CONSIDERATIONS IN THE PRACTICAL IMPLEMENTATION OF A TRAVELLING WAVE COCHLEAR IMPLANT PROCESSOR

by

Christiaan Cronjé Du Preez

Submitted in partial fulfillment of the requirements for the degree

Master of Engineering (Bioengineering)

in the

Department of Electrical, Electronic and Computer Engineering

Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

April 2012

SUMMARY

CONSIDERATIONS IN THE PRACTICAL IMPLEMENTATION OF A TRAVELING WAVE COCHLEAR IMPLANT PROCESSOR

by

Christiaan Cronjé Du Preez

Supervisor : Prof J.J. Hanekom

Department : Electrical, Electronic and Computer Engineering

University : University of Pretoria

Degree : Masters in Engineering (Bioengineering)

Keywords : Cochlear implants, travelling wave, processing strategy, stimulation rate, channel number, neural spike train patterns, temporal, spatial

Speech processing in the human cochlea introduces travelling waves on the basilar membrane. These travelling waves have largely been ignored in most processing strategies. This study implements a hydrodynamical model in a speech processing strategy in order to investigate the neural spike train patterns for a travelling wave processing strategy. In cochlear implants a trade-off remains between the simulation rate and the number of electrode channels. This trade-off was investigated in the proposed travelling wave strategy. Taking into consideration existing current spread and electrical stimulation models, predicted neural spike train responses have shown that stimulating fewer channels (six and four) at stimulation rates of 2 400 pps and 3 600 pps gives better approximations of predicted normal hearing responses for input frequencies of 200 Hz, 600 Hz and 1 kHz, compared to stimulating more chan-

nels at lower channel stimulation rates. The predicted neural spike train patterns suggest that these resulting neural patterns might contain both spatial and temporal information that could be extracted by the auditory system. For a frequency of 4 kHz the predicted neural patterns for a channel-number stimulation-rate configuration of 2 - 7 200 pps suggested that although there is no travelling wave delay information, the predicted neural patterns still contain temporal information. The predicted ISI histograms show peaks at the input tone period and multiples thereof, with clusters of spikes evident around the tone period in the predicted spatio-temporal neural spike train patterns. Similar peaks at the tone period were observed for calculated ISI histograms for predicted normal hearing neural patterns and measured neural responses. The predicted spatio-temporal neural patterns for the input frequency of 200 Hz show the travelling wave delay with clusters of spikes at the tone period. This travelling wave delay can also be seen from predicted normal hearing neural responses. The current spread, however, shows a significant distortion effect around the characteristic frequency place where the travelling wave delay increases rapidly. Spacing electrodes more closely results in an increase in this distortion, with the nerve fibre threshold decreasing in adjacent populations of nerve fibres, increasing the probability of firing. The current spread showed a more limited distortion effect on travelling wave delays when electrodes were spaced across the cochlea, at an electrode spacing of 6.08 mm. ISI histogram results also showed increased peaks around the tone period and multiples thereof. These predicted neural spike train patterns suggest that travelling waves in processing strategies, although mostly ignored, might provide the auditory system with both the spatial and temporal information needed for better pitch perception.

OPSOMMING

OORWEGINGS IN DIE PRAKTIESE IMPLEMENTERING VAN ’N VOORTPLANTINGSGOLF KOGLEÊRE IMPLANTINGSPROSSESEERDER

deur

Christiaan Cronjé du Preez

Studieleier	:	Prof J.J. Hanekom
Departement	:	Elektriese, Elektroniese en Rekenaar-Ingenieurswese
Universiteit	:	Universiteit van Pretoria
Graad	:	Magister in Ingenieurswese (Bio-Ingenieurswese)
Sleutelwoorde	:	Kogleêre inplantings, voorplantingsgolf, prosesseringstrategie, stimulasietempo, aantal kanale, neurale aktiveringspatrone, tyd-afhanklike, ruimte-afhanklike

Spraakprosessesering in die menslike koglea maak gebruik van voortplantingsgolwe op die basilêre membraan. Hierdie voorplantingsgolwe is groot en deels geïgnoreer in prosesseringstrategieë tot dusver. Hierdie studie implementeer ’n hidrodinamiese model in ’n spraakprosesseseringstrategie om die neurale aktiveringspatrone vir ’n voorplantingsgolfprosesseseringstrategie te ondersoek. Daar bestaan ’n kompromis in kogleêre inplantings tussen die stimulasietempo en die aantal elektrodekanale. Hierdie kompromis word in the voorgestelde voorplantingsgolfprosesseseerder ondersoek. Deur gebruik te maak van huidige stroomverspreiding en elektiese stimulasiemodelle, toon die neurale aktiveringspatrone dat stimulasie met minder kanale (vier en ses) by stimulasietempos van 2 400 pps en 3 600 pps

beter voorspelde neurale aktiveringspatrone vir normale gehoor by insetfrekwensies van 200 Hz, 600 Hz en 1 kHz gee, as stimulasie met meer kanale by laer stimulasietempos. Die voorspelde neurale aktiveringspatrone dui daarop dat hierdie aktiveringspatrone dalk tyd- en ruimteafhanklike inligting mag bevat wat deur die gehoorstelsel onttrek kan word. Die neurale aktiveringspatrone vir ’n 4 kHz-frekwensie met die aantal-kanale, stimulasie-tempokonfigurasie van 2 - 7 200 pps, dui daarop dat daar geen voorplantingsgolfvertraging voorkom in the aktiveringspatrone nie, maar dat daar wel tyd-afhanklike inligting voorkom. Die voorspelde interaktiveringsinterval- (IAI) histogramme wys wel pieke by die insettoon-periode en veelvoude van die toonperiode, met groepe aktiverings wat gesien kan word rondom die toonperiode in die ruimte-tyd neurale aktiveringspatrone. Dieselfde tipe pieke word gesien in berekende IAI histogramme vir voorspelde normale gehoor neurale aktiveringspatrone en gemete neurale aktiveringspatrone. Die voorspelde neurale aktiveringspatrone vir die insetfrekwensie van 200 Hz wys die voorplantingsgolfvertraging met groepe neurale aktiverings by die toonperiode. Hierdie voorplantingsgolfvertraging kan ook gesien word in neurale aktiveringspatrone in normale gehoor. Die stroomverspreiding, aan die ander kant, wys ’n noemenswaardige vervormingeffek rondom die plek van die karakteristieke frekwensie waar die voortplantingsgolfvertraging spoedig vermeerder. Indien elektrodes nader aan mekaar gespasieer word, verminder die neurale vesels se drempel in die omstaande veselpopulasies, wat ’n vermeerdering in die aktiveringswaarskynlikheid tot gevolg het en sodoende die hoeveelheid vervorming vermeerder. Die neurale aktiveringspatrone wys dat hierdie stroomverspreiding ’n laer effek het wanneer die elektrodes versprei word oor die lengte van die koglea, met ’n spasieëringslengte van 6.08 mm. IAI-histogramme wys groter pieke by die toonperiode en veelvoude daarvan. The neurale aktiveringspatrone dui daarop dat alhoewel voorplantingsgolwe in prosesseringstrategieë meestal geïgnoreer is, hierdie golwe die gehoorstelsel dalk mag voorsien van die nodige tyd- en ruimteafhanklike inligting wat nodig is vir verbeterde toonhoogtewaarneembaarheid.

