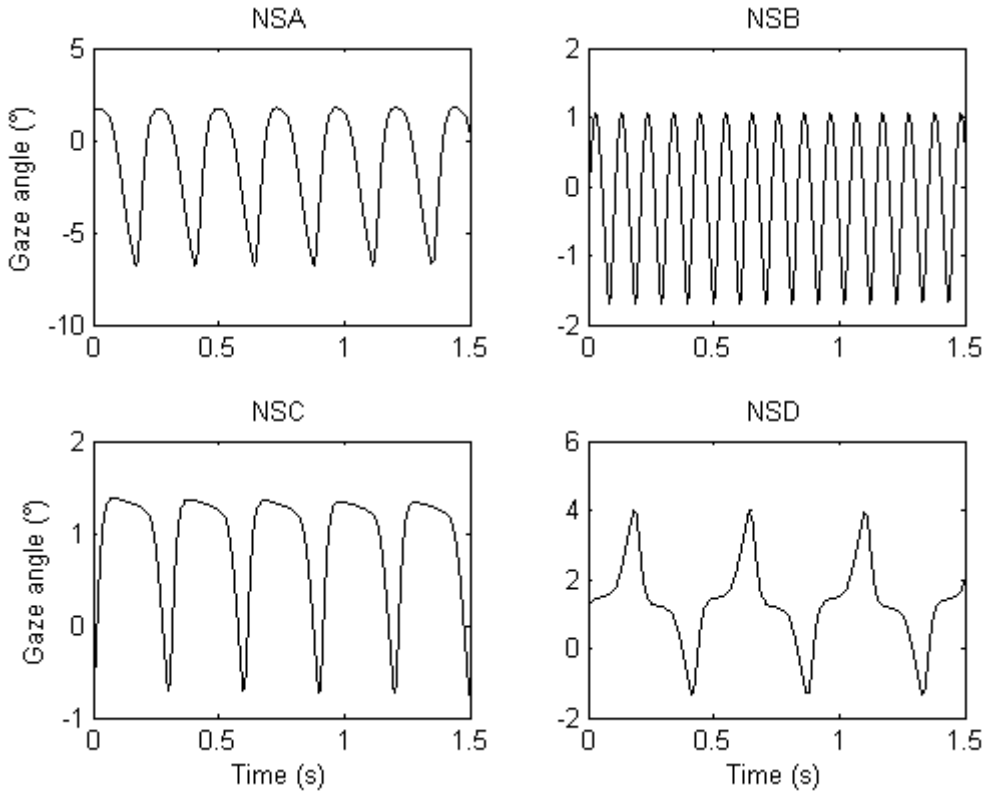

Figure 5.3: Parts of the synthetic nystagmus time series generated by the Broomhead *et al.* [1] model. The corresponding parameter values are shown in Table 5.2. The simulated waveform types are: NSA - asymmetric pseudo-cycloid; NSB - pseudo-cycloid; NSC - jerk; NSD - bidirectional jerk. For each parameter combination, 6 s simulation was used to extract the single-period target waveform on which the different NSGA-II parameter combinations were tested. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

The box plots in Figure 5.5 and Figure 5.6 show the optimised parameter values for the different population sizes. The results for the first four deterministic parameters α, β, ε, and γ confirm that the NSGA-II can get a good convergence with population size of 4000. However, in the case of the parameters α' and β', there is no improvement in convergence with increasing population size (Figure 5.6). The bar plot in Figure 5.15 shows the coefficient of variation of each parameter for each waveform target with population size 4000. The fitness function is less sensitive to the values of the parameters α' and β' (controlling the saccadic on-response – recall section 2.2), as indicated by their significantly

larger coefficients of variation, and more sensitive to the values of the parameters α and β (controlling the saccadic off-response), as indicated by their significantly smaller coefficients of variation. Figure 5.4 compares the waveforms generated by the NSGA-II solutions with population size 4000 to the target waveforms. This shows that the optimised waveforms are almost identical to the target ones.

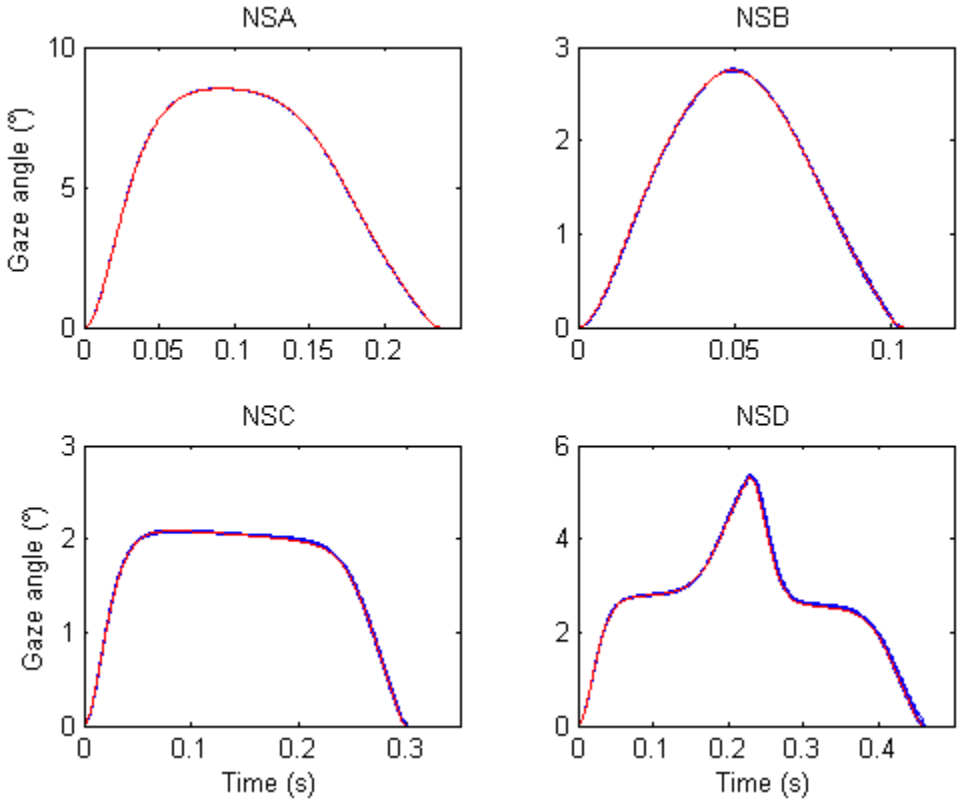

Figure 5.4: Waveforms generated from the solutions found by the NSGA-II for each synthetic nystagmus waveform target. The target waveforms are plotted in red; the 16 optimised waveforms are plotted in blue and were chosen from each NSGA-II run and selected using the method described in section 5.2.1. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). The simulated nystagmus waveform types are: NSA - asymmetric pseudo-cycloid; NSB - pseudo-cycloid; NSC - jerk; NSD - bidirectional jerk.

Table 5.3: Table with the target parameter values, found mean and coefficient of variation (CV) of the optimised parameter values for each target waveforms (NSA, NSB, NSC and NSD). The target waveforms are shown in Figure 5.3 and they were generated using the Broomhead *et al.* [1] model, with the parameter values shown in Table 5.2. The found mean and CV were calculated by the solutions selected using the method described in section 5.2.1. This was performed for 16 NSGA-II runs end with population size of 4000.

Target ID	Value type	α	β	ε	γ	α'	β'
NSA	target	270	3.5	0.0035	0.06	600	10
	mean	269.85	3.51	0.0037	0.06	652.79	11.19
	CV	0.0029	0.00701	0.0703	0.02	0.0846	0.1134
NSB	target	210	1.5	0.0020	0.03	380	6
	mean	209.71	1.49	0.0016	0.033	357.14	5.49
	CV	0.00463	0.01444	0.3359	0.09219	0.49607	0.65141
NSC	target	110	1.5	0.0035	0.05	600	9
	mean	111.46	1.35	0.00445	0.05168	680.558	9.9
	CV	0.0157	0.0625	0.097	0.1753	0.3914	0.411
NSD	target	110	1.5	0.0065	0.07	550	9
	mean	111.3	1.53	0.0065	0.068	891.59	15.04
	CV	0.0025	0.0047	0.0034	0.02145	0.1259	0.1326

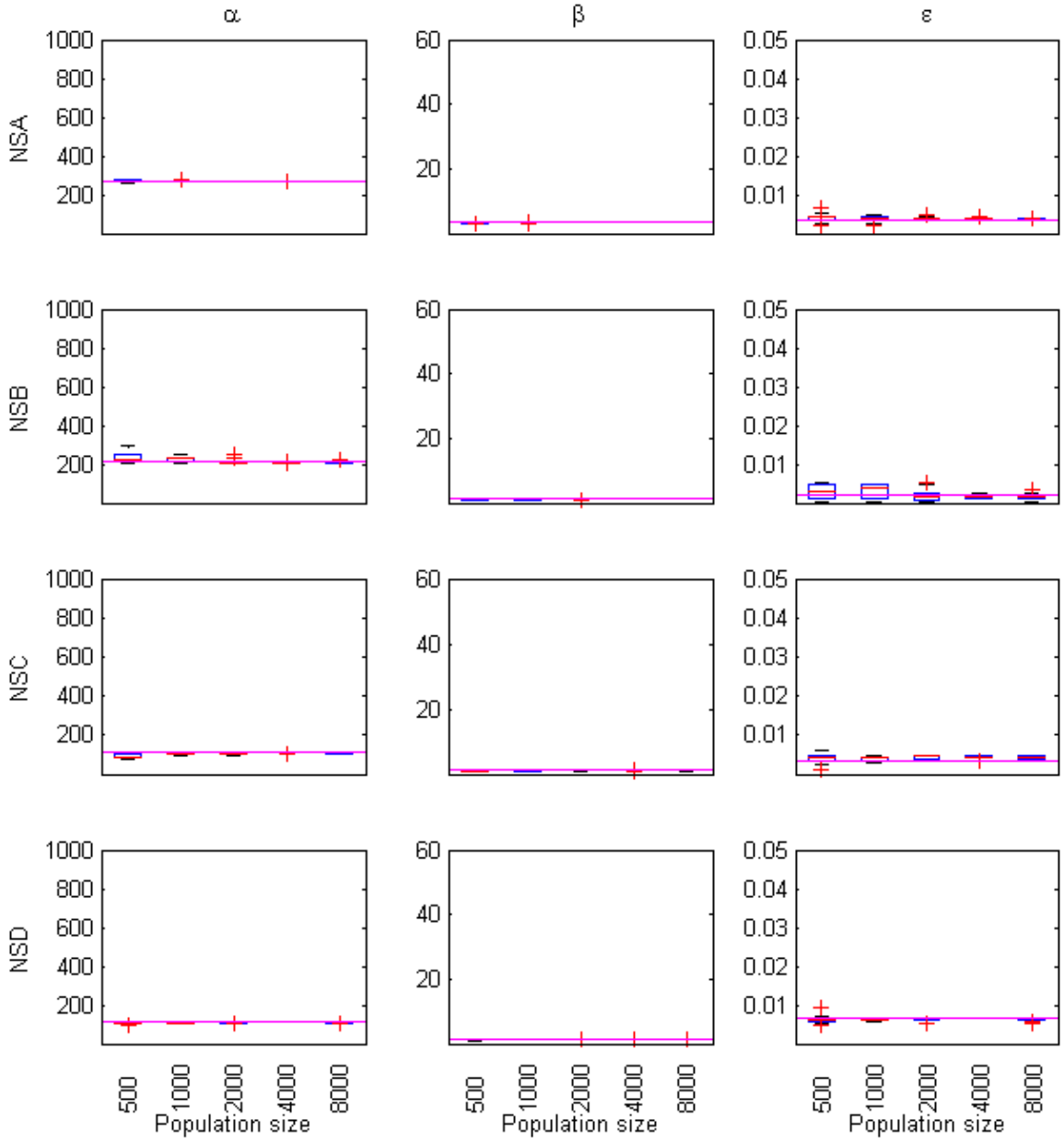

Figure 5.5: Boxplots showing the optimised Broomhead *et al.* [1] model parameter values for each synthetic nystagmus waveform target versus population size (target waveforms are shown in Figure 5.4). The first line of plots shows the optimised values of α, β and ε for the synthetic waveform target NSA, whereas those for NSB, NSC, and NSD are shown by the second, third and fourth lines of plots, respectively. In each plot, the magenta line represents the parameter value used to generate the target waveform. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

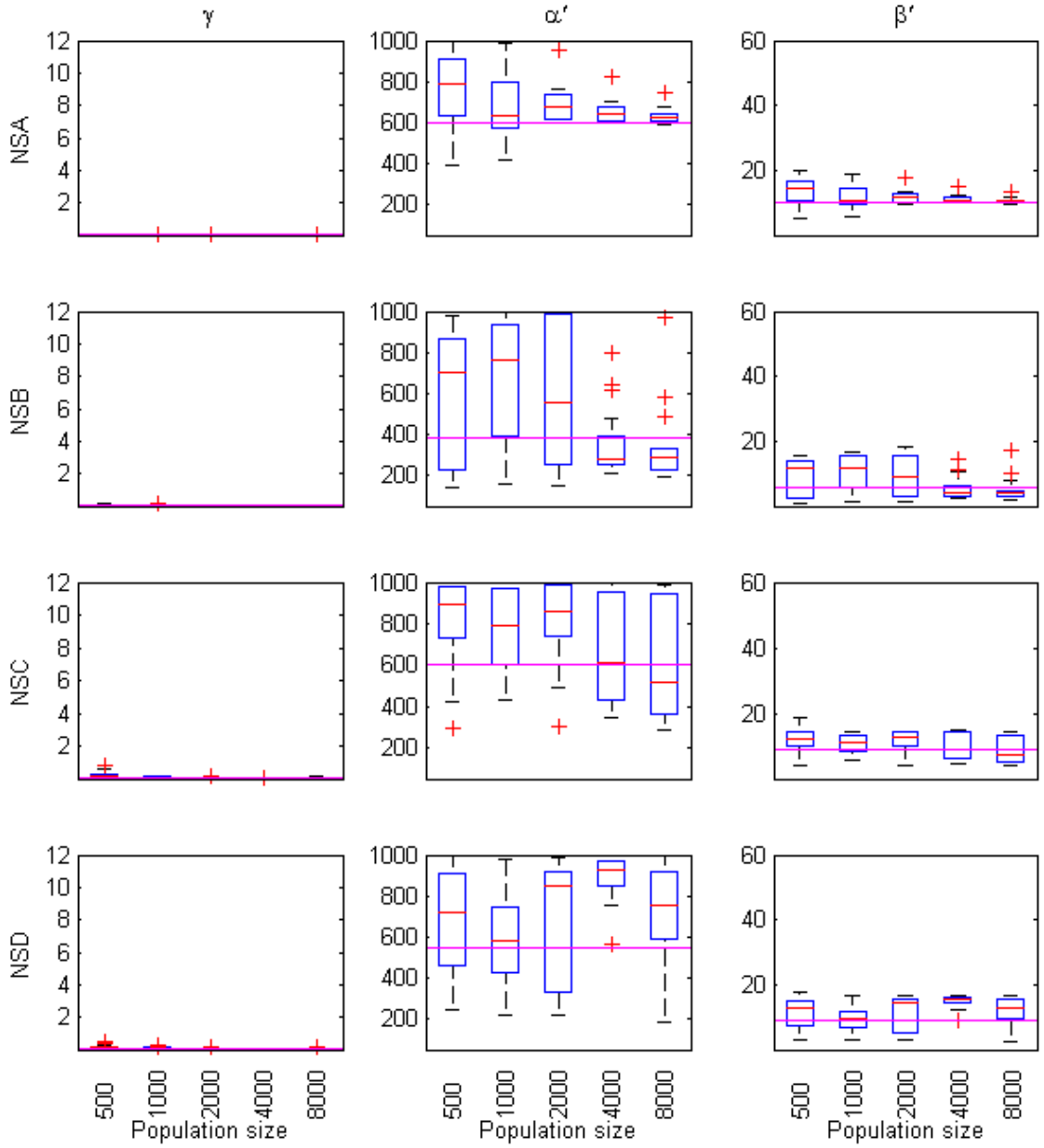

Figure 5.6: Boxplots showing the optimised Broomhead *et al.* [1] model parameter values for each synthetic nystagmus waveform target versus population size (target waveforms are shown in Figure 5.4). The first line of plots show the optimised values of γ, α' and β' for the synthetic waveform target NSA, whereas those for NSB, NSC, and NSD are shown by the second, third and fourth lines of plots, respectively. In each plot, the magenta line represents the parameter value used to generate the target waveform. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

Table 5.4: Variation with population size of the mean and coefficient of variation (in parenthesis) of the nystagmus shape difference objective for each synthetic waveform target (NSA, NSB, NSC and NSD) shown in Figure 5.4. 16 NSGA-II runs were carried out for each population size (500, 1000, 2000, 4000 and 8000). The best solution from each NSGA-II run was selected using the method described in section 5.2.1.

Target ID	500	1000	2000	4000	8000
NSA	176.65 (2.5003)	23.061 (1.7073)	7.641 (1.0079)	4.420 (0.9515)	1.877 (1.3628)
NSB	24.258 (0.8653)	17.001 (0.6073)	6.228 (1.4130)	0.155 (0.7827)	1.012 (1.5797)
NSC	1154.788 (1.0813)	138.079 (0.8308)	67.299 (1.0530)	21.405 (0.9786)	43.979 (1.1256)
NSD	2587.758 (1.1069)	550.819 (1.3653)	146.421 (1.9628)	19.741 (0.2195)	223.742 (1.6406)

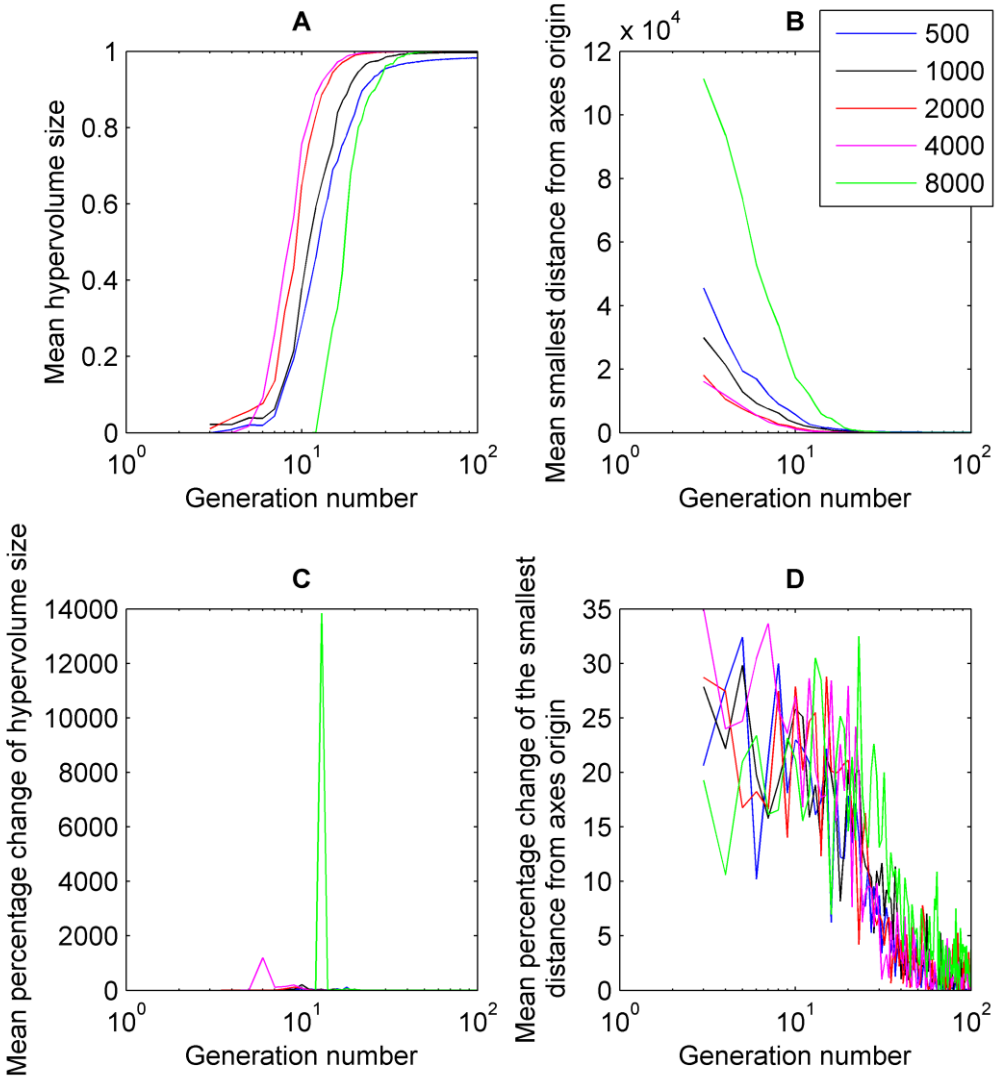

Figure 5.7: Mean values and mean percentage change of the hypervolume size and smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population sizes (in logarithmic scale) for the target NSA. The mean values were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSA shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000).

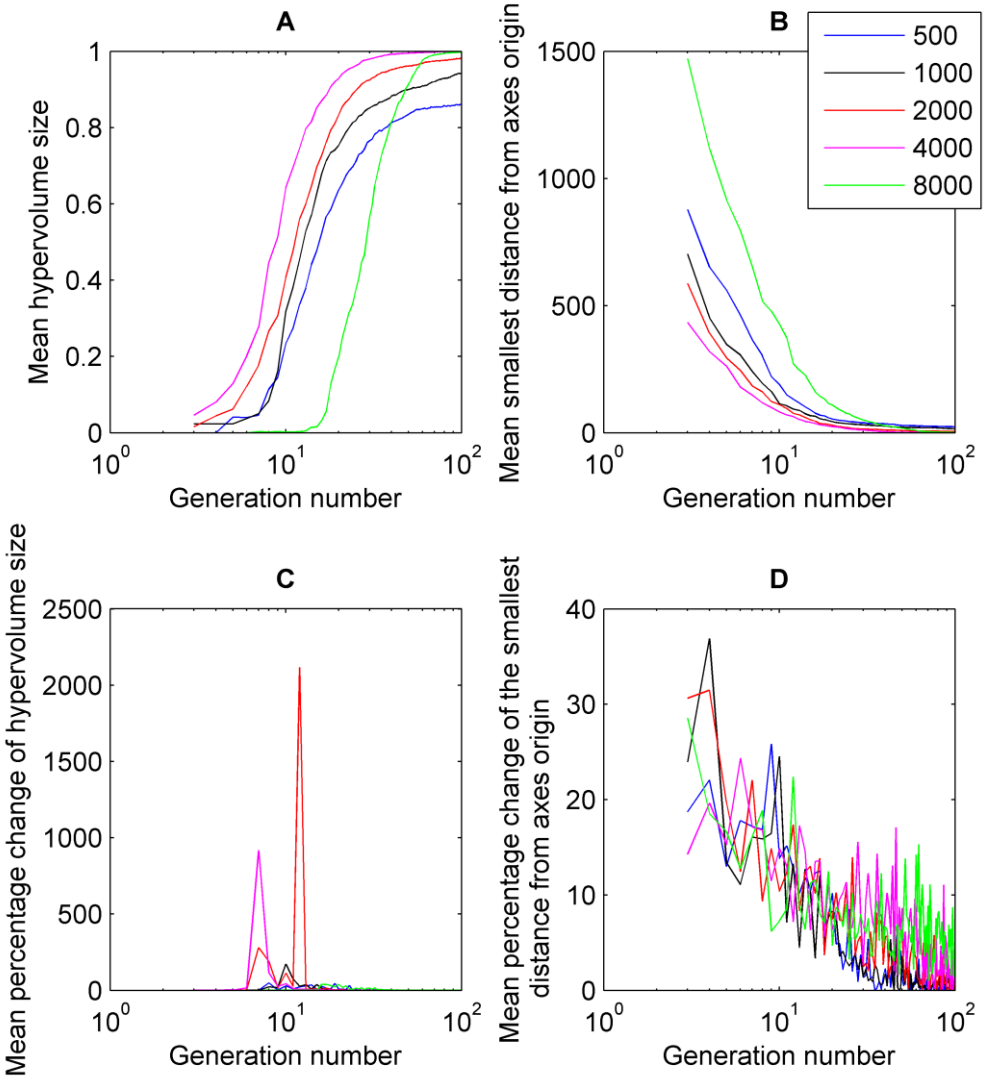

Figure 5.8: Mean values and mean percentage change of the hypervolume size and smallest distance of the selected individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population sizes (in logarithmic scale) for the target NSB. The mean values were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSB shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000).

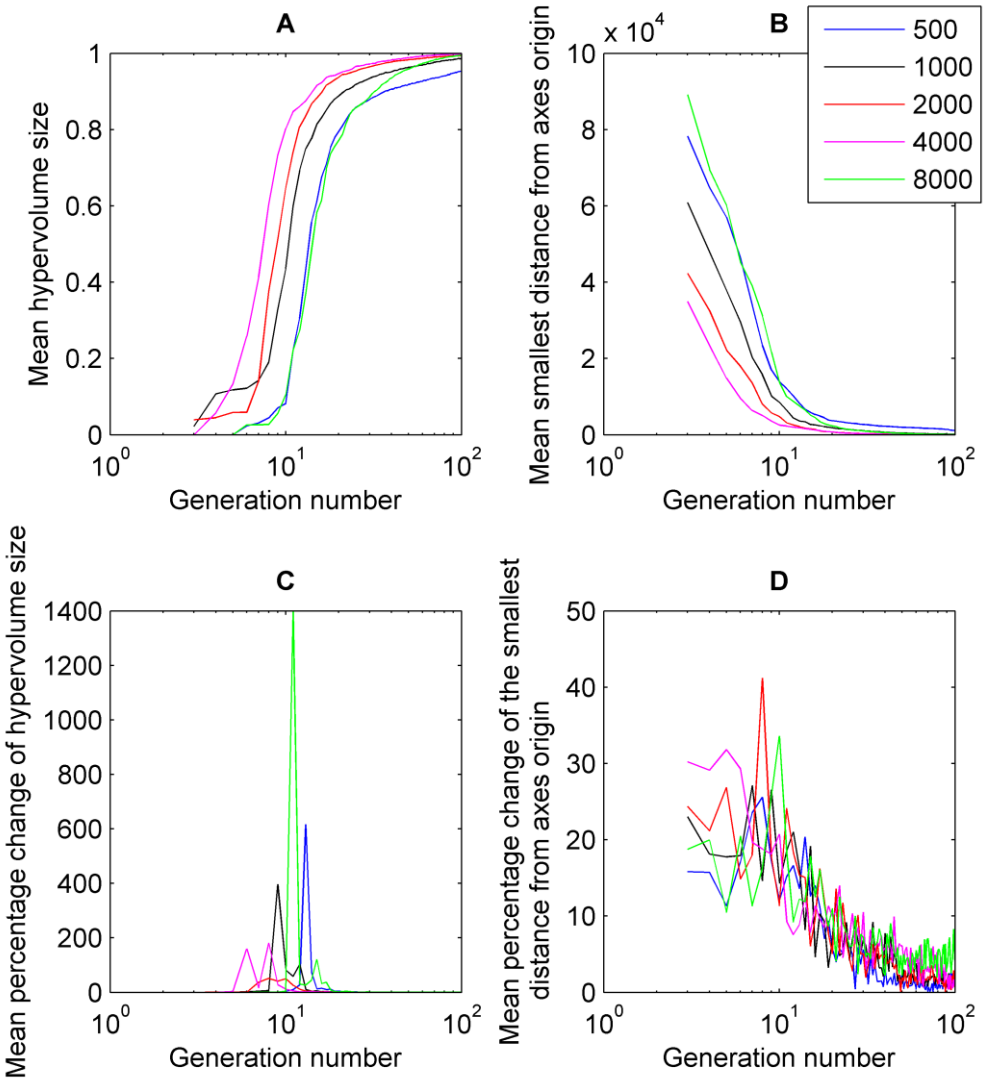

Figure 5.9: Mean values and mean percentage change of the hypervolume size and smallest distance of the selected individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population sizes (in logarithmic scale) for the target NSC. The mean values were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSC shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000).

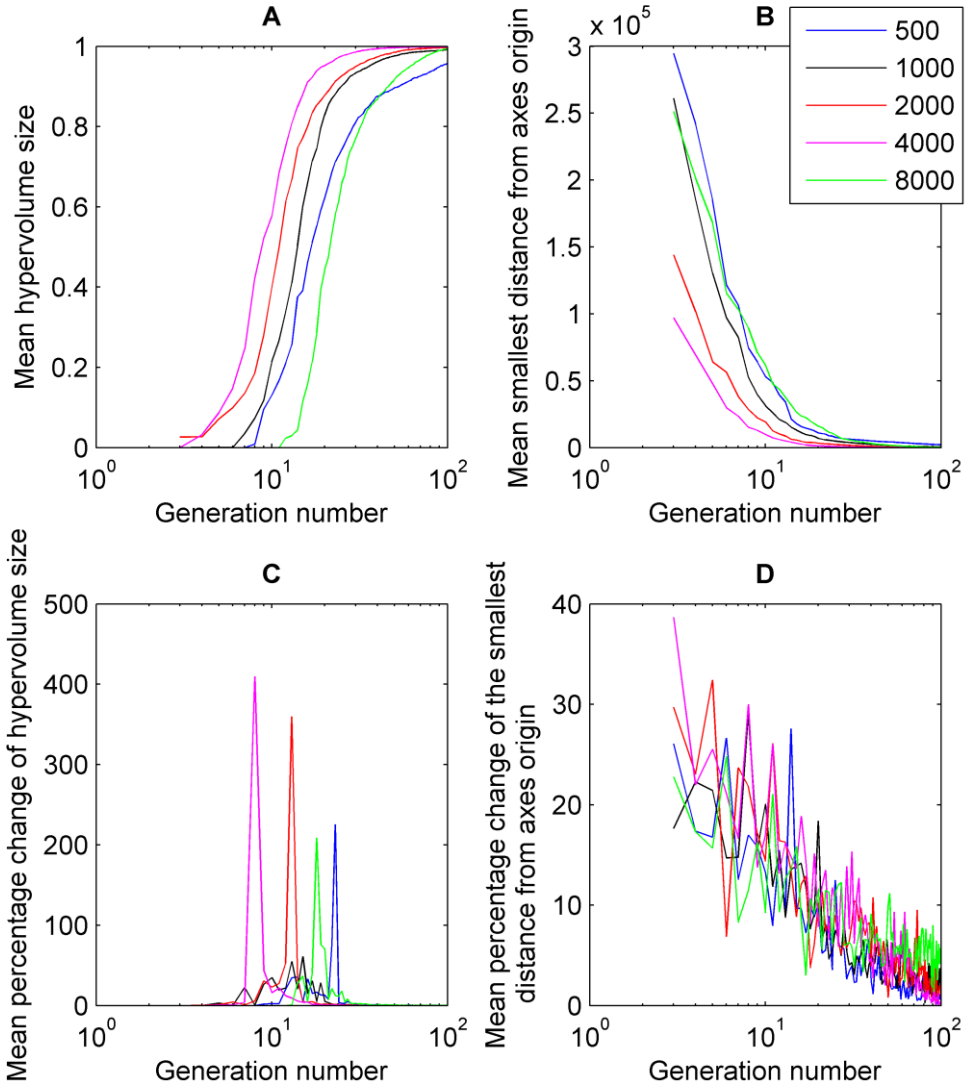

Figure 5.10: Mean values and mean percentage change of the hypervolume size and smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population sizes (in logarithmic scale) for the target NSD. The mean values were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSD shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000).

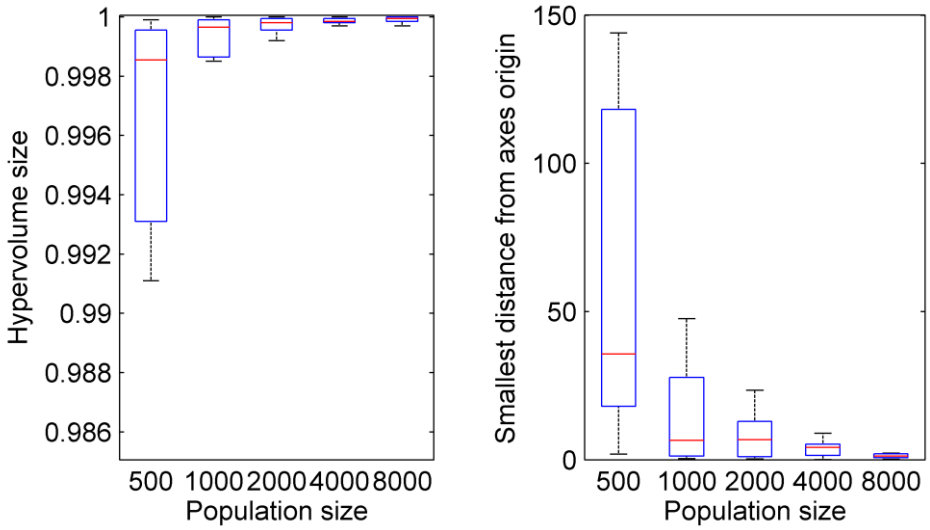

Figure 5.11: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSA shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

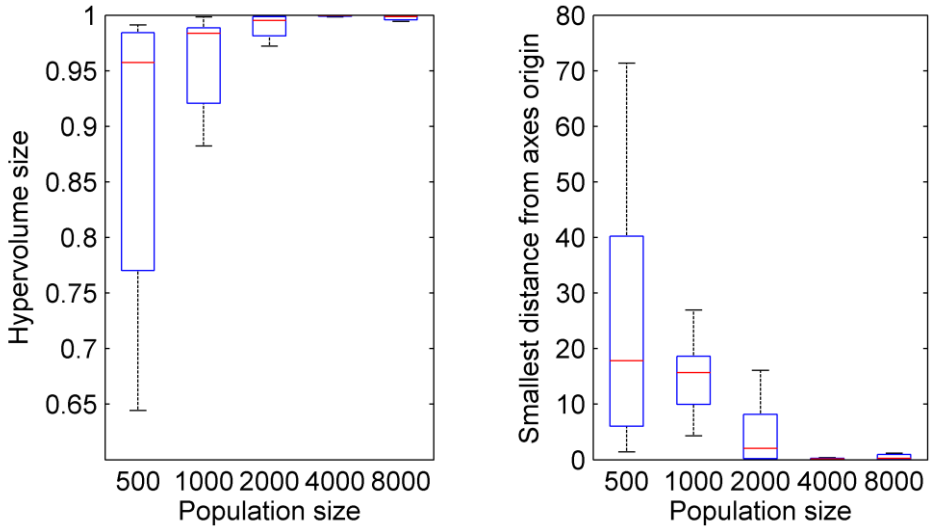

Figure 5.12: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSB shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

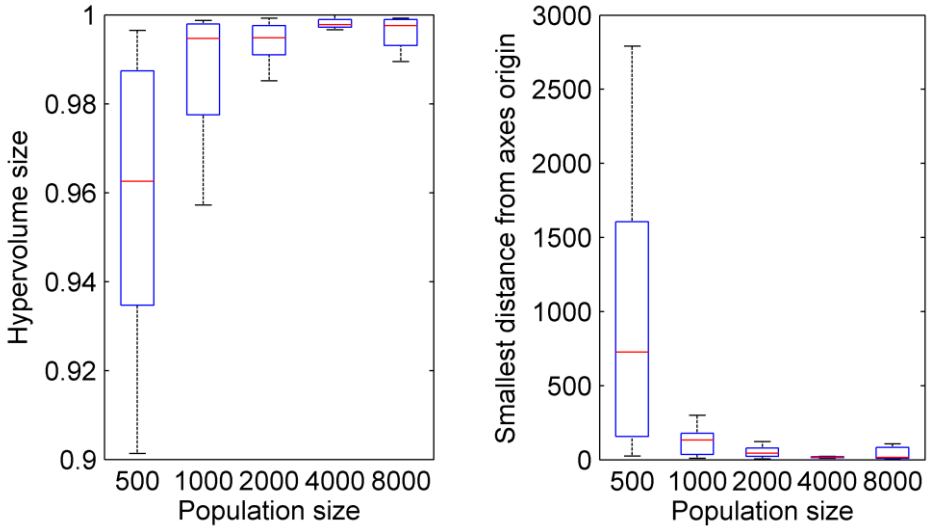

Figure 5.13: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSC shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

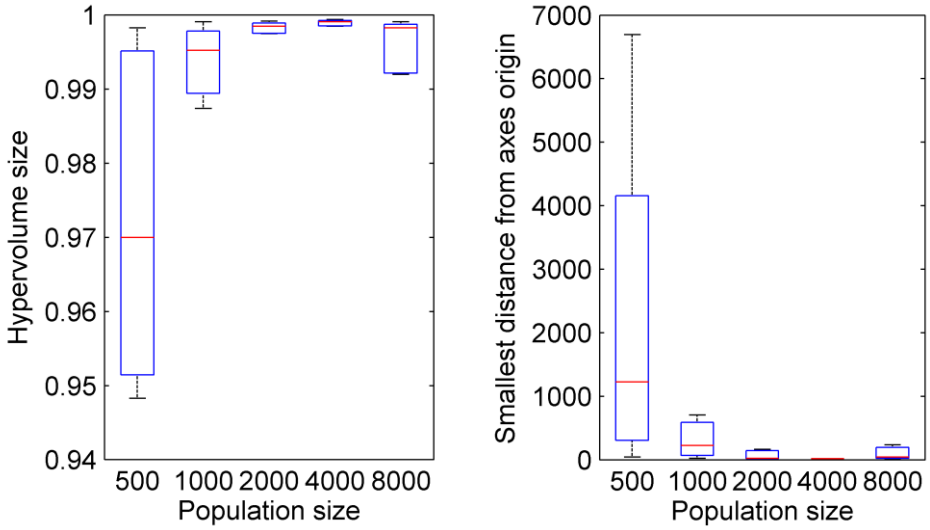

Figure 5.14: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals of the estimated *Pareto* front against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic waveform target NSD shown in Figure 5.4, for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25^{th} and 75^{th} percentiles. The whiskers extend to the interquartile range.

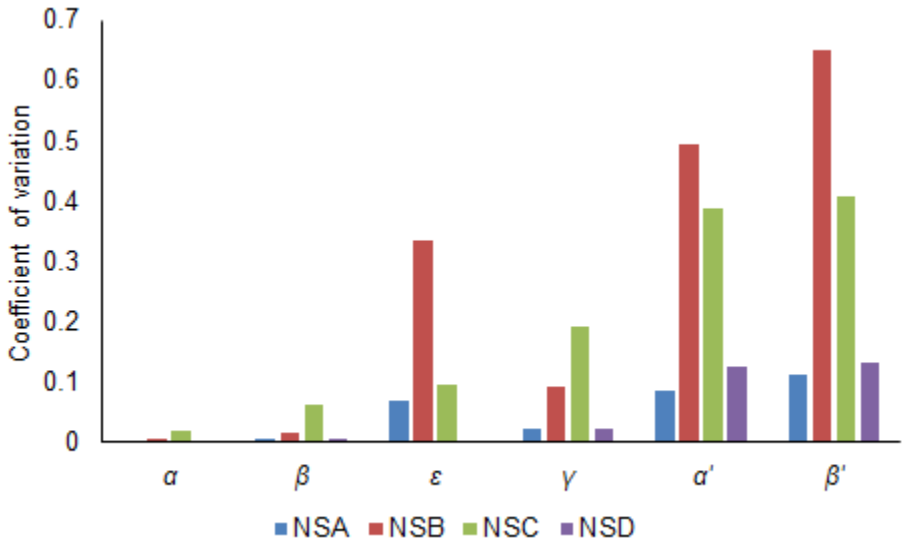

Figure 5.15: Coefficient of variation of the optimised Broomhead *et al.* [1] model parameters for each target synthetic nystagmus waveform shown in Figure 5.4. The coefficient of variation values were calculated from 16 NSGA-II runs with a population size of 4000 in each case. The best solution from each NSGA-II run was selected using the method described in section 5.2.2.