Acknowledgements

- Thank You Lord Jesus that You are my supply of Life. You have supplied me through these last few years. Thank You for still transforming my heart into a heart of flesh, which is more open to You. You have gently revealed more of Yourself to me. You have encouraged me, and have given me rest! Thank You that You are good and righteous and that You are God, the King of heaven and earth.

- To the most beautiful girl, Pou, you are awesome! The times with you, I treasure so much. They are so special to me. Thank you for your wonderful attitude, and all your encouragement. Thank you for understanding when I had to work. Thank you for all our couch conversations, for all the hugs, and all that times you were thinking of me. You have blessed me this year and half in so many ways and I am crazy about you. You are the girl and princess of my heart.

- To all the Bio-group members (Piti, Tiaan, Marianne, Ilze, Dirk, Robert, Stuart, Suzanne, Rikus, Linda, Kobie, Ramana, Pieter and Renier), thank you for all the good and awesome times that we had. Thank you for all the motivation, the coffees, the good conversations and the encouragement. A special thanks to the lunch club. Eating together with you guys were a lot of fun.

- A special thanks to Prof J.J. Hanekom for all the countless hours of revision, meetings and all the help, leading and encouragement through these past few years. Thank you prof for all the good conversations about random topics. Thank you for an open door, and an open ear.

- I also want to thank Prof T. Hanekom for the quick internal revision and for her encouragement and motivation through the years.

- I also want to thank my parents for all their prayer, their caring hearts and for all their financial support. How can I ever repay you, mom and dad, for that which you have done for me. I really appreciate you so, so much and I love you dearly.

- To my two brother, Niclaas and Cronje. You guys are awesome! Thank you for your friendship, for your brotherly love and for all your support.

- Then all of my friends: Ina, Christa, Carla, Richa, Tiaan, Piti and Christel, Marianne, Rikus and Albe, Cornelis and Anna, Trienko, Cornelia, Lielie, Elmien, Dawid, Carina, Gerda, Jacobus and Jannike, AB, Lize and FD, Christene, Fred and Gene, Albert and Jolene, Tinus, Ruan, Viaanus, Bevan, Halmar, Savo and Maria, Schantal, Clechia, Pieter and Carina, Elaine, Marge, Rudo, Tommie, Hano, Lana, Charl, Paulus, Paul, Rudo, Janneman, Jacques, Karin, Stefan and Gorethi, Dirk, Dave and Duan. You are all awesome! You have blessed me so, so much through all these years and I have learned so much from you. Thank you for your friendship, for all your motivation, encouragement, your caring and interest and just for being there.

- Thank you uncle Frank and anti Pollie for all the lunches, an open home and open hearts. I really appreciate you so much. Also to Carl and Lee, and Lili and Werner, thank you for your interest and encouragement.

- Thank you to the National Research fund for their financial support.

LIST OF ABBREVIATIONS

ACE	:	Advanced Combination Encoder
BM	:	Basilar membrane
BF	:	Best frequency
CF	:	Characteristic frequency
CIS	:	Continuous Interleaved Sampling
CI	:	Cochlear implant
CL	:	Current level
cps	:	cycles per second
DOF	:	Degrees of freedom
FFT	:	Fast Fourier transform
ISI	:	Inter-spike interval
LGF	:	Loudness growth function
LIN	:	Lateral Inhibitory Network
pps	:	pulses per second
PST	:	Post-stimulus time
SMSP	:	Spectral Maxima Sound Processor
SFOAE	:	Stimulus-frequency Otoacoustic Emissions
TW	:	Traveling wave

Contents

Chapter 1

Introduction

1.1 PROBLEM STATEMENT

Cochlear implants are implantable prosthetic devices that provide previously deaf individuals with the ability to regain their hearing ability partially. Although much progress has been made in cochlear implant research, there are still many questions and limitations regarding the implementation of cochlear implants. One of these areas of uncertainty is the sound processing strategy, which replaces the function of the outer, middle and part of the inner ear. Processing strategies are mainly aimed at transferring as much information as possible to the auditory system. Strategies such as SPEAK and Continuous Interleaved Sampling (CIS) focus on increasing either the spectral information via place coding or the temporal information conveyed via rate coding. The Advanced Combined Encoder (ACE) strategy of Cochlear [1] aims at improving both spectral and temporal information, by combining these two coding techniques (Kim, Shim, Chung and Lee, 2000). Although the auditory system uses this spatial and coarse temporal information provided by electrical stimulation, it seems

[1] www.cochlear.com

that the fine spatio-temporal information needed for high-fidelity sound remains inadequate (Clark, 2004), resulting in rather poor overall hearing ability compared to normal-hearing individuals. One of the possible contributing factors is that information is not presented in the same way as in a normal hearing auditory system, where travelling waves on the basilar membrane convey the sound information (Clark, 2004).

It has been known since von Békésy (1947) published his findings that travelling waves exist in the cochlea. He discovered that with sinusoidal vibrations on the stapes, the phase difference between the stapes and a specific place on the basilar membrane (BM) differed along the length of the cochlear partition, indicating a travelling wave, travelling from the base to the apex. As the travelling wave approaches the specific place related to the tonotopicity of the cochlea, its wavelength shortens, causing the wave amplitude to reach a maximum peak at that place along the cochlea. The frequency related to this place is called the characteristic frequency (CF). Greenwood (1990) derived equations for this frequency-place relationship from measured data for several species. Because of the anatomical characteristics of the cochlear structures, the travelling wave's velocity is much faster in the base, resulting in a latency of less than a millisecond. As the wave travels towards the apex, this latency increases dramatically, up to a few milliseconds. The latency seen in tone frequencies in the auditory cortex, measured by Roberts, Ferrari and Poeppel (1998) might be an expansion of the travelling wave latency (Greenberg, Poepel and Roberts, 1998). Greenberg and colleagues further argue that the travelling wave might be the first step in temporally pre-formatting the activity pattern entering the cochlear nuclei (Greenberg, 1997).

Travelling waves have been largely ignored in cochlear implant (CI) processing strategies. Only a few examples of attempts to incorporate characteristics of the travelling wave in CI processing appear to be available in the literature. Recent research by Taft, Grayden and Burkitt (2009) incorporates a set of calibrated travelling wave delays into a sound processing

strategy. Each band-pass filter output was shifted in time relative to the travelling wave delay. Implementing these delays into six subjects' own processing strategies yielded a significant improvement in sentence recognition in noise for all the subjects. They concluded that one explanation for the improvement might be that the de-synchronising in timing information minimises contention across frequency bands when the maxima is calculated. This improves the maxima selection process. In a later study, Taft, Grayden and Burkitt (2010) tried to explain this improvement in speech perception (Taft *et al.*, 2009). In this study three of the eight subjects showed an improvement in word recognition in quiet conditions. Speech perception results were similar, whether low frequency reversed or random cochlear delays were used. They concluded that the results in this study showed that these improvements are merely an improvement in the processing strategy rather than a cochlear wave mechanism.

Clark (2004) discusses the implementation of a travelling wave processor where overlapping filter bands with added phase delay describe the BM travelling wave. He mentions that a preliminary study showed that five subjects found distinguishing high-pitched sounds from low-pitched sounds easier. Further observations from this study showed better pitch ranking between 150 Hz and 250 Hz, compared to ACE, with subjects finding it easier to recognise musical instruments. Blamey, Swanson, McDermott, Patrick and Clark (2006) developed a real time travelling wave sound processor based on psychophysical tuning curves in normal hearing individuals. This processor not only implements a travelling wave delay with band-pass filter outputs, but also excites a broad range of overlapping regions, aiming to approximate the spatio-temporal neural excitation patterns induced by the travelling wave. This work has been patented, but at the time of writing the present document, no publications have appeared in open literature that evaluate this strategy.

The travelling wave introduces a fine spatio-temporal resolution, with a broad spatial range of activation (Rhode, 1971), producing stimulation patterns very different from those of current

processing strategies. Although spatial and coarse temporal pitch is represented relatively well by current processing strategies, it is probably this fine spatio-temporal coding that is needed for high fidelity sound (Clark, 2003:788). As mentioned in the previous paragraph, a number of travelling wave aspects have been implemented in sound processors, but none of the studies showed compelling experimental results. The present work considers the practical issues and attempts to predict potential outcomes of travelling wave speech processing in CIs. Specifically, the study investigates the result of the trade-off between stimulation rate and number of channels stimulated in a travelling wave implementation. By implementing a hydrodynamic basilar membrane motion model in a processing strategy, the effect of different temporal and spatial configurations on neural spiking patterns could be investigated.

1.2 RESEARCH QUESTIONS

The following research questions are addressed in this study:

What is the stimulation rate versus number of channels trade-off that needs to be considered when implementing the travelling wave in the processing strategy of a cochlear implant?

When observing the neural spike train patterns for various stimulation rate versus number of channel configurations for the travelling wave strategy, is the electrical auditory system presented with enough information to enable it to extract travelling wave information?

Most processing strategies only consider the place of maximum activation of the BM, corresponding to the tonotopical arrangement of the cochlea. As a result, travelling waves have been largely ignored in processing strategies. However, the relationship between the spectral and temporal properties seems to be complex, when one observes travelling waves in the normal auditory system. The analysis in this study does not attempt to decode the travelling

wave but considers the extent to which it is represented in predicted neural spike patterns, when comparing these patterns to modelled spike patterns for normal hearing. The inherent assumption is that if information is available in the nerve firing pattern, the auditory system will be able to extract and use this information. Thus in this work, the problem is reduced to discovering under which conditions the travelling wave is best reflected in CI-mediated neural spike patterns.

1.3 OVERVIEW OF PROCESSING STRATEGIES

The following two sections serve as introduction to the literature study. A separate literature study chapter has not been included, as each of chapters 2 and 4 includes a literature study. This section gives an overview of the different types of processing strategies in cochlear implants. The next section gives an introduction on how travelling waves are excited and how they travel along the basilar membrane. It also considers the phase shifts related to travelling waves and the effect of these shifts on spiking patterns.

Although travelling waves have been known to exist for a long time (von Békésy, 1947) and are widely accepted (Zwislocki, 1950; Peterson and Bogert, 1950; Duifhuis, 1988; Ruggero, 1994; Ruggero and Temchin, 2007) they have been largely ignored in processing strategies. Commercial processing strategies only consider the tonotopicity of the cochlea, corresponding to the place of maximum excitation of the BM for a pure tone input. The first kind of strategies extracted features such as fundamental frequency information (F0/F1/F2) from the signal and stimulated the appropriate electrodes. These strategies also included the MULTIPEAK (MPEAK) strategy. The CIS strategy, which introduced a set of bandpass filters, was introduced later. Each frequency band is associated with an electrode related to the tonotopical arrangement of the cochlea. Aiming at increasing temporal information, this strategy stimulates fewer channels at higher stimulation rates. The pulses are presented in

a non-overlapping sequence to the electrodes, decreasing cross-interference between electrodes. Strategies such as SPEAK and Spectral Maxima Sound Processor (SMSP) followed, which introduced a larger set of bandpass filters. From these filter outputs a set of maxima channels is determined and then stimulated. Recent strategies, such as the ACE from Cochlear, aim at improving spectral and temporal information. By filtering the signal into a large set of channels and giving the option of stimulating at high pulse rates, spectral information and temporal information are combined for better speech processing (Kim *et al.*, 2000). The ACE and CIS strategies seem to represent rapid envelope variations much better, providing better encoding of temporal information. Although recent strategies have come a long way in improving speech perception in cochlear implantees, speech and especially musical perception compared to that of normal-hearing individuals is far from adequate. (Loizou, 1997; Loizou, 1998; Loizou, 2006) gives a good detailed review of the various implant processors and strategies.

1.4 INTRODUCTION TO TRAVELLING WAVES

In the normal auditory system, the outer ear receives a sound wave, which is channelled by the pinna and translated into a vibration by the tympanic membrane (eardrum). The eardrum is attached to the maleus, incus and stapes. These three ossicles match the high impedance of the perilymph (fluid in the scala tympani and scala vestibuli) to the low impedance of the air. This impedance matching is introduced by the area ratio difference between the eardrum and the stapes and to a smaller extent, the lever ratio of the ossicles (Moller, 1963). This high value of cochlear impedance explains the necessity of an impedance transformation in the middle ear, suggesting that only a small part of the sound energy is lost in the middle ear, leaving most of the energy to be absorbed by the cochlear duct (Zwislocki, 1950). The fluid in the scala vestibuli (top cochlear duct) is compressed through a translational

and “pistonlike” motion by the stapes, which is connected to the oval window (elliptical window)(Guinan and Peake, 1967). This fluid is connected to the fluid in the scala tympani with an opening (helicotrema) at the apex. Thus a compression of the oval window through motion from the stapes is translated into fluid pressure, which is absorbed by the membrane in the round window, connected to the scala tympani (bottom cochlear duct). The deflection in the BM is a result of a difference in pressure between the scala vestibuli and the scala tympani (Peterson and Bogert, 1950).