5.3.3 Synthetic saccadic velocity profiles - fitting results

As in the case of synthetic nystagmus targets, we used four combinations of model parameter values (Table 5.5) to generate different saccadic velocity profiles and those were used as targets in the parameter estimation method. The parameter combinations were chosen to set the Broomhead *et al.* [1] model to generate saccades of different velocity profiles. For each parameter combination, we generated velocity profiles for amplitudes of 5, 10 and 20 deg, which, as described in section 4.2.3, are the most frequent saccade amplitudes. In order to determine the endpoint of each saccade in the experimental data, a velocity threshold is used, with the saccade considered to have ended once the velocity $v(t)$ drops below 2 deg/s. The extracted velocity profiles are derived from the start of the saccade to the found endpoint.

Table 5.5: Parameter combinations of the Broomhead *et al.* [1] model used to generate the four target synthetic saccadic velocity profiles shown in Figure 5.16. The four waveforms are designated as SSA, SSB, SSC and SSD.

Target ID	α	β	ε	γ	α'	β'
SSA	650	15	0.0035	0.05	380	1.5
SSB	10	0.5	0.0065	0.5	480	9
SSC	100	50	0.009	4	380	9
SSD	15	5	0.005	5	600	10

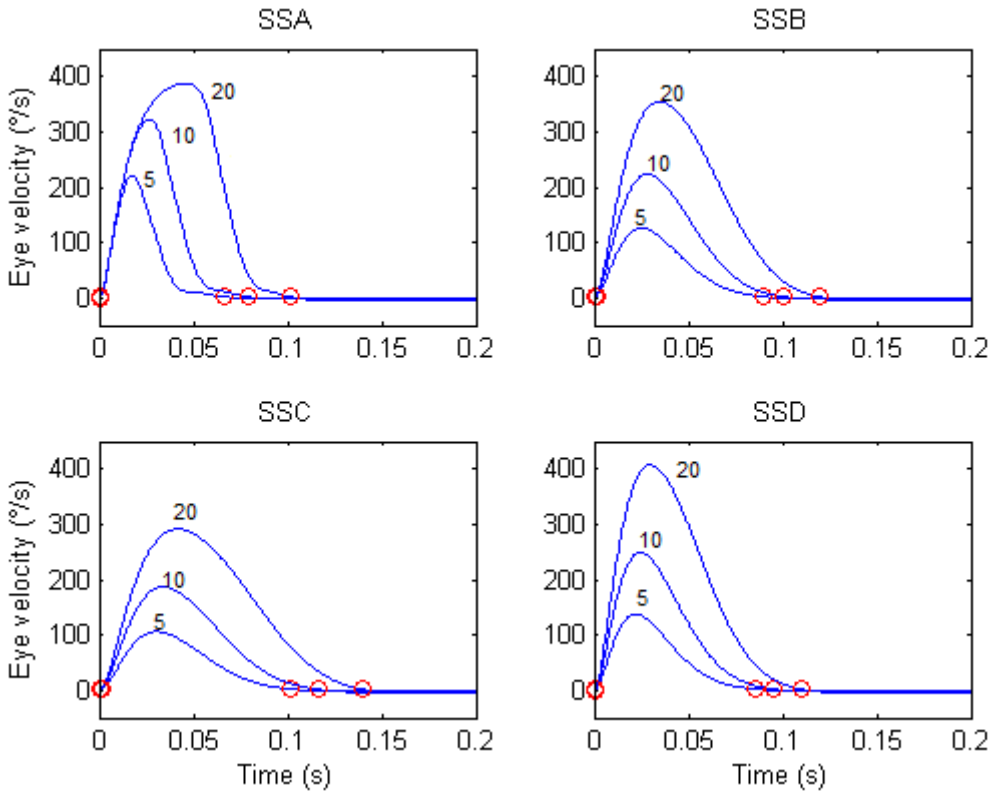

Figure 5.16: Synthetic velocity profiles simulated by the Broomhead *et al.* [1] model using the parameters from Table 5.5 for amplitudes of 5, 10 and 20 degs. The "red" circles represent the initiation and termination of the saccades calculated by applying a velocity threshold of 2 degs/s. On each plot, the vertical axis represents the horizontal eye velocity in degs/s (°/s), with positive values denoting rightward motion. Time is in seconds (s).

The hypervolume indicator shows that with population size 8000, the NSGA-II returns solutions with the best convergence compared to the other population sizes. Figures 5.17-5.20 show the hypervolume size for the different population sizes against generation number for all the four velocity profile targets, whereas Figures 5.21-5.24 show the hypervolume size of the last generation for the different population sizes for all the four velocity profile targets. Moreover, the means and coefficients of variation of all the objectives (Table 5.6) for the best selected individuals using the method described in section 5.2.2, show that a population size of 8000 is sufficient to provide the best convergence compared to the other population sizes. The box plots in Figures 5.25 and 5.26 show the optimised parameter values for the different population sizes (results for population 500 is not shown as it provided extremely inaccurate results). For all parameters, convergence to the targets' parameter values is observed with population size 8000.

The NSGA-II can get very good convergence for the target parameter values of α', β' and ε even with small population sizes (refer to Figures 5.25 and 5.26 and Table 5.6). This suggests that the Broomhead *et al.* [1] model is more sensitive to the values of these parameters when generating saccades. The velocity profiles generated using the optimised parameters are not shown, as they are the same as the target velocity profiles (see Figure 5.16).

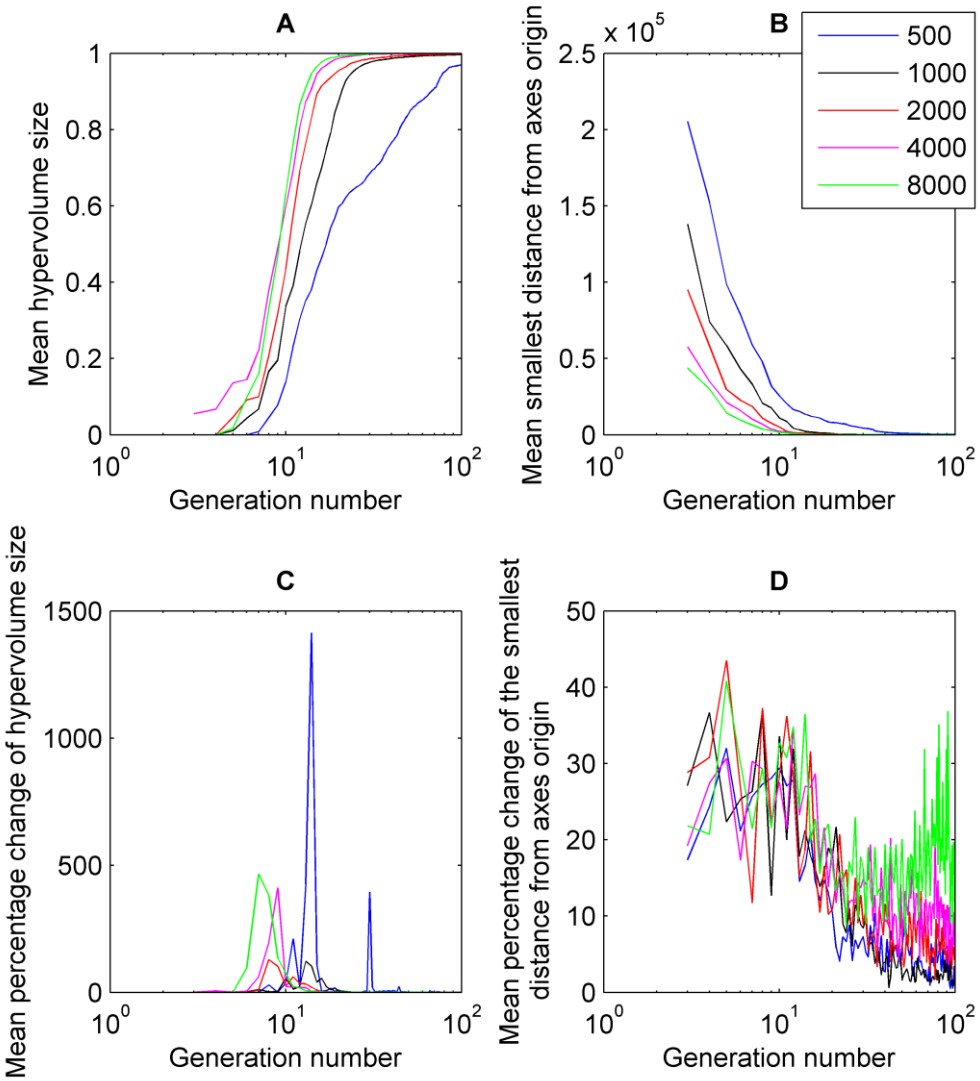

Figure 5.17: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target SSA. The mean were calculated from 16 runs of the NSGA-II for the experimental velocity profiles shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000).

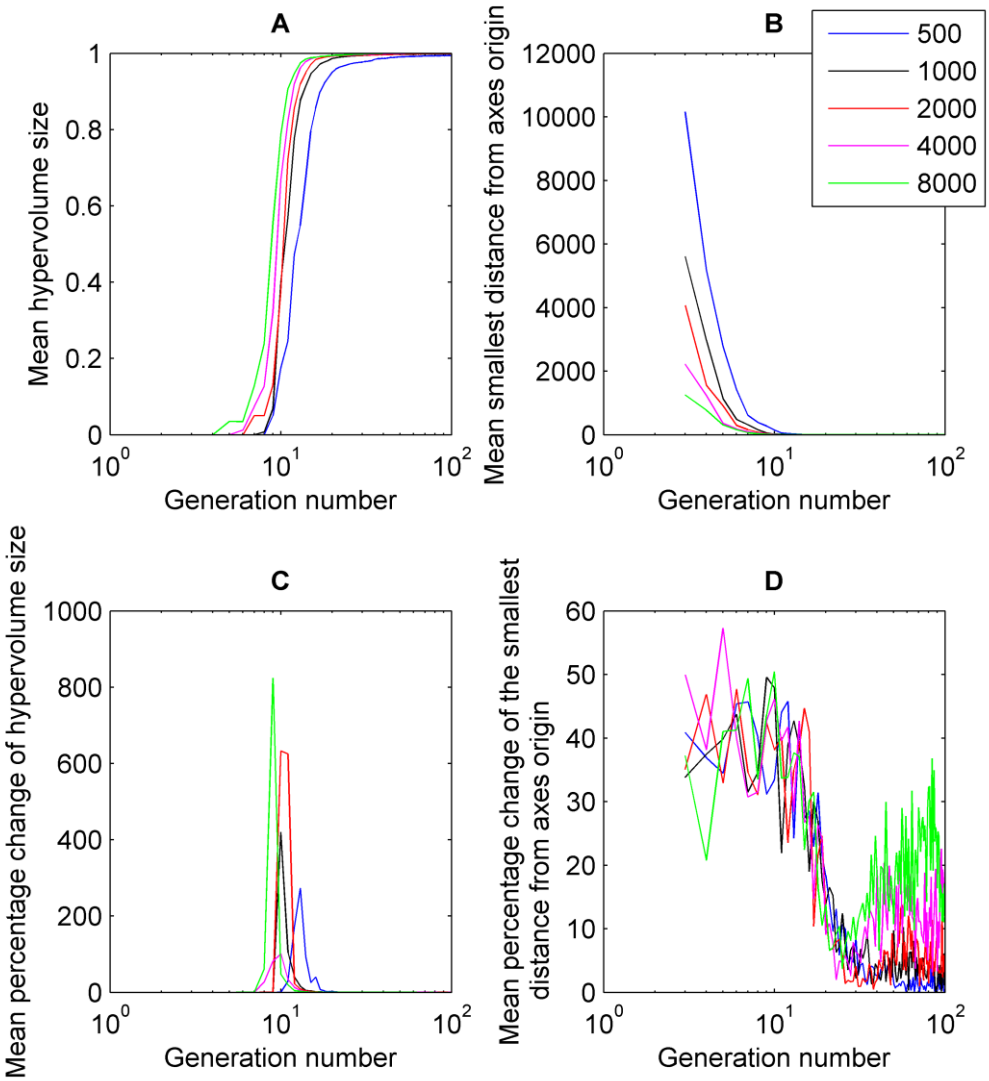

Figure 5.18: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target SSB. The mean were calculated from 16 runs of the NSGA-II for the experimental velocity profiles shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000).

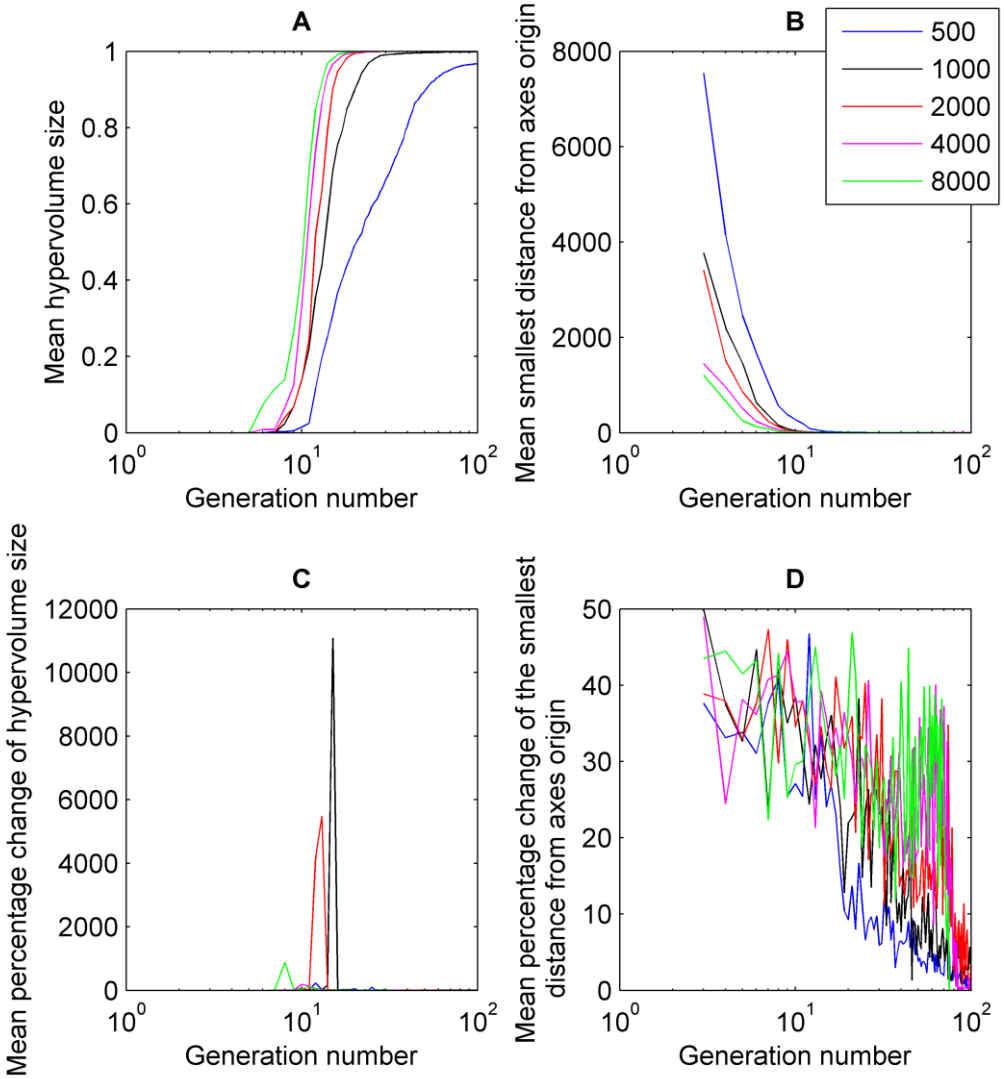

Figure 5.19: Mean values and mean change of the hypervolume size and smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target SSC. The mean were calculated from 16 runs of the NSGA-II for the experimental velocity profiles shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000).

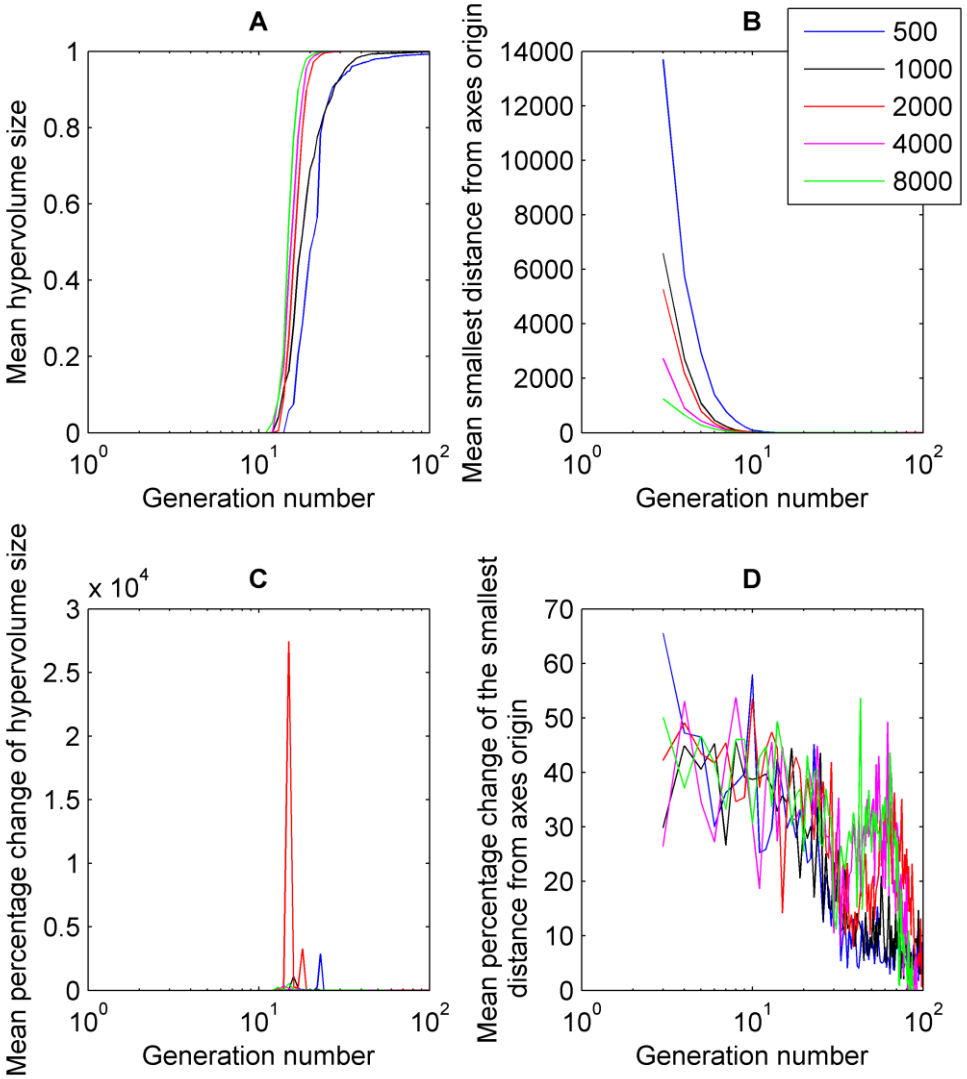

Figure 5.20: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target SSD. The mean were calculated from 16 runs of the NSGA-II for the experimental velocity profiles shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000).

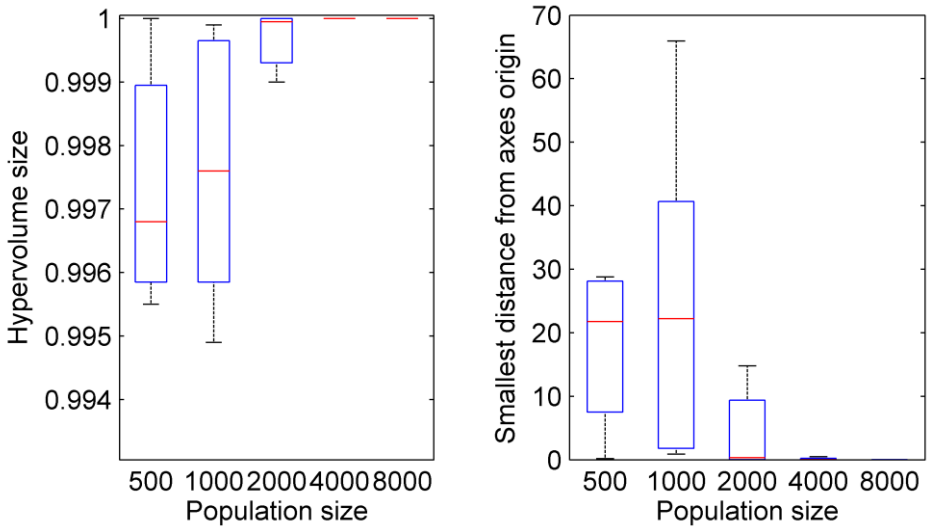

Figure 5.21: The box plots of the left plot shows the change of the hypervolume size versus population size. The box plots of, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic velocity profiles targets SSA shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25^{th} and 75^{th} percentiles. The whiskers extend to the interquartile range.

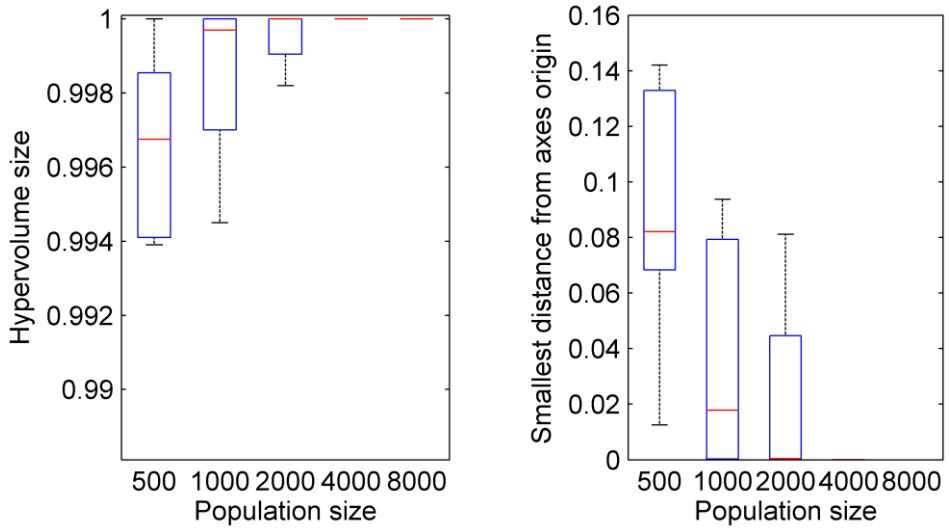

Figure 5.22: the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic velocity profiles targets SSB shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

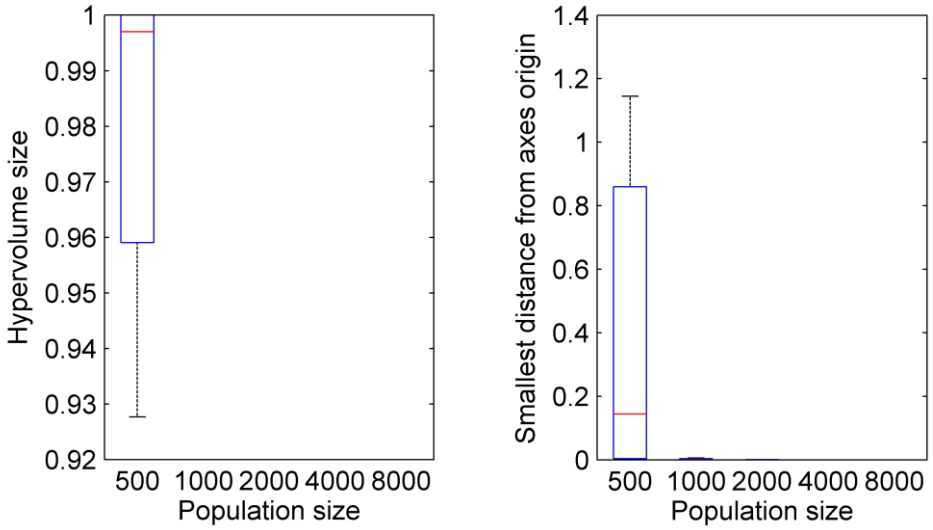

Figure 5.23: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic velocity profiles targets SSC shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

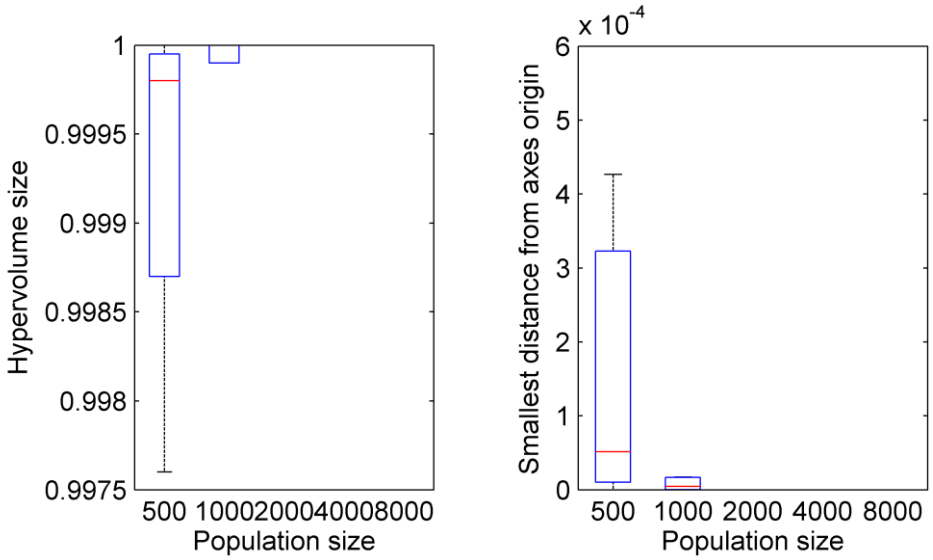

Figure 5.24: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the synthetic velocity profiles targets SSD shown in Figure 5.16 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

Table 5.6: Variation with population size of the mean and coefficient of variation (in parenthesis) of the objectives for each synthetic saccadic velocity profile target (SSA, SSB, SSC and SSD) shown in Figure 5.16. The objectives measure the fitness for saccades of amplitude 5, 10 and 20 degrees (refer to section 5.1.2). 16 NSGA-II runs were carried out for each population size (500, 1000, 2000, 4000 and 8000). The best solution from each NSGA-II run was selected using the method described in section 5.2.2.

Target ID	Objective	500	1000	2000	4000	8000
SEA	5 deg	172.91 (1.693)	1926.57 (2.698)	12.98 (2.385)	0.39 (3.868)	0.000232 (3.998)
	10 deg	5042.55 (3.0357)	34734.65 (3.407)	71.84 (2.079)	9.19 (3.924)	0.003609 (3.998)
	20 deg	708564.91 (3.394)	5491173 (3.541)	4011.04 (2.108)	37.16 (2.482)	0.979408 (3.997)
SEB	5 deg	0.0235 (2.8797)	0.000679 (1.271)	0.00012 (2.226)	9.243×10^{-8} (3.983)	1.4397×10^{-12} (3.972)
	10 deg	10.661 (2.311)	9.8115 (1.898)	0.01405 (3.5759)	7.7983×10^{-7} (3.9189)	7.5745×10^{-10} (3.999)
	20 deg	6858.56 (2.424)	6650.729 (1.843)	18.4023 (3.691)	1.3241×10^{-5} (3.332)	1.4898×10^{-7} (3.999)

Continued on the next page

Target ID	Objective	500	1000	2000	4000	8000
SEC	5 deg	0.078 (2.5722)	6.334 $x10^{-8}$ (3.3355)	2.166 $x10^{-10}$ (3.2918)	6.65 $x10^{-10}$ (3.222)	0 (-)
	10 deg	120.24 (2.751)	1.157 $x10^{-6}$ (3.456)	5.376 $x10^{-10}$ (3.523)	4.159 $x10^{-10}$ (3.062)	0 (-)
	20 deg	71467.93 (2.7542)	0.0011 (3.8624)	2.605 $x10^{-8}$ (2.6247)	9.557 $x10^{-9}$ (2.7477)	0 (-)
SED	5 deg	3.8167 $x10^{-5}$ (3.98)	5.3009 $x10^{-10}$ (3.695)	1.634 $x10^{-12}$ (3.999)	0 (-)	0 (0)
	10 deg	0.00239 (3.969)	3.854 $x10^{-9}$ (3.612)	1.93 $x10^{-11}$ (3.999)	0 (-)	0 (0)
	20 deg	1.8029 (3.972)	5.0799 $x10^{-8}$ (2.691)	1.7111 $x10^{-11}$ (3.999)	0 (-)	0 (0)
	10 deg	0.00239 (3.969)	3.854 $x10^{-9}$ (3.612)	1.93 $x10^{-11}$ (3.999)	0 (-)	0 (0)
	20 deg	1.8029 (3.972)	5.0799 $x10^{-8}$ (2.691)	1.7111 $x10^{-11}$ (3.999)	0 (-)	0 (0)

Table 5.7: Target parameter values, found mean and coefficient of variation (CV) of the optimised parameter values for each target waveforms (SSA, SSB, SSC and SSD). The target waveforms are shown in Figure 5.16 and were generated using the Broomhead *et al.* [1] model with the parameter values shown in Table 5.5. The found mean and CV were calculated by the solutions selected using the method described in section 5.2.2 from 16 NSGA-II runs and using population size of 8000.

Target ID	Value type	α	β	ε	γ	α'	β'
SSA	target	650	15	0.0035	0.05	380	1.5
	mean	643.77	15.341	0.0035	0.053	380.02	1.5
	CV	0.0753	0.07045	0.00222	0.14889	0.00036	0.00128
SSB	target	10	0.5	0.0065	0.5	480	9
	mean	11.5984	0.5844	0.00649	0.5031	479.99	8.99
	CV	0.428	0.449	0.00003	0.0161	0.00002	0.00002
SSC	target	100	50	0.009	4	380	9
	mean	96.4407	47.75	0.00899	3.978	379.99	8.999
	CV	0.10548	0.1313	0.000006	0.015	0.00001	0.00001
SSD	target	15	5	0.005	5	600	10
	mean	14.99	4.99	0.005	5	600	10
	CV	0.00002	0.000005	0.000005	0.000006	0.000004	0.00002

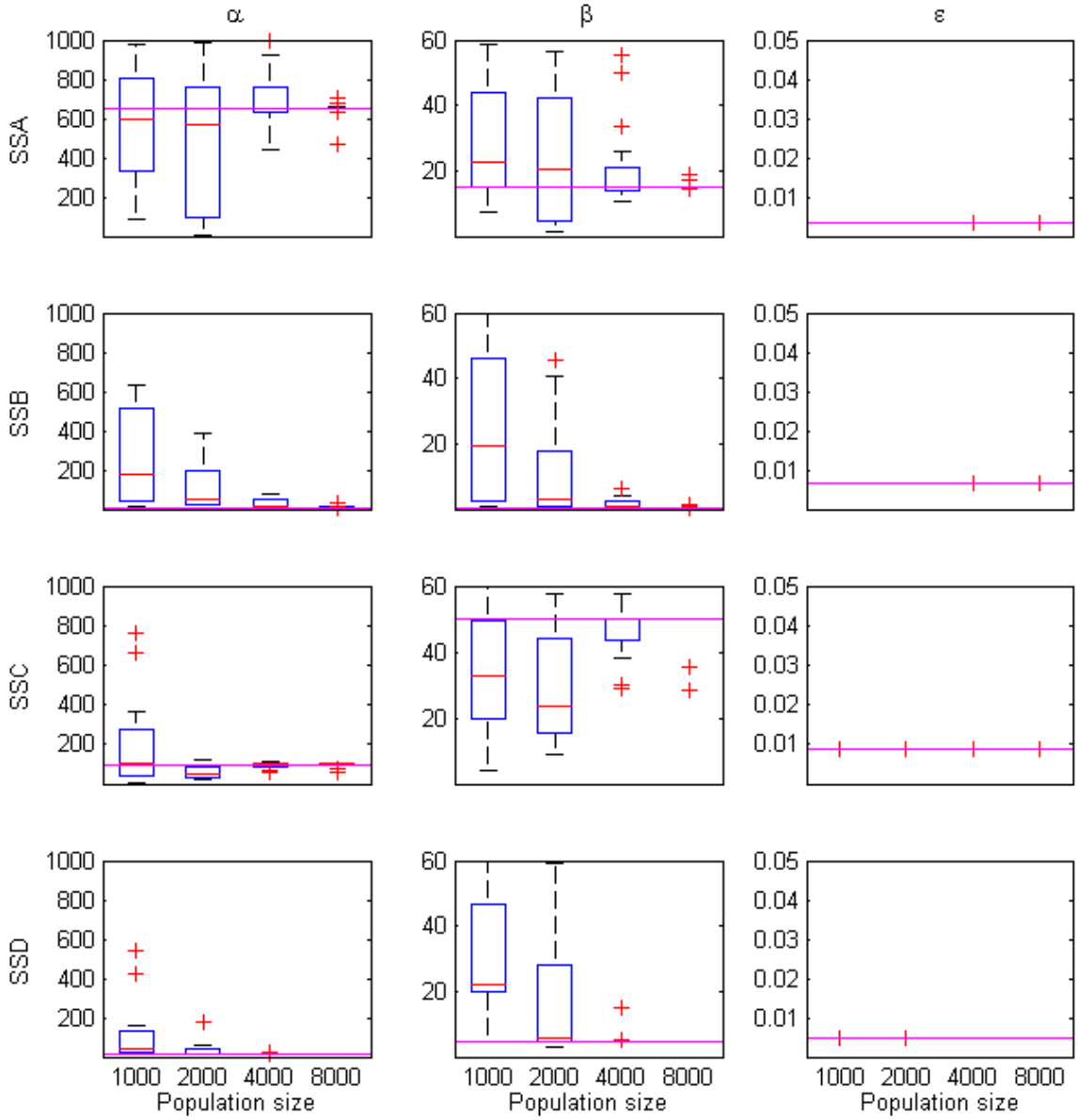

Figure 5.25: Optimised Broomhead *et al.* [1] model parameter values for each synthetic saccadic velocity profile target versus population size (target waveforms are shown in Figure 5.16). The first line of plots shows the optimised values of α, β and ε for the synthetic saccadic velocity profile target set SSA, whereas those for SSB, SSC and SSD are shown by the second, third and fourth lines of plots, respectively. In each plot, the magenta line represents the parameter value used to generate the target waveform. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

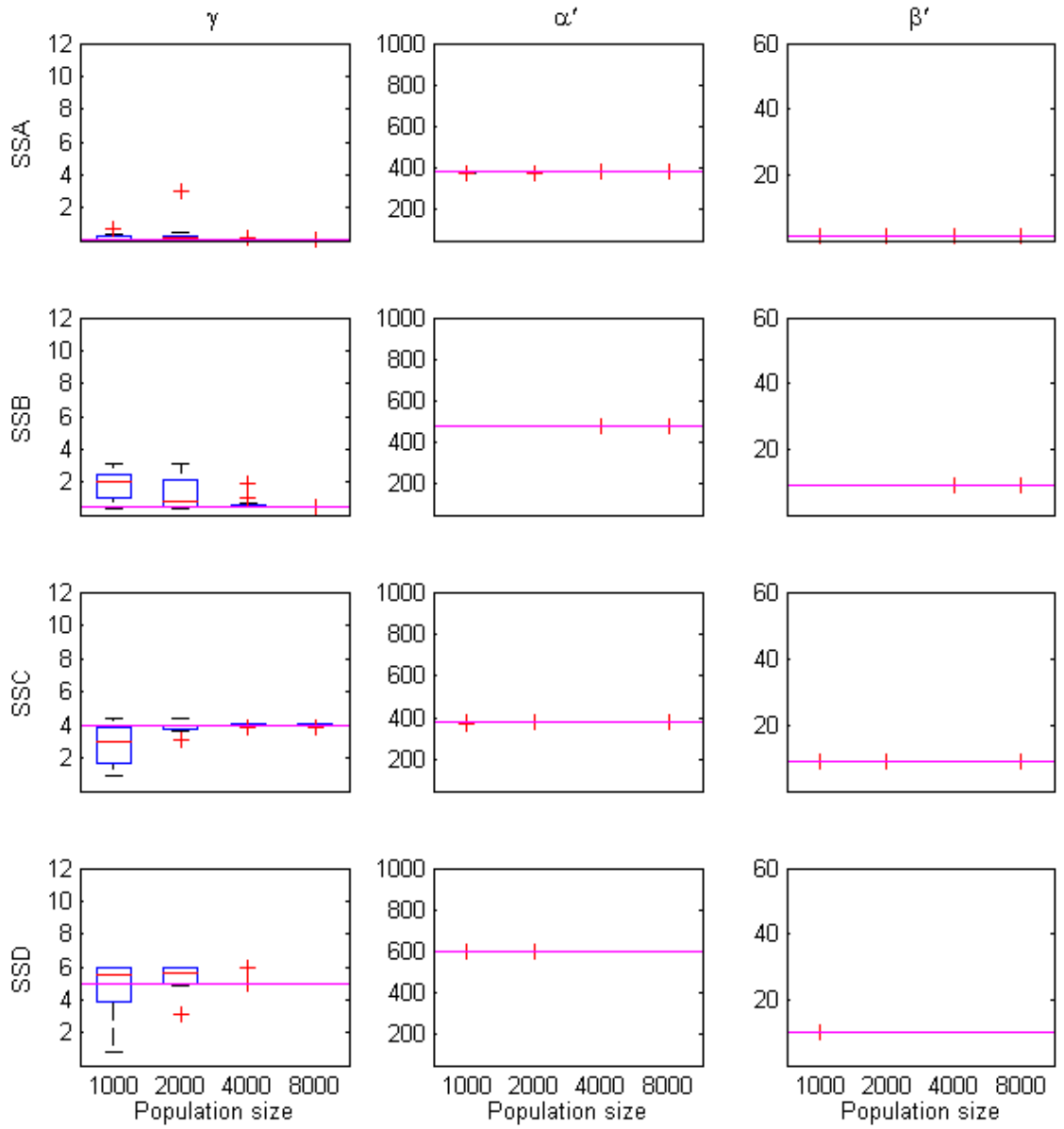

Figure 5.26: Optimised Broomhead *et al.* [1] model parameter values for each synthetic saccadic velocity profile target versus population size (target waveforms are shown in Figure 5.16). The first line of plots shows the optimised values of γ, α' and β' for the synthetic saccadic velocity profile target SSA, whereas those for SSB, SSC and SSD are shown by the second, third and fourth line of plots, respectively. In each plot, the magenta line represents the parameter value used to generate the target waveform. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

5.3.4 Experimental nystagmus targets - fitting results

The experimental nystagmus time series recordings from which we extracted the unstable periodic orbits (UPOs) to use them as targets for fitting the Broomhead *et al.* [1] model, were acquired from the database of [96] (see section 4.2.1). Figure 5.27 shows segments of the used four experimental time series and the extracted UPOs. Each experimental time series was recorded from a different participant. These particular time series parts were chosen as their UPOs shapes were more likely to be able to be generated by the Broomhead *et al.* [1] model. To extract the UPOs, we applied the method described in section 4.2.2. The sampling frequency of the time series recordings was 250 Hz. As discussed in Chapter 4, the best results were found when the sampling frequency was 2500Hz. Therefore, we up-sampled the extracted UPOs to 2500 Hz by applying cubic interpolation. Figure 5.28 shows extracted UPOs that we modified so they begin at the fast phase of the nystagmus waveform, and which were used in the parameter estimation method (see section 4.2.2).

The application of the hypervolume indicator described in section 4.3.1 shows that - as was observed for synthetic nystagmus targets - a population size of 4000 is sufficient for the NSGA-II to return solutions with the best convergence compared to all the population sizes that we examined. Figures 5.29-5.32 show the hypervolume size for different population sizes against generation number for the four waveform targets, whereas Figures 5.33-5.36 show the hypervolume size of the last generation for the different population sizes for the four waveform targets. Furthermore, the variation of the mean and CV of the shape fitness with population size shown in Table 5.9 also indicates that sufficient accuracy and precision can be obtained with this population size.