Von Békésy (1947) discovered that the phase difference between the stapes and a specific place on the BM differed with sinusoidal vibrations on the stapes, indicating a definite progression of travelling waves from the base to the apex. As the travelling wave approaches the specific place related to the tonotopical arrangement, its wavelength shortens, causing the wave amplitude to reach a maximum peak at that place along the cochlea. The frequency associated with this place is called the CF. Beyond this place the wavelength is associated with a high loss and quickly dies out. Shamma (1985a) examined response patterns for populations of auditory nerve fibres in cat and observed that spectral parameters are reflected in the relative phases of synchronised activity in the populations of nerve fibres. Further observations showed small phase shifts for the region basal to the CF, reflecting the fast propagation of the travelling wave. These relative phase shifts increase near the region of the CF, owing to the slowing down of the travelling wave, accumulating rapidly. This rapid accumulation of phase shifts can be seen as steep and localised spatial discontinuities in the otherwise relatively uniform instantaneous response pattern. These discontinuities served as the input to a lateral inhibitory network (LIN) where (Shamma, 1985b) showed that it was possible to extract the formants and low-order harmonics with such a network. It can be concluded that the change in phase shifts close to the CF carries useful temporal information, which could be important in the perception of pitch.

Phase locking to the driving stimulus has been observed from auditory nerve fibre discharges. A study (Kiang, Watanabe, Thomas and Clark, 1965) showed phase locking of cat nerve fibres for 4 000 - 5 000 cycles per second (cps). Squirrel monkey fibre responses showed phase locking to frequencies as high as 4 500 cps (Rose, Brugge, Anderson and Hind, 1967). Another later study on cats (Johnson, 1980), showed synchronised auditory nerve fibre responses at frequencies below 6 kHz. Measured nerve responses in the inferior colliculus of guinea pig showed that locking occurred up to a frequency of 1 kHz (Liu, Palmer and Wallace, 2006). Implementing a BM model that could convey this phase information to CI users might aid their ability to discriminate differences in pitch.

1.5 OBJECTIVES

This study had three primary objectives. The first was to develop a model which would approximate the BM motion with respect to time across the length of the cochlea. The second was to integrate this model into a processing strategy, which could be used in an experimental type setup. The third was to integrate this model with an existing electrical stimulation model to consider the various trade-off limitations in implementing a travelling wave processor.

1.6 APPROACH

BM measured responses (von Békésy, 1960) showed that a travelling wave travels across the BM when the auditory system is stimulated by a certain tone input. Since then, numerous models approximating the BM response have been developed. These models are solved using either a time domain solution (Allen and Sondhi, 1979; Neely, 1981b; Diependaal and Viergever, 1983; Duifhuis, Hoogstraten, Van Netten, Diependaal and Bialek, 1986; Diepen-

daal, Duifhuis, Hoogstraten and Viergever, 1987; Diependaal and Viergever, 1989; Kates, 1991; Van der Raadt and Duifhuis, 1990; Xin, Qi and Deng, 2003) or approximating the BM response with a frequency domain solution (Neely, 1981b; Neely, 1993; Kanis and De Boer, 1993). Implementing a travelling wave processing strategy requires an approximated time domain model response. The model therefore has to be a space-time model where the BM displacement for each time interval can be calculated. Considering the BM motion model, a trade-off has to be made between simplicity, accuracy, and computational time in implementing such a model. Although the cochlear BM motion shows definite nonlinear behaviour (Rhode, 1971), this was not considered.

A passive one-dimensional time domain model (Diependaal and Viergever, 1983; Diependaal *et al.*, 1987) was implemented because of its simplicity and relatively fast computational time. The accuracy of this model implementation was verified with time analysis data from (Diependaal *et al.*, 1987). A middle ear model proposed in (Diependaal *et al.*, 1987; Neely, 1981b) was also implemented along with the travelling wave model to obtain a more accurate approximation of the response of the BM to an input signal. The one-dimensional model solution, however, does not consider the movement of fluid along the width and height axes of the cochlea. As a result the transfer function fails to approximate the measured transfer functions accurately. However, the purpose of this study was not to approximate the transfer functions accurately, but to investigate the effect of travelling waves in a processing strategy and therefore the one-dimensional solution was considered adequate.

The one-dimensional model was adapted into a processing strategy. Approximating a very simple hair cell model, the BM model's displacement is rectified and resampled both in space and time. To account for the sequential stimulation phase shift, the active channels were further resampled at a frequency of 14 400 Hz, corresponding to the maximum stimulation rate for the activated channels in the Nucleus device. The model results were further

resampled in space, corresponding to the 22 electrodes in the Nucleus electrode array. As the maximum frequency is divided between the number of activated channels, a trade-off remains between the number of channels and the stimulation rate. The resampled values were compressed within a loudness range, ranging between the threshold and comfortable current levels for each electrode. The compression function was implemented using a nonlinear loudness growth curve (LGF). These current levels were transformed to current amplitudes spanning the electrode stimulation current range in microamperes.

The trade-off was investigated by integrating the travelling wave processing strategy with an electrical stimulation model (Bruce, White, Irlicht, O'leary, Dynes, Javel and Clark, 1999a; Bruce, Irlicht, White, O'leary, Dynes, Javel and Clark, 1999b; Smith, 2011). The current amplitude (μA) at every time instant serves as the input to the electrical stimulation model (Smith, 2011). The first part of the model describes the excitation spread across the nerves due to the current flow through the electrodes (Smith, 2011; Strydom and Hanekom, 2011). The second part implements an electrical stimulation model, which was adapted from the models of (Bruce *et al.*, 1999a; Bruce *et al.*, 1999b).

Finally, spatio-temporal spike train patterns, inter-spike interval (ISI) histograms and post-stimulus spike (PST) histograms for various electrode and stimulation rate configurations were generated. These configurations were investigated in comparison to a proposed normal hearing model.

1.7 CONTRIBUTION

Current processing strategies attempt to provide the auditory system with as much information as possible. When travelling waves are observed, there seems to be a close relationship between temporal and spatial information, which has not been incorporated in processing

strategies. Because of the various (current, electrode, stimulation etc.) limitations in CIs, the implementation of travelling waves in a processor remains questionable. However, research into topics such as current steering, creating virtual channels and higher stimulation rates, might increase the resolution seen in the cochlea.

This study does not present the travelling wave processor as the better or worse type of processor, but aims at investigating the possibility of improved encoding with such a processor. The study investigates the implementation of a travelling wave strategy, with reference to channel-number stimulation-rate configurations, using a existing electrical stimulation model.