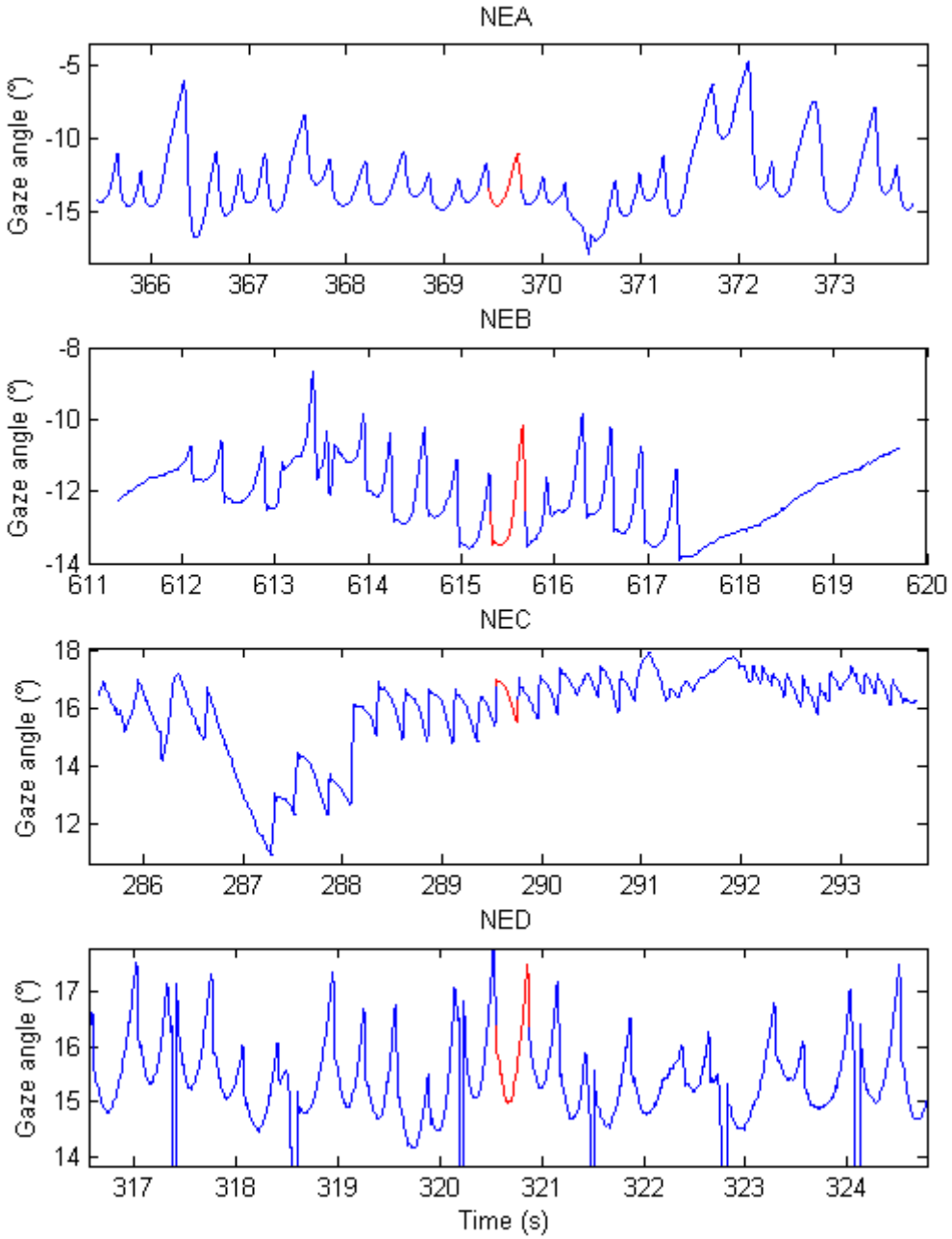

Figure 5.27: Parts of the four experimental time series (shown in blue) used to extract the UPOs for target fitting. From top to bottom, the nystagmus waveform types are: NEA - asymmetric pseudo-cycloid; NEB - jerk with extended foveation; NEC - jerk; NED - asymmetric pseudo-cycloid. The UPOs (shown in red) were extracted using the method described in section 4.2.2. Before being used as targets for model fitting, the UPOs were modified so they begin at the start of the nystagmus fast phase (see Figure 5.14). On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). In the bottom plot, the vertical lines show the locations of eye blinks.

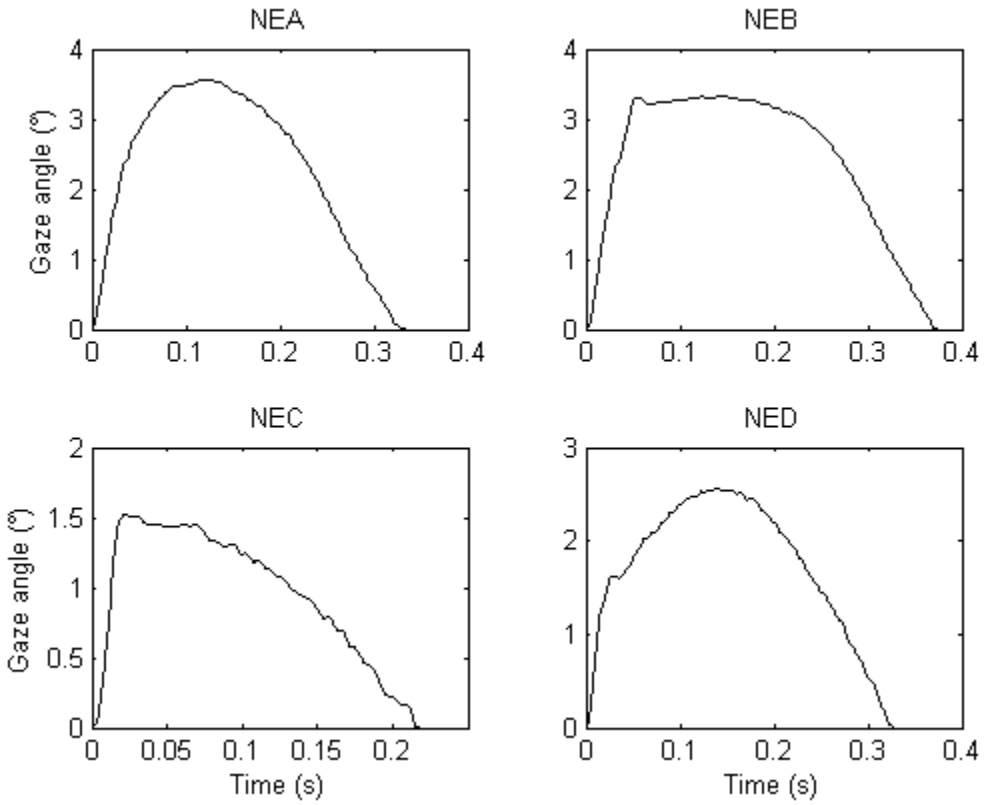

Figure 5.28: UPOs extracted from the experimental time series show in Figure 5.27 and modified so they begin at the start of the nystagmus fast phase. These UPOs were used as waveform targets for fitting the Broomhead *et al.* [1] model. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). The nystagmus waveform types are: NEA - asymmetric pseudo-cycloid; NEB - jerk with extended foveation; NEC - jerk; NED - asymmetric pseudo-cycloid.

The box plots in Figures 5.37 and 5.38 show the optimised deterministic parameter values for different population size. These provide further evidence that the NSGA-II convergences sufficiently with population size 4000. The bar plot in Figure 5.39 shows the coefficient of variation of each parameter for each waveform target with population size 4000. The high coefficient of variation of the ε values for the NEA and NED waveforms may perhaps be due to the simulated oscillation type (asymmetric pseudo-cycloid). Based on these results, it seems that these particular oscillations are not constrained by specific ε values. The coefficients of variation for the other parameters mirror those observed for synthetic nystagmus waveforms (see Figure 5.7). In particular, the fitness

function is sensitive to the parameters α and β that control the saccadic off-response (the saccadic braking signal – refer to section 2.2). These results are therefore consistent with the hypothesis of van Gisbergen *et al.* [25], Broomhead *et al.* [1] and Akman *et al.* [62] stating that the dynamics of the braking signal is the cause of the instability. Figure 5.40 compares the waveforms generated by the NSGA-II solutions with population size of 4000 for the target waveforms. The optimised waveforms for NEA, NEB and NED are very close to the target waveforms (Figure 5.40). In the case of the NEC waveform, a poorer fit to the target waveform is obtained, as the Broomhead *et al.* [1] model cannot reproduce the very rapid fast phase.

In the case of the NEC oscillation, we increased the upper constraint of parameter γ to 128 because the parameter values using constraint value 12 were found to be very close to 12. Parameter values found very close to a constraint value could mean that the optimal solution is beyond the constraint value. We performed more exploratory runs with even higher constraint γ value than 128 and found that the found γ values had a mean value of 90. However, Broomhead *et al.* [1] could not provide a better fit to the NEC nystagmus oscillation that the one shown in this work.

Moreover, we see that the found values for the parameter α' are concentrated close to the constraint value of 1000. We performed exploratory runs with higher α' values but we did not acquired sufficiently better fits to the examined data. This is interesting as this constraint was set to that specific value based on the experimental measurements of van Gisbergen *et al.* [25]. This could mean that the Broomhead *et al.* [1] model has to be modified in order to account for these results or that in the real system the burst neurons are set to fire maximally based on these high α' values (close to 1000 spikes/s).

Table 5.8: Table with the found mean and coefficient of variation (CV) of the optimised parameter values for each target waveforms (NEA, NEB, NEC and NED). The target waveforms are shown in Figure 5.28. The found mean and CV were calculated by the solutions selected using the method described in section 5.2.1 from 16 NSGA-II runs and with population size 4000.

Target ID	Value type	α	β	ε	γ	α'	β'
NEA	mean	72.624	1.46183	0.00051	2.83363	895.749	19.0602
	CV	0.007	0.0105	0.835	0.13846	0.1207	0.1273
NEB	mean	80.0815	1.28473	0.00305	2.70308	932.272	15.8947
	CV	0.02324	0.05330	0.18582	0.23501	0.10745	0.15980
NEC	mean	53.891	0.4137	0.00131	95.31	826.149	6.8257
	CV	0.0308	0.03	0.211	0.2062	0.16314	0.1737
NED	mean	55.2688	1.66647	0.00009	0.52871	896.392	28.722
	CV	0.01135	0.03285	1.05951	0.14658	0.10156	0.11368

Table 5.9: Variation with population size of the mean and coefficient of variation (in parenthesis) of the nystagmus shape difference objective for each experimental waveform target (NEA, NEB, NEC and NED) shown in Figure 5.28. 16 NSGA-II runs were carried out for each population size (500, 1000, 2000, 4000 and 8000). The best solution from each NSGA-II run was selected using the method described in section 5.2.1.

Target ID	500	1000	2000	4000	8000
NEA	3480.644 (0.3154)	2726.983 (0.1491)	2449.465 (0.0699)	2316.418 (0.0402)	2270.522 (0.0261)
NEB	22922.767 (2.4636)	5920.202 (0.4963)	4757.846 (0.0901)	4498.202 (0.0730)	4400.1543 (0.0255)
NEC	6992.538 (0.0532)	6401.237 (0.0514)	6321.503 (0.0309)	6235.154 (0.019)	6189.581 (0.0133)
NED	12130.327 (0.5662)	7737.276 (0.1948)	7218.755 (0.1203)	6703.085 (0.0121)	6705.796 (0.0136)

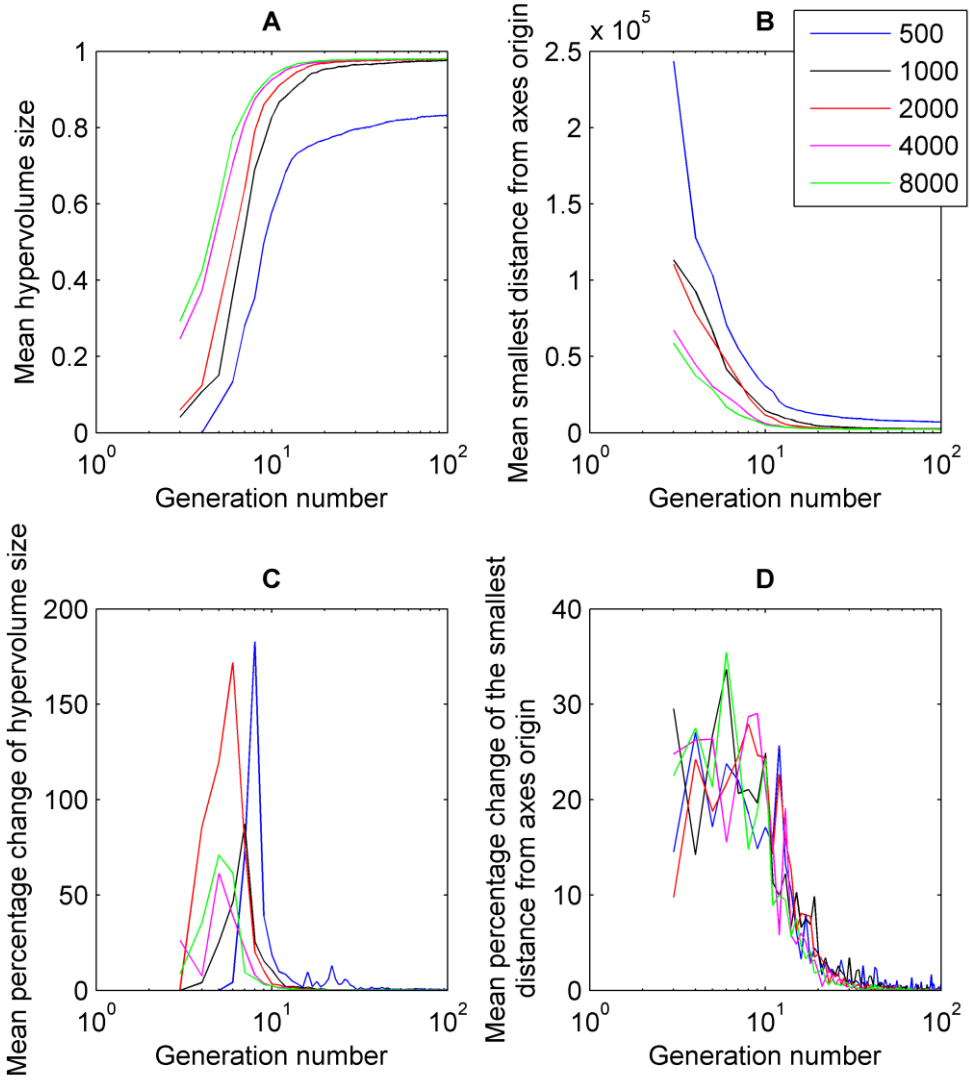

Figure 5.29: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target NEA. The mean were calculated from 16 runs of the NSGA-II for the synthetic waveform target NEA shown in Figure 5.28 and for different population sizes (500, 1000, 2000, 4000 and 8000).

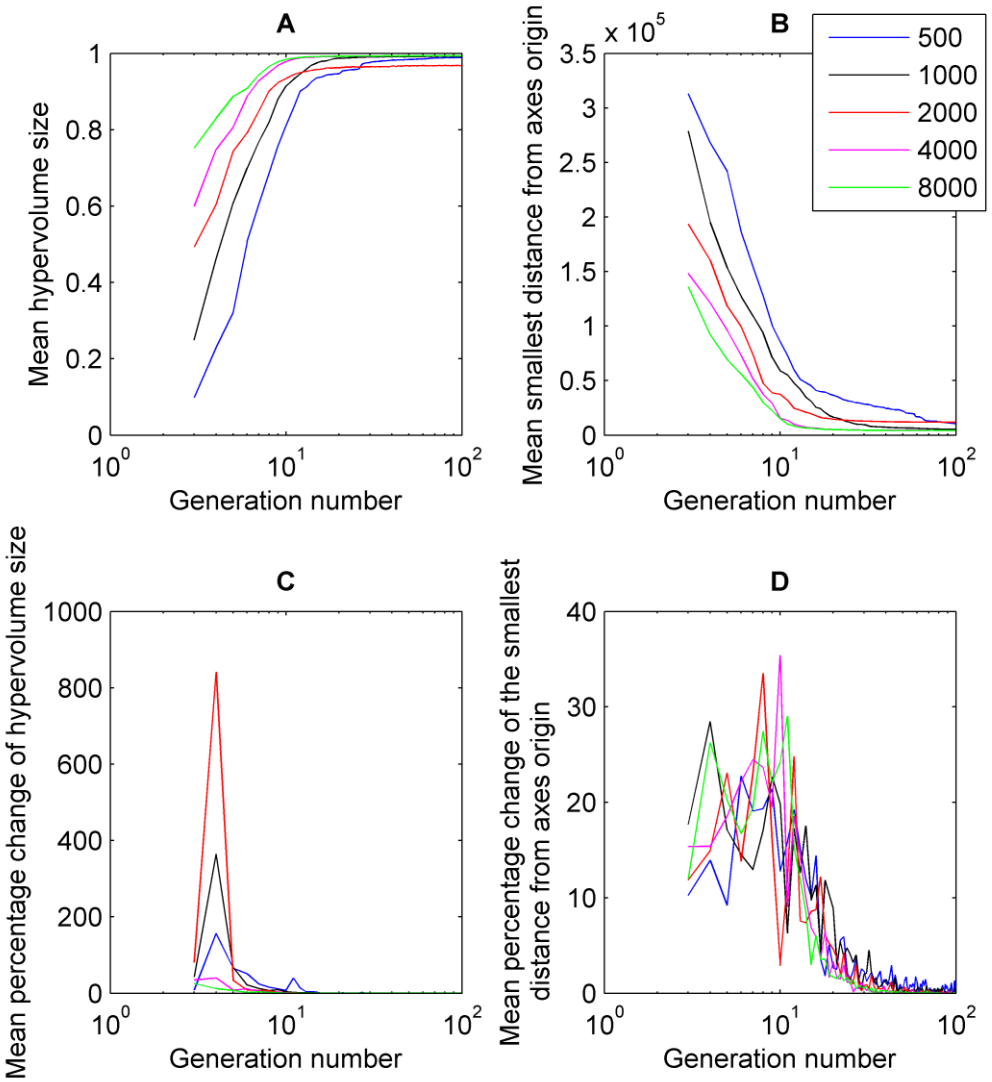

Figure 5.30: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target NEB. The mean were calculated from 16 runs of the NSGA-II for the synthetic waveform target NEB shown in Figure 5.28 and for different population sizes (500, 1000, 2000, 4000 and 8000).

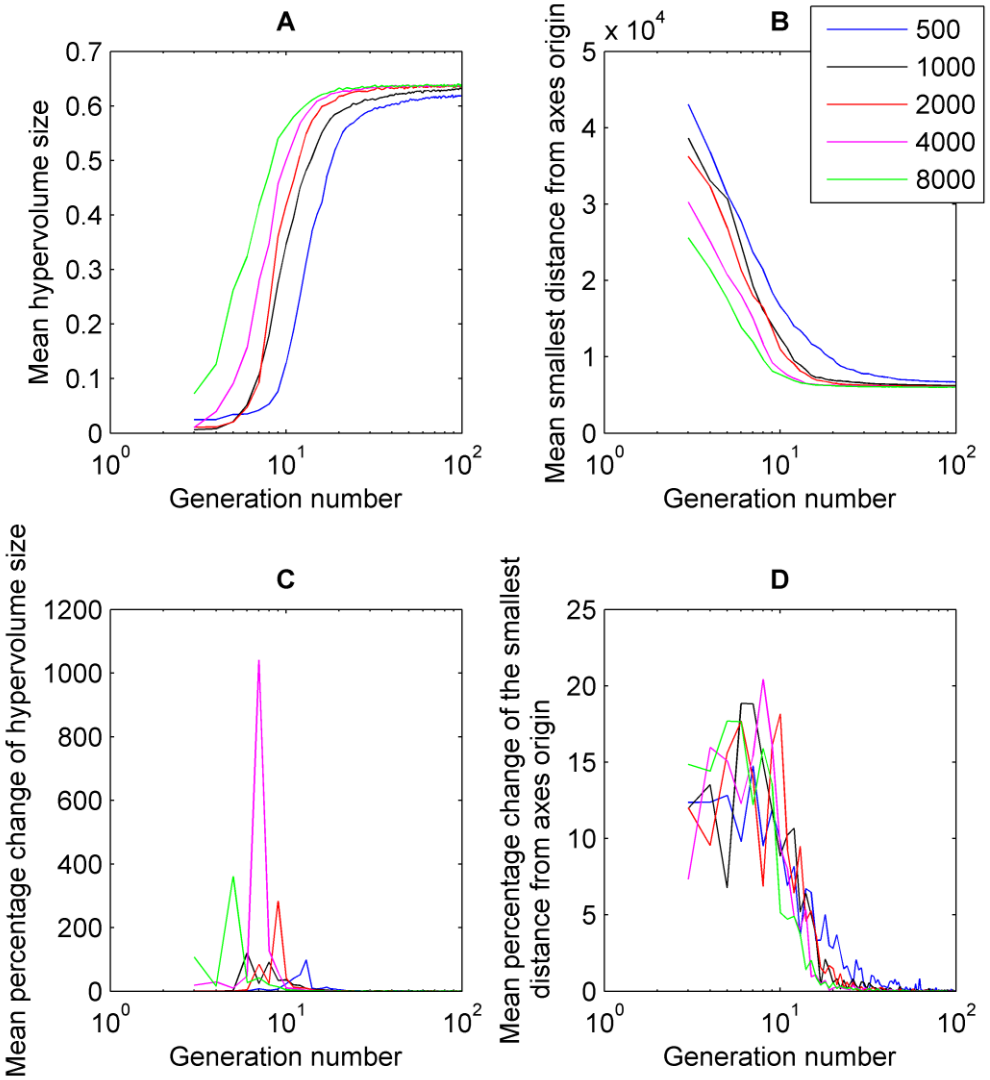

Figure 5.31: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target NEC. The mean were calculated from 16 runs of the NSGA-II for the synthetic waveform target NEC shown in Figure 5.28 and for different population sizes (500, 1000, 2000, 4000 and 8000).

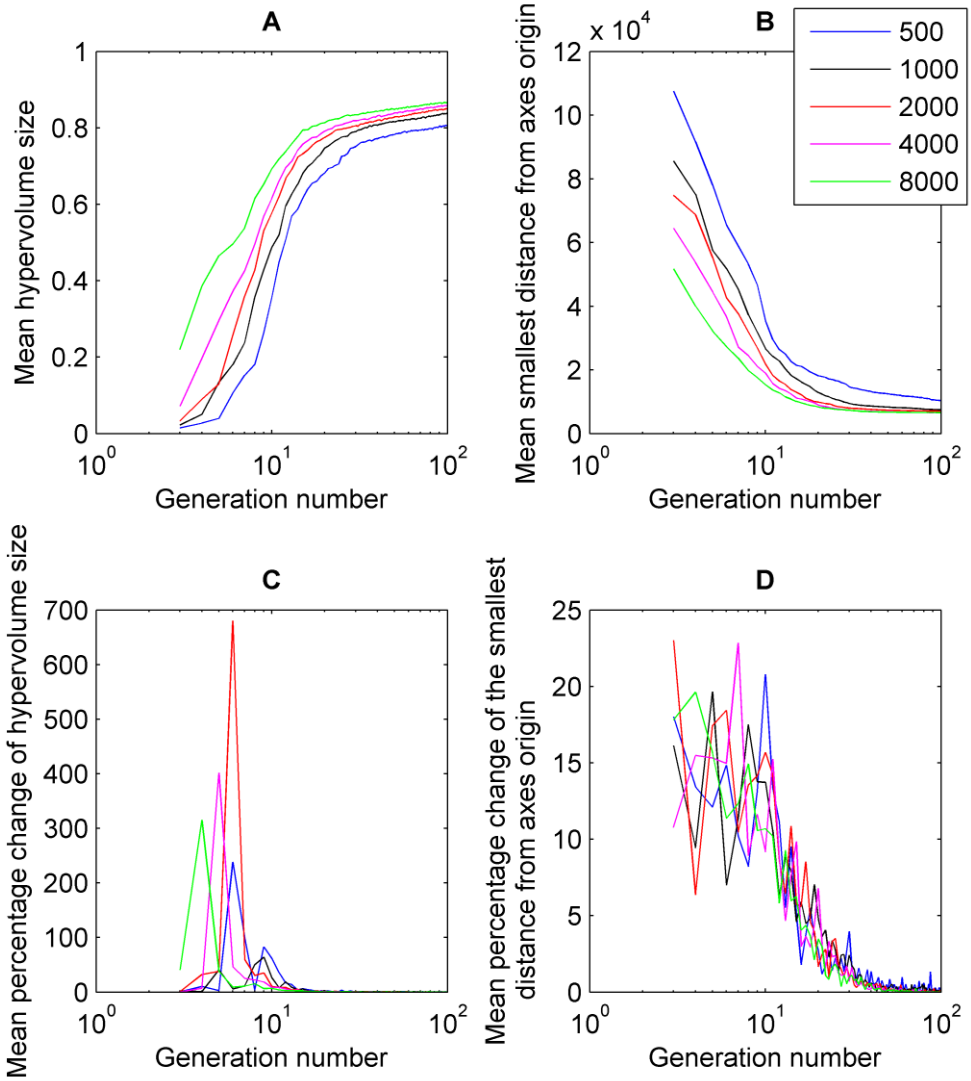

Figure 5.32: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population size (in logarithmic scale) for the target NED. The mean were calculated from 16 runs of the NSGA-II for the synthetic waveform target NED shown in Figure 5.28 and for different population sizes (500, 1000, 2000, 4000 and 8000).

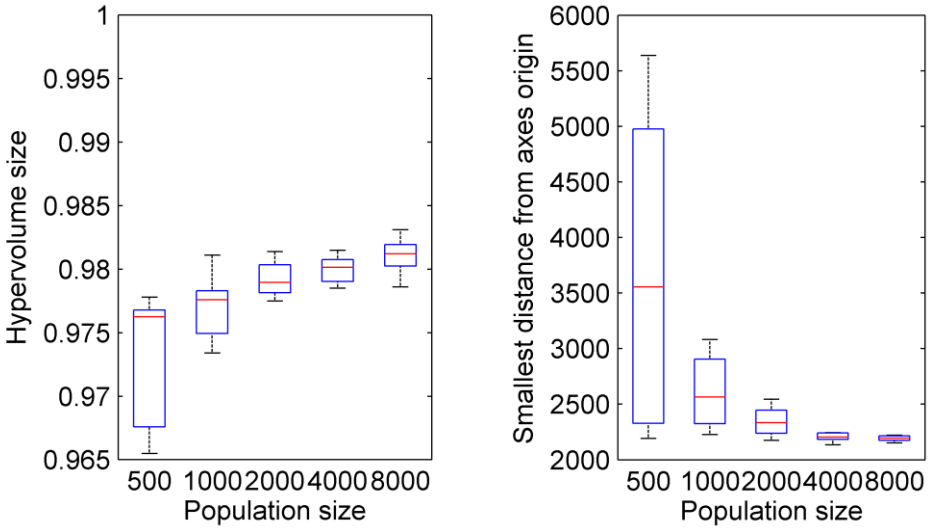

Figure 5.33: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the experimental waveform target NEA shown in Figure 5.28 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25^{th} and 75^{th} percentiles. The whiskers extend to the interquartile range.

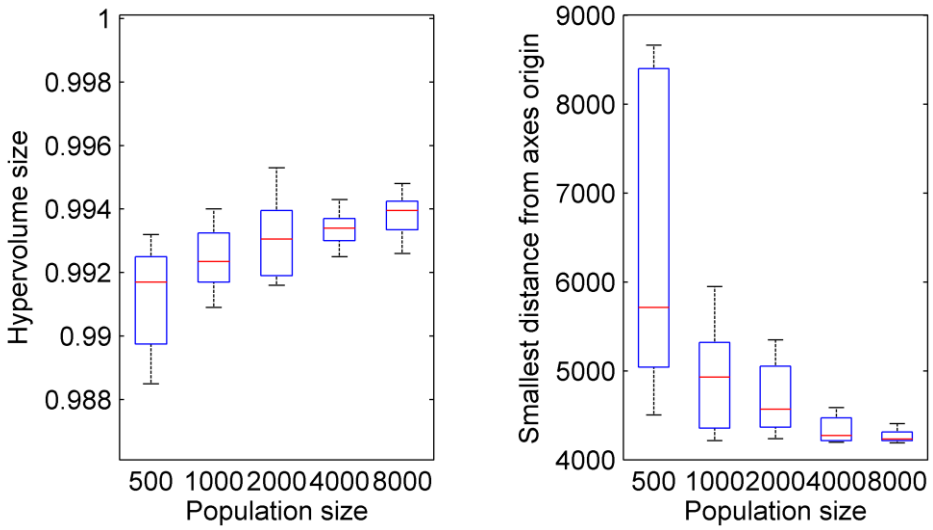

Figure 5.34: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the experimental waveform target NEB shown in Figure 5.28 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

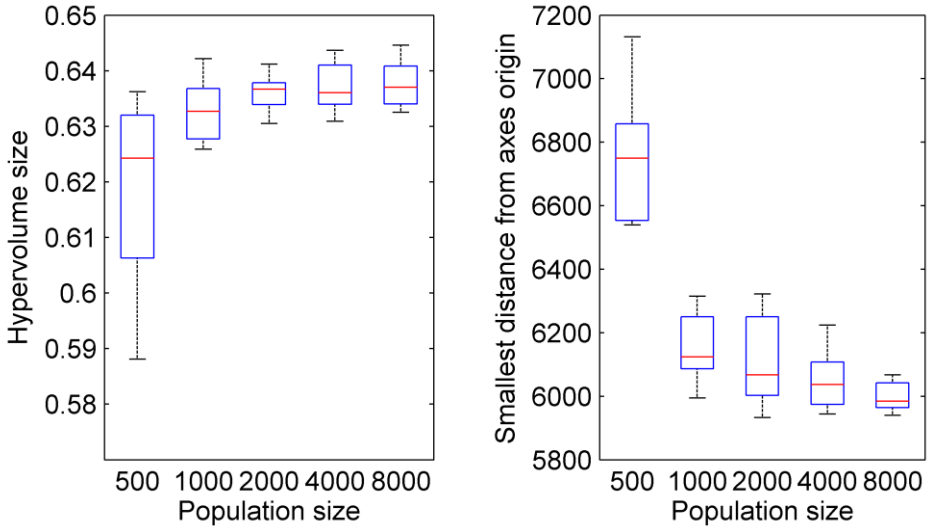

Figure 5.35: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the experimental waveform target NEC shown in Figure 5.28 for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

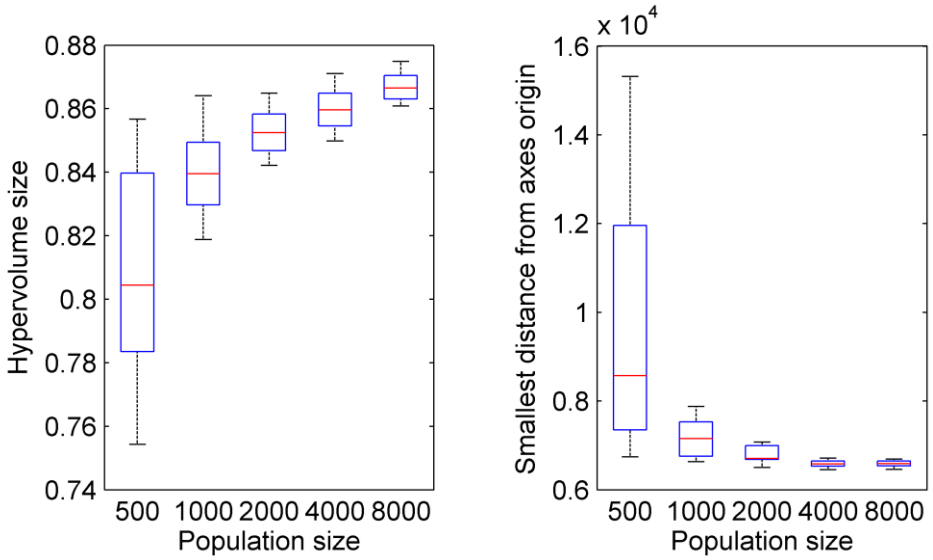

Figure 5.36: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot shows the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the experimental waveform target NED shown in Figure 5.28 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

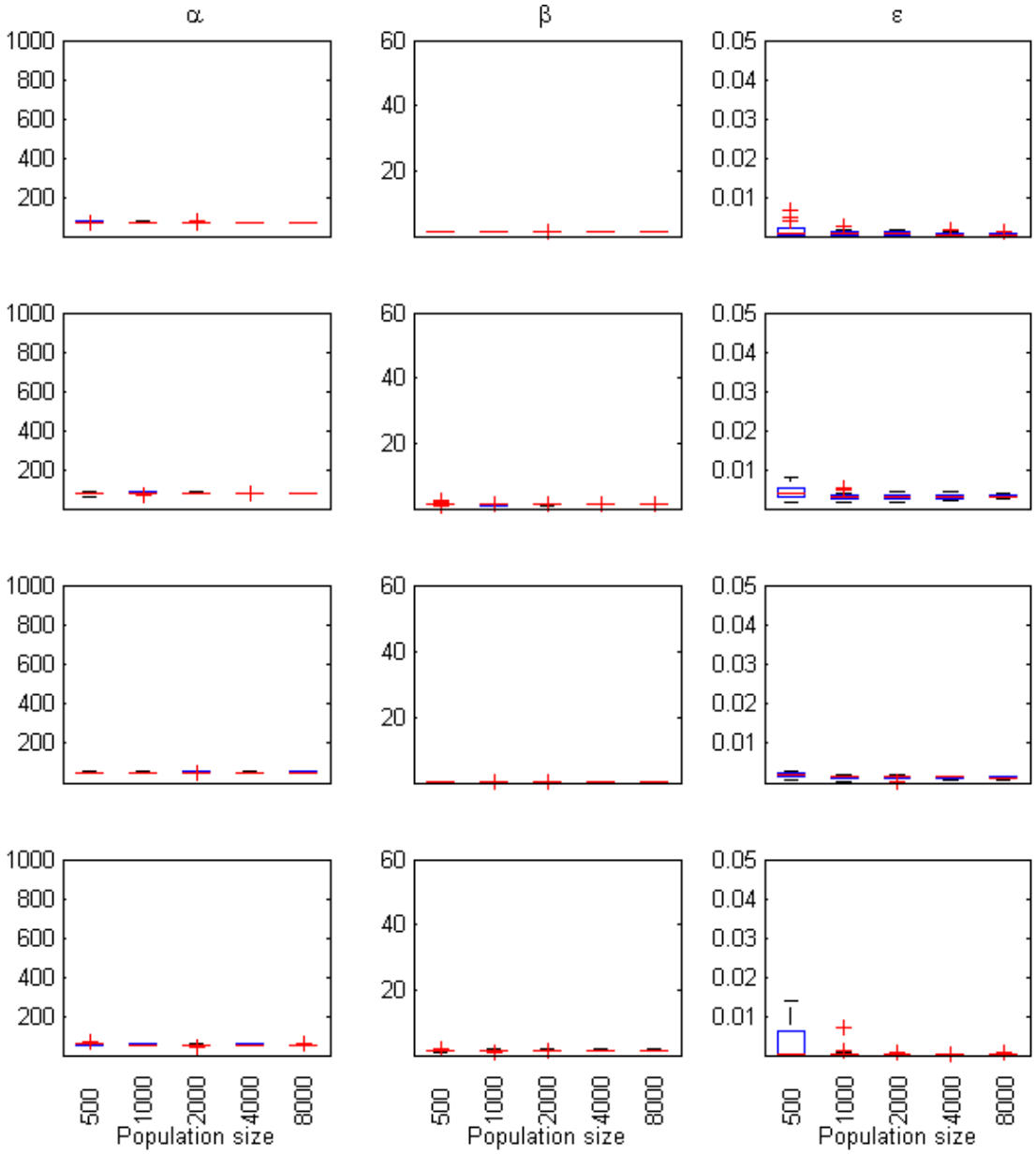

Figure 5.37: Optimised Broomhead *et al.* [1] model parameter values for each experimental nystagmus waveform target versus population size (target waveforms are shown in Figure 5.28). The first line of plots shows the optimised values of α, β and ε for the experimental waveform target NEA, whereas those for NEB, NEC and NED are shown by the second, third and fourth lines of plots, respectively. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

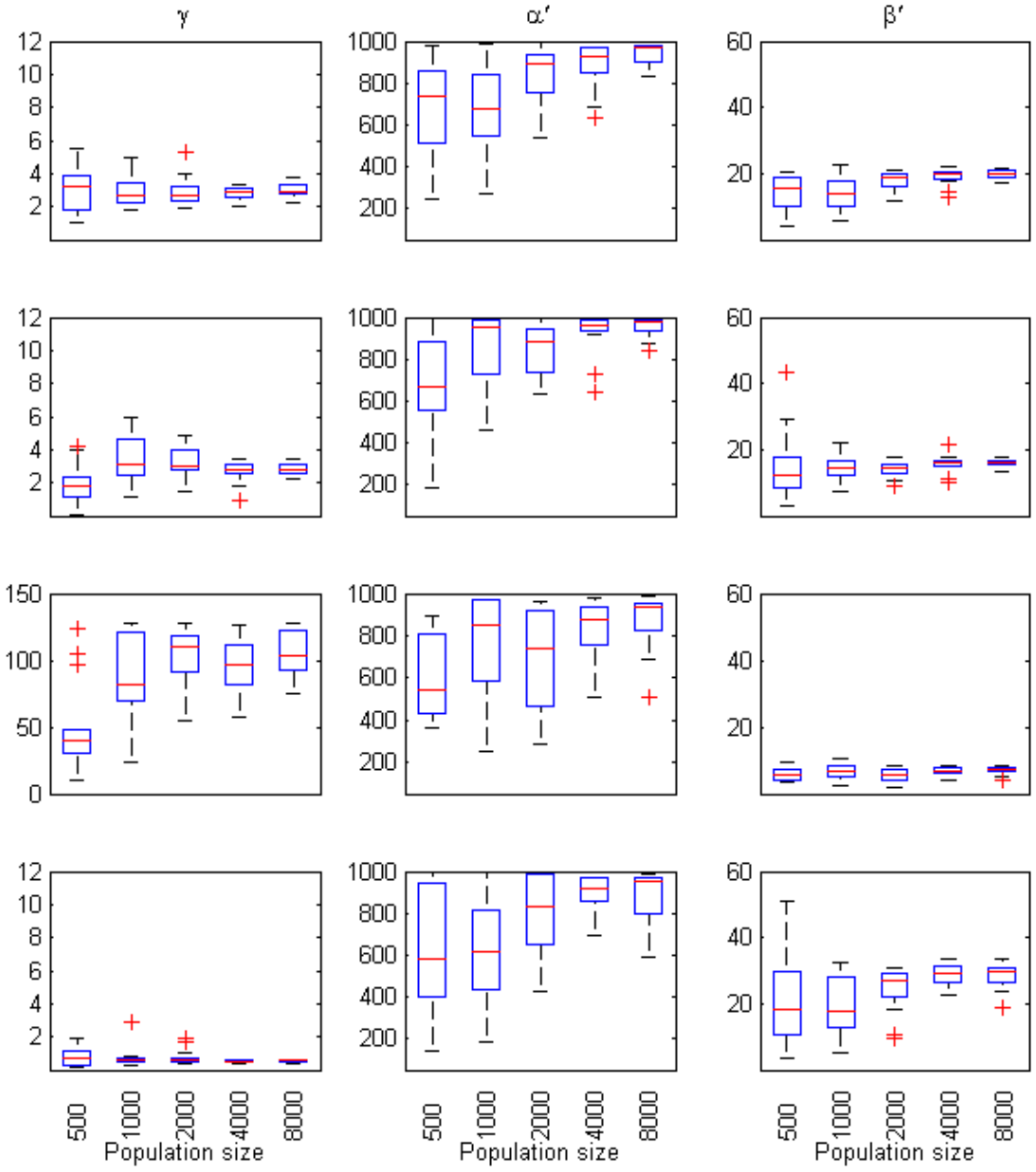

Figure 5.38: Optimised Broomhead *et al.* [1] model parameter values for each experimental nystagmus waveform target versus population size (target waveforms are shown in Figure 5.28). The first line of plots shows the optimised values of γ, α' and β' for the experimental waveform target NEA, whereas for NEB, NEC and NED are shown by the second, third and fourth lines of plots, respectively. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

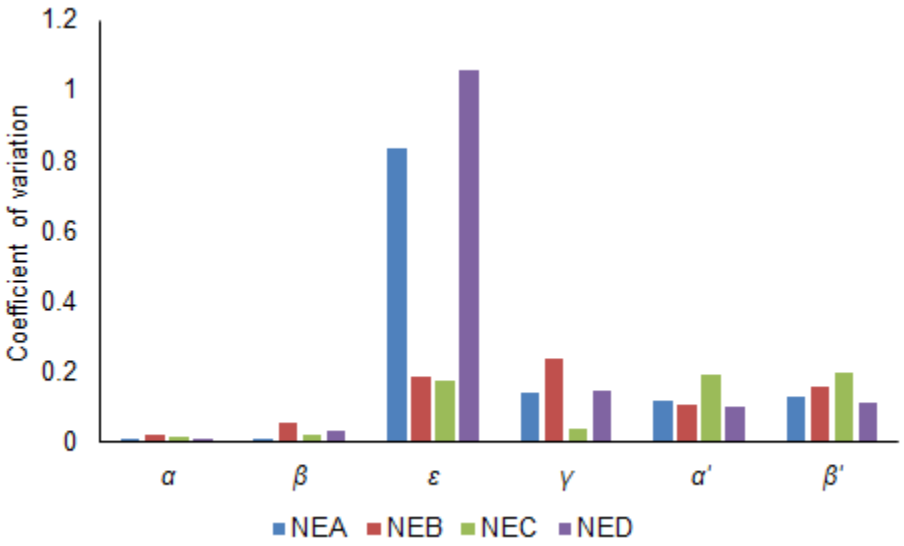

Figure 5.39: Coefficient of variation of the optimised Broomhead *et al.* [1] model parameters for each target experimental nystagmus waveform shown in Figure 5.28. The coefficient of variation values were calculated from 16 NSGA-II runs with a population size of 4000 in each case. The best solution from each NSGA-II run was selected using the method described in section 5.2.2**.**

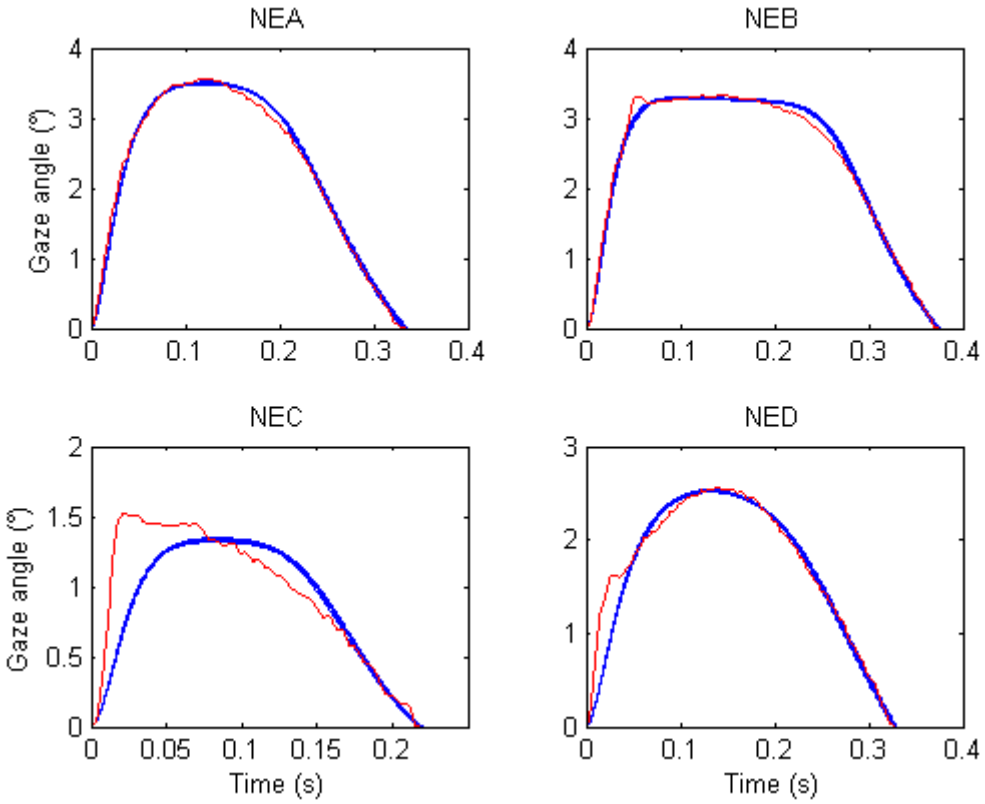

Figure 5.40: Waveforms generated from the solutions found by the NSGA-II for each experimental nystagmus waveform target. The target waveforms are plotted in red; the 16 optimised waveforms are plotted in blue and were chosen from each NSGA-II run with population size of 4000. They were selected using the method described in section 5.2.1. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). The experimental nystagmus waveform types are: NEA - asymmetric pseudo-cycloid; NEB - jerk with extended foveation; NEC - jerk; NED - asymmetric pseudo-cycloid.