1.8 LAYOUT OF DISSERTATION

Chapter 2: This chapter gives background on the different functional blocks of the travelling wave model. The implementation of travelling waves in a processor requires a time domain model motion. Model dimensionality and its importance are discussed, with a thorough literature study of the various BM motion models. The focus, however, remains on a time-space model where the BM displacement for each time interval can be described. Although the aim is to provide a relative good approximation of the BM motion, the middle ear provides impedance matching between the sound waves from the eardrum and the cochlea and also needs investigation. This chapter continues to introduce the fundamental equation for the fluid dynamics in a one-dimensional model, giving background on these equations. The various implementations of these equations and the optimisation techniques are described, with reference to the complexity of implementation and computational time.

Chapter 3: The results of the model implementations in chapter 2 are presented and discussed, with reference to the frequency response transfer functions, latency and the frequency-

to-place mapping. The time analyses for various pure tones are compared and presented. These data from the model are compared with measured squirrel monkey (Rhode, 1971) and cat data (Guinan and Peake, 1967; Moller, 1963).

Chapter 4: This chapter presents the implementation of the travelling wave model in a processing strategy. Temporal and spectral limitations are investigated, with reference to literature. This implementational model is integrated with current excitation and electrical models, to investigate the effect of various channel-number stimulation-rate configurations in electrical stimulation on nerve response patterns.

Chapter 5: The chapter considers the results for a proposed normal hearing model. These results are compared to results obtained with the electrical model, to study the effect of various channel-number stimulation-rate configurations in travelling wave processors.

Chapter 6: The results in chapter 5 are discussed, with reference to the channel-number stimulation-rate trade-off encountered when implementing a travelling wave processor.

Chapter 7: The dissertation is concluded with a brief overview of how the study answered the research questions, with reference to the significance to the wider research field. Along with this, possible future research and the limitations of this study are also briefly discussed.

Chapter 2

Implementation of a travelling wave model

2.1 INTRODUCTION

Basic anatomy shows that the normal human auditory system is divided into four functional blocks, namely the outer ear, the middle ear, the cochlea and the hair-cell transduction region between the hair cells and nerves. The first part of this study was undertaken to find an adequate hydrodynamical model of the cochlea that could be implemented within a CI processing algorithm. This hydrodynamical model should describe the transformation of the BM in the cochlea with respect to time.

2.2 BACKGROUND

2.2.1 Dimensionality

The simplest implementation of the BM movement, the one-dimensional model, assumes that the biggest percentage of the fluid follows a laminar flow. The result of this assumption is that the fluid is displaced in only one direction, the longitudinal direction (along the length of the cochlea) owing to uniform pressure distribution along the length of the cochlea. This further assumes that the cross-section of the cochlear canal is small compared to the propagating wave's wavelength ($\lambda >> 2\pi h$, h being the cochlear channel height and λ the wavelength of the propagating wave), resulting in the upper wall sound pressure being only slightly smaller than the pressure just above the basilar membrane (De Boer, 1997). This is also known as the long wave theory. The transmission line model introduced by Zwislocki (1948) incorporated the travelling wave experimental results (von Békésy, 1928; von Békésy, 1944). This model describes the transversal and longitudinal coupling in the cochlear system as an electrical network. Peterson and Bogert (1950), and Zwislocki (1950) introduced a hydrodynamical one-dimensional model, which described the difference in pressure between the two cochlear channels. Several models followed (Fletcher, 1951; Diependaal and Viergever, 1983; Neely and Kim, 1986; Wickesberg and Geisler, 1986; Diependaal *et al.*, 1987), which implemented the fundamental differential equation that (Peterson and Bogert, 1950; Zwislocki, 1950) used. Results from these models, however, only showed a qualitative correlation to the measured results of (Rhode, 1971), which showed a much narrower frequency response. This is a result of the long wave approximation, where the cochlear canal cross-section is small compared to the propagating wave's wavelength. Near the place of resonance where the wavelength is smallest, this assumption does not hold. At this place the wavelength becomes approximately equal to the canal cross-section (λ is approximately equal to h)

(De Boer, 1980; Viergever, 1986). A more extensive study of the place of maximum BM excitation is necessary to understand cochlear sharpening, frequency selectivity and nonlinear effects. Therefore fluid movement needs to be investigated in other directions apart from the longitudinal. These directions are considered in two- and three-dimensional models, the difference between these two models being that two-dimensional models consider the pressure change in the longitudinal x and transversal y direction (y being perpendicular on the basilar membrane), whereas three-dimensional models also include the pressure across the width of the membrane (Viergever, 1986).

Ranke (1950) was the first investigator to consider a two-dimensional model. He made numerous assumptions in order to solve his boundary-value problems. One of these assumptions was that the cochlea was not treated as a closed cavity. These assumptions might be the reason why his work was mostly ignored. Lesser and Berkley (1972) introduced a complete partially numerical solution for an enclosed two-dimensional cavity model where the cochlea is represented as a rectangular box. Allen (1977) improved their numerical solution by deriving an integral equation using the Green's function method, reducing computational time with increased accuracy. Sondhi (1978) approximated the transformed integral formulation of (Allen, 1977), obtaining a much faster solution with the same accuracy. Making use of fast fourier transform (FFT) convolution methods, Allen and Sondhi (1979) introduced a finite difference method based on the methods of Allen (1977) and Sondhi (1978). Finite difference methods are used to simplify the implementation of differential equations and require less computing power. Other models also implemented such methods (Neely, 1981a; Steele and Taber, 1979).

Diependaal and Viergever (1989), Viergever (1980) and Viergever and Kalker (1975) followed a similar method to (Sondhi, 1978) by transforming the boundary value problem into an equivalent integral having only one independent variable. Diependaal and Viergever

(1989) managed to solve this equation for a time domain solution as opposed to (Sondhi, 1978; Viergever, 1980; Viergever and Kalker, 1975), who solved this integral equation for a frequency domain solution. The results of these two-dimensional models showed much more sharpened frequency responses (Rhode, 1971) compared to one-dimensional models (Allen, 1985).

To consider the fluid flow variations in the direction lateral to the BM, it was necessary to consider three-dimensional models. Although various three-dimensional models have been developed over the years, these models have not been considered in this study, since a one- or two-dimensinal model was found adequate in describing the travelling wave movement in a way that could be implemented in a speech processor. For a further review on cochlear models and three-dimensional models see (Lighthill, 1981; Kim, 1986; Viergever, 1986).