5.3.5 Experimental saccadic velocity profiles – fitting results

In order to fit the Broomhead *et al.* [1] model to experimental saccadic velocity profiles, the used data set was based on the typical experimental horizontal saccadic velocity profiles taken from [23] (shown in Figure 4.4). As for the synthetic data fits, we used profiles for saccades of amplitude 5, 10 and 20 degs, which are in the interval of the most frequent saccade amplitudes (discussed in section 4.2.3). The mean hypervolume indicator (Figures 5.41 and 5.42), shows that a population size 8000 is sufficient to provide good convergence. The same

population size has been previously shown to provide sufficient convergence and accuracy when fitting synthetic targets (refer to section 5.3.3).

The heat map in Figure 5.43 shows the seriated solutions of the approximated *Pareto* front from a single representative run of the NSGA-II, with a population size of 8000. The seriation is based on the objective rank values. The seriated solutions show that there is a trade-off between the three objectives. Moreover, the model cannot find solutions satisfying all the objectives simultaneously, *i.e.* not all the objectives can be fitted perfectly with the same Broomhead *et al.* [1] model parameter values. Thus, the model cannot generate saccadic velocity profiles that exactly match the experimental data for all saccade amplitudes.

The boxplots in Figure 5.44 confirm that the method used for choosing the final solution from the estimated *Pareto* front has a strong influence on the obtained parameter values. In the first selection method, method A (described in section 5.2.2), the best trade-off between the three objectives is selected by choosing the solution which minimises the sum of the normalised objectives. Selection method B chooses the solution with the lowest value in objective 1 (*i.e.* the best fit to a 5 deg saccade), whereas selection methods C and D choose the solutions with the lowest value in objective 2 and 3, respectively (*i.e.* the best fit to a 10 and 20 deg saccade, respectively). The parameters ε, α' and β' are very sensitive to the selection method, exhibiting the greatest variation across the methods used. This further illustrates that the model cannot generate the dynamics of the experimental saccades when using one set of deterministic parameters alone. Figure 5.45 compares the optimised saccadic velocity profiles obtained with different selection methods against the corresponding experimental data. The results show that, in order to get better fits for each of the saccade amplitudes, different parameter values have to be used for each saccade amplitude.

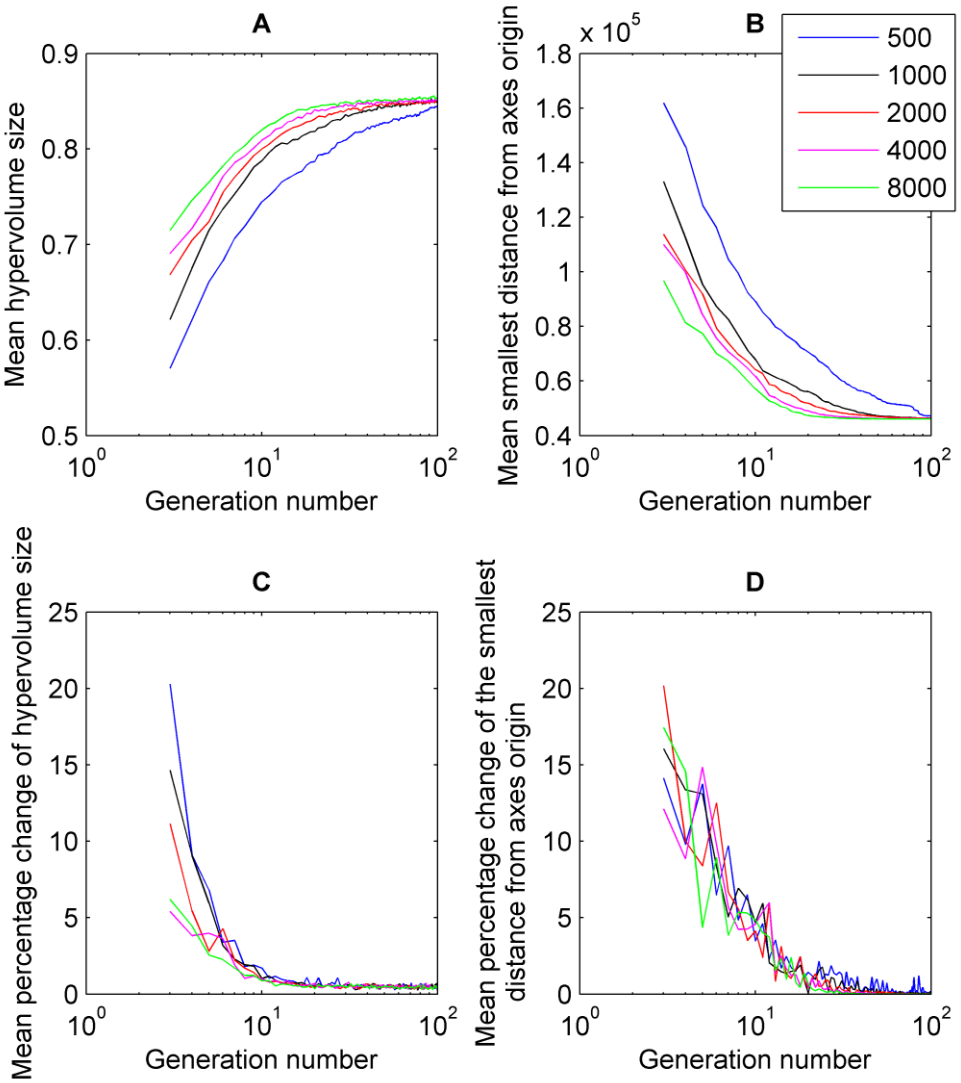

Figure 5.41: Mean values and mean change of the hypervolume size and the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the generation number for different NSGA-II population sizes (in logarithmic scale) for the experimental velocity profiles targets shown in Figure 5.45. The mean were calculated from 16 runs of the NSGA-II for the experimental velocity profiles shown in Figure 5.45 and for different population sizes (500, 1000, 2000, 4000 and 8000).

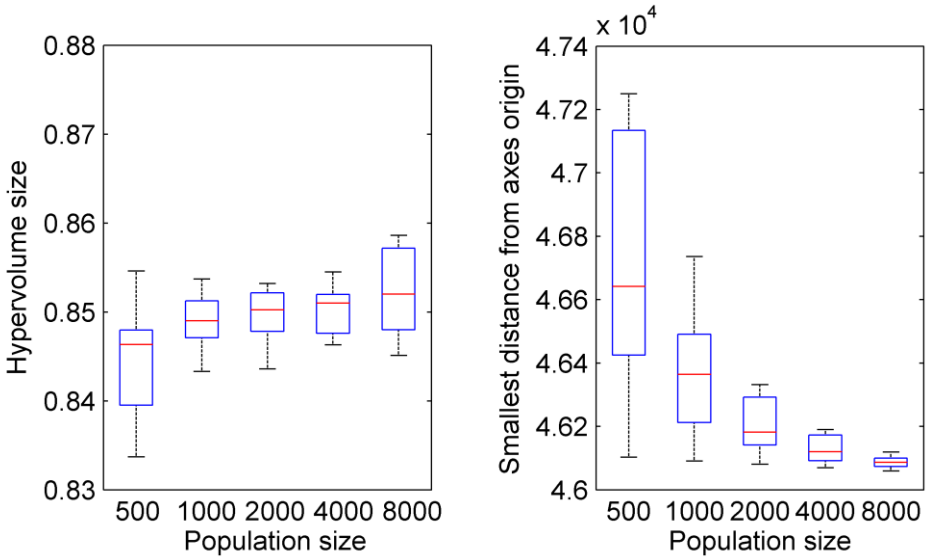

Figure 5.42: The box plots of the left plot shows the change of the hypervolume size versus population size, whereas the right plot show the change of the smallest distance of the individuals from the estimated *Pareto* front to the axes origin against the population size. The box plots were calculated from 16 runs of the NSGA-II for the experimental velocity profiles targets shown in Figure 5.45 and for different population sizes (500, 1000, 2000, 4000 and 8000). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range.

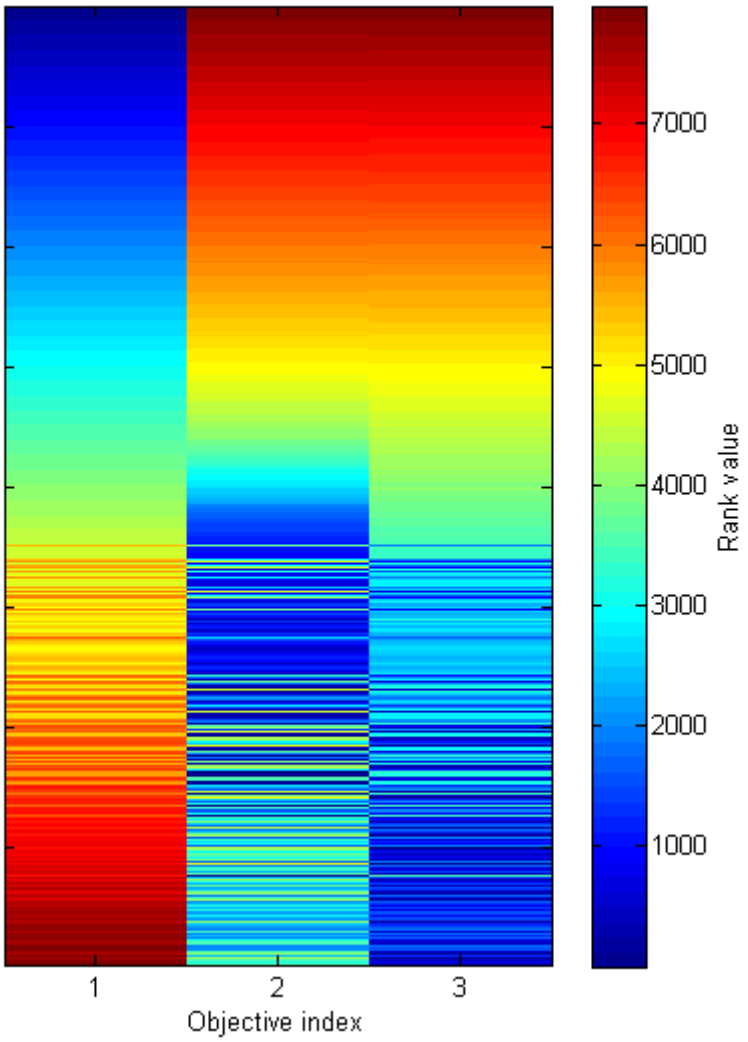

Figure 5.43: Objective trade-offs in fitting the Broomhead *et al.* [1] model to experimental saccadic velocity profiles. The heat map shows the seriated solutions of the estimated *Pareto* front from a single representative run of the NSGA-II, with a population size of 8000. The objective indices 1, 2 and 3 denote fits to saccades of 5, 10 and 20 degs, respectively. Each objective value is represented by its rank (refer to section 3.4 for rank calculation method). Red indicates high rank values, whereas blue represents low rank values.

Table 5.10: Table with the found mean and coefficient of variation (CV) of the optimised parameter values for each experimental saccadic velocity profiles (5, 10, and 20) for the four different final solution selection method. The experimental velocity profiles are shown in Figure 4.4. The found mean and CV were calculated by the solutions selected using the method described in section 5.2.2 from 16 NSGA-II runs with population size 8000.

Selection method	Value type	α	β	ε	γ	α'	β'
A	mean	14.47	40.14	0.0058	0.00025	385.49	1.8106
	CV	1.88	0.378	0.011	0.3605	0.0161	0.0254
B	mean	131.6	0.1152	0.0028	0.4323	250.03	0.316
	CV	0.46	0.172	0.058	2.0552	0.016	0.1525
C	mean	44.29	35.94	0.0043	0.0006	343.35	1.3159
	CV	0.958	0.436	0.0117	0.1759	0.0032	0.021
D	mean	110.74	37.17	0.0043	0.00007	425.24	2.8201
	CV	0.777	0.397	0.0448	2.11411	0.0129	0.0462

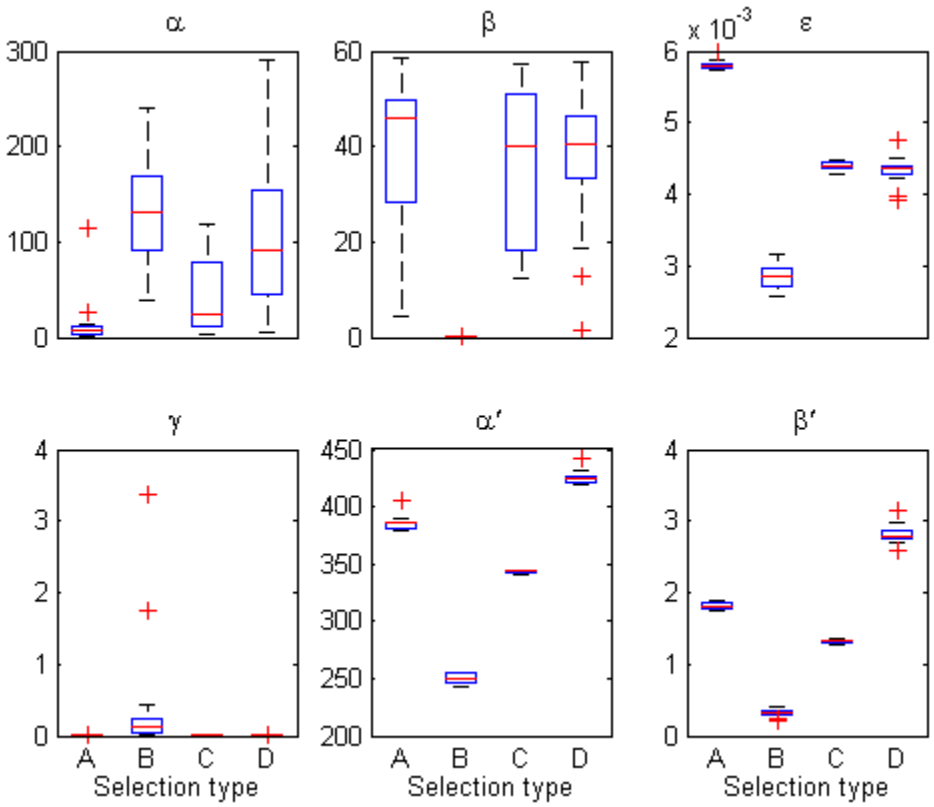

Figure 5.44: Optimised Broomhead *et al.* [1] model parameter values for experimental saccadic velocity profile targets (shown in Figure 5.45) versus the method used for selecting the best solution. Method A selects the solution with the smallest distance to the axes origin. Method B selects the best fit to a 5 deg saccade (objective 1), method C selects the best fit to a 10 deg saccade (objective 2) and method D selects the best fit to a 20 deg saccade (objective 3). The top plots (from left to right) show values of parameters α, β and ε, respectively, whereas bottom plots (from left to right) show values of parameters γ, α' and β', respectively. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

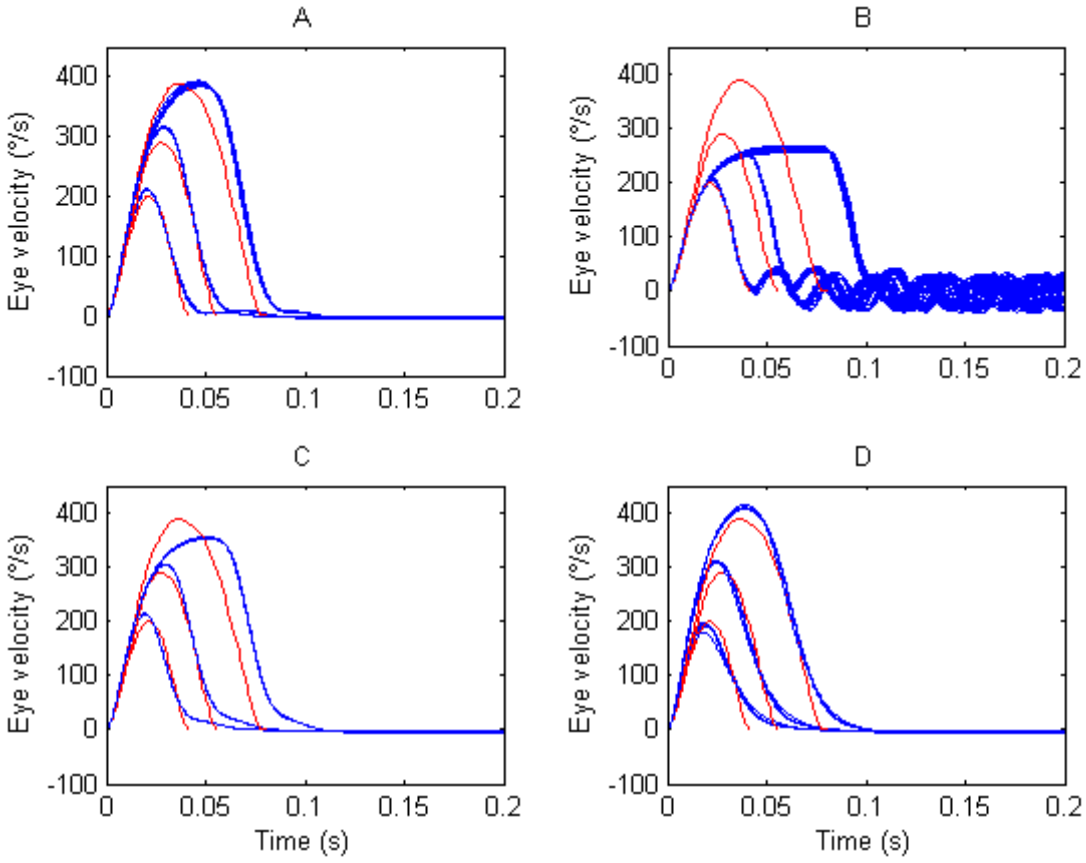

Figure 5.45: Saccadic velocity profiles generated by the Broomhead *et al.* [1] model from the solutions obtained with 16 NSGA-II runs with a population size of 8000. Each plot shows the optimised profiles obtained for saccade amplitudes of 5, 10 and 20 degrees using different final solution selection methods. **A**: Model fits obtained using selection method A (choosing the solution with the smallest distance to the axes origin). **B**: Model fits obtained using selection method B (best fit to a saccade of amplitude 5 degs). **C**: Model fits obtained using selection method C (best fit to a saccade of amplitude 10 degs). **D**: Model fits obtained using selection method D (best fit to a saccade of amplitude 20 degs). The target saccadic velocity profiles are plotted in red; the 16 optimised saccadic velocity profiles are plotted in blue. The target profiles were taken from [23] (see Figure 5.20). On each plot, the vertical axis represents the horizontal eye velocity in degrees per second (°/s), with positive values denoting rightward eye velocities. Time is in seconds (s).

5.4 Discussion

In this Chapter, we have presented the results of finding the NSGA-II parameter values which yield sufficiently accurate and precise fits of the Broomhead *et al.* [1] model to synthetic and experimental data, for both nystagmus oscillations and normal saccades. The accuracy and precision of the results was evaluated using the hypervolume indicator and the smallest distance of the individuals of the estimated *Pareto* front to the axes origin against the population size. We presented the fitness functions that we have developed to measure the fit of the model output to the target time series. These were based on characteristic properties of the nystagmus waveforms and saccades, such as the shape and period size, in the case of the nystagmus data, and the velocity profile in the case of the saccadic data (discussed in section 4.4).

The fitting results show both the advantages and disadvantages of the Broomhead *et al.* [1] model. In the case of the nystagmus fitting, the model can replicate a number of oscillations (*e.g.* jerk, asymmetric pseudo-cycloid and pseudo-cycloid; see Figure 5.40). Some oscillations (*e.g.* jerk with a very rapid fast phase, as in the NEC waveform of Figure 5.40) and some characteristics of specific oscillations (such as the transition between the fast phase and slow phase in the NEB waveform of Figure 5.40) cannot be reproduced by the Broomhead *et al.* [1] model. In the case of the saccadic data, in order for the model to get better fits for each of the saccade amplitudes, either different parameter values have to be used for each saccade amplitude or the model itself has to be modified.

The fitting method (*i.e.* fitness functions and NSGA-II parameter values) presented in this Chapter can be used as a benchmark and analysis method to determine which nystagmus waveforms and saccadic velocity profiles can be generated by a given oculomotor model. This might provide directions on how to further modify the model in order for it to be able to fit a broader range of waveforms and saccadic velocity profiles.

6 Quantifying the effect of noise in the oculomotor system

In this Chapter, we will present the stochastic model and the results from fitting its parameters (both deterministic and stochastic) to experimental nystagmus and saccadic data sets. We fit the deterministic and stochastic parameters separately (see Chapter 4) and we will show that the noise processes do not change the nystagmus waveforms and saccadic velocity profiles, compared to those generated by the zero noise model, thereby justifying the bipartite optimisation approach.

The Chapter is organised as follows. In section 6.1 we will describe the stochastic version of the Broomhead *et al.* [1] model. In section 6.2.1 we will show the experimental nystagmus time series that was used to fit the deterministic parameters of the stochastic model. Section 6.2.2 will present the four segments of the experimental time series and the UPOs extracted from these which will be used as waveform targets, together with the results of fitting the deterministic parameters to these UPOs. In section 6.2.3 we will show the local dimensions pattern of the four experimental time series segments, while in section 6.2.4 the effects of different noise types on the local dimensions pattern generated by the stochastic model will be presented and compared with the experimental data. In section 6.2.5 we will present results supporting our split fitting method (see section 4.1) for the nystagmus waveforms fitting, showing that the noise processes do not affect the shape, period and amplitude of the generated nystagmus waveforms. In section 6.2.6 we will present the results of fitting the stochastic model to experimental saccadic endpoint variability data. Finally, in section 6.2.7, we will present results supporting our split fitting method (see section 4.1) for saccadic fitting. Similarly to the nystagmus fitting, we will show that the noise processes do not affect the shape of the saccade velocity profiles generated by the stochastic model.

6.1 A stochastic model of the saccadic system

Here we propose a modified version of the Broomhead *et al.* [1] model (see equations (2.1)-(2.6)), which incorporates both signal dependent noise (SDN) and constant noise (CN) in the neural control signals. Based on the work of Harris and Wolpert [30] and van Beers [33], we hypothesise that the noise is generated by the neural control signals from the burst neuron populations and the neural integrator. These two signals are the command signals that drive the motor neurons which send signals to the eye muscles (refer to section 2.2). We believe that this model constitutes an initial approach which involves the less possible assumptions to add noise to the command signals. Accordingly, we have expanded equation (2.2) into a stochastic version, where SDN and CN are added to the neural integrator signal n and the burst signal $(r-l)$. This model includes the deterministic parameters (shown in Table 2.1) and the stochastic parameters (shown in Table 6.1). In the expanded model, the velocity equation (2.2) was replaced with the following stochastic differential equation (SDE):

$$\frac{dv}{dt} = -\left(\frac{1}{T_1}+\frac{1}{T_2}\right)v - \frac{1}{T_1T_2}g + \frac{1}{T_1T_2}n(1+k_{sdn}^{n}\xi_t^{n}) + \left(\frac{1}{T_1}+\frac{1}{T_2}\right)(r-l)\left(1+k_{sdn}^{p}\xi_t^{p}\right)+k_{cn}\eta_t \tag{6.1}$$

where ξ_t^p, ξ_t^n and η_t are zero-mean white noise processes with unit variance, k_{sdn}^n describes the magnitude of SDN on the neural integrator signal, k_{sdn}^p is the magnitude of signal-dependent noise on the burst neural signal and k_{cn} is the magnitude of the combined CN on the neural integrator and burst signals (Table 6.1). The remaining variables from equation (6.1) are described in section 2.2. The solver, which we used to numerically solve the stochastic model, was the stochastic Improved Euler (Heun) method [111] with time step 10^{-8} s.

Table 6.1: Stochastic model parameters (stochastic parameters).

Parameter name	Description
k_{sdn}^n	Strength of the SDN on the neural integrator signal
k_{sdn}^p	Strength of the SDN on the burst neural signal
k_{cn}	Strength of the CN on the neural integrator and burst signals

6.2 Results

In this section, we will present the results acquired by using the parameter estimation method described in Chapters 4 and 5. The method involves fitting the stochastic model's deterministic parameters (*i.e.* the Broomhead *et al.* [1] model parameters, see Table 2.1) to the experimental nystagmus waveforms and saccadic velocity profiles. Subsequently, these values were fixed and the noise intensity parameters (stochastic parameters, see Table 6.1) were fitted to set the stochastic model to generate similar local dimensions pattern to the experimental nystagmus data, and to reproduce the experimental saccadic endpoint variability data.

6.2.1 Experimental nystagmus time series

The experimental time series that we used is shown in Figure 6.1 and is derived from a single participant of the experimental database retrieved from [96] (see section 4.2.1). By using recordings from other participants, very similar local dimensions pattern was observed. Four parts of the time series were examined, namely those where there are no blinks for at least 10 s, as blinks could interfere with the analysis process. These time series parts are shown in red in Figure 6.1 and in detail in Figure 6.2. Each time series part is 10 s long and was subjected to analysis using the method of Akman *et al.* [15].

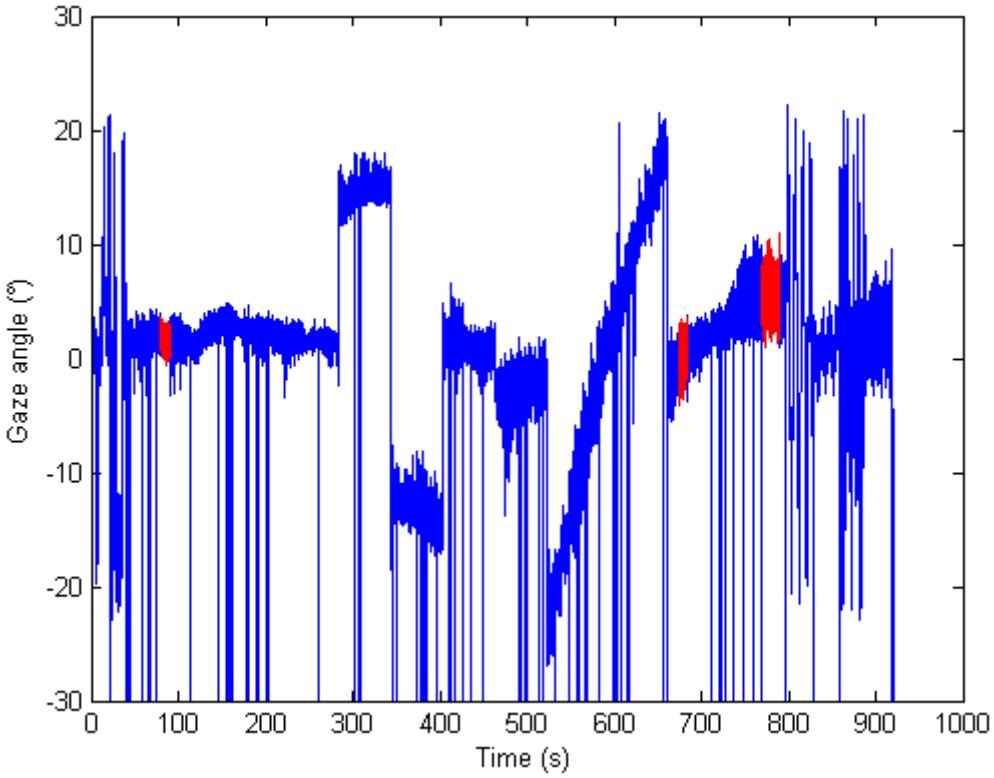

Figure 6.1: Horizontal gaze angle recording (blue line) from a single participant derived from [96] study. Details of the study can be found in section 4.2.1. The four extracted parts of the time series that we used to measure the local dimensions pattern are shown in red. The 3rd and 4th parts are next to each other. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is given in seconds (s).

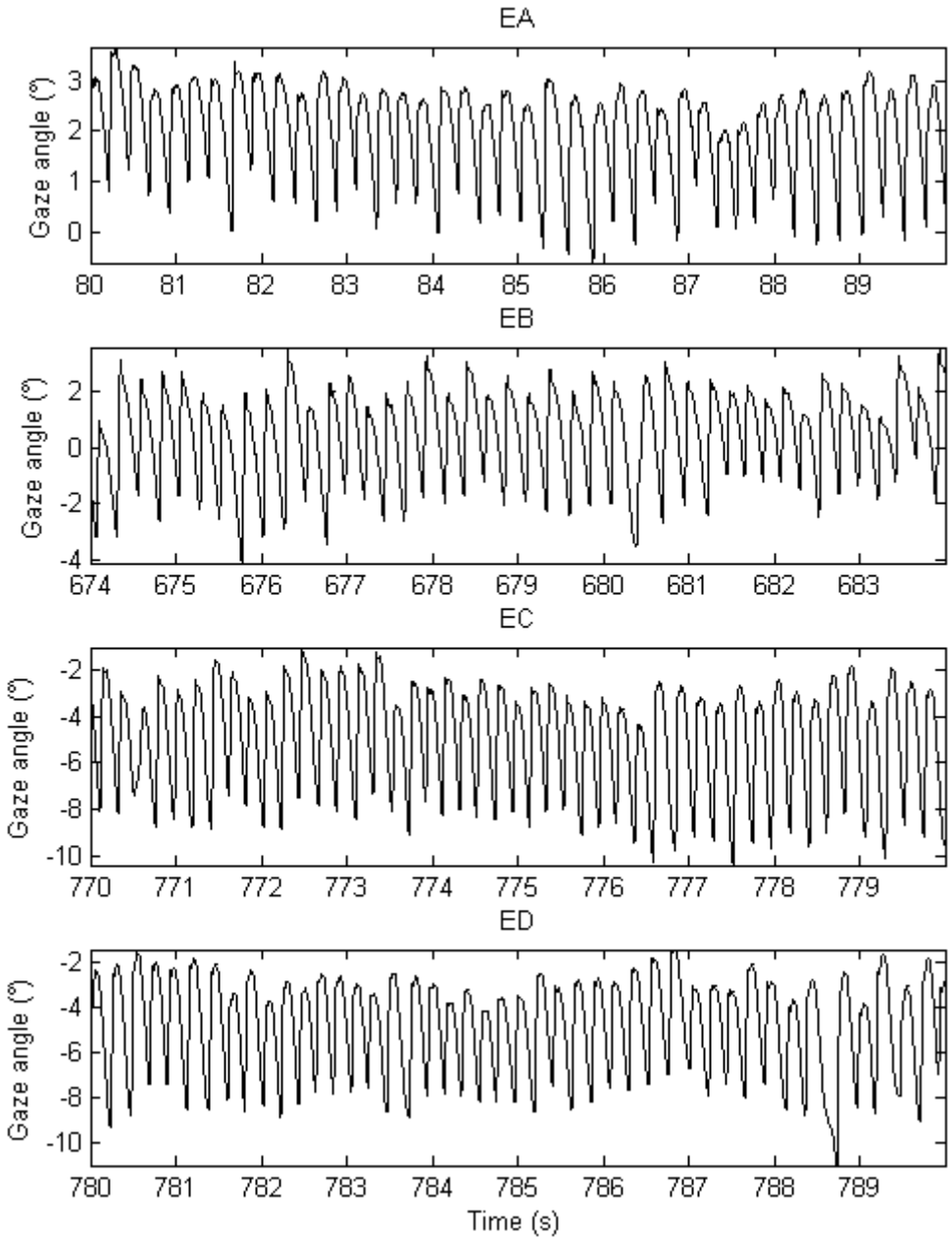

Figure 6.2: The four extracted parts from the time series shown in Figure 6.1. From top to bottom, the time series are designated as EA, EB, EC and ED. The waveform type in each case is a jerk nystagmus. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

6.2.2 Fitting the stochastic model's deterministic parameters to the experimental UPOs

The methodology presented in section 4.2.2 allowed extraction of the unstable periodic orbit from each experimental time series in Figure 6.2. The extracted

UPOs are shown in Figure 6.3. Subsequently, the deterministic parameters of the stochastic model were fitted to the extracted UPOs using the methods presented in Chapter 5. The results of the fitting, using the NSGA-II with a population size of 4000, are shown as box plots in Figures 6.4 and 6.5. The population size choice was based on the results shown in Chapter 5, which showed that the population size of 4000 is sufficient for the NSGA-II to provide sufficiently converged results. The waveforms obtained from the parameter optimisation procedure are plotted against the target waveforms in Figure 6.6.

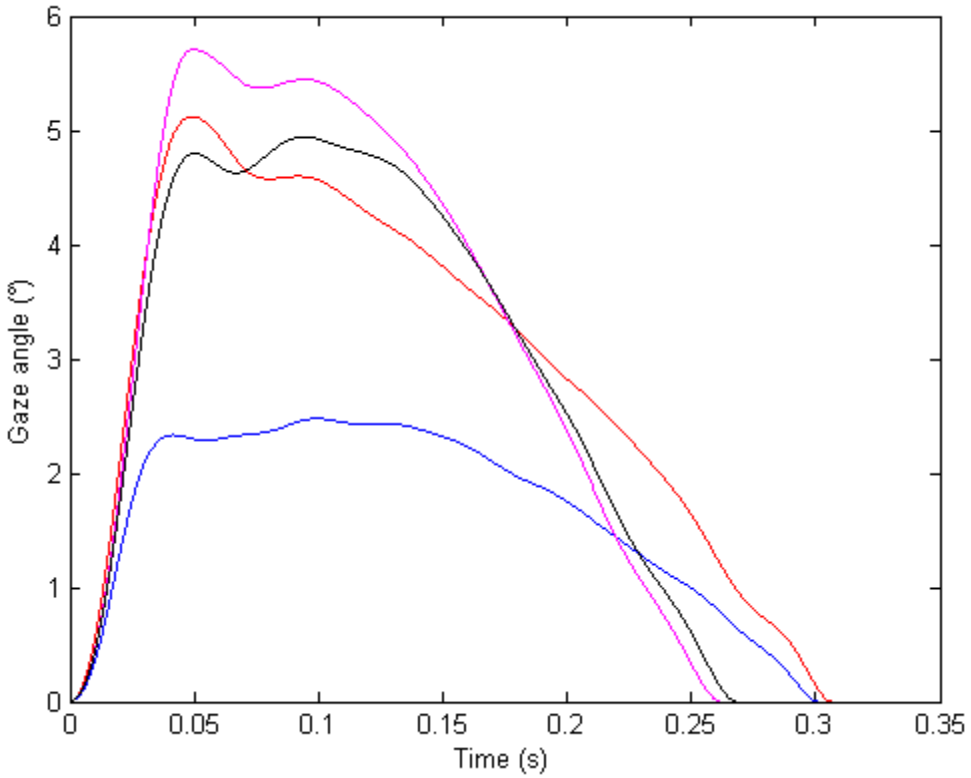

Figure 6.3: UPOs extracted from the experimental time series of Figure 6.2 using the method described in section 4.2.2. The UPOs were modified so they began at the start of the nystagmus fast phase, in order to be used as waveform targets for fitting the deterministic parameters of the stochastic model. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. The UPO in blue is the EA waveform target, red colour denotes the EB waveform target, magenta represents the EC waveform target and black denotes the ED waveform target. Time is in seconds (s). The UPO nystagmus waveform types are all of jerk type.

Although the extracted UPOs are derived from a single participant during a single recording, the extracted UPOs are different. Therefore, the Broomhead *et al.* [1] model needs different parameter values for each UPO. This is very interesting as it validates that there could be variations to the waveform characteristics in a single recording of a single participant. This perhaps could mean that the values, which set the behaviour of the burst neurons, change as they are influenced by other oculomotor or other systems (such as the cortex).

Table 6.2: Found mean and coefficient of variation (CV) of the optimised deterministic parameter values for each target waveforms (EA, EB, EC and ED). The target waveforms are shown in Figure 6.3. The found mean and CV were calculated by the solutions selected using the method described in section 5.2.1 from 16 NSGA-II runs and with population size 4000.

Target ID	Value type	α	β	ε	γ	α'	β'
NEA	mean	62.198	0.8799	0.00197	8.957	860.85	13.11
	CV	0.01921	0.02622	0.16	0.17158	0.1221	0.13026
NEB	mean	112.814	1.36894	0.00259	9.91697	933.157	13.0417
	CV	0.01033	0.0079	0.1212	0.07779	0.06964	0.07332
NEC	mean	148.287	1.79229	0.00171	2.45331	848.207	12.434
	CV	0.03876	0.03865	0.90835	0.19692	0.14338	0.18058
NED	mean	124.946	1.82110	0.00096	1.72702	922.725	14.9749
	CV	0.03486	0.0331	1.41783	0.19882	0.0725	0.08443

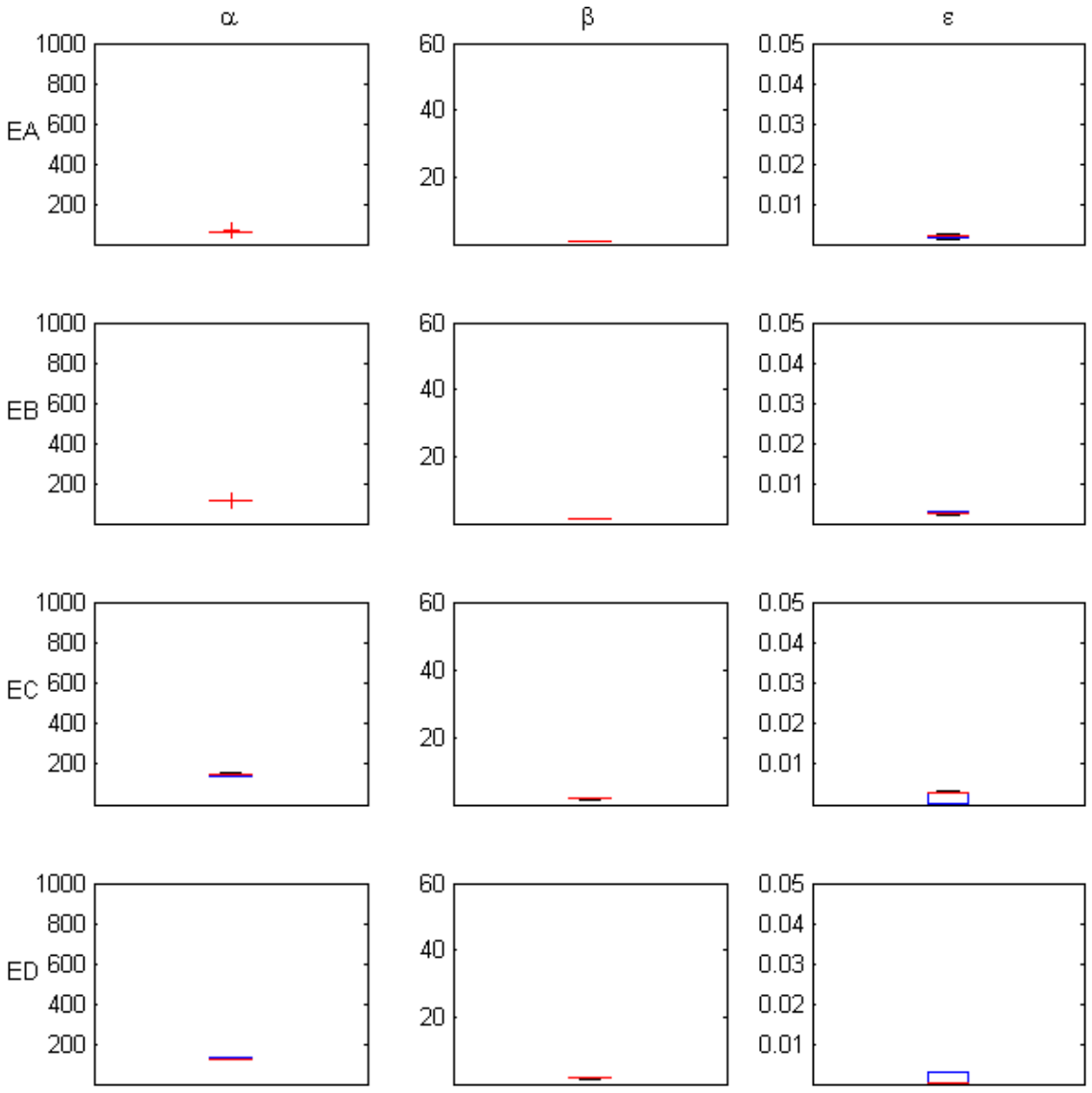

Figure 6.4: Optimised deterministic parameter values of the stochastic model obtained from 16 NSGA-II runs with population size 4000 for each experimental nystagmus waveform target (target waveforms are shown in Figure 6.3). The first line of box plots shows the optimised values of α, β and ε for the experimental waveform target EA, whereas those for EB, EC and ED are shown by the second, third and fourth lines of plots, respectively. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

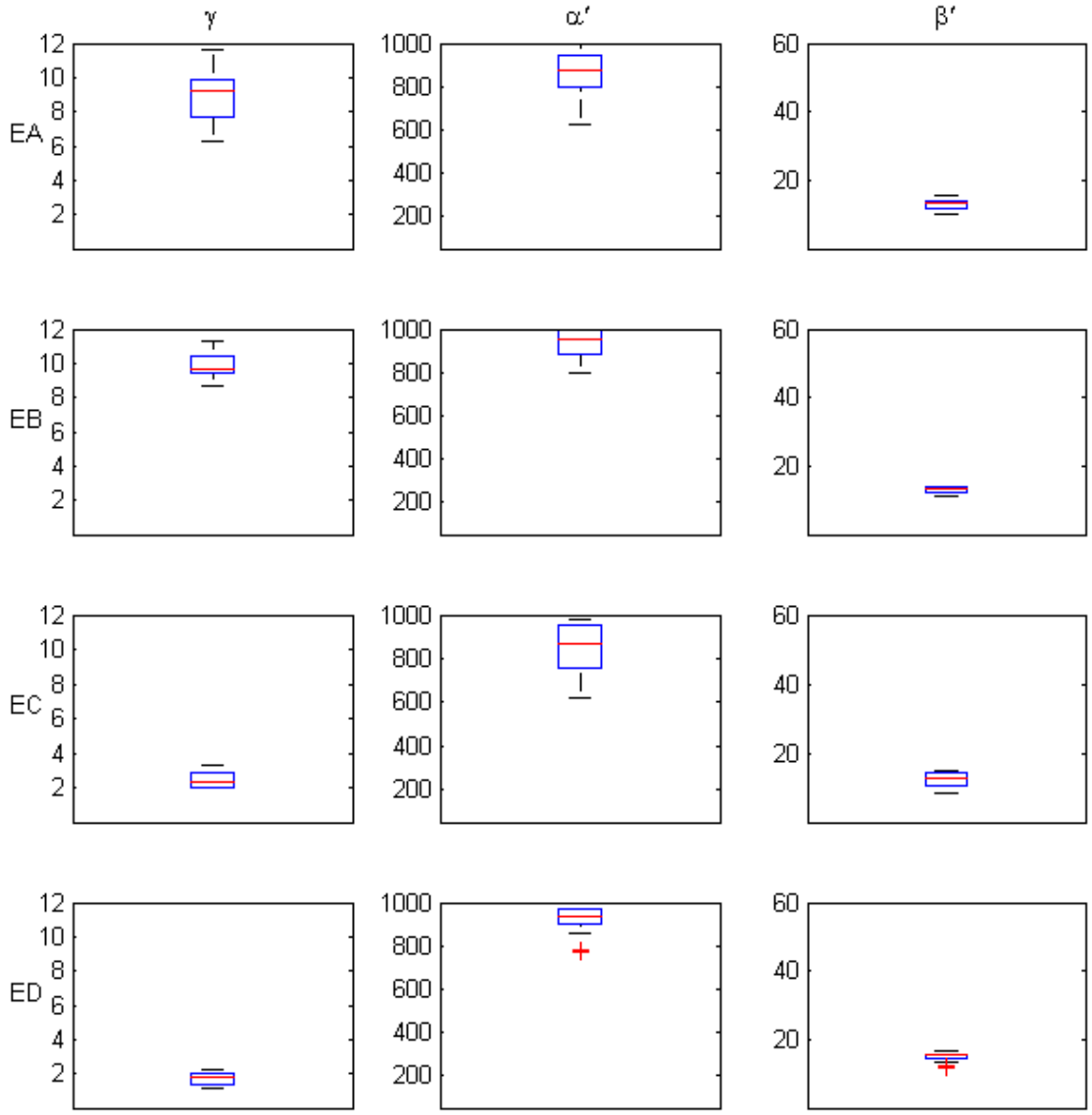

Figure 6.5: Optimised deterministic parameter values of the stochastic model obtained from 16 NSGA-II runs with population size 4000 for each experimental nystagmus waveform target (target waveforms are shown in Figure 6.3). The first line of box plots shows the optimised values of γ, α' and β' for the experimental waveform target EA, whereas those for EB, EC and ED are shown by the second, third and fourth lines of plots, respectively. The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

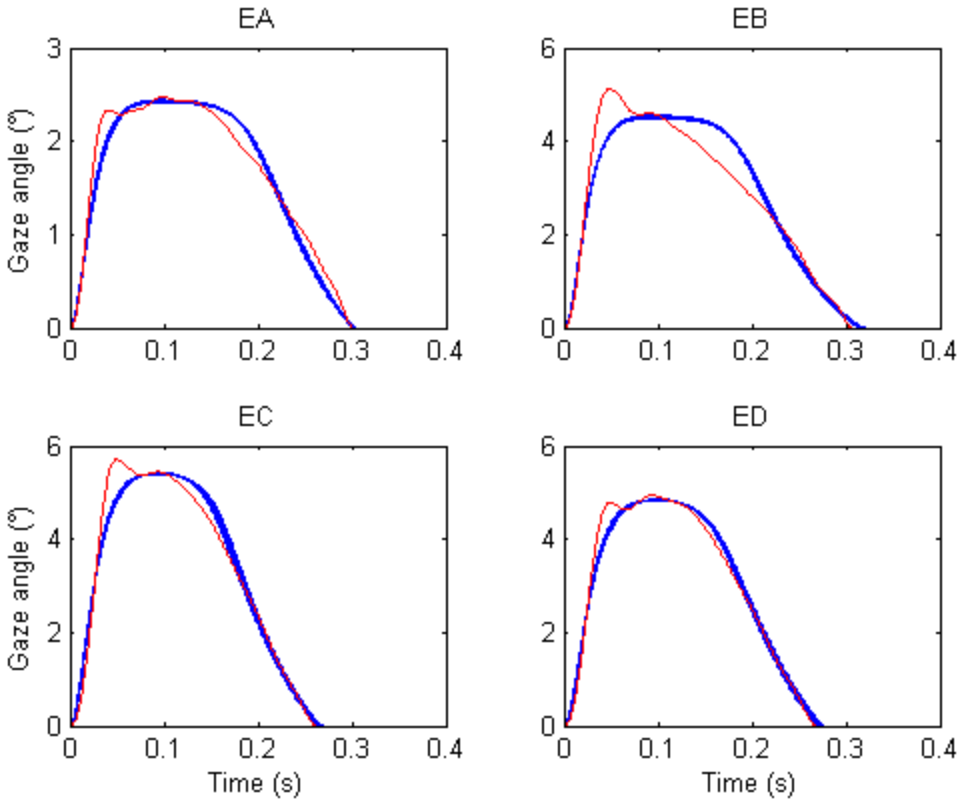

Figure 6.6: Waveforms generated by fitting the deterministic parameters of the stochastic model to the experimental nystagmus waveform targets shown in Figure 6.3. The target waveforms are plotted in red; the 16 optimised waveforms are plotted in blue. The waveforms were chosen from each NSGA-II run and they were selected using the method described in section 5.2.1. Starting from top left and going clockwise, the target identifiers are EA, EB, EC and ED. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is given in seconds (s). The experimental nystagmus waveform types are all of jerk type.

6.2.3 Experimental time series - local dimensions pattern

The main objective of this analysis was to demonstrate the relative differences between the local dimensions values at different points on the waveform for the four experimental time series parts from Figure 6.2, specifically at the foveation window (where the eye velocity is the lowest), and during the slow and fast phases.

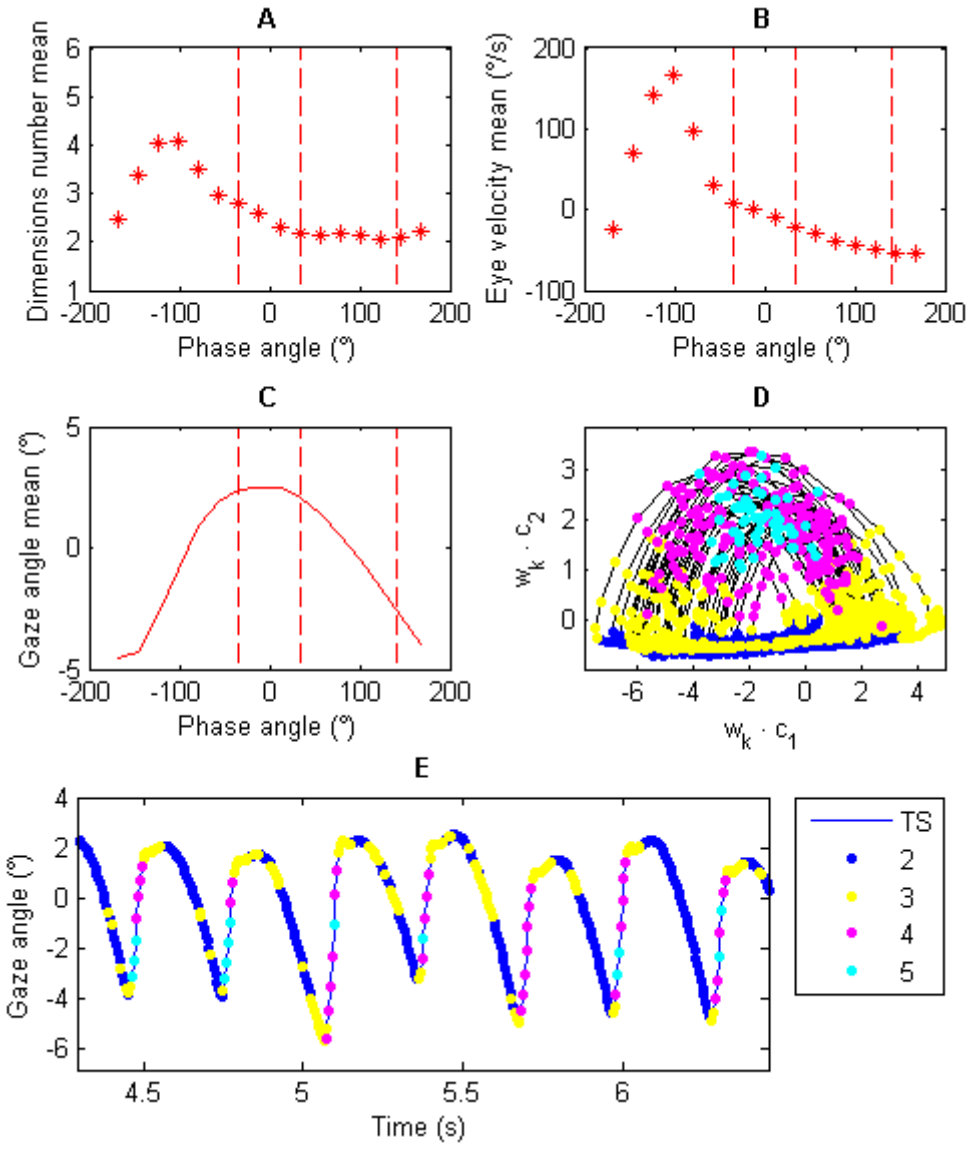

Figure 6.7: Nonlinear time series analysis of experimental jerk nystagmus waveform EA (see Figure 6.2). **A**: Mean local dimensions versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle part denotes fixation, the third part denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity versus phase angle. **C**: Mean gaze angle versus phase angle. **D**: Global dynamics reconstructed from the time series using the method of delays. **E**: A portion of the time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). In D and E, the colour of each circle represents the local dimensions at that particular point.

Figure 6.7A shows the mean local dimensions versus the phase angle area (size of each area $\pi/8$). The number of local dimensions was calculated using the method of Akman *et.al.* [15] (described in section 2.4). The mean phase angle was computed using the Hilbert transform of the time series, where the phase angle of each point is calculated and the points are grouped in areas of $\pi/8$. The vertical dashed lines divide the phase plot into the different waveform parts. The first part denotes the fast phase, the middle part determines the fixation point (in which the eye is mostly stable), the third part shows the slow phase in which the eye gains distance from the fixed point and the last part denotes the start of the fast phase. Figure 6.7B shows the mean velocity versus the phase angle, Figure 6.7C shows the mean gaze versus the phase angle, whereas Figure 6.7D shows the global dynamics reconstructed from the experimental time series generated by the method of delays. The reconstructed delay vectors $\boldsymbol{w_k}$ are projected to the first singular vectors (c_1 and c_2) of the local trajectory matrix (refer to section 2.4). The coloured circles show the mean local dimensions of the points at that particular phase angle area. Figure 6.7E shows a part of the time series. The colours (blue, yellow, magenta, cyan and black) in subplots D and E designate the local dimensions. The subplots A, B, D and E demonstrate that there is a specific pattern in the local dimensions. The fast phase has the highest number of local dimensions, while the slow phase has the lowest. However, at the fixed point the size of the local dimensions ranges between the values of the fast and slow phase dimensions. Interestingly, the number of local dimensions is correlated with the velocity of the waveform versus the phase angle. Therefore, during the fast phase the burst neurons fire maximally. This increase of the firing rate enables the SDN noise to cause more variability having an effect on the local dimensions (*i.e.* they increase).

Figures 6.8-6.10 show the results of the local dimensions analysis for the experimental time series EB, EC and ED, respectively. Moreover, it is very interesting that they show the same pattern as experimental time series EA, with the local dimensions increasing to a maximum of 4-5 during the fast phase, decreasing to around 3 close to the fixed point, and further decreasing to a minimum of 2 during the slow phase, before increasing again as the slow phase ends.

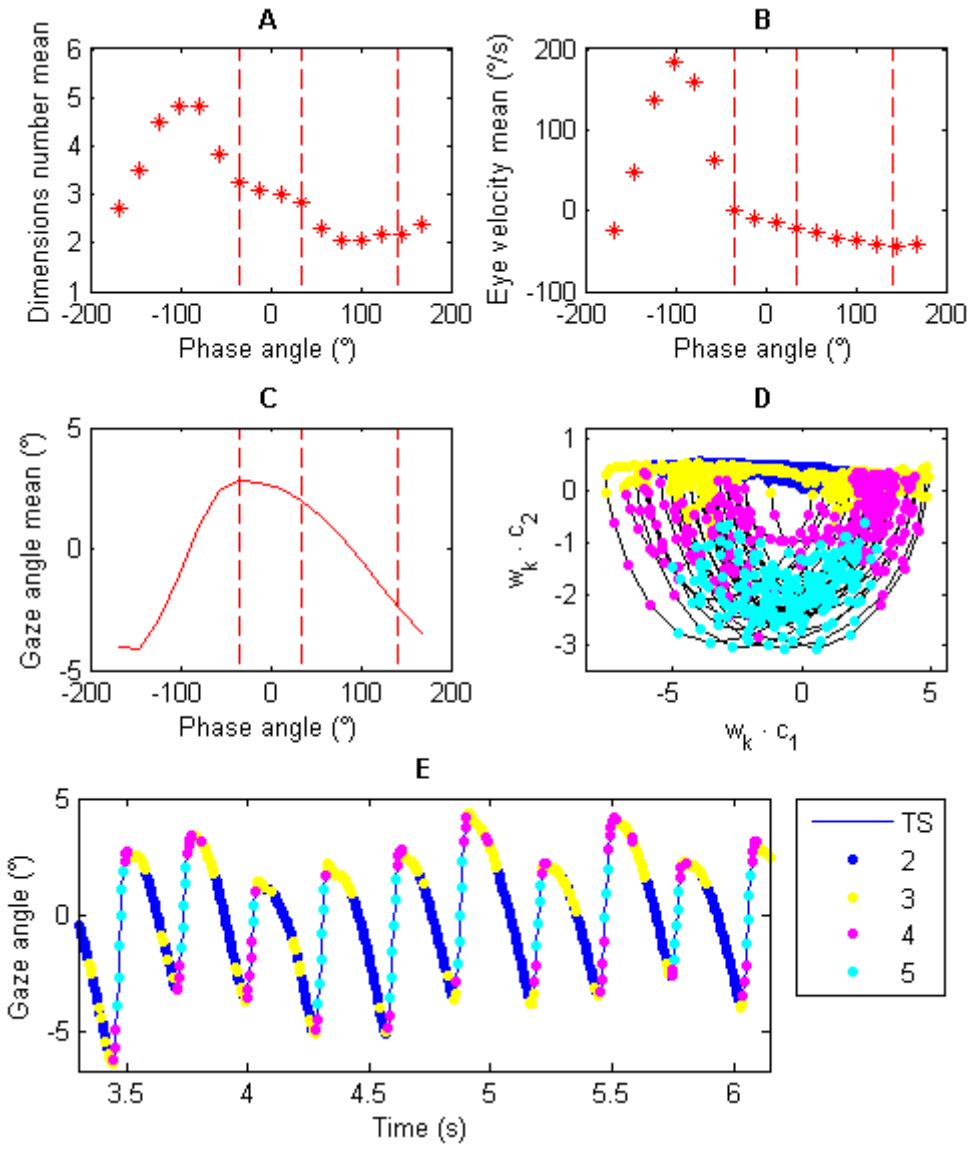

Figure 6.8: Nonlinear time series analysis of experimental jerk nystagmus waveform EB (see Figure 6.2). **A**: Mean local dimensions versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle part denotes fixation, the third part denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity versus phase angle. **C**: Mean gaze angle versus phase angle. **D**: Global dynamics reconstructed from the time series using the method of delays. **E**: A portion of the time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). In D and E, the colour of each circle represents the local dimensions at that particular point.

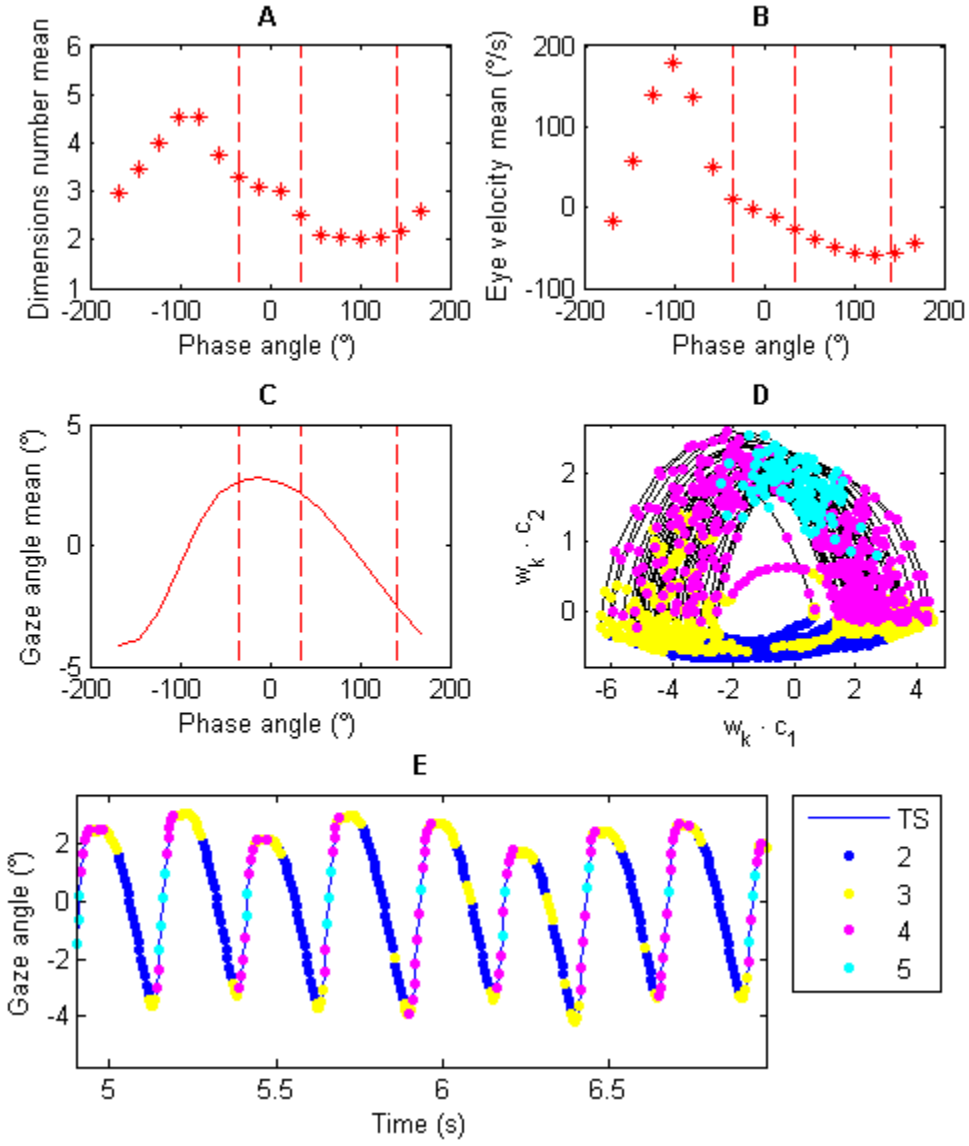

Figure 6.9: Nonlinear time series analysis of experimental jerk nystagmus waveform EC (see Figure 6.2). **A**: Mean local dimensions versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle part denotes fixation, the third part denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity versus phase angle. **C**: Mean gaze angle versus phase angle. **D**: Global dynamics reconstructed from the time series using the method of delays. **E**: A portion of the time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). In D and E, the colour of each circle represents the local dimensions at that particular point.

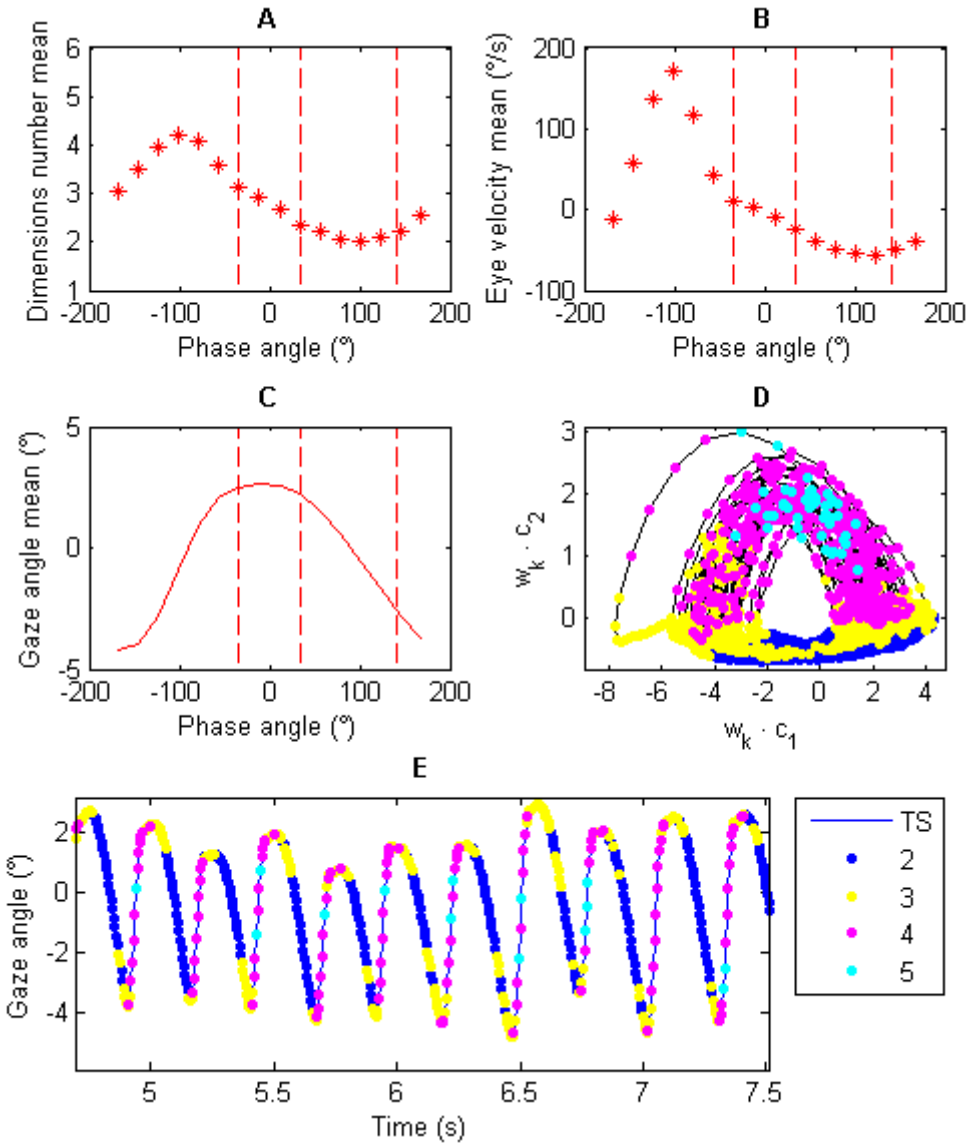

Figure 6.10: Nonlinear time series analysis of experimental jerk nystagmus waveform ED (see Figure 6.2). **A**: Mean local dimensions versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle part denotes fixation, the third part denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity versus phase angle. **C**: Mean gaze angle versus phase angle. **D**: Global dynamics reconstructed from the time series using the method of delays. **E**: A portion of the time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s). In D and E, the colour of each circle represents the local dimensions at that particular point.

6.2.4 Effects of different noise types and magnitudes on the simulated dynamics

In this section, we will show how different noise types (*i.e.* different combinations of SDN and CN on the burst and neural signals) in the stochastic model can affect the local dimensions pattern at different waveform phases (such as fast phase, fixed point and slow phase). The simulated data used for analysis was generated by fixing the deterministic parameter combination SD shown in Table 6.3, and by exploring the stochastic parameter space. This allowed to estimate the values that allow the model to have similar local dimensions pattern as that of the corresponding experimental jerk nystagmus waveform ED. The deterministic parameter values SD were chosen due to the good qualitative fit they gave to waveform ED. The Broomhead *et al.* [1] model can generate waveforms that are more similar to the fast phase, fixed point and slow phase locations of the target waveforms (see Figure 6.6). In order to get the simulated time series for the analysis, we utilised the part of the oscillation that follows the post-saccadic drift. This drift ends after around 140 seconds and may have an effect on the analysis results (see Figure 5.1). Therefore, we ran the deterministic model (*i.e.* the stochastic model without noise) for 160 seconds using the Implicit Euler method and used the last values as initial conditions for the stochastic version of the model. Subsequently, the stochastic model was run for 15 seconds and the last 10 seconds were used for the local dimensions analysis.

By performing a grid search for the values of the stochastic parameters, we found that there is a variation in local dimensions with waveform phase. However, the simulation is not exactly the same as that measured in the experimental data. This confirms the assumption that a stochastic model based on the Broomhead *et al.* [1] model could generate similar behaviour. In the following subsections we will present the results for a range of noise types and magnitudes.

Table 6.3: Representative parameter combinations (SA, SB, SC and SD) obtained by fitting the deterministic parameters of the stochastic model to the four UPOs shown in Figure 6.16 (refer to section 6.2.2).

Target	α	β	ε	γ	α'	β'
SA	60.68	0.90	0.001806	7.37	695.91	11.19
SB	112.54	1.35	0.002965	9.55	880.86	12.38
SC	155.44	1.69	0.003165	2.34	959.09	14.19
SD	129.80	1.75	0.002822	1.26	989.52	16.74

6.2.4.1 Signal dependent noise on the burst signal

We added signal dependent noise only to the signal generated by the burst neurons ($r - l$) (see equation (6.1)) by using the following stochastic parameter values: $k_{sdn}^{n} = 0$, $k_{sdn}^{b} = 0.005$ and $k_{cn} = 0$. The stochastic model's behaviour is similar to that of the experimental time series ED (Figure 6.11). Although, there is an increase in local dimensions at the fast phase, the variation in local dimensions at the slow phase differs significantly from that observed in the experimental time series ED. In particular, the simulated local dimensions reaches its minimum at the end of the fast phase, rather than during the slow phase. Moreover, the maximum dimension appears earlier in the fast phase compared with the experimental time series (Figure 6.11A and D). The simulation does demonstrate, however, that signal dependent noise on the burst signal can create a local dimensions increase during the fast phase of the nystagmus oscillation.

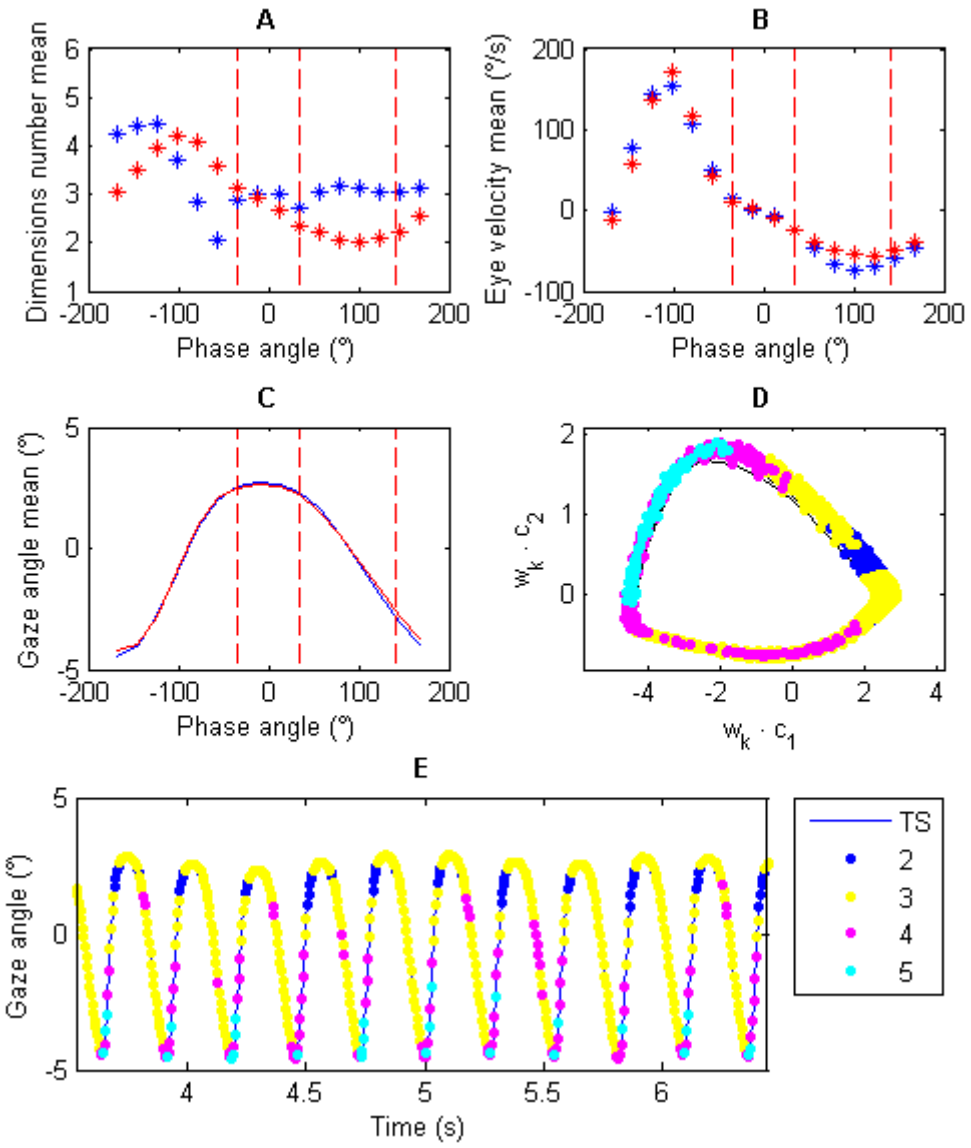

Figure 6.11: Nonlinear time series analysis of the simulated nystagmus oscillation generated by the stochastic model with SDN only in the burst signal**.** The waveform was generated using the deterministic parameter set SD shown in Table 6.3 and the following stochastic parameter values: $k_{sdn}^{n} = 0$, $k_{sdn}^{b} = 0.005$ and $k_{cn} = 0$. **A**: Mean local dimension of the simulated time series (blue) and the experimental time series ED (red) versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle denotes the fixed point, the third denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity of the simulated time series (blue) and the experimental time series (red) versus phase angle. **C**: Mean gaze angle of the simulated time series (blue) and the experimental time series (red) versus phase angle. **D**: Global dynamics reconstructed from the simulated time series using the method of delays. **E**: A

portion of the simulated time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. The time is in seconds (s). In D and E, the colour of each circle denotes the local dimension at that point.

6.2.4.2 Constant noise on the burst and neural integrator signals

We added constant noise to the signal generated by the burst neurons ($r - l$) and the neural integrator signal n (see equation (6.1)) by using the following stochastic parameter values: $k_{sdn}^{n} = 0$, $k_{sdn}^{b} = 0$ and $k_{cn} = 50$. For this noise model, although the simulated waveform quite closely reproduces the local dimensions at the fixed point, the variations in local dimensions during the fast and slow phases do not match those of the experimental oscillation. The simulated and experimental dimension-phase profiles during these phases of the waveform are roughly anti-correlated. In particular, during the fast phase the local dimensions of the simulated waveform increases when that of the experimental waveform decreases, and vice versa (Figure 6.12A).