2.2.2 Cochlear mechanics in cochlear models

Having considered the dimensionality of cochlear models, Viergever (1986) stated that macro-mechanical model results could not be matched adequately to experimental results (Rhode, 1971; Sellick, Patuzzi and Johnstone, 1982). These results showed sharpened tuning responses, with active and nonlinear processes evident in the cochlear mechanics. Considering nonlinearity in measured results, experiments by von Békésy (1947) showed an apical phase shift for increasing loudness. These experiments, however, were conducted at above normal hearing thresholds. No change in the pattern of vibration could be observed for intensities in the normal hearing threshold range, therefore von Békésy (1960:481) concluded that the cochlea was a linear system. Several models have been developed based on the assumption that the cochlea is a linear system (Lesser and Berkley, 1972; Diependaal and Viergever, 1983; Neely and Kim, 1986; Mammano and Nobili, 1993). Rhode (1971) did measurements on the vibrations of the BM in squirrel monkeys and showed a definite me-

chanical nonlinearity in the cochlea. Several models (Hubbard and Geisler, 1972; Kim, Molnar and Pfeiffer, 1973; Allen and Sondhi, 1979) aimed at explaining this nonlinear behaviour, suggesting that nonlinearity in the damping term of the cochlear-partition dynamics might be the reason for Rhode's findings. Zweig (1976) stated that a response is linear when the amplitude and phase are related to each other for a certain point on the membrane. His results showed a disagreement for high frequencies. He concluded that nonlinearities might be responsible for this disagreement. Following Rhode's measurements Sellick *et al.* (1982) measured BM responses in guinea pigs using the same technique and concluded that there might be an active process that amplifies energy around the place of the CF. This process is also called the "cochlear amplifier", resulting in high amplification and frequency selectivity Skrodzka (2005). He gives a good description of this active process. In the passive case however BM displacements are smaller and excitation of inner hair cells only results from BM motions (Skrodzka, 2005). He developed a macro-mechanical active model based on a passive BM model with physical parameters chosen as close to human cochlea parameters as possible. Input-output velocity functions for the active model showed compressibility due to loudness and additional amplification, whereas the passive model showed a linear pattern.

Kim (1986) later concluded that for a model to describe cochlear dynamics adequately, it should incorporate bidirectional coupling between the cochlear partition mechanics and the hair-cell transduction processes (Neely and Kim, 1983). Van Netten and Duifhuis (1983) incorporated active and nonlinear behaviour into the damping of the membrane. They implemented a model where the local impedance is described by a Van der Pol oscillator. Diependaal and Viergever's (1983) numerical analysis for the partial differential equation for the BM motion was very different from the results obtained by Van Netten and Duifhuis (1983). Duifhuis *et al.* (1986) investigated this difference in result but concluded that a gap remains between data and analysis. They stated that the unexpected results that the Van der Pol oscillators produced could be expected from complex nonlinear systems. Van der Raadt and

Duifhuis (1990) observed the effects of adding one strong active oscillator in a passive and active enviroment. The first case investigates a passive model with an active oscillator and the second case investigates the effect of a more significantly active oscillator in an active environment. Their preliminary results showed that for both cases the dominant oscillator affects the response pattern basalward from its location.

In view of the complexity of nonlinear models, the numerical analysis of these models in the time domain is very time-consuming. Asymptotic methods are computationally faster and usually lead to better understanding of physical mechanisms and nonlinearity (Diependaal and Viergever, 1983; Kanis and De Boer, 1996). However, most of these methods transform the analysis to the frequency domain. Kanis and De Boer (1996) compared a time domain model (Diependaal *et al.*, 1987) with a quasilinear frequency domain approximation method (Kanis and De Boer, 1993). The frequency domain analysis showed a computational advantage compared to the time domain analysis. This advantage becomes even more apparent as the number of computational steps increases. The results for the two models compared well, except for a mismatch found with the computation of combination tones, which they reported as a possible error in the quasilinear analysis.

Implementing a BM displacement model in a travelling wave processor algorithm results in a displacement solution that needs to be calculated for every time interval, therefore the resulting solution would be in the time domain. Allen and Sondhi (1979) used a finite difference method and FFT convolution in analysing a two-dimensional model in the time domain. Another time domain model (Diependaal and Viergever, 1983) used Galerkin's method to reduce the partial differential equation for the fluid dynamics to a set of ordinary differential equations. Diependaal and Viergever developed a numerical solution for a cochlear partition having nonlinear and active mechanical properties. As part of their analysis they also developed an analysis for a partition having linear and passive properties, which compared well

with a frequency domain finite element solution (Borsboom, 1979).

Although literature results show the importance of active and nonlinear processes in the cochlea, these processes increase the complexity of the models approximating the cochlear dynamics. For the purpose of this first investigative study a less complex, passive linear model (Diependaal and Viergever, 1983) was adequate to approximate the cochlear dynamics as part of the speech processing strategy. In their proposed solution they also described an active nonlinear solution, which might be implemented in a later study for a better approximation of the BM displacement. (Diependaal and Viergever, 1989) also included a middle ear model which was added to the hydrodynamical model approximating the BM displacement.

This next section supplies background to the underlying equation used in the BM models proposed by Diependaal and Viergever (1983) and Diependaal *et al.* (1987). It also provides an overview of the equations used in these models and their respective boundary conditions.

2.2.3 Foundational equations

Zwislocki (1950) described two ways in which the dynamics of fluid in the cochlea can be solved mathematically. The two solutions involve either treating the problem as a boundary value problem or expressing the whole system as a set of differential equations, setting boundary conditions only at the base and apex. The first method proves to be more difficult when simplifying methods, whereas the second solution can only be applied to long waves in relation to the cochlear duct cross-section. This solution's calculation can be made independent of the shape of the cochlear duct and the bending pattern of the cochlea. Following the second method, Zwislocki introduced the differential equation (eq. 2.1) that describes the movement of the BM. This equation is based on a number of fundamental fluid dynamic

conditions and is independent of cochlear duct impedance variations, canal shape or time. The dynamic differential equation (Zwislocki, 1950) is:

$$\frac{\partial^2 p}{\partial x^2} = 2\frac{\rho}{F_0 J}\frac{\partial p}{\partial t} + 2\frac{R_1}{F_0 J}p \tag{2.1}$$

where x - distance from the stapes, t - coordinate in time, p - sound pressure, ρ - fluid density, R_1 - coefficient of the perilymph per unit length of the canal, F_0 - average cross-section of the canal, J - impedance of the cochlear duct per unit length.