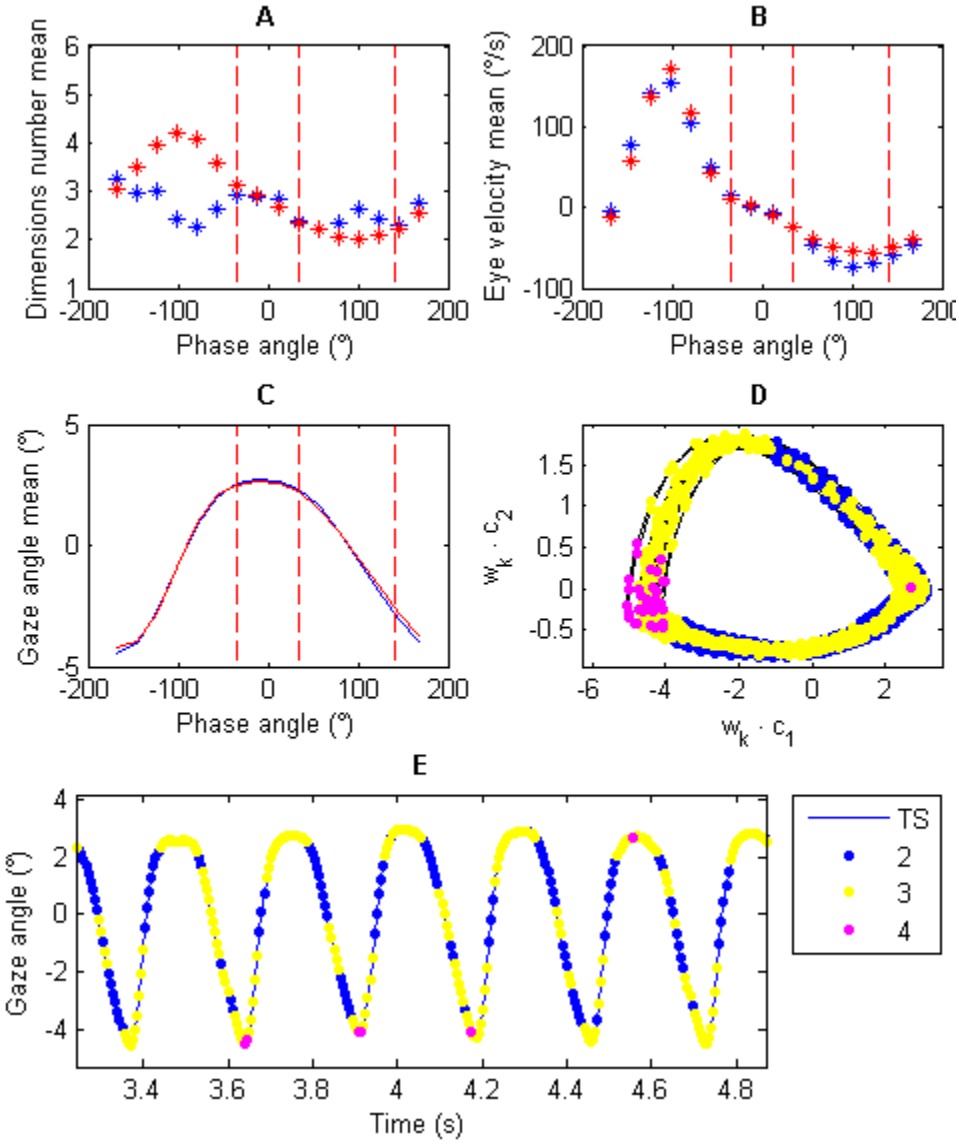

Figure 6.12: Nonlinear time series analysis of the simulated nystagmus oscillation generated by the stochastic model with CN only on the burst and neural integrator signals. The waveform was generated using the deterministic parameter set SD shown in Table 6.3 and the following stochastic parameter values: $k_{sdn}^{n} = 0$, $k_{sdn}^{b} = 0$ and $k_{cn} = 50$. **A**: Mean local dimension of the simulated time series (blue) and the experimental time series ED (red) versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle denotes the fixed point, the third denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity of the simulated time series (blue) and the experimental time series (red) versus phase angle. **C**: Mean gaze angle of the simulated time series (blue) and the experimental time series (red) versus phase angle. **D**: Global dynamics reconstructed from the simulated time series using the method of delays. **E**: A

portion of the simulated time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. The time is in seconds (s). In D and E, the colour of each circle denotes the local dimension at that point.

6.2.4.3 Signal dependent noise on the neural integrator signal only

We added signal dependent noise to the neural integrator signal n (see equation 6.1) by using the following stochastic parameter values: $k_{sdn}^{n} = 0.3$, $k_{sdn}^{b} = 0$ and $k_{cn} = 0$. For this stochastic parameter values, the variation in local dimensions is closer to that observed experimentally than to the previous two stochastic parameter values sets (sections 6.2.4.1 and 6.2.4.2). The local dimensions increase to a maximum value during the fast phase, then decrease at the fixed point and decreases more at the start of the slow phase. However, there is an increase of the local dimensions before and at the start of the fast phase, which decreases later in the fast phase (Figure 6.13). This behaviour is qualitatively different from the one found in the experimental data, nevertheless it exhibits a close fit to the experimental data.

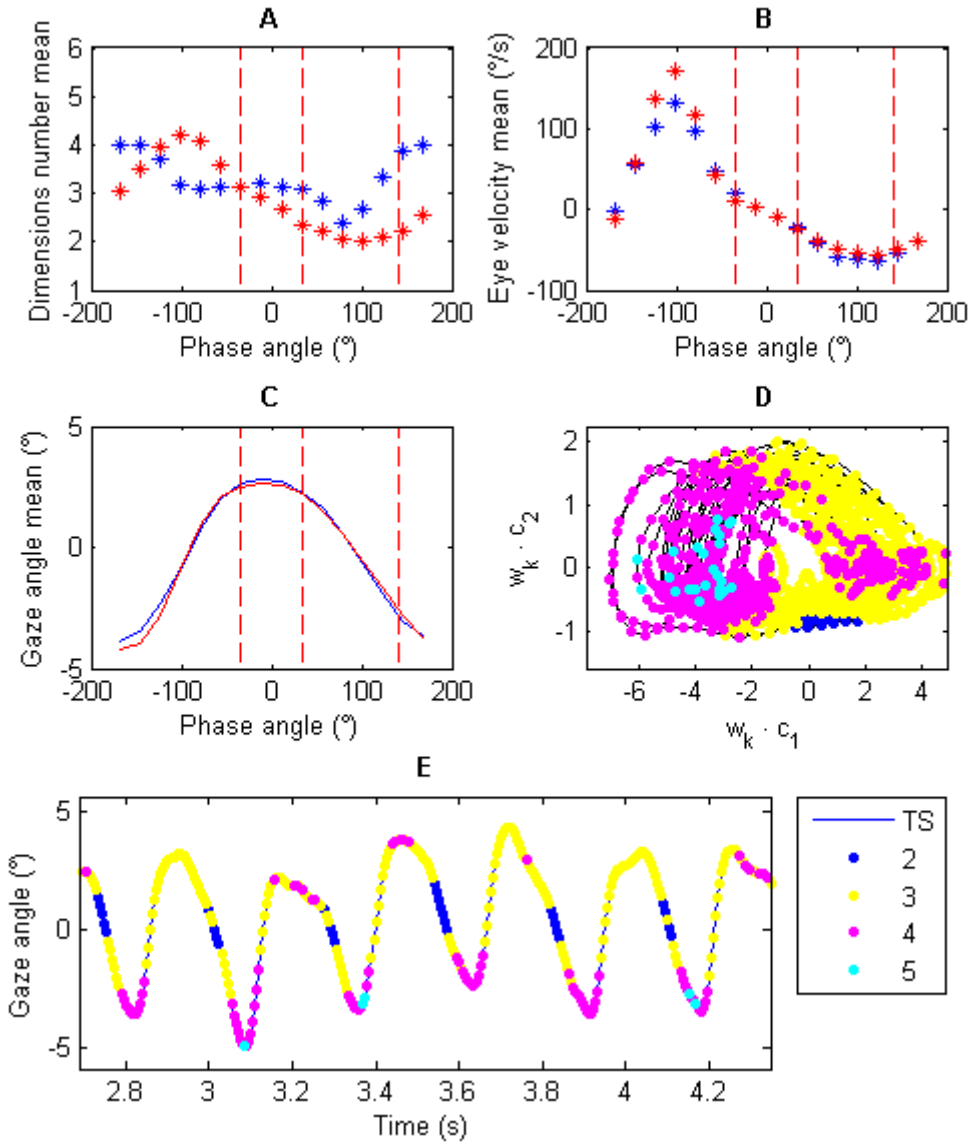

Figure 6.13: Nonlinear time series analysis of the simulated nystagmus oscillation generated by the stochastic model with SDN added only on the neural integrator signal only**.** The waveform was generated using the deterministic parameter set SD shown in Table 6.3 and the following stochastic parameter values: $k_{sdn}^{n} = 0.3$, $k_{sdn}^{b} = 0$ and $k_{cn} = 0$. **A**: Mean local dimension of the simulated time series (blue) and the experimental time series ED (red) versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle denotes the fixed point, the third denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity of the simulated time series (blue) and the experimental time series (red) versus phase angle. **C**: Mean gaze angle of the simulated time series (blue) and the experimental time series (red) versus phase angle. **D**: Global dynamics reconstructed from the simulated time series using the method of delays. **E**: A

portion of the simulated time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. The time is in seconds (s). In D and E, the colour of each circle denotes the local dimension at that point.

6.2.4.4 Signal dependent and constant noise on burst and neural integrator signals

We added SDN and CN to the signal derived from the burst neurons ($r - l$) and neural integrator n (see equation 6.1) by using the following stochastic parameter values: $k_{sdn}^{n} = 0.1$, $k_{sdn}^{b} = 0.005$ and $k_{cn} = 50$. This behaviour is qualitatively different from the one found in the experimental data, nevertheless it exhibits the closest fit to the experimental data of the all stochastic parameter combinations we have examined.

This stochastic model generates the closest fits to the experimental data of all the stochastic parameter combinations we examined. Although - as for the model with SDN on the neural integrator signal only - the minimum and maximum local dimensions are observed at earlier phases compared with the experimental data, the decrease in dimensions near the fixed point and the first section of the slow phase is closer to that observed experimentally (in particular, the minimum dimensions of the simulated and experimental waveforms are equal) (Figure 6.14).

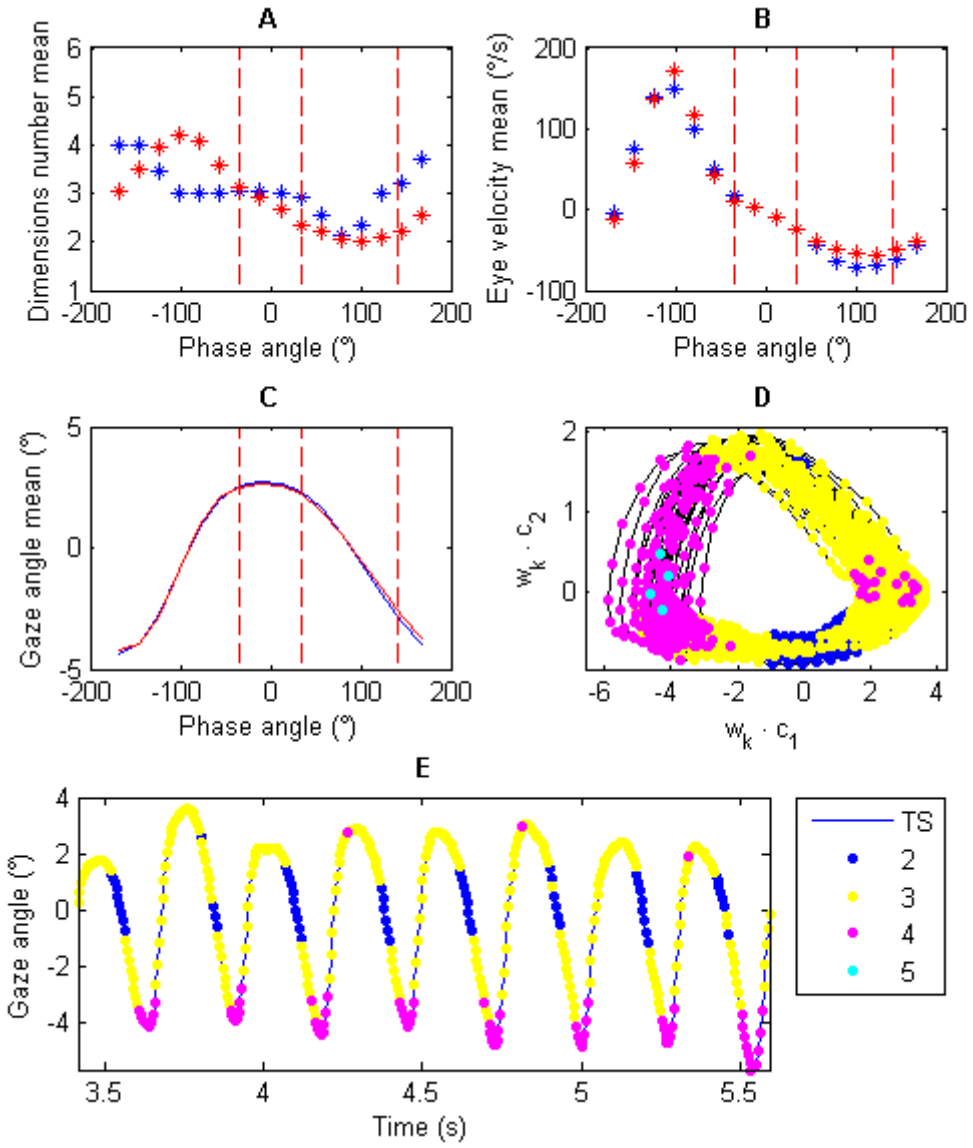

Figure 6.14: Nonlinear time series analysis of the simulated nystagmus oscillation generated by the stochastic model with SDN and CN on the burst and neural integrator signals**.** The waveform was generated using the deterministic parameter set SD shown in Table 6.3 and the following stochastic parameter values $k_{sdn}^{n} = 0.1$, $k_{sdn}^{b} = 0.005$ and $k_{cn} = 50$. **A**: Mean local dimensions of the simulated time series (blue) and the experimental time series ED (red) versus phase angle. The vertical dashed lines divide the phase plot into different waveform parts. The first part denotes the fast phase, the middle denotes the fixed point, the third denotes the slow phase and the last part denotes the start of the fast phase. **B**: Mean velocity of the simulated time series (blue) and the experimental time series (red) versus phase angle. **C**: Mean gaze angle of the simulated time series (blue) and the experimental time series (red) versus phase angle. **D**: Global dynamics reconstructed from the simulated time series using the

method of delays. **E**: A portion of the simulated time series. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. The time is in seconds (s). In D and E, the colour of each circle denotes the local dimensions at that point.

6.2.5 The noise processes do not affect the nystagmus waveform shape generated by the stochastic model

In this section, we will present results showing that the noise processes do not change the key characteristics of the waveform (shape, period and amplitude) generated by the stochastic model. In order to demonstrate this, we generated four time series (Figure 6.15) using the stochastic model with the deterministic parameter values SD shown in Table 6.3 and the following stochastic parameter values: k_{sdn}^{p}= 0.03, k_{sdn}^{n}= 0.1 and k_{cn} = 30. Figure 6.16 compares the four UPOs (in blue) extracted from the time series shown in Figure 6.15 against the nystagmus waveform generated using the same deterministic parameters, but with no noise (*i.e.* with the stochastic parameters set to 0). It can be seen that the UPOs extracted from the synthetic time series are similar to the deterministic waveform.

We then used these UPOs as target waveforms to fit the stochastic model's deterministic parameters as described in section 4.3. The box plots in Figure 6.17 and Figure 6.18 show that the optimised parameter values are similar to the target values. Figure 6.19 compares the waveforms generated by the optimised deterministic parameters against the target UPOs and the deterministic waveform. Again, these waveforms are similar.

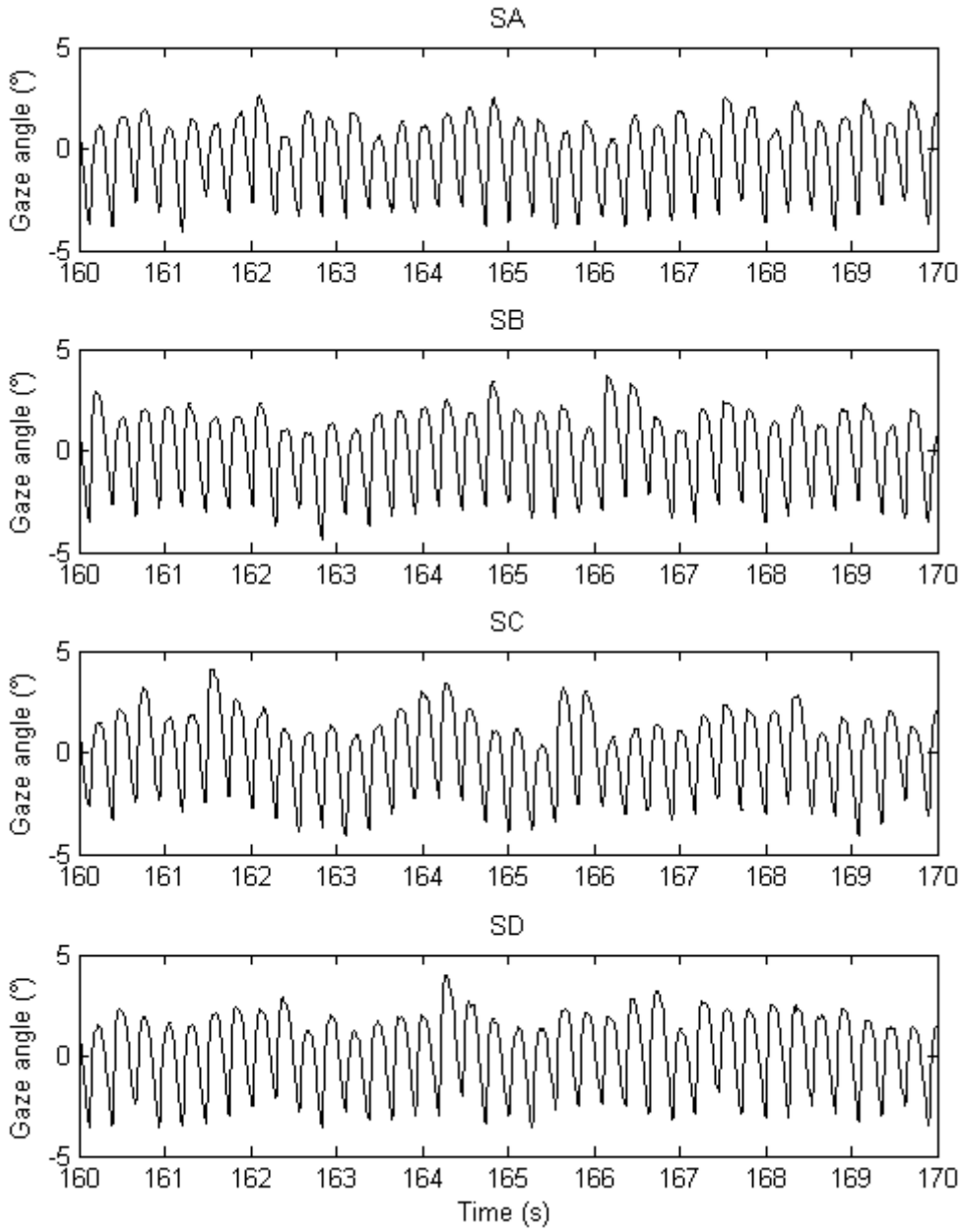

Figure 6.15: Four synthetic time series generated by the stochastic model with the deterministic parameter values SD shown in Table 6.3 and stochastic parameter values k_{sdn}^{p}= 0.03, k_{sdn}^{n}= 0.1 and k_{cn} = 30. From top to bottom, the time series are labelled SA, SB, SC and SD. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is given in seconds (s).

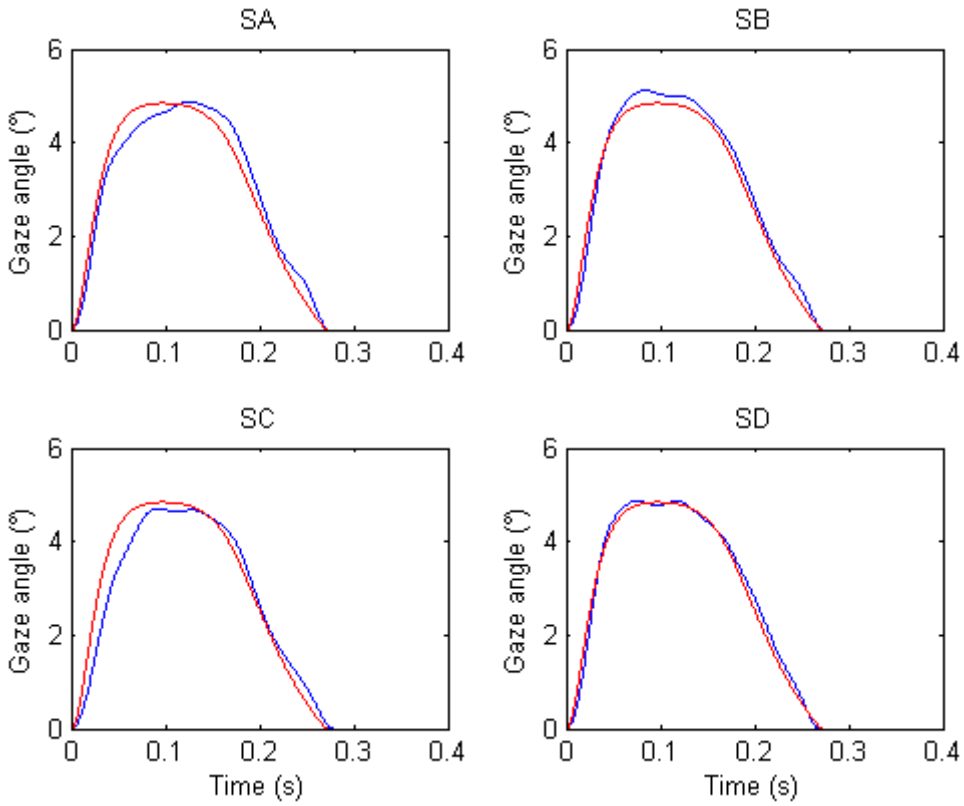

Figure 6.16: UPOs extracted from the synthetic time series SA, SB, SC and SD shown in Figure 6.15 (blue lines). The waveform generated from the stochastic model using the parameter set SD from Table 6.3 with no noise (*i.e.* all stochastic parameters set to 0) is shown in red. The UPOs were extracted using the method described in section 4.2.2. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

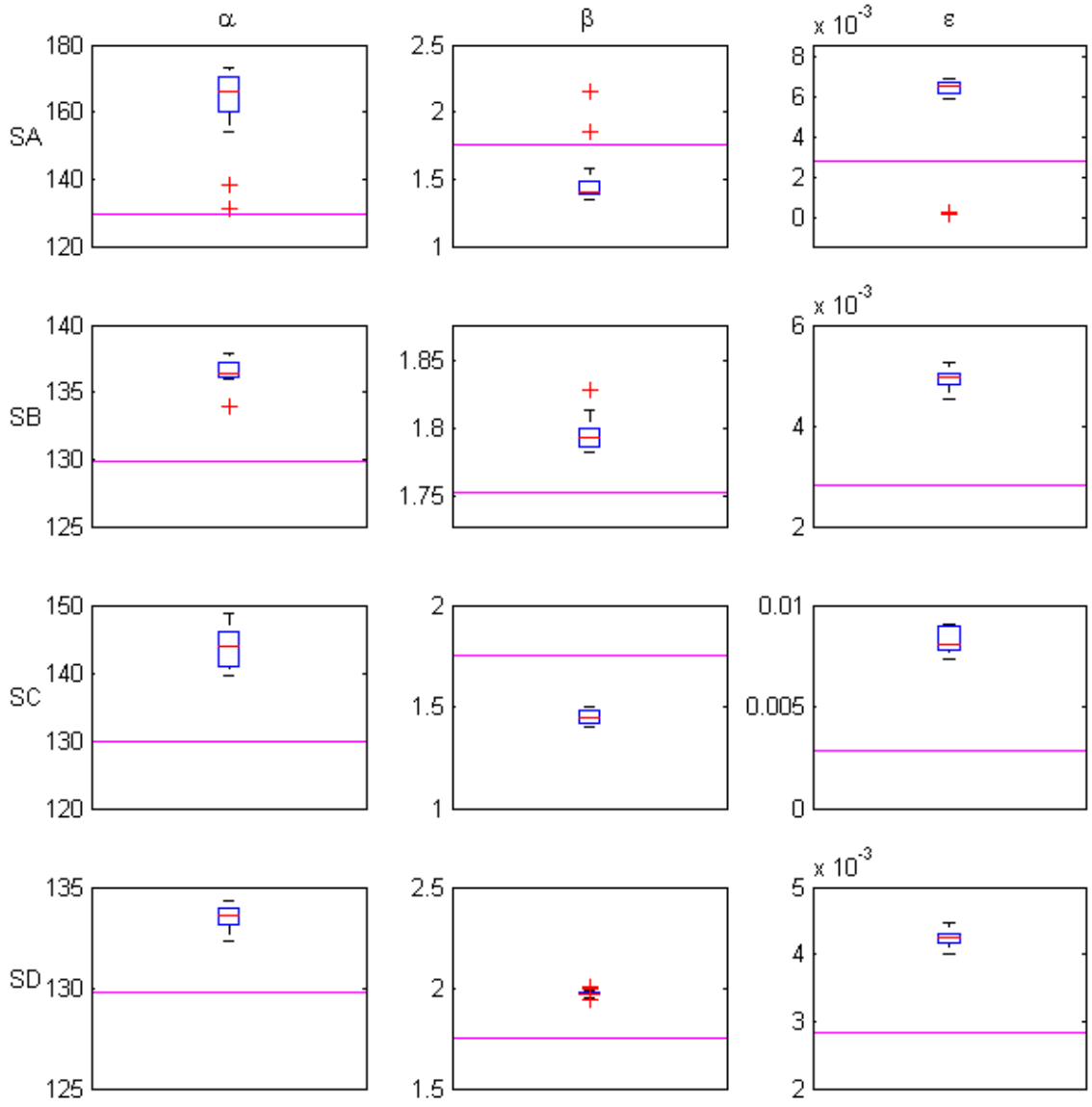

Figure 6.17: Optimised deterministic parameter values for each of the UPOs shown in Figure 6.16. For each UPO, the NSGA-II was run using a population size of 4000. The first line of plots shows the optimised parameter values of α, β and ε for the UPO extracted from the synthetic time series SA, whereas those for SB, SC and SD are shown by the second, third and fourth lines of plots, respectively. The magenta lines in the boxplots show the deterministic parameter values from SD which were used to generate the target waveform in each case (see Table 6.3; the stochastic parameters are given in the legend of Figure 6.15). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

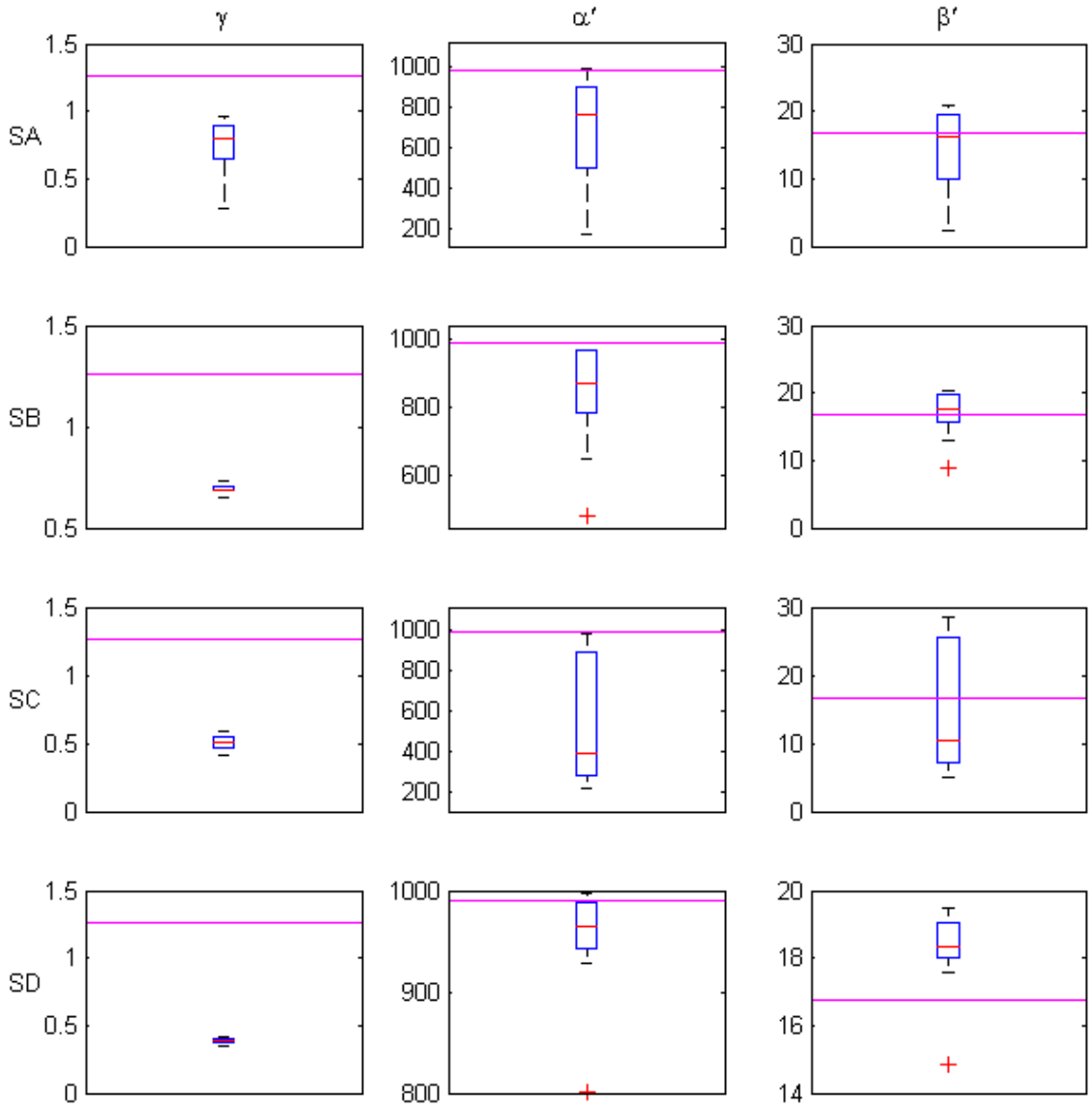

Figure 6.18: Optimised deterministic parameter values for each of the UPOs shown in Figure 6.16. For each UPO, the NSGA-II was run using a population size of 4000. The first line of plots shows the optimised parameter values of γ, α' and β' for the UPO extracted from the synthetic time series SA, whereas those for SB, SC and SD are shown by the second, third and fourth lines of plots, respectively. The magenta lines in the boxplots show the deterministic parameter values from SD which used to generate the target waveform in each case (see Table 6.3; the stochastic parameters are given in the legend of Figure 6.15). The red line in each boxplot denotes the median of the solutions. The edges of each box are the 25th and 75th percentiles. The whiskers extend to the interquartile range and data points outside this range are depicted as red crosses.

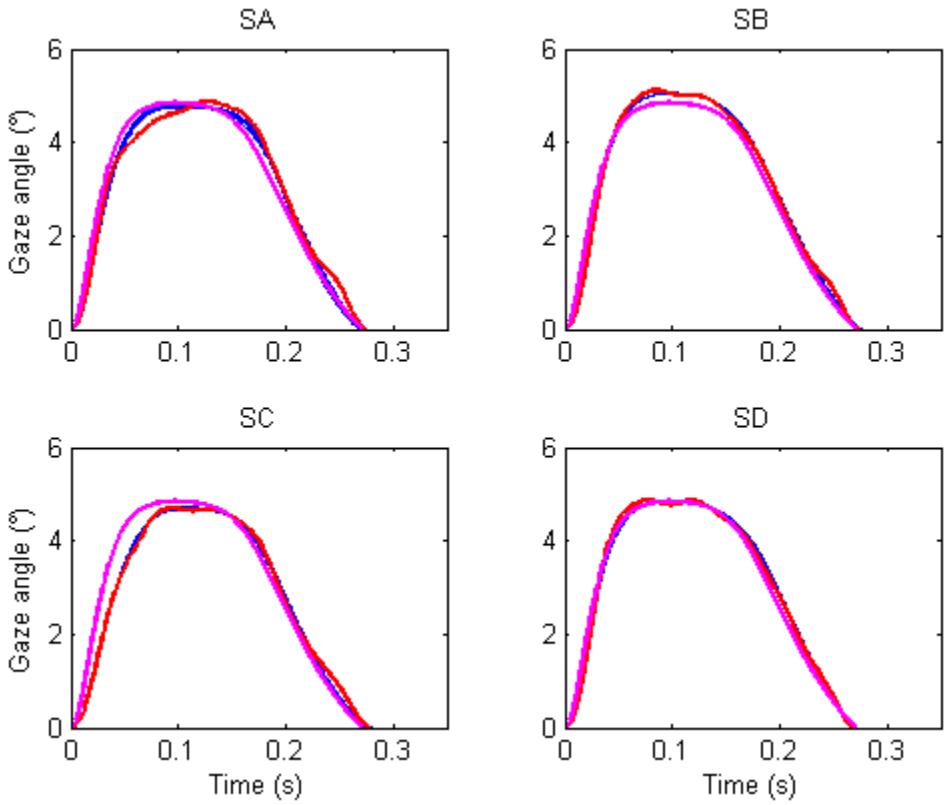

Figure 6.19: Waveforms generated by optimising the deterministic parameters of the stochastic model to the UPOs (shown in Figure 6.16). The target UPO waveforms are plotted in red; the 16 optimised waveforms are plotted in blue. The magenta line represents the waveform generated using the parameter set SD shown in Table 6.3 with no noise (*i.e.* all stochastic parameters set to 0). Starting from top left and going clockwise, the plots show the results obtained for the UPO targets extracted from synthetic time series SA, SB, SC and SD. On each plot, the vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

6.2.6 Saccadic noise fitting

In this section, we will describe the results of fitting the stochastic model to experimental saccadic data. The target data in this case was the coefficient of variation of saccadic endpoints recorded experimentally by [30] for horizontal saccades of amplitude 5, 10 and 15 degrees (see section 4.2.4).

In order to fit the model parameters to the experimental data from [30], we used the optimised deterministic parameters for saccadic velocity profiles presented in

section 5.3.5. From the parameter sets obtained, we used the values shown in the following Table 6.4.

Table 6.4: Deterministic parameter values used in the stochastic model fitting.

Deterministic parameter	α	β	ε	γ	α'	β'
Value	14.32	58.43	0.005826	0.00022	386.07	1.856

These parameters are one of the 16 optimised parameter combinations and they were selected from the found estimated *Pareto* front by selecting the individual with the lowest distance to the axes origin. Having fixed the deterministic parameters, we performed a grid search for the values of the stochastic parameters using the function described in section 4.4.1, to measure the goodness-of-fit.

In order to calculate the coefficient of variation of the simulated saccade, we ran the stochastic model 640 times for each amplitude and stochastic parameter combination. The use of GPUs allowed us to extensively test a wide range of parameter combinations. The optimised stochastic parameter values used in the stochastic model fitting are shown in Table 6.5.

Table 6.5: Stochastic parameter values found by optimising the stochastic model.

Stochastic parameter	k_{sdn}^{p}	k_{sdn}^{n}	k_{cn}
Value	0.0021	0.014	74.2

Figure 6.20 compares the simulated saccadic endpoint coefficient of variation (shown in blue) for saccades of 5, 10 and 15 degrees against the experimental data from [30] (shown in red). The simulated and experimental data almost superimpose. Figure 6.21 shows the saccades generated using the optimised parameter values.

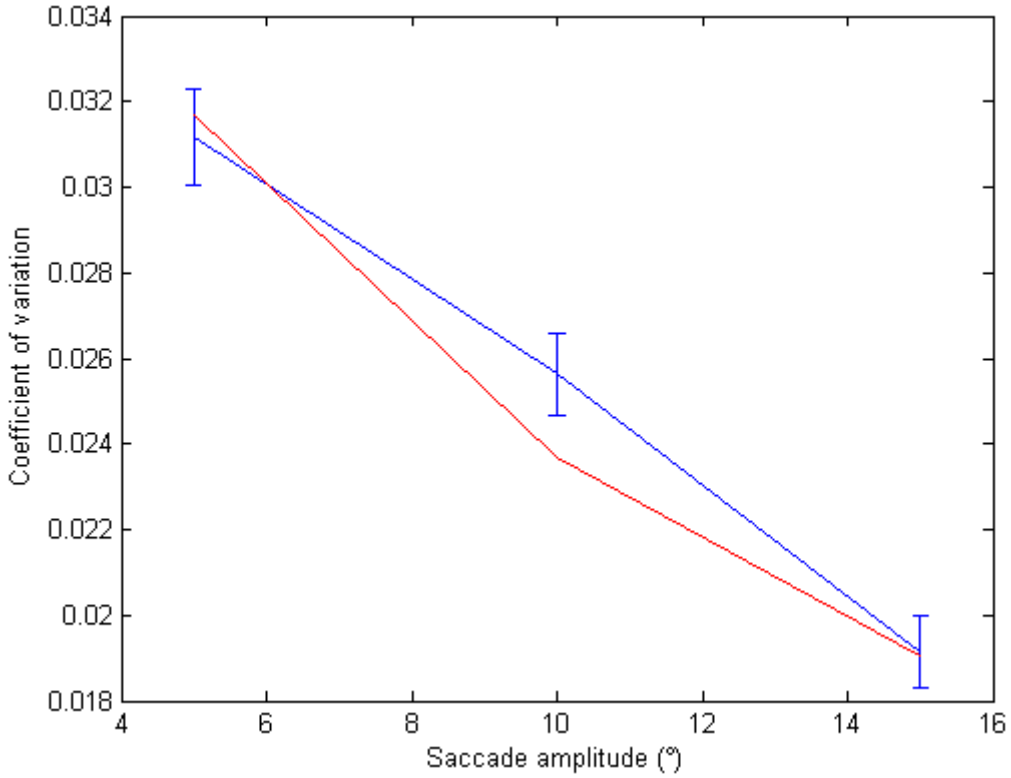

Figure 6.20: Comparison of the experimental saccadic endpoint variability and the simulated endpoint variability generated by the stochastic model for saccades amplitudes of 5, 10 and 15 degs. The endpoint variability is quantified by the coefficient of variation. The experimental data are the mean coefficient of variation (shown in red) calculated across a number of participants from [30] (refer to section 4.2.4). For the model, the mean coefficient of variation (shown in blue) was calculated from 640 simulated saccades for each amplitude (bars show the standard deviation). The deterministic parameter values used for simulating the saccades are shown in Table 6.4. The stochastic parameter values are shown in Table 6.5. On each plot, the horizontal axis represents the saccade amplitude in degrees (°).