Through further mathematical analysis, Zwislocki (1950) proved that as a result of friction, maximum oscillations cannot occur at the place of resonance, but occur before this site. This means that travelling waves are eminent in the cochlea. Peterson and Bogert (1950) defined two differential equations for the two channels of the cochlea. They further defined the pressure difference as the force that causes the deflection of the membrane. This pressure difference was defined as a function of the velocity of the wave and the stiffness, mass and resistance of the membrane (eq. 2.2).

$$p_v - p_t = m\dot{v} + rv + k\int^t vdt \tag{2.2}$$

where $v(x,t)$ is the vertical membrane velocity, and $m(x)$, $r(x)$ and $s(x)$ are the mass, resistance and the stiffness per unit area respectively. Viergever (1980) defined a differential equation for transmembrane pressure, following the formulation of models such as Zwislocki (1950) and Peterson and Bogert (1950). The geometry of the model of Diependaal and Viergever (1983) and Diependaal *et al.* (1987) was similar to that of (Viergever, 1980). The partial differential equation describing the BM response due to stapes movement $u(x,t)$ is:

$$p'' - A(x)\,\ddot{u}(x,t) = 0, \tag{2.3}$$

where the double prime is the double differentiation with respect to space and the double dot is the double differentiation with respect to time. The transmembrane pressure $p(x,t)$ is defined by

$$p(x,t) = m(x)\,\ddot{u}(x,t) + r(x,t)\,\dot{u}(x,t) + s(x,t)\,u(x,t). \tag{2.4}$$

The mass, $m(x)$ is taken to be time-independent, whereas the BM resistance, $r(x,t)$ and the stiffness of the BM $s(x,t)$ are time-dependent, being functions of $\dot{u}(x,t)$ and $u(x,t)$. These time-dependent functions make the model nonlinear. For the implementation used in the processing algorithm, the BM resistance and the stiffness were independent of time, resulting in a passive linear model. The model was implemented with the following initial and boundary conditions:

$$u(x,0) = 0, \qquad 0 \le x \le l, \tag{2.5}$$

$$\dot{u}(x,0) = 0, \qquad 0 \le x \le l, \tag{2.6}$$

$$p(l,t) = 0, \qquad t \ge 0. \tag{2.7}$$

Eq. (2.7) takes into account that the transmembrane pressure disappears at the helicotrema $(x = l)$. At the base $(x = 0)$ the simplest boundary condition is a prescribed movement of the stapes:

$$p'(0,t) = f(t), \qquad t \geq 0, \tag{2.8}$$

where prime is the differentiation with respect to space and $f(t)$ is the acceleration of the displacement of the stapes[1]:

$$f(t) = \ddot{u}_{stapes}(t), \qquad t \geq 0. \tag{2.9}$$

The next section describes different methods for solving these one-dimensional model equations. A comparison of these different methods with reference to their processing time gave a good indication of which method to implement.

2.3 METHODS FOR SOLVING ONE-DIMENSIONAL MODEL EQUATIONS

2.3.1 Finite difference iterative method

The hyperbolic equation (eq. 2.3) can be solved using an iterative method with the derivation of difference equations. This algorithm iterates the (*space* x *time*) matrix, where *space* is defined as the number of intervals of the BM length and *time* is the number of time increments.

[1]Note that in (Diependaal *et al.*, 1987), $f(t)$ is described as the movement of the stapes. This description might be confusing, as $f(t)$ is the acceleration of the stapes as mentioned in (Wickesberg and Geisler, 1986).

2.3.1.1 Method

The methodology for the iterative procedure is:

1. Derive the equations for the boundary conditions eqs 2.18, 2.31 and 2.32 at the boundaries.
2. Derive the pressure distribution equation, eq. 2.20.
3. Define the initial conditions for the starting time, eqs 2.12 and 2.13.
4. Assume a starting value for every index in the (*space* x *time*) matrix, where *space* is defined as the number of intervals of the BM length and *time* is the number of time increments.
5. Determine the value for every index from the boundary and pressure distribution equations.
6. Determine the residual value. This value is determined by subtracting the new index value from the previous one.
7. Steps 5 and 6 are repeated until the residual value for every index is smaller than the residual error. As the error decreases, the iterations and the processing time also increase. The optimal error was chosen to be 0.0001.

From the boundary condition equations, eq.s 2.3 - 2.9, one can derive boundary conditions by using difference equations:

$$\dot{u}\left(x_i,t_j\right)=\frac{u\left(x_i,t_{j+1}\right)-u\left(x_i,t_{j-1}\right)}{2\Delta t},\qquad \dot{u}\left(x_i,t_j\right)=\frac{u\left(x_i,t_{j+1}\right)-u\left(x_i,t_j\right)}{\Delta t}, \tag{2.10}$$

and

$$\ddot{u}\left(x_i,t_j\right)=\frac{u\left(x_i,t_{j-1}\right)-2u\left(x_i,t_j\right)+u\left(x_i,t_{j-1}\right)}{\left(\Delta t\right)^2}. \tag{2.11}$$

Thus the first initial condition remains unchanged

$$u\left(x,0\right)=0, \qquad 0\leq x\leq l. \tag{2.12}$$

Using the difference equations, the second initial condition becomes

$$\ddot{u}\left(x_i,t_j\right)=\frac{u\left(x_i,t_{j+1}\right)-2u\left(x_i,t_j\right)}{\left(\Delta t\right)^2}=0 \quad \Rightarrow \quad u\left(x_i,t_1\right)=u\left(x_i,0\right)=0. \tag{2.13}$$

Substituting the difference equations into eq. 2.4, one obtains

$$\ddot{p}\left(x_i,t_j\right)=\left[Dx_i\right]u\left(x_i,t_{j-1}\right)+\left[Ex_i\right]u\left(x_i,t_j\right)+\left[Cx_i\right]u\left(x_i,t_{j+1}\right) \tag{2.14}$$

where

$$Cx_i=\frac{m\left(x_i\right)}{\left(\Delta t\right)^2}+\frac{r\left(x_i,t_j\right)}{2\Delta t}, \tag{2.15}$$

$$Dx_i = \frac{m(x_i)}{(\Delta t)^2} - \frac{r(x_i,t_j)}{2\Delta t}, \tag{2.16}$$

and

$$Ex_i = -\frac{2m(x_i)}{(\Delta t)^2} + s(x_i,t_j). \tag{2.17}$$

Then the first boundary condition at $(x = l)$, $p(l,t) = 0$ can be written as:

$$\ddot{u}(x_l,t_{j+1}) = -\frac{1}{[Cx_l]}\left[[Dx_l]\,u(x_l,t_{j-1}) + [Ex_l]\,u(x_l,t_j)\right]. \tag{2.18}$$