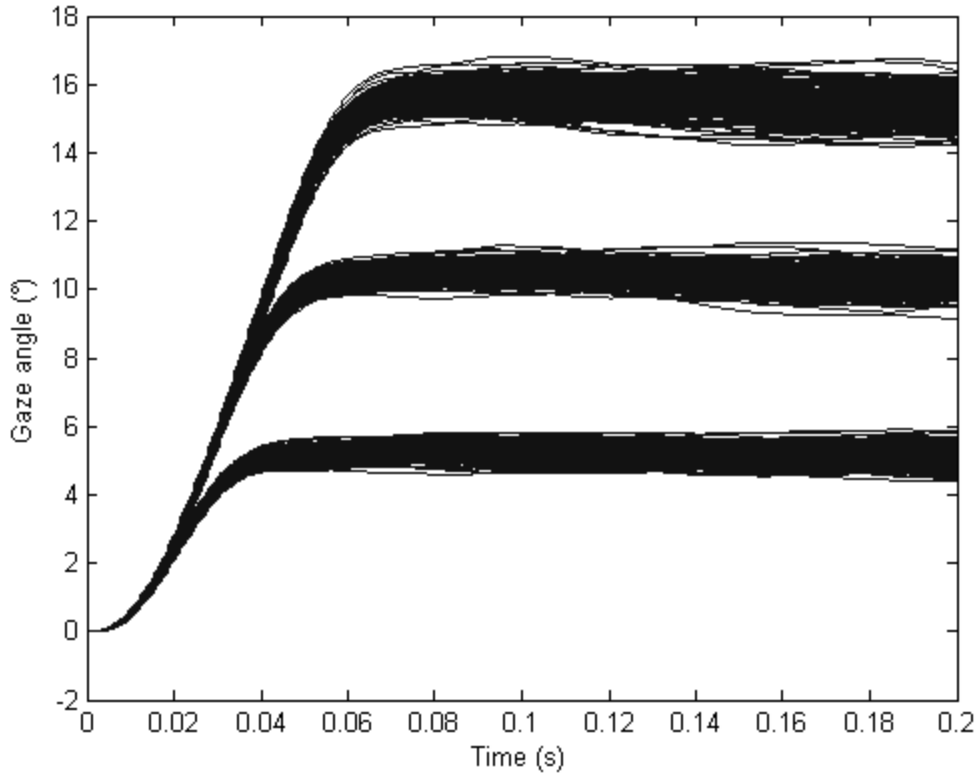

Figure 6.21: Saccades generated by the stochastic model for amplitudes 5, 10 and 15. The deterministic parameter values used for simulating the saccades are shown in Table 6.4. The stochastic parameter values are shown in Table 6.5. The vertical axis represents the horizontal eye position in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

6.2.7 The noise does not affect the saccadic velocity profile generated by the stochastic model

In this section, we will show that the noise processes do not change the characteristics of the saccadic velocity profiles generated by the stochastic model. In order to verify this, we generated 640 simulated saccades for amplitudes of 5, 10 and 15 degs. We used the stochastic model with the deterministic parameter values shown in Table 6.4 and stochastic parameter values used are shown in Table 6.6. These values were chosen as they are high and in case the noise processes affect the saccadic velocity profiles generated by the stochastic model, they would be more likely to do so than lower values. The high stochastic parameter values where chosen to be considerably higher than the values which set the model to generate variability found in the experimental data.

Table 6.6: Stochastic parameter values used to test whether the noise processes affect the velocity profile generated by the stochastic model.

Stochastic parameter	k_{sdn}^{p}	k_{sdn}^{n}	k_{cn}
Value	0.03	0.1	30

The mean gaze angle time series and saccadic velocity profiles calculated from the ensemble of simulated saccades were found to be same as those generated by the stochastic model with the stochastic parameter values set to zero (Figures 6.22 and 6.23). Our results confirm that the noise processes do not change the saccadic velocity profile.

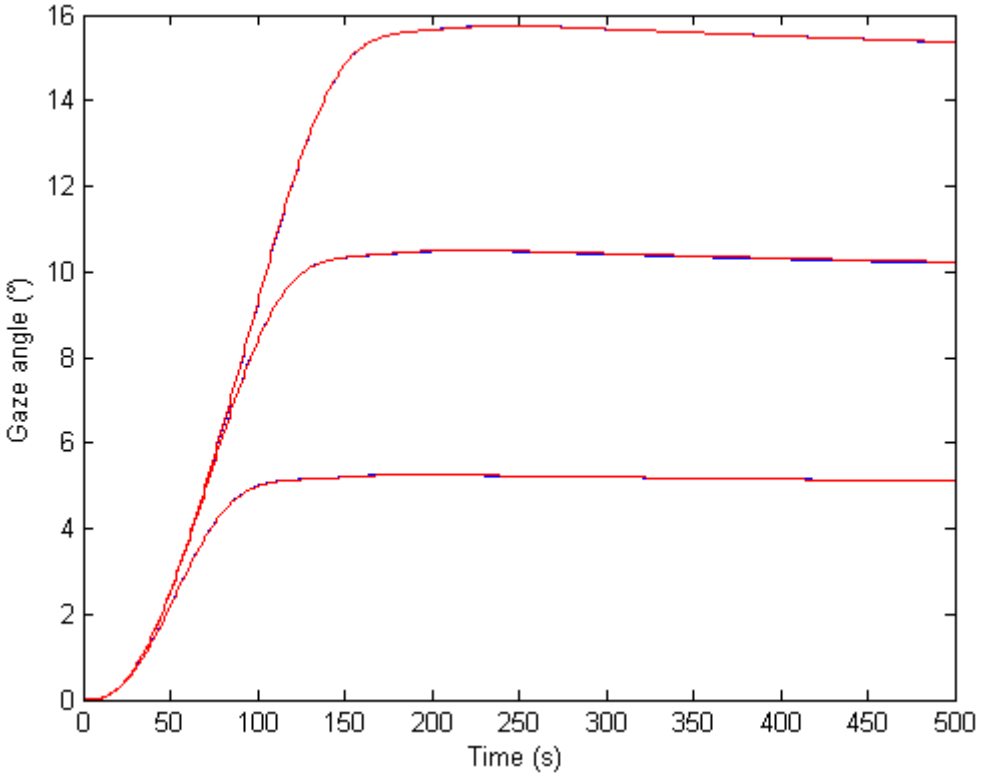

Figure 6.22: Saccades generated by the stochastic model without noise (red) versus the mean gaze angle of the saccades generated with noise (blue). The mean saccadic trajectories were calculated from 640 simulated saccades for each amplitude. The deterministic parameter values used are shown in Table 6.4. The stochastic parameter values that were used are shown in Table 6.6. The vertical axis represents the horizontal eye position in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

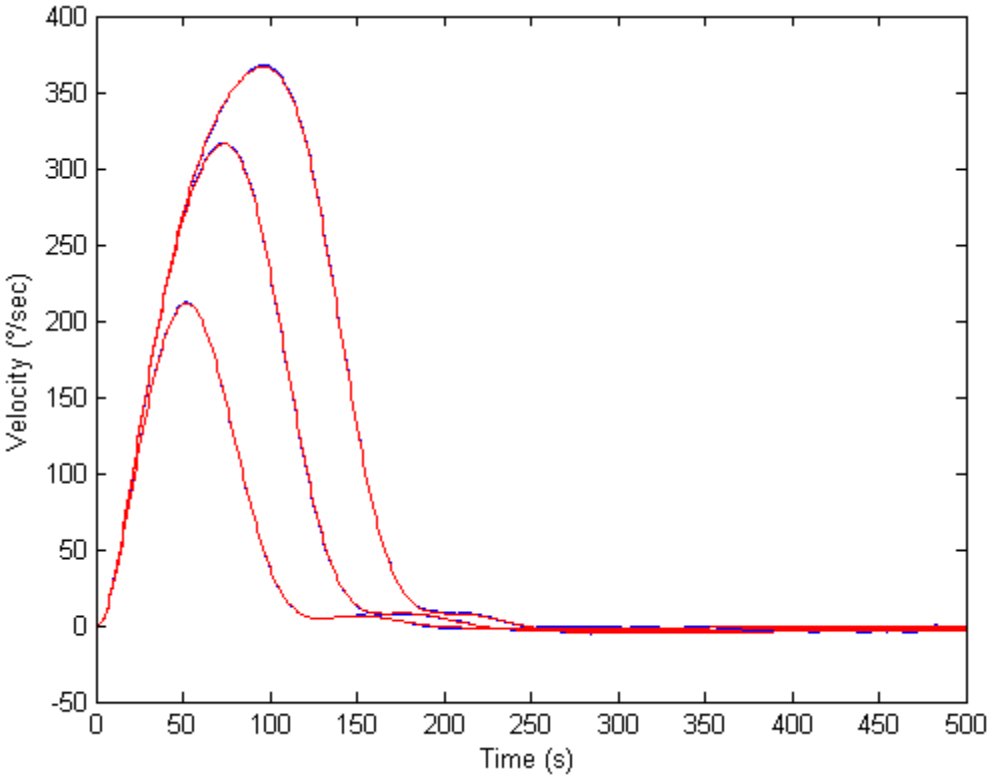

Figure 6.23: Saccadic velocity profiles generated by the stochastic model without noise (red) versus the mean gaze angle of the saccades generated with noise (blue). The mean velocity profiles were calculated from 640 simulated saccades for each amplitude. The deterministic parameter values used are shown in Table 6.4. The stochastic parameter values that were used are shown in Table 6.6. The vertical axis represents the horizontal eye velocity in degrees per seconds (°/s), with positive values denoting rightward eye velocities. Time is in seconds (s).

6.3 Discussion

In this Chapter, we have presented the stochastic saccadic model, which is based on the Broomhead *et al.* [1] model. The model incorporated SDN and CN in the control signals generated by the neural integrator and burst neurons (equation 6.1), and we explored different noise models by varying the magnitude of SDN and CN on the control signals.

We have demonstrated that in the case of nystagmus waveforms, the stochastic model can cause variations in the local dimensions across different parts of the waveform (fast phase, fixation or slow phase). Moreover, the simulations show that SDN and CN on both the neural integrator and burst neurons signals set the

stochastic model to generate time series with similar local dimensions pattern to that measured in the experimental data (Figure 6.13). The principal difference between the simulated and the fast phase (at the end of the slow phase) start to decrease sooner in the fast phase (Figure 6.13).

We hypothesise that one of the factors underlying this disparity between the simulated and experimental data is the inability of the stochastic model to generate the exact experimental waveform shape. This is demonstrated in Figure 6.24.

The velocity at the start and at the end of the fast phase (points 1 and 2 in Figure 6.24) is different in the simulated and experimental waveforms. In particular, the velocity during the fast phase of the simulated waveform starts to decrease much sooner compared to the experimental UPO (point 2 in Figure 6.24). This may cause the earlier decrease in the simulated local dimensions due to the effect of SDN. Moreover, at the fixed point there are oscillations that are not reproduced in the simulated waveform. These oscillations are consistent across all the time series cycles, suggesting that they should not be attributed to noise. Furthermore, during the slow phase of the experimental UPO, the eye exits the fixed point with larger velocity compared to the simulated waveform (point 3 in Figure 6.24). Lastly, at the end of the slow phase, the simulated waveform shows a decrease in velocity which is not observed in the experimental waveform (point 4 in Figure 6.24). This may cause the greater increase in local dimensions at the end of the slow phase compared to the experimental data (Figure 6.13). We propose that these difference in the characteristics of the experimental waveform could be generated by oculomotor processes that are absent from the stochastic saccadic model, such as the smooth pursuit system, a visual target feedback system, different neural integrator properties or different inhibition types between the right and left burst neurons. However, this has not been explored in this work and requires further investigation.

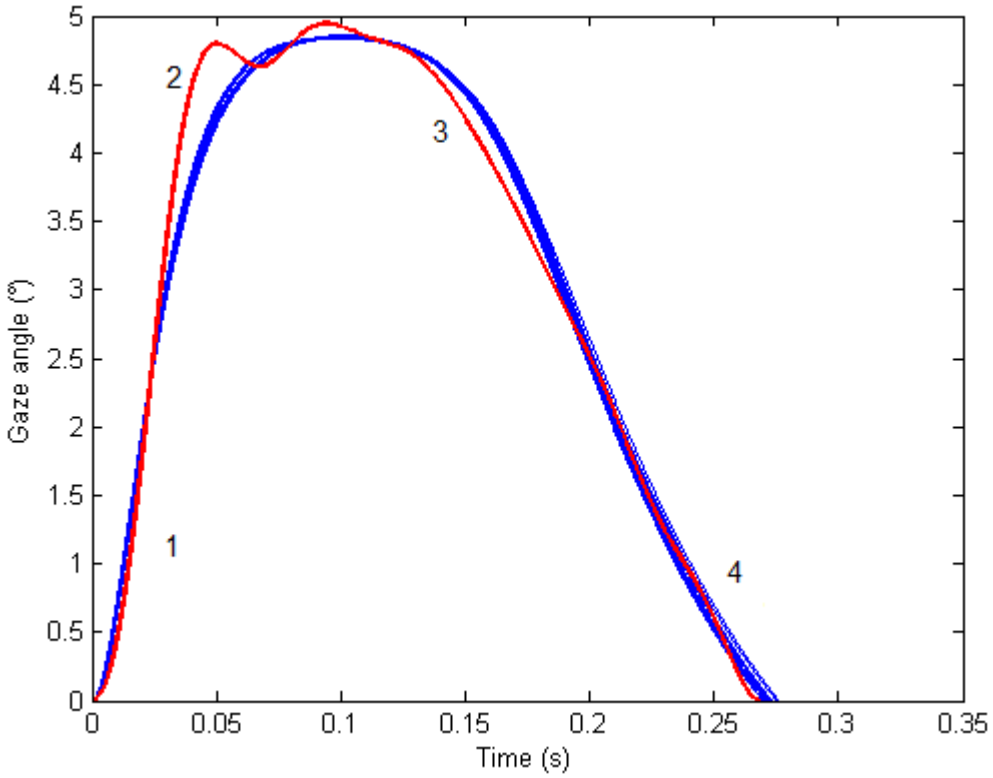

Figure 6.24: Waveforms generated by optimising the deterministic parameters of the stochastic model to UPO extracted from the experimental nystagmus waveform shown in Figure 6.3. The target UPO is plotted in red; the 16 optimised solution waveforms are plotted in blue. The areas labelled 1 to 4 denote waveform phases at which the fit of the stochastic model to the experimental data is least accurate. The vertical axis represents the horizontal eye gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

Moreover, the generated time series in Figure 6.15 show the non-deterministic complex refixation mechanism that is found in experimental data [15]. Furthermore, the same time series demonstrates a general oscillation of the time series that resembles the baseline oscillation (refer to section 1.8.2 and see Figure 1.8).

In the case of normal saccades, we have shown that the model is able to generate similar saccadic endpoint variance to that measured experimentally. We suggest that a direct comparison between the stochastic parameter values obtained for nystagmus and saccade fitting cannot be performed, as the results presented in this work show that the model cannot reproduce both the exact nystagmus oscillations (see Figure 6.24) and the exact saccadic velocity profiles with a single

set of deterministic parameters (see section 5.3.5). We hypothesise that such a comparison could be made after further development of the stochastic model, enabling the model to generate more accurate fits to experimental data (nystagmus oscillations and saccadic metrics) than those shown here.

Lastly, we have shown that the noise processes do not affect the key characteristic of the nystagmus waveforms and saccadic velocity profiles (*i.e.* those used for parameter optimisation). This confirms that the method we used to fit the stochastic model's deterministic and stochastic parameters separately can provide sufficiently accurate results. Moreover, these results are consistent with the form of the stochastic model. The equations that drive the system in the case of the Broomhead *et al.* [1] model are (2.4)-(2.6) (refer to section 2.2). These equations are the response of the right and left population of neurons, and the feedback of the motor error. We assume that the noise processes would change the dynamics of the model if they were a part of the feedback of the stochastic model. However, the noise processes are in the stochastic model equation (6.1) which is the extension of equation (2.1) and that is not a part of the feedback mechanism of the Broomhead *et al.* [1] model (refer to sections 2.2 and 6.1).

7 Conclusions and future work

The study of normal and abnormal eye movements is a non-invasive way to decode the functioning of the brain. The use of nonlinear dynamics techniques to model and analyse the oculomotor system provide a complementary approach to control theory. Moreover, the use of optimisation methods for parameter estimation allows model testing, further analysis of the behaviour of the modelled system and the formulation of predictions.

In this Chapter, we will present the conclusions of the study, and propose further directions for the work presented here.

7.1 Optimisation of the Broomhead *et al.* model

The parameter optimisation method involved the parallel runs of independent NSGA-IIs (independent parallel NSGA-II model, refer to sections 3.5 and 4.3.5), where the integration of the oculomotor model is calculated by the compute units in a GPU (“master-slave” NSGA-II model, refer to sections 3.5 and 4.3.2).

There are two main advantages of using an NSGA-II implemented on a CPU-GPU architecture. Firstly, the use of MOGAs allows one to address some of the difficulties arising from optimisation of biological models (refer to section 3.2). These include the existence of multiple local optima, difficulty in finding an exact rather than approximate solution, the non-smoothness of the objective function and the need to simultaneously optimise multiple objectives. The NSGA-II is a widely used MOGA optimisation method that addresses all of these difficulties. Secondly, the use of GPUs to integrate the model for very large number of parameter sets (>30000) provides a speedup of ~30 compared to a mid-range CPU. Moreover, as a population of 30000 is not required for the NSGA-II to converge to the global optimum, the developed parallel model allows one to run multiple NSGA-II instances using only one GPU (refer to sections 4.3.5 and 4.3.6). This allows full use of the GPU with considerable speedup, depending on NSGA-II population size. When the NSGA-IIs run serially, the speedup is relatively low.

In order to use the GPU for the integration of the model, we implemented the Implicit Euler method in OpenCL (refer to section 4.3.3). The use of the GPU provided processing power that considerably decreased the model optimisation time (refer to section 4.3.4). That translates into more time available for developing the appropriate fitness function and testing the NSGA-II on a larger number of experimental datasets. For example, the time required to test one fitness function for five different populations (500, 1000, 2000, 4000 and 8000) and for four different waveform targets is ~99 hours (4.2 days) using one GPU card, whereas one mid-range CPU requires ~2970 hours (124 days). Another advantage of the GPU technology is that up to four GPUs can be fitted into an off-the-self PC at a lower cost compared to an HPC cluster.

7.1.1 Fitting nystagmus waveforms

In the case of experimental nystagmus waveform fitting, our optimisation method included the extraction of the UPOs from the experimental time series (refer to section 4.2.2). The extracted UPOs were used as target waveforms for the parameter estimation. The fitness function was based on the shape and period size of the waveforms (refer to section 5.1.1). The selection of the most suitable solution from the approximation to the *Pareto* front was carried out by normalising the objectives values to be within the interval [0,1]. Subsequently, we added the normalised values for each objective and chose the individual with the lowest sum. The selected solution has the best trade-off between the difference in period and the difference in shape between the target and simulated waveform (refer to section 5.2.1).

The results of using this particular optimisation method on experimental waveforms showed that the most influential parameters are α and β. These regulate the off-response (saccadic braking signal) of the burst neurons, which supports the original hypothesis of van Gisbergen *et al*. [25] that irregularities in the braking signal can cause oscillations. Moreover, optimised values of the parameter γ, which controls the strengths of mutual inhibition between the left and right burst neural populations, were much larger (values 2-10) than the value used by Broomhead *et al.* [1], which was 0.05.

Interestingly, fitting the Broomhead *et al.* [1] model to UPOs extracted from different parts of a single experimental times series recording from a single

participant yielded quite different parameter values. This shows that the parameters that set the dynamics of the burst neurons could change evenly over small time frames. Perhaps, this could mean that the dynamics of the burst neurons is modulated by different brain areas (*e.g.* premotor cortical areas) [9].

Furthermore, our results show that jerk, jerk with extended foveation, asymmetric pseudo-cycloid and pseudo-cycloid oscillations can be simulated by the Broomhead et al. [1] model (see Figure 5.40 and Figure 7.1). Moreover, the fitting results showed which oscillations and waveform characteristics the model can simulate. In the case of jerk oscillations, we found that there are key differences between the simulated and experimental nystagmus waveforms at the fast phase, at the foveation point, and at the start and end of the slow phase (refer to section 6.3, Figure 6.24). This provides future directions according to which the model could be extended or a new model developed. This methodology is applicable to any oculomotor model, enabling the predictive capacity of the model to be quantified and highlighting possible modifications that would enable a broader class of behaviours to be reproduced. In the case of the stochastic model presented in this work, the model could incorporate extra modifications that could allow it to generate the jerk waveform characteristics which, in its current form, is unable to do so (see section 7.4.2).

7.1.2 Fitting saccadic velocity profiles

In the case of experimental saccadic velocity profile fitting, our method applied velocity thresholding to extract the sections of the velocity profiles in order to use them as targets in the parameter optimisation method (refer to section 4.2.3). The multi-objective fitness function employed evaluated differences between the simulated and experimental velocity profiles for saccade amplitudes of 5, 10 and 20 degrees (refer to section 5.1.2).

The findings from the application of our method to experimental saccadic velocity profiles showed that the most influential parameters are α' and β', which regulate the on-responses of the burst neurons, and the parameter ε, which regulates the speed of the neurons' response to the motor error. Moreover, our results show that the model cannot reproduce the experimentally measured velocity profiles for all three amplitudes simultaneously with one set of parameter values. This confirms the hypothesis of van Opstal and van Gisbergen [22] that the

parameters affecting burst neuron dynamics can change based on the amplitude of the saccade.

7.2 Optimisation of the stochastic model

The nonlinear time series analysis of Akman *et al.* [15] on experimental jerk nystagmus data showed that there is a specific relationship between local dimensions and waveform phase (refer to section 6.2.3). Akman *et al.* [15] found that the local dimensions is highest at the fast phase of the waveform and lowest at the slow phase. However, at the fixed point corresponding to foveation, the local dimensions number is intermediate between the fast phase and slow phase values. Akman *et al.* [15] hypothesised that this pattern could be either due to the existence of signal dependent noise in the neuronal signal or the activation of an additional oculomotor system.

It is known from previous studies [30], [31], [34], [112], [113] that SDN and CN play an important role in the dynamics of saccades. Based on these studies, this work tested the hypothesis that an extended version of the Broomhead *et al.* [1] model, which includes SDN and CN on the burst neuron and neural integrator signals to the muscle plant (refer to section 6.1, equation 6.1), could generate similar local dimensions variability pattern as found in experimental jerk nystagmus.

As stated at the beginning of this chapter, having optimised the deterministic parameters of the extended model to UPOs extracted from experimental jerk oscillations, we then optimised the stochastic parameters by comparing the pattern of simulated local dimensions changes with those observed in the experimental data (refer to section 6.2.4). The results indicated that the stochastic model can generate similar patterns as those measured experimentally, particularly when there is both SDN and CN on the burst and neural integrator control signal. In this case, the simulated local dimensions increase during the fast phase, but decrease much sooner compared to the dimensions calculated from the experimental data (see Figure 6.14). Moreover, although there is a substantial decrease in local dimensions at the start of the slow phase which matches the experimental data, the simulated local dimensions increase much faster at the end of the slow phase, just before the oscillations enters the fast phase (see Figure 6.14). This may be due to the inability of the model to generate

all the characteristics of the experimental jerk waveforms shapes. The major differences between the simulated and experimental oscillations occur at four locations on the waveform (see Figure 5.45 points 1-4). The first location is during the fast phase, where in the case of the experimental data, the eye begins to decelerate closer to the fixed point, whereas the stochastic model decelerates much earlier (see Figure 5.45 point 1). The second location is at the fixed point. The experimental data shows much more variability in the eye position, which might not be generated by noise, but caused instead by an overshoot of the eye (Figure 5.45 point 2). The third different location is during the slow phase, where in the experimental data, the eye moves away from the fixed point in a more linear path compared to the model (Figure 5.45 point 3). The last location is just before the start of the fast phase where, as for point 3, the experimental data follows a more linear path compared to the model (Figure 5.45 point 4).

We hypothesise that different oculomotor processes need to be included in the model in order to reproduce the experimental jerk waveform characteristics mentioned above and corresponding local dimensions pattern. A visual target feedback system or smooth pursuit system could be included as extra components. Moreover, the neural integrator could play a role in the dynamics of the oscillations and could be modelled in a more sophisticated manner that matched experimental eye movements.

Using the stochastic model to simulate normal saccades resulted in saccadic endpoint variabilities similar to those measured in experimental data (refer to section 6.2.6).

Lastly, we have shown that the stochastic model can: (i) generate nystagmus waveforms yielding UPOs that are very similar to the solutions of the corresponding deterministic model; and (ii) generate saccades with the same mean velocity profile as the corresponding deterministic model. This suggests that noise does not significantly affect the key nystagmus waveform characteristics (refer to section 6.2.5) or the shape of the velocity profile (refer to section 6.2.7); hence, the bipartite method is a viable approach for fitting the deterministic and stochastic parameter separately.

7.3 Comparison of the optimised stochastic parameter values for saccades and jerk nystagmus waveforms

For saccades, the stochastic parameter values that allowed the model to generate saccadic endpoint variances similar to those measured from experimental data were found to be: k_{sdn}^{b}= 0.0021, k_{sdn}^{n}= 0.014 and k_{cn} = 74.2 (refer to section 6.2.6). For nystagmus oscillations, the stochastic parameter values that provided the closest match to experimental data were found to be: $k_{sdn}^{b} = 0.005$, $k_{sdn}^{n} = 0.1$ and $k_{cn} = 50$ (refer to section 6.2.4.4). Despite the fact that there are significant differences in the stochastic parameter values for nystagmus and normal saccadic eye movements, strong conclusions cannot be made on this occasion. The reason for that is that the stochastic model cannot generate the exact nystagmus waveforms and saccadic velocity profiles observed experimentally using one set of deterministic parameter values. Thus, it does not allow for direct comparison between the stochastic parameter values.

In the next section, we will outline future work such as further developments of the Broomhead *et al.* [1] model and model fitting methods which may contribute to gaining a deeper insight into the functioning of the oculomotor system. Thereafter, we will provide a number of preliminary results for a model that incorporates visual target information, which could provide behaviour closer to that observed in the experimental data.

7.4 Future development of the model

The future work we propose includes: (i) further development of the parameter optimisation methods; (ii) further development of the Broomhead *et al.* [1] model; (iii) an oculomotor model incorporating learning processes; and (iii) further development of the OpenCL ODE solvers.

7.4.1 Further development of the parameter optimisation methods

During the development of the nystagmus fitness function, we developed a new method for classifying the waveform shapes that can be generated by the Broomhead *et al.* [1] model. This classification method is based on the number and location of the local minima and maxima of the 1st and 2nd derivatives of one period from the nystagmus oscillation (Figure 7.1). These local minima and maxima are the points in the oscillation where there is a **change in velocity**

direction (CVD) and a **change in acceleration direction** (CAD). Figure 7.1 shows the CVD and CAD points for four waveforms generated by the Broomhead *et al.* [1] model. The number of CVD and CAD can classify the waveform, allowing to determine whether two waveforms are qualitatively the same. The first oscillation in Figure 7.1 is pendular, the second and third are jerk, and the fourth is bias reversal. The pendular oscillation has two CVD points and seven CAD points, the jerk oscillations have two CVD points and five CAD points, whereas the bias reversal oscillation has six CVD points and seven CAD points. In order to compare the similarity between two waveforms with the same number of CVDs and CADs, we use the following cost equation:

$$f = \frac{\sum_{i=1}^{W} \frac{\|msv_i - mtv_i\|_1}{\max(\|msv_i\|_1, \|mtv_i\|_1)}}{\mathrm{W}} + \frac{\sum_{j=1}^{K} \frac{\left\|msv_j - mtv_j\right\|_1}{\max\left(\left\|msa_j\right\|_1, \left\|mta_j\right\|_1\right)}}{\mathrm{K}}, \tag{7.1}$$

where: (i) W and K are the number of local minima and maxima for the first and second derivatives, respectively; (ii) mtv_i and msv_i specify the location and magnitude of the first derivative local minima/maxima for the target and synthetic waveforms, respectively; and (iii) mta_j and msa_j specify the location and magnitude of the local minima/maxima for the second derivative. By using this graded scoring system, we were able to rate a different number of waveforms in terms of similarity. Unfortunately, initial applications of the classification scheme to experimental data showed that it cannot be used for noisy experimental waveforms. Nevertheless, we could use the scheme to systematically differentiate between the different waveform types in a model's parameter space. The use of the GPU would allow us to integrate the model for numerous (tens of thousands) parameter combinations in a relatively short period of time. By using methods which allow us to visualise multiple dimensions (such as seriated heat maps), a better understanding of how different waveform types are distributed in the model's parameter space could be acquired. Equation (7.1) could also be used for seriation of the solutions and model parameters (in terms of similarity), and could group combinations of parameters that generate similar waveforms. This may potentially show which parameters affect the qualitative behaviour of the model.

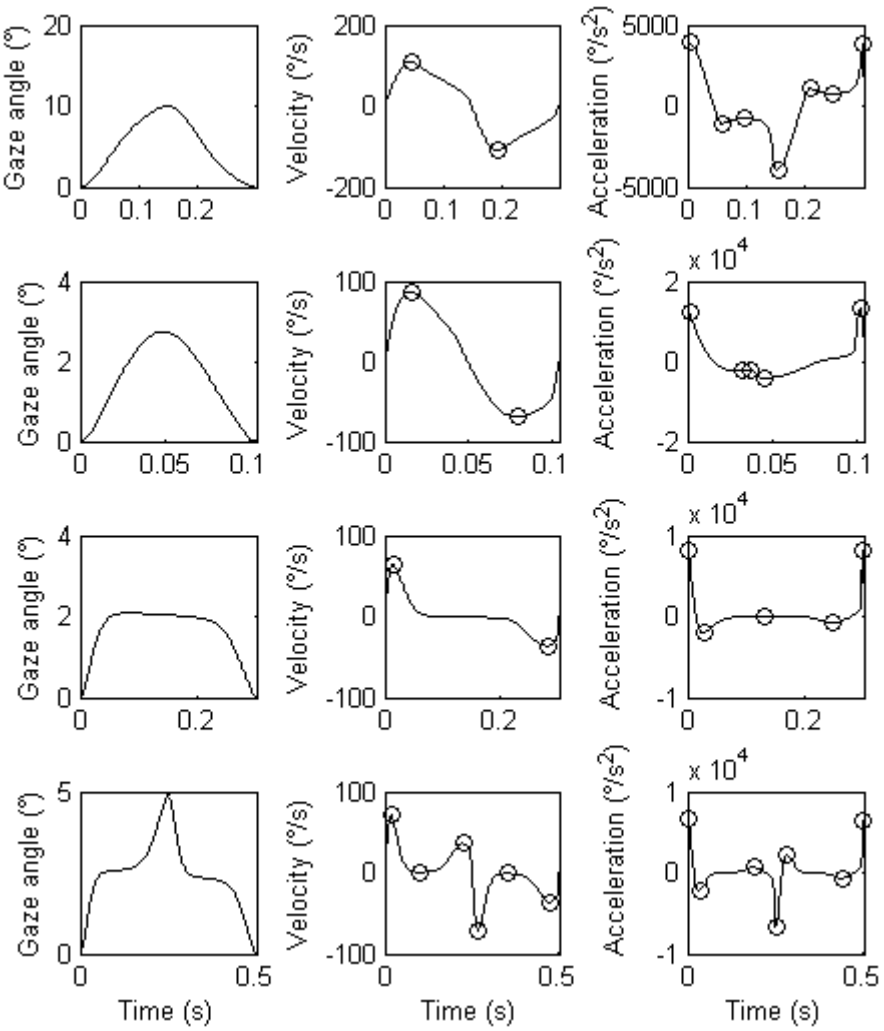

Figure 7.1: Synthetic nystagmus waveforms (eye position, velocity and acceleration) generated by the Broomhead *et al.* [1] model. The left column shows the gaze position, the middle column shows the first derivative (velocity) and the right column shows the second derivative (acceleration). The circles represent the local minima and maxima of the first and second derivatives (CVDs and CADs, respectively). The parameter values in each case are: **First row**: Pendular, CVDs: 2, CADs: 7. $\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.05$, $\gamma = 0.05$, $\alpha' = 600$ and $\beta' = 9$. **Second row**: Jerk, CVDs: 2, CADs: 5. $\alpha = 210$, $\beta = 1.5$, $\varepsilon = 0.002$, $\gamma = 0.03$, $\alpha' = 380$ and β'=6. **Third row**: Jerk, CVDs: 2, CADs: 5. $\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.0035$, $\gamma = 0.05$, $\alpha' = 600$ and $\beta' = 9$. **Fourth row**: Bias reversal, CVDs: 6, CADs: 7. $\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.0065$, $\gamma = 0.07$, $\alpha' = 550$ and $\beta' = 9$. On each plot, the vertical axis represents the horizontal eye gaze angle in degrees (°), with positive values denoting rightward eye positions. The time is given in seconds (s).

In the stochastic model that we have presented in this work (refer to section 6.1), we showed that for the noise levels that we have considered, the noise does not significantly affect the shape of the nystagmus waveforms or saccadic velocity profiles (refer to sections 6.2.5 and 6.2.7), justifying the separate optimisation of the stochastic and deterministic parameters. However, other potential formulations of the stochastic model may not have this characteristic. Hence, an extension of the optimisation method presented here should be developed which would enable all the parameters of a stochastic model of the oculomotor system to be fitted simultaneously. It has been shown that noise in optimisation causes uncertainties to the process [114]. To avoid such problems in optimisation using MOGAs, a resampling and a probabilistic ranking system has been proposed [115]. Moreover, a MOGA developed for stochastic multi-objective problems could also be applied [116]. However, regardless of the optimisation method used, the UPO from each stochastic time series has to be extracted and compared with the UPO extracted from the experimental time series in order to calculate fitness. In such cases, the method of So *et al.* [98] could be applied. This would require the model to be integrated for more than 6 s, causing an increase in computation time. To decrease the integration time, a more elaborated stochastic solver could be used such as with step size control [117]. Such a solver would potentially decrease the integration time step without causing numerical instability.

Finally, the optimisation methods could be used to examine how the parameters of the model evolve over the duration of a nystagmus time series recording. During a single experimental recording, the nystagmus waveform characteristics can change due to, for example, variations in different attention levels and gaze angle [58]. The use of optimisation methods on such recordings could reveal how the parameter values change, forming a path in the model's parameter space, and the relation between these changes and the model's bifurcation structure. The association of specific model parameter paths to nystagmats groups could potentially link underlying causes to the formulation of nystagmus in each group.

7.4.2 Further development of the Broomhead *et al.* model

The Broomhead *et al.* [1] model could be extended by incorporation of other oculomotor systems, such as smooth pursuit and/or by the addition of a system

which provides the location of the visual target during the saccade. Moreover, a more precise version of the neural integrator could be modelled, that better matches experimental findings.

7.4.2.1 Incorporation of a visual target feedback system

It would be anticipated that incorporation of a visual target system should allow the model to generate different behaviour, based on the gaze angle and saccade amplitude, enabling a wider range of oscillation shapes to be simulated. We consider a simple expanded version of the model incorporating the target, in which the motor error equation (2.6) (refer to section 2.2) is replaced by the following:

$$\frac{dm}{dt} = -\left(\frac{m}{T_{rb}}\right) - (r - l) + \frac{T - g}{T_{rb}}, \tag{7.2}$$

where, T represents the visual target's location in degrees, $\frac{T-g}{T_{rb}}$ describes the visual feedback of the target position relatively to the eye position g and the time constant T_{rb} sets the feedback signal's rate of deterioration.

We have explored a small number of different parameter values and generated a number of preliminary results (Figure 7.2-7.5). The parameter values we used to generate the time series were as follows:

$\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.0065$, $\gamma = 0.05$, $\alpha' = 600$, $\beta' = 9$ and $T_{rb} = 0.5$.

The initial condition in each case was

$$g_0 = v_0 = n_0 = r_0 = l_0 = m_0 = 0,$$

which allowed to simulate a saccade from rest.

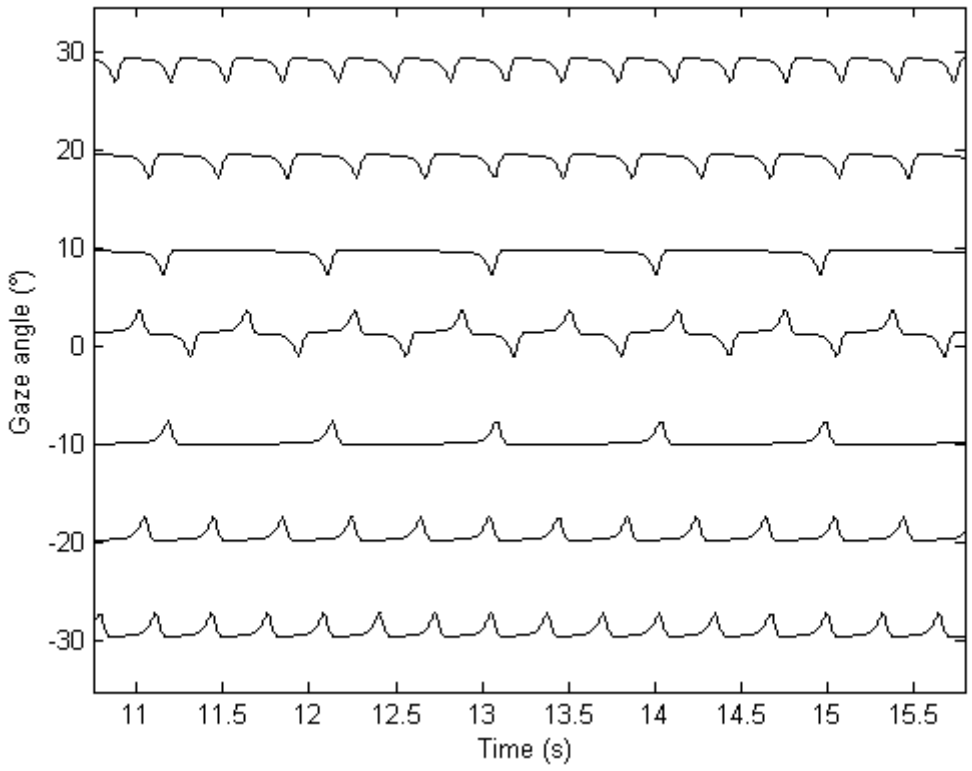

Figure 7.2: Nystagmus oscillations for gaze angles -30, -20, -10, 0, 10, 20 and 30 degrees generated using the proposed visual target model (see equation (7.2)). The parameter values used were: $\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.0065$, $\gamma = 0.05$, $\alpha' = 600$, $\beta' = 9$ and $T_{rb} = 0.5$. The simulation for a given gaze angle was obtained by fixing the target position *T* at that value -30, -20, -10, 0, 10, 20 and 30 degrees). The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

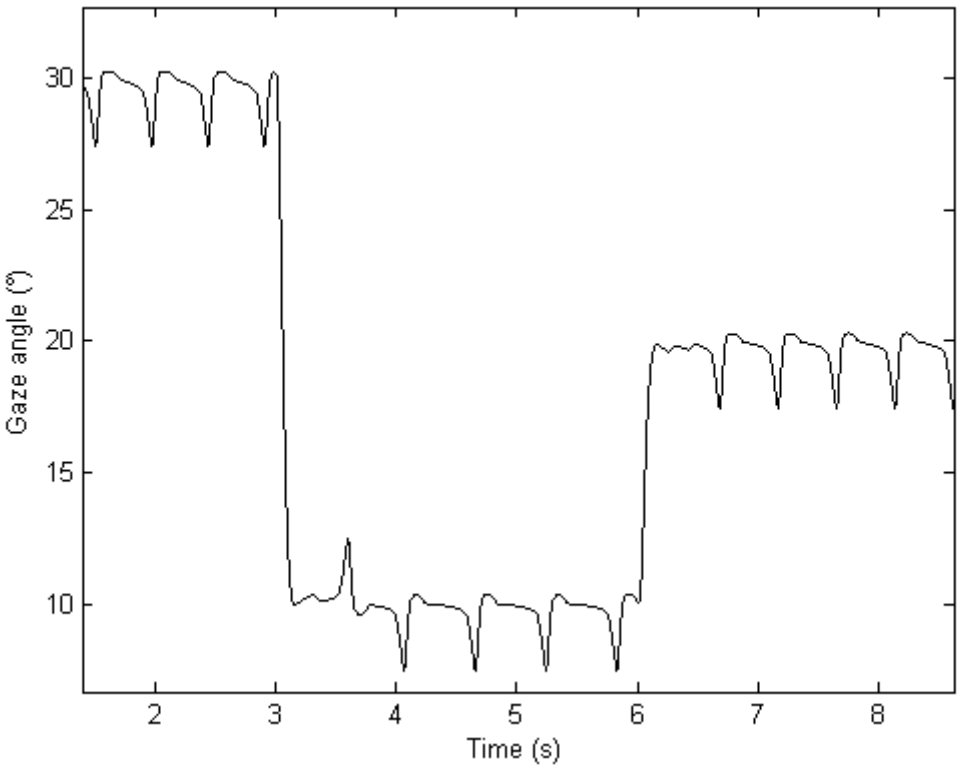

Figure 7.3: Simulated change in target position from 30 to 10 to 20 degs obtained using the proposed visual target model (see equation (7.2)). The parameters values used were: $\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.0065$, $\gamma = 0.05$, $\alpha' = 600$, $\beta' = 9$ and $T_{rb} = 0.5$. The vertical axis represents the horizontal eye gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

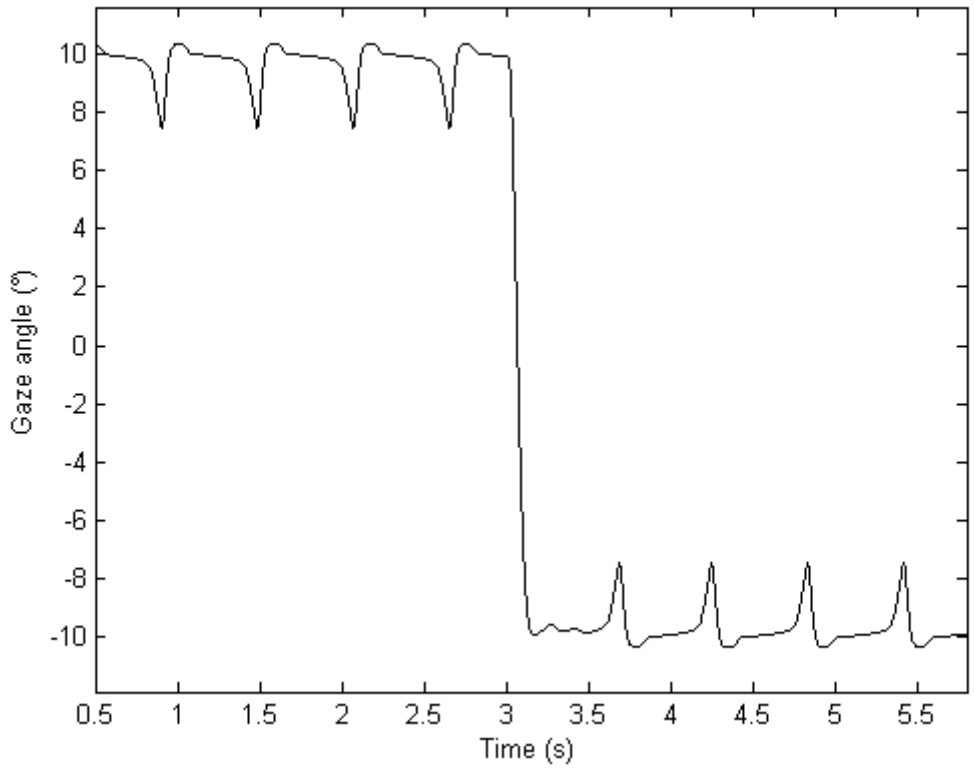

Figure 7.4: Simulated change in target position from 10 to -10 degs obtained using the proposed visual target model (see equation (7.2)). The parameter values used were: $\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.0065$, $\gamma = 0.05$, $\alpha' = 600$, $\beta' = 9$ and $T_{rb} = 0.5$. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

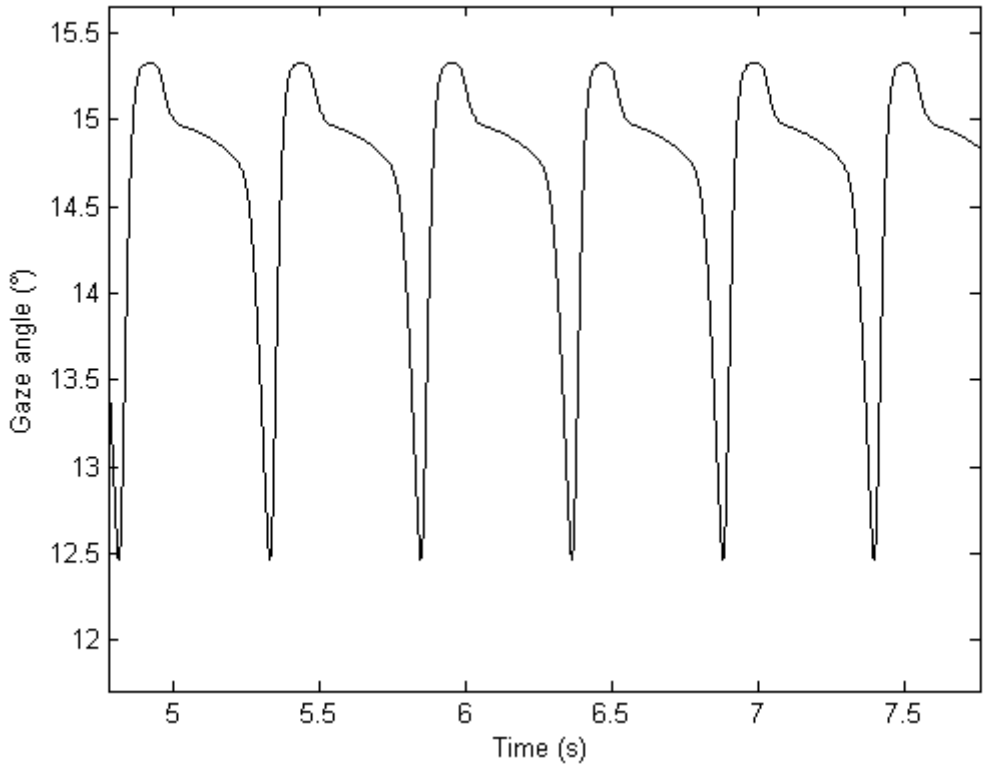

Figure 7.5: Simulated waveform shape for gaze angle 15 degs obtained using the proposed visual target model (see equation (7.2)). The parameter values used were: $\alpha = 110$, $\beta = 1.5$, $\varepsilon = 0.0065$, $\gamma = 0.05$, $\alpha' = 600$, $\beta' = 9$ and $T_{rb} = 0.05$. The vertical axis represents the horizontal gaze angle in degrees (°), with positive values denoting rightward eye positions. Time is in seconds (s).

The preliminary results indicate that the model can generate different oscillation shapes based on the gaze angle (Figure 7.2), as demonstrated in experimental data studies [58]. Moreover, during changes of gaze angle on the same visual side there is no flip in the oscillation direction (Figure 7.3), and this again is consistent with the experimental recordings [118]. Interestingly, the model can generate extra complexity at the fixed point (Figure 6.24), namely an overshooting eye movement is generated at the end of the fast phase (Figure 7.5). These preliminary results show that this is an improvement to the original Broomhead *et al.* [1] model that allows the simulation of additional experimentally recorded nystagmus oscillations. Future experiments will aim to fit this model to the experimental data using the methodology developed in this work. It will be

interesting to find out whether the addition of SDN and CN to this model would yield local dimension variations closer to those seen in the experimental data.

7.4.2.2 Neural integrator modelling

The neural integrator equation could be further developed to model the experimental findings of Khojasteh *et al.* [29]. Khojasteh *et al.* [29] have shown that there is a large variability in values of the time constant and the location of the null point of the neural integrator in horizontal eye movements. Initially, further fits of the model could include the neural integrator time constant as a parameter. Subsequently, a better equation describing the experimental data in more detail could be included, which would also simulate the variability in null zone between individuals.

7.4.2.3 Incorporation of a smooth pursuit system

Our proposed model incorporating the smooth pursuit system is based on the assumption that the smooth pursuit neurons can be modelled in the same way as the saccadic neurons [119]. The left and right smooth pursuit populations of neurons inhibit each other and they are also inhibited by the antagonistic saccadic population. Likewise, the saccadic neurons are inhibited by the antagonistic smooth pursuit neurons. The model comprises the following nine ordinary differential equations (ODEs):

$$\frac{dg}{dt} = v \tag{7.3}$$

$$\frac{dv}{dt} = -\left(\frac{1}{T_1} + \frac{1}{T_2}\right)v - \frac{1}{T_1 T_2}g + \frac{1}{T_1 T_2}n + \left(\frac{1}{T_1} + \frac{1}{T_2}\right)(r_B - l_B + r_P - l_P) \tag{7.4}$$

$$\frac{dn}{dt} = -\frac{1}{T_N}n + r_B - l_B + r_P - l_P \tag{7.5}$$

$$\frac{dr_B}{dt} = \frac{1}{\varepsilon_B}\left(-r_B - \gamma_B r_B l_B^2 - \gamma_{BP} r_B l_P^2 + F_B(m_B)\right) \tag{7.6}$$

$$\frac{dl_B}{dt} = \frac{1}{\varepsilon_B}\left(-l_B - \gamma_B l_B r_B^2 - \gamma_{BP} l_B r_P^2 + F_B(-m_B)\right) \tag{7.7}$$

$$\frac{dr_P}{dt} = \frac{1}{\varepsilon_P}\left(-r_P - \gamma_P r_P l_P^2 - \gamma_{PB} r_P l_B^2 + F_P(m_P)\right) \tag{7.8}$$

$$\frac{dl_P}{dt} = \frac{1}{\varepsilon_P}\left(-l_P - \gamma_P l_P r_P^2 - \gamma_{PB} l_P r_B^2 + F_P(-m_P)\right) \quad (7.9)$$

$$\frac{dm_B}{dt} = -\frac{1}{T_{RB}} m_B - (r_B - l_B) + \frac{T - g}{T_{RB}} \quad (7.10)$$

$$\frac{dm_P}{dt} = -\frac{1}{T_{RP}} m_P - (r_P - l_P) + \frac{\dot{T} - \dot{g}}{T_{RP}} \quad (7.11)$$

Compared with the saccadic model, the control signal sent to the muscle plant (equations (7.3) and (7.4)) is now a weighted sum of the net velocity command $r_B - l_B + r_P - l_P$ generated by the saccadic burst and smooth pursuit neurons, and the position command n obtained by integrating $r_B - l_B + r_P - l_P$ (equation (7.5)). Equations (7.6)-(7.9) model the activities of the burst and smooth pursuit neurons. Each neuron type is split into two populations: right (r_B: right burst and r_P: right smooth pursuit) and left (l_B: left burst and l_P: left smooth pursuit). As before, the right and left burst neurons suppress each other's firing and the parameter γ_B in equations (7.6) and (7.7) quantifies the strength of this mutual inhibition; the parameter γ_{BP} quantifies the inhibition of the saccadic neurons by the antagonistic smooth pursuit neurons. Similarly, in equations (7.8) and (7.9), the parameter γ_P quantifies the strength of the mutual inhibition between the right and left smooth pursuit neurons, while the parameter γ_{PB} quantifies the inhibition by the antagonistic saccadic burst neurons.

Equations (7.10) and (7.11) model the motor error for each neuron type: for the saccadic neurons, this is the difference between the target position T and the current eye position g; for the smooth pursuit neurons, this is the difference between the target velocity $\dot{T}$ and the current eye position v. In these equations, the time constants T_{RB} and T_{RP} describe the rate of deterioration of the feedback signals $T - g$ and $\dot{T} - v$, respectively. The functions $F_B(m_b)$ and $F_P(m_p)$ specifying how the burst and smooth pursuit neuron firing rates change with motor error are defined by:

$$F_B(m_B) = \begin{cases} a'_B\left(1 - e^{-m_B/\beta'_B}\right) \; if \; m_B \geq 0 \\ -\dfrac{\alpha_B}{\beta_B} m_B e^{m_B/\beta_B} \quad if \; m_B < 0 \end{cases} \quad (7.12)$$

$$F_P(m_P) = \begin{cases} a'_P\left(1 - e^{-m_P/\beta'_P}\right) \; if \; m_P \geq 0 \\ -\dfrac{\alpha_P}{\beta_P} m_P e^{m_P/\beta_P} \quad if \; m_P < 0 \end{cases} \tag{7.13}$$

The smooth pursuit neural response function $F_P(m_P)$ is thus assumed to have the same form as the saccadic function $F_B(m_B)$ used in the original Broomhead *et al.* [1] model (*cf.* equation (2.7)).

Using the methodology developed in this work, future fitting of this model can be performed to experimental recordings of normal saccades, nystagmus oscillations and smooth pursuit eye movements. It would be interesting to find out whether the addition of smooth pursuit system could enable the model to simulate a wider spread of jerk nystagmus characteristics, such as overshooting of the eye at the end of the fast phase (see point 2 in Figure 6.24).

7.5 Oculomotor learning model

A stochastic oculomotor learning model could be developed based on the mechanism that visual feedback “guides” changes to the model parameters and the hypothesis that degraded visual acuity can cause the model parameters to drift to the unstable area of the parameter space [48] could be tested. This could explain why degraded visual acuity is linked to the development of infantile nystagmus [48].

7.6 OpenCL ODE solvers

The OpenCL ODE solvers could be further developed by including more solvers (*e.g.* variable step and multi-step methods), allowing a broader range of users inexperienced in GPU programming to take advantage of this technology.

A better interface of the OpenCL solvers to MATLAB could be implemented using a *mex* MATLAB file. This file would enable direct call of the GPU without the need of running an external executable file. This would also allow users to run the solvers without the requirement for a RAM disk in order to save the results. In particular, further development of the OpenCL ODE solvers would be beneficial in the optimisation of the current and future oculomotor models considered here.

7.7 Final conclusions

The aetiology of infantile nystagmus is not yet well understood and a number of computational models have been suggested in order to get a deeper insight into this disorder. Here, we have presented methods for fitting the model of Broomhead *et al.* [1] to the experimental infantile nystagmus and saccadic data sets. We believe that these methods are applicable to any model, allowing its predictive capacity to be systematically explored. Moreover, we accelerated these methods by implementing OpenCL ODE solvers which can utilise the parallel capabilities of a GPU, thereby providing a speedup of 30 compared to a midrange CPU.

In addition, we formulated a stochastic model based on the Broomhead *et al.* [1] model by adding SDN and CN to the model's burst and neural integrator signals. Application of this stochastic model allowed detailed examination of whether SDN and CN could be the cause of the local dimension variations measured in experimental jerk nystagmus time series by Akman *et al.* [15]. This involved utilisation of the parameter optimisation methods that we developed and the application of nonlinear time series analysis methods based on the method of delays.

Our results suggest that SDN and CN could be the cause of the local dimension variation measured in experimental data. Furthermore, the results showed the advantages and disadvantages of the Broomhead *et al.* [1] model in simulating normal and abnormal oculomotor behaviour. Future work has to be performed and this could include robust methods for optimising more complex stochastic oculomotor models, the incorporation of more oculomotor systems, the development of an oculomotor learning model and the further development of OpenCL ODE solvers.

In closing, the work presented in this study is beneficial for further development of the oculomotor models and it has shown that parameter optimisation methods can be a valuable tool in the analysis of these models.

Appendix A

Table A.1: The computer hardware used in this study.

Part Type	Part Name
CPU	Intel® Core™ i5-3550
GPU	XFX AMD Radeon HD 7970 Black Edition 3GB GDDR5
Memory	32GB Kit (4 x 8GB) Corsair PC3-12800 (1600MHz)
Motherboard	Gigabyte GA-Z77-D3H (BIOS version F21)
Hard disk	1TB Seagate ST1000DM003 Barracuda, rotation speed 7200rpm, Interface SATA III 6Gb/s, cache 64MB, access time 8ms

Table A.2: The software used in this study.

Software Type	Software Name
Operating system	Microsoft Windows 8.1 x64
GPU display driver	Display Driver ver. 13.251 OpenCL(tm) Driver ver. 10.0.1348.5 Catalyst Control Center ver. 2013.1206.1602.28764
Software Development Kit	AMD APP SDK v2.9.1
Helper C++ libraries	Boost library version 1.55.0
RAM disk	SoftPerfect RAM Disk version version 3.4.2. 64-bit
Optimization Toolbox	MATLAB R2012B

Appendix B

The following titles of articles are in preparation and are based on the work presented in this thesis:

1. Fitting experimental nystagmus waveforms to computational models using multi-objective genetic algorithms executed on a GPU-CPU combination: application to an exemplar motor control system

2. Signal-dependent and constant noise as a source of variability in infantile nystagmus oscillations

3. Exploring the predictive capacity of oculomotor models

4. A model for the oculomotor control system incorporating visual target

Bibliography

[1] Broomhead, D. S., Clement, R. A., Muldoon, M. R., Whittle, J. P., Scallan, C., and Abadi, R. V. (2000) Modelling of congenital nystagmus waveforms produced by saccadic system abnormalities. *Biological Cybernetics*, ***82***(5), 391-399.

[2] Leigh, R. J., and Zee, D. S. (2015) *The neurology of eye movements*. Oxford University Press.

[3] Wade, N. J., and Tatler, B. W. (2011) Origins and applications of eye movement research. In: *Oxford handbook on eye movements*. Edited by: Liversedge, S., Gilchrist, I., and Everling, S., Oxford University Press. pp. 17-46.

[4] Hayhoe, M., and Ballard, D. (2005) Eye movements in natural behavior. *Trends in Cognitive Sciences*, ***9***(4), 188-194.

[5] Schütz, A. C., Braun, D. I., and Gegenfurtner, K. R. (2011) Eye movements and perception: A selective review. *Journal of Vision*, ***11***(5):9, 1-30.

[6] Akman, O. E., Broomhead, D. S., Abadi, R. V., and Clement, R. A. (2012) Components of the neural signal underlying congenital nystagmus. *Experimental Brain Research*, ***220***(3-4), 213-221.

[7] Scudder, C. A., Kaneko, C. R., and Fuchs, A. F. (2002) The brainstem burst generator for saccadic eye movements. *Experimental Brain Research*, ***142***(4), 439-462.

[8] Robinson, D. A. (1986) Is the oculomotor system a cartoon of motor control? In: *The oculomotor and skeletal motor systems: Differences and similarities.* Edited by: Freund, H. J., Büttner, U., Cohen, B., and Noth, J., Elsevier Science Publishers B.V. (Biomedical Division). ***64***, pp. 411-417.

[9] Girard, B., and Berthoz, A. (2005) From brainstem to cortex: computational models of saccade generation circuitry. *Progress in Neurobiology*, ***77***(4), 215-251.

[10] Scudder, C. A. (1988) A new local feedback model of the saccadic burst generator. *Journal of Neurophysiology*, ***59***(5), 1455-1475.

[11] Young, L. R., and Stark, L. (1963) Variable feedback experiments testing a sampled data model for eye tracking movements. In: *Human Factors in Electronics, IEEE Transactions*, IEEE, ***4***(1), 38-51.

[12] Van Opstal, A. J., Van Gisbergen, J. A. M., and Eggermont, J. J. (1985) Reconstruction of neural control signals for saccades based on an inverse method. *Vision Research*, ***25***(6), 789-801.

[13] Jacobs, J. B., and Dell'Osso, L. F. (2004) Congenital nystagmus: hypotheses for its genesis and complex waveforms within a behavioral ocular motor system model. *Journal of Vision*, ***4***(7), 604–625.

[14] Clement, R. A., Abadi, R. V., Broomhead, D. S., and Whittle, J. P. (2002) A new framework for investigating both normal and abnormal eye movements. *Progress in Brain Research*, ***140***, 499-505.

[15] Akman, O. E., Broomhead, D. S., Clement, R. A., and Abadi, R. V. (2006) Nonlinear time series analysis of jerk congenital nystagmus. *Journal of Computational Neuroscience*, ***21***(2), 153-170.

[16] Moschovakis, A. K., Scudder, C. A., and Highstein, S. M. (1996) The microscopic anatomy and physiology of the mammalian saccadic system. *Progress in Neurobiology*, ***50***(2-3), 133-254.

[17] Bahill, A. T., Clark, M. R., and Stark, L. (1975) The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, ***24***(3-4), 191-204.

[18] Fukushima, K., Fukushima, J., Warabi, T., and Barnes, G. R. (2013) Cognitive processes involved in smooth pursuit eye movements: behavioral evidence, neural substrate and clinical correlation. *Frontiers in Systems Neuroscience*, ***7***, 4.

[19] Cullen, K. E. (2012) The vestibular system: multimodal integration and encoding of self-motion for motor control. *Trends in Neurosciences*, ***35***(3), 185-196.

[20] Konen, C. S., Kleiser, R., Seitz, R. J., and Bremmer, F. (2005) An fMRI study of optokinetic nystagmus and smooth-pursuit eye movements in humans. *Experimental Brain Research*, ***165***(2), 203-216.

[21] Cullen, K. E., and Van Horn, M. R. (2011) The neural control of fast vs. slow vergence eye movements. *European Journal of Neuroscience*, ***33***(11), 2147-2154.

[22] Van Opstal, A. J., and Van Gisbergen, J. A. M. (1987) Skewness of saccadic velocity profiles: a unifying parameter for normal and slow saccades. *Vision Research*, ***27***(5), 731-745.

[23] Collewijn, H., Erkelens, C. J., and Steinman, R. M. (1988). Binocular co-ordination of human horizontal saccadic eye movements. *The Journal of Physiology*, ***404***(1), 157-182.

[24] Sparks, D. L. (2002) The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience*, ***3***(12), 952-964.

[25] Van Gisbergen, J. A. M., Robinson, D. A., and Gielen, S. T. A. N. (1981) A quantitative analysis of generation of saccadic eye movements by burst neurons. *Journal of Neurophysiology*, ***45***(3), 417-442.

[26] Liversedge, S., Gilchrist, I., and Everling, S. (2011) *The Oxford handbook of eye movements*. Oxford University Press.

[27] Cullen, K. E., and Van Horn, M. R. (2011) Brainstem pathways and premotor control. In: *Oxford handbook of eye movements.* Edited by: Liversedge, S., Gilchrist, I., and Everling, S., Oxford University Press. pp. 151-172.

[28] Goossens, H. H. L. M., and Van Opstal, A. J. (2006) Dynamic ensemble coding of saccades in the monkey superior colliculus. *Journal of Neurophysiology*, ***95***(4), 2326-2341.

[29] Khojasteh, E., Bockisch, C. J., Straumann, D., and Hegemann, S. C. (2012) A re-examination of the time constant of the oculomotor neural integrator in human. In: *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, IEEE. pp. 4780-4783.

[30] Van Beers, R. J. (2007) The sources of variability in saccadic eye movements. *The Journal of Neuroscience*, ***27***(33), 8757-8770.

[31] Van Beers, R. J. (2008) Saccadic eye movements minimize the consequences of motor noise. *PLoS One*, ***3***(4), e2070.

[32] Van Beers, R. J. (2009) Motor learning is optimally tuned to the properties of motor noise. *Neuron*, ***63***(3), 406-417.

[33] Harris, C. M., and Wolpert, D. M. (1998) Signal-dependent noise determines motor planning. *Nature*, ***394***(6695), 780-784.

[34] Harris, C. M., and Wolpert, D. M. (2006) The main sequence of saccades optimizes speed-accuracy trade-off. *Biological Cybernetics*, ***95***(1), 21-29.

[35] Yap, Y. L., Levi, D. M., and Klein, S. A. (1987) Peripheral hyperacuity: isoeccentric bisection is better than radial bisection. *Journal of the Optical Society of America A*, ***4***(8), 1562-1567.

[36] White, J. M., Levi, D. M., and Aitsebaomo, A. P. (1992) Spatial localization without visual references. *Vision Research*, ***32***(3), 513-526.

[37] Westheimer, G. (2005) Anisotropies in peripheral vernier acuity. *Spatial Vision*, ***18***(2), 159-167.

[38] Westheimer, G. (2001) Relative localization in the human fovea: radial–tangential anisotropy. *Proceedings of the Royal Society of London B: Biological Sciences*, ***268***(1471), 995-999.

[39] Whitaker, D., and Latham, K. (1997) Disentangling the role of spatial scale, separation and eccentricity in Weber's law for position. *Vision Research*, ***37***(5), 515-524.

[40] Aitsebaomo, A. P., and Bedell, H. E. (1992) Psychophysical and saccadic information about direction for briefly presented visual targets. *Vision Research*, ***32***(9), 1729-1737.

[41] Abadi, R. V. (2002) Mechanisms underlying nystagmus. *Journal of the Royal Society of Medicine*, ***95***(5), 231-234.

[42] Ehrt, O. (2012) Infantile and acquired nystagmus in childhood. *European Journal of Paediatric Neurology*, ***16***(6), 567-572.

[43] Abadi, R. V., and Bjerre, A. (2002) Motor and sensory characteristics of infantile nystagmus. *British Journal of Ophthalmology*, ***86***(10), 1152-1160.

[44] Harris, C., and Berry, D. (2006) A developmental model of infantile nystagmus. *Seminars in Ophthalmology*, ***21***(2), 63-69.

[45] Abadi, R. V., Forster, J. E., and Lloyd, I. C. (2006) Ocular motor outcomes after bilateral and unilateral infantile cataracts. *Vision Research*, ***46***(6), 940-952.

[46] Felius, J., Busettini, C., Lynn, M. J., Hartmann, E. E., and Lambert, S. R. (2014) Nystagmus and related fixation instabilities following extraction of unilateral infantile cataract in the Infant Aphakia Treatment Study (IATS). *Investigative Ophthalmology and Visual Science*, ***55***(8), 5332-5337.

[47] Holmström, G., Bondeson, M. L., Eriksson, U., Åkerblom, H., and Larsson, E. (2014) ‘Congenital’ nystagmus may hide various ophthalmic diagnoses. *Acta Ophthalmologica*, ***92***(5), 412-416.

[48] Harris, C. M. (2011) Oculomotor developmental pathology: an evo-devo perspective. In: *Oxford handbook on eye movements*. Edited by: Liversedge, S., Gilchrist, I., and Everling, S., Oxford University Press. pp. 663-686.

[49] Sarvananthan, N., Surendran, M., Roberts, E. O., Jain, S., Thomas, S., Shah, N., Proudlock, F. A., Thompson, J. R., McLean, R. J., Degg, C., Woodruff, G., and Gottlob, I. (2009) The prevalence of nystagmus: the Leicestershire nystagmus survey. *Investigative Ophthalmology and Visual Science*, ***50***(11), 5201-5206.

[50] Abadi, R. V., and Worfolk, R. (1989) Retinal slip velocities in congenital nystagmus. *Vision Research*, ***29***(2), 195-205.

[51] Papageorgiou, E., McLean, R. J., and Gottlob, I. (2014) Nystagmus in childhood. *Pediatrics and Neonatology*, ***55***(5), 341-351.

[52] McLean, R. J., Windridge, K. C., and Gottlob, I. (2012) Living with nystagmus: a qualitative study. *British Journal of Ophthalmology*, ***96***(7), 981-986.

[53] Pilling, R. F., Thompson, J. R., and Gottlob, I. (2005) Social and visual function in nystagmus. *British Journal of Ophthalmology*, ***89***(10), 1278-1281.

[54] Gottlob, I., and Proudlock, F. A. (2014) Aetiology of infantile nystagmus. *Current Opinion in Neurology*, ***27***(1), 83-91.

[55] Thomas, M. G., Crosier, M., Lindsay, S., Kumar, A., Araki, M., Leroy, B. P., McLean, R. J., Sheth, V., Maconachie, G., Thomas, S., Moore, A. T., and Gottlob, I. (2014) Abnormal retinal development associated with FRMD7 mutations. *Human Molecular Genetics*, ***23***(15), 4086-4093.

[56] Chen, C. C., Bockisch, C. J., Olasagasti, I., Weber, K. P., Straumann, D., and Huang, M. Y. (2014) Positive or negative feedback of optokinetic signals: degree of the misrouted optic flow determines system dynamics of human ocular motor behavior. *Investigative Ophthalmology and Visual Science*, ***55***(4), 2297-2306.

[57] Pasquariello, G., Cesarelli, M., Bifulco, P., Fratini, A., La Gatta, A., and Romano, M. (2009) Characterisation of baseline oscillation in congenital nystagmus eye movement recordings. *Biomedical Signal Processing and Control*, ***4***(2), 102-107.

[58] Abadi, R. V., and Dickinson, C. M. (1986) Waveform characteristics in congenital nystagmus. *Documenta Ophthalmologica*, ***64***(2), 153-167.

[59] Dell'Osso, L. F., and Daroff, R. B. (1975) Congenital nystagmus waveforms and foveation strategy. *Documenta Ophthalmologica*, ***39***(1), 155-182.

[60] Optican, L. M., and Zee, D. S. (1984) A hypothetical explanation of congenital nystagmus. *Biological Cybernetics*, ***50***(2), 119-134.

[61] Brodsky, M. C., and Dell'Osso, L. F. (2014) A unifying neurologic mechanism for infantile nystagmus. *JAMA Ophthalmology*, ***132***(6), 761-768.

[62] Akman, O. E., Broomhead, D. S., Abadi, R. V., and Clement, R. A. (2005) Eye movement instabilities and nystagmus can be predicted by a nonlinear dynamics model of the saccadic system. *Journal of Mathematical Biology*, ***51***(6), 661-694.

[63] Abadi, R. V., Broomhead, D. S., Clement, R. A., Whittle, J. P., and Worfolk, R. (1997) Dynamical systems analysis: a new method of analysing congenital nystagmus waveforms. *Experimental Brain Research*, ***117***(3), 355-361.

[64] Broomhead, D. S., and King, G. P. (1986) Extracting qualitative dynamics from experimental data. *Physica D: Nonlinear Phenomena*, ***20***(2-3), 217-236.

[65] Banga, J. R. (2008) Optimization in computational systems biology. *BMC Systems Biology*, ***2***(1), 47.

[66] Boyd, S. and Vandenberghe, L. (2004) Convex Optimization, Cambridge University Press.

[67] Forrest, S. (1993) Genetic algorithms: principles of natural selection applied to computation. *Science*, ***261***(5123), 872-878.

[68] Sutherland, W. J. (2005) The best solution. *Nature*, ***435***(7042), 569.

[69] Lillacci, G., and Khammash, M. (2010) Parameter estimation and model selection in computational biology. *PLoS Computational Biology*, ***6***(3), e1000696.

[70] Luque, G., and Alba, E. (2011) Parallel genetic algorithms: theory and real world applications (Vol. 367). Springer.

[71] Baker, S. M., Schallau, K., and Junker, B. H. (2010) Comparison of different algorithms for simultaneous estimation of multiple parameters in kinetic metabolic models. *Journal of Integrative Bioinformatics*, ***7***(3), 1–9.

[72] Sun, J., Garibaldi, J. M., and Hodgman, C. (2012) Parameter estimation using metaheuristics in systems biology: a comprehensive review. In: *Computational Biology and Bioinformatics, IEEE/ACM Transactions*, IEEE, ***9***(1), 185-202.

[73] Carvalho, A. R., Ramos, F. M., and Chaves, A. A. (2011) Metaheuristics for the feedforward artificial neural network (ANN) architecture optimization problem. *Neural Computing and Applications*, ***20***(8), 1273-1284.

[74] Caramia, M., and Dell'Olmo, P. (2008) *Multi-objective management in freight logistics: Increasing capacity, service level and safety with optimization algorithms*. Springer London.

[75] Zitzler, E., and Thiele, L. (1999) Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. In: *Evolutionary computation, IEEE Transactions*, IEEE, ***3***(4), 257-271.

[76] Deb, K. (2011) Multi-objective optimisation using evolutionary algorithms: an introduction. In: *Multi-objective evolutionary optimisation for product design and manufacturing*. Edited by: Wang, L., Ng, A. H. C., and Deb, K., Springer London. pp. 3-34.

[77] Walker, D. J., Everson, R. M., and Fieldsend, J. E. (2013) Visualizing mutually non-dominating solution sets in many-objective optimization. In: *Evolutionary Computation, IEEE Transactions*, IEEE, ***17***(2), 165-184.

[78] Zavala, G. R., Nebro, A. J., Luna, F., and Coello, C. A. C. (2014) A survey of multi-objective metaheuristics applied to structural optimization. *Structural and Multidisciplinary Optimization*, ***49***(4), 537-558.

[79] Mitchell, M. (1998) *An introduction to genetic algorithms*. MIT Press.

[80] Lichocki, P., Floreano, D., and Keller, L. (2014) Selection methods regulate evolution of cooperation in digital evolution. *Journal of the Royal Society Interface*, ***11***(90), 20130743.

[81] Goldberg, D. E., and Deb, K. (1991) A comparative analysis of selection schemes used in genetic algorithms. In: *Foundations of Genetic Algorithms*. Edited by: Rawlins, G. J. E., Morgan Kaufmann Publishers, Inc. pp. 69-93.

[82] Bäck, T. (1994) Selective pressure in evolutionary algorithms: A characterization of selection mechanisms. In: *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence, Proceedings of the First IEEE Conference*, IEEE, ***1***, 57-62.

[83] Simoncini, D., Collard, P., Verel, S., and Clergue, M. (2007) On the influence of selection operators on performances in cellular genetic algorithms. In: *Evolutionary Computation, 2007. CEC 2007. IEEE Congress,* IEEE. pp. 4706-4713.

[84] Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J. A., and Blom, J. G. (2009) Systems biology: parameter estimation for biochemical models. *FEBS Journal*, ***276***(4), 886-902.

[85] Poon, P. W., and Carter, J. N. (1995) Genetic algorithm crossover operators for ordering applications. *Computers and Operations Research*, ***22***(1), 135-147.

[86] Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. A. M. T. (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. In: *Evolutionary Computation, IEEE Transactions*, IEEE, ***6***(2), 182-197.

[87] Deb, K. (2001) Multi-objective optimization using evolutionary algorithms (Vol. 16). John Wiley and Sons.

[88] Deb, K., Anand, A., and Joshi, D. (2002). A computationally efficient evolutionary algorithm for real-parameter optimization. *Evolutionary Computation*, ***10***(4), 371-395.

[89] Nowostawski, M., and Poli, R. (1999) Parallel genetic algorithm taxonomy. In: *Knowledge-Based Intelligent Information Engineering Systems, 1999. Third International Conference*, IEEE. pp. 88-92.

[90] Pospíchal, P., and Jaros, J. (2009) Gpu-based acceleration of the genetic algorithm. *Genetic and Evolutionary Computation Conference Competition GECCO,* pp. 2–3.

[91] Demidov, D., Ahnert, K., Rupp, K., and Gottschling, P. (2013) Programming CUDA and OpenCL: a case study using modern C++ libraries. *SIAM Journal on Scientific Computing*, ***35***(5), C453-C472.

[92] Cavuoti, S., Garofalo, M., Brescia, M., Longo, G., and Ventre, G. (2013) Genetic algorithm modeling with GPU parallel computing technology. In: *Neural Nets and Surroundings*, SIST 19. Edited by: Apollini, B., Bassis,

S., Esposito, A., and Morabito, F. C., Springer Berlin Heidelberg. pp. 29-39.

[93] Johar, F. M., Azmin, F. A., Suaidi, M. K., Shibghatullah, A. S., Ahmad, B. H., Salleh, S. N., Aziz, M. Z. A. A., and Md Shukor, M. (2013) A review of genetic algorithms and parallel genetic algorithms on Graphics Processing Unit (GPU). In: *Control System, Computing and Engineering (ICCSCE), 2013 IEEE International Conference*, IEEE. pp. 264-269.

[94] Kolonias, V., Goulas, G., Gogos, C., Alefragis, P., and Housos, E. (2014) Solving the examination timetabling problem in GPUs. *Algorithms*, ***7***(3), 295-327.

[95] Ben-Shalom, R., Aviv, A., Razon, B., and Korngreen, A. (2012) Optimizing ion channel models using a parallel genetic algorithm on graphical processors. *Journal of Neuroscience Methods*, ***206***(2), 183-194.

[96] McLean, R., Proudlock, F., Thomas, S., Degg, C., and Gottlob, I. (2007) Congenital nystagmus: Randomized, controlled, double-masked trial of memantine/gabapentin. *Annals of Neurology*, ***61***(2), 130-138.

[97] So, P., Ott, E., Schiff, S. J., Kaplan, D. T., Sauer, T., and Grebogi, C. (1996) Detecting unstable periodic orbits in chaotic experimental data. *Physical Review Letters*, ***76***(25), 4705-4708.

[98] So, P., Ott, E., Sauer, T., Gluckman, B. J., Grebogi, C., and Schiff, S. J. (1997) Extracting unstable periodic orbits from chaotic time series data. *Physical Review E*, ***55***(5), 5398-5417.

[99] So, P., Francis, J. T., Netoff, T. I., Gluckman, B. J., and Schiff, S. J. (1998) Periodic orbits: a new language for neuronal dynamics. *Biophysical Journal*, ***74***(6), 2776-2785.

[100] Clement, R. A., Whittle, J. P., Muldoon, M. R., Abadi, R. V., Broomhead, D. S., and Akman, O. (2002) Characterisation of congenital nystagmus waveforms in terms of periodic orbits. *Vision Research*, ***42***(17), 2123-2130.

[101] Theodorou, M., and Clement, R. A. (2007) Fixed point analysis of nystagmus. *Journal of Neuroscience Methods*, ***161***(1), 134-141.

[102] Takens, F. (1981) Detecting strange attractors in turbulence. In: *Dynamical Systems and Turbulence, Warwick 1980.* Edited by: Rand, D., and Young, L-S. Springer Berlin Heidelberg. pp. 366-381.

[103] O'Regan, K. (1979) Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics*, ***25***(6), 501-509.

[104] Von Wartburg, R., Wurtz, P., Pflugshaupt, T., Nyffeler, T., Luthi, M., and Muri, R. M. (2007) Size matters: Saccades during scene perception. *Perception*, ***36***(3), 355-365.

[105] MATLAB, *version R2012b*. Natick, Massachusetts: The MathWorks Inc., 2012.

[106] Roeva, O., Fidanova, S., and Paprzycki, M. (2013) Influence of the population size on the genetic algorithm performance in case of cultivation process modelling. In: *Computer Science and Information Systems (FedCSIS), 2013 Federated Conference,* IEEE. pp. 371-376.

[107] Knowles, J., and Corne, D. (2002) On metrics for comparing non-dominated sets. In: *Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress,* IEEE. pp. 711-716.

[108] Bishop, C. M. (2006) *Pattern recognition and machine learning*. Springer-Verlag New York.

[109] Atkins, J. E., Boman, E. G., and Hendrickson, B. (1998) A spectral algorithm for seriation and the consecutive ones problem. *SIAM Journal on Computing*, ***28***(1), 297-310.

[110] Fieldsend, J., and Everson, R. (2013) Visualising high-dimensional Pareto relationships in two-dimensional scatterplots. In: *Evolutionary Multi-Criterion Optimization.* Edited by: Purshouse, R. C., Fleming, P. J., Fonseca, C. M., Greco, S., and Shaw, J., Springer Berlin Heidelberg. pp. 558-572.

[111] Roberts, A. J. (2012) Modify the Improved Euler scheme to integrate stochastic differential equations. *arXiv preprint, arXiv:1210.0933.*

[112] Van Beers, R. J., Haggard, P., and Wolpert, D. M. (2004) The role of execution noise in movement variability. *Journal of Neurophysiology*, ***91***(2), 1050-1063.

[113] Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008) Noise in the nervous system. *Nature Reviews Neuroscience*, ***9***(4), 292-303.

[114] Breeden, J. L. (1995) Optimizing stochastic and multiple fitness functions. In: *Evolutionary Programming*. Edited by: McDonnell, J. R., Reynolds, R. G., and Fogel, D. B. MIT Press. pp. 127-134.

[115] Zhou, A., Qu, B. Y., Li, H., Zhao, S. Z., Suganthan, P. N., and Zhang, Q. (2011) Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, ***1***(1), 32-49.

[116] Fieldsend, J. E., and Everson, R. M. (2014) The rolling tide evolutionary algorithm: A multiobjective optimizer for noisy optimization problems. In: *Evolutionary Computation, IEEE Transactions*, IEEE, ***19***(1), 103-117.

[117] Mauthner, S. (1998) Step size control in the numerical solution of stochastic differential equations. *Journal of Computational and Applied Mathematics*, ***100***(1), 93-109.

[118] Abadi, R. V., and Dickinson, C. M. (1987) Waveform characteristics in congenital nystagmus. *Documenta Ophthalmologica*, ***64***(2), 153-167.

[119] Keller, E. L., and Missal, M. (2003) Shared brainstem pathways for saccades and smooth-pursuit eye movements. *Annals of the New York Academy of Sciences*, ***1004***(1), 29-39.