Taking the partial differential equation (eq. 2.3) and writing this equation into difference equations where time differentiation can be written in the form of eq. 2.13 and space differentiation can be written as

$$p''(x_i,t_j) = \frac{p(x_{i-1},t_j) - 2p(x_i,t_j) + p(x_{i+1},t_j)}{(\Delta x)^2} \tag{2.19}$$

one obtains the equation

$$u\left(x_{i+1},t_{j+1}\right)=\frac{-1}{\psi_I}\Big[\Psi_B u\left(x_i,t_{j-1}\right)+\Psi_E u\left(x_i,t_j\right)+\Psi_H u\left(x_i,t_{j+1}\right)+\Psi_A u\left(x_{i-1},t_{j-1}\right)+\Psi_D u\left(x_{i-1},t_j\right)+\Psi_G u\left(x_{i-1},t_{j+1}\right)+\Psi_C u\left(x_{i+1},t_{j-1}\right)+\Psi_F u\left(x_{i+1},t_j\right)\Big] \tag{2.20}$$

where

$$\Psi_A\left(x_i,t_j\right) = m\left(x_{i-1}\right)-\frac{\Delta t}{2}r\left(x_{i-1},t_j\right), \tag{2.21}$$

$$\Psi_B\left(x_i,t_j\right) = -2m\left(x_i\right)+\left(\Delta t\right)r\left(x_i,t_j\right)-A(\Delta x)^2, \tag{2.22}$$

$$\Psi_C\left(x_i,t_j\right) = m\left(x_{i+1}\right)-\frac{\Delta t}{2}r\left(x_{i+1},t_j\right), \tag{2.23}$$

$$\Psi_D\left(x_i,t_j\right) = -2m\left(x_{i-1}\right)+\left(\Delta t\right)^2 s\left(x_{i-1},t_j\right), \tag{2.24}$$

$$\Psi_E\left(x_i,t_j\right) = 4m\left(x_i\right)-2(\Delta t)^2 s\left(x_i,t_j\right)+2A(\Delta x)^2, \tag{2.25}$$

$$\Psi_F\left(x_i,t_j\right) = -2m\left(x_{i+1}\right)+\left(\Delta t\right)^2 s\left(x_{i+1},t_j\right), \tag{2.26}$$

$$\Psi_G\left(x_i,t_j\right) = m\left(x_{i-1}\right)+\frac{\Delta t}{2}r\left(x_{i-1},t_j\right), \tag{2.27}$$

$$\Psi_H\left(x_i,t_j\right) = -2m\left(x_i\right)-\left(\Delta t\right)r\left(x_i,t_j\right)-A(\Delta x)^2 \tag{2.28}$$

$$\Psi_I\left(x_i,t_j\right) = m\left(x_{i+1}\right)+\frac{\Delta t}{2}r\left(x_{i+1},t_j\right). \tag{2.29}$$

The difference equation for the transmembrane pressure is

$$p^{'}\left(x_i,t_j\right)=\frac{p\left(x_{i+1},t_j\right)-u\left(x_{i-1},t_j\right)}{2\Delta x}. \tag{2.30}$$

When substituting the above equation and eq. 2.4 into the second boundary condition at $(x=0)$, $p^{'}(0,t)=f(t)$, the following equation describing the displacement at interval x_0 is obtained:

$$\begin{aligned}u\left(x_0,t_{j+1}\right)=\frac{1}{Cx_0}\Big[&[Dx_1]\,u\left(x_1,t_{j-1}\right)+[Ex_1]\,u\left(x_1,t_j\right)\\&+[Cx_1]\,u\left(x_1,t_{j+1}\right)-[Dx_0]\,u\left(x_0,t_{j-1}\right)\\&+[Ex_0]\,u\left(x_0,t_j\right)-\Delta x f\left(t_j\right)\Big].\end{aligned} \tag{2.31}$$

If eq. 2.31 is substituted into eq. 2.20 one obtains the equation at space interval x_1:

$$\begin{aligned}u\left(x_1,t_{j+1}\right)=\frac{1}{\left[\Psi_H+\Psi_G\frac{Cx_1}{Cx_0}\right]}\Bigg[&\left[\Psi_B+\Psi_G\frac{Dx_1}{Cx_0}\right]u\left(x_1,t_{j-1}\right)\\&+\left[\Psi_E+\Psi_G\frac{Ex_1}{Cx_0}\right]u\left(x_1,t_j\right)+\left[\Psi_A-\Psi_G\frac{Dx_0}{Cx_0}\right]u\left(x_0,t_{j-1}\right)\\&+\left[\Psi_D-\Psi_G\frac{Ex_0}{Cx_0}\right]u\left(x_0,t_j\right)+\Psi_C u\left(x_2,t_{j-1}\right)+\Psi_F u\left(x_2,t_{j-1}\right)\\&+\Psi_I u\left(x_2,t_{j+1}\right)-\frac{\Psi_G}{Cx_0}\Delta x f\left(t_j\right)\Bigg].\end{aligned} \tag{2.32}$$

2.3.1.2 Model outputs

The results of the predicted travelling wave BM displacement for a 2 kHz pure tone input are shown for a specific time interval in fig. 2.1.

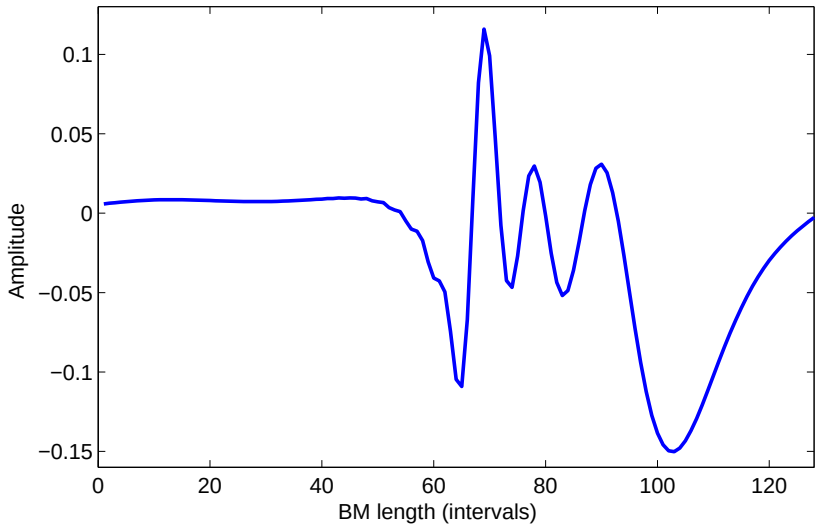

Figure 2.1: Output of the FDI (Finite Difference Iterative) model for a 2 kHz pure tone.

The transient seen is due to the switching on of the input signal. When a ramped signal is used, the transient disappears. Although the maximum peak compares well with the results obtained by Diependaal *et al.* (1987) fig. 1, the processing time of 29 hours, 36 minutes and 36 seconds (Table 2.3) is much too long and would not be feasible to implement in real time, i.e. in a speech processing algorithm. Therefore it was necessary to obtain a faster computational method for implementation of the travelling wave. Fig. 2.2 shows the spatial and time matrix for the predicted BM displacement.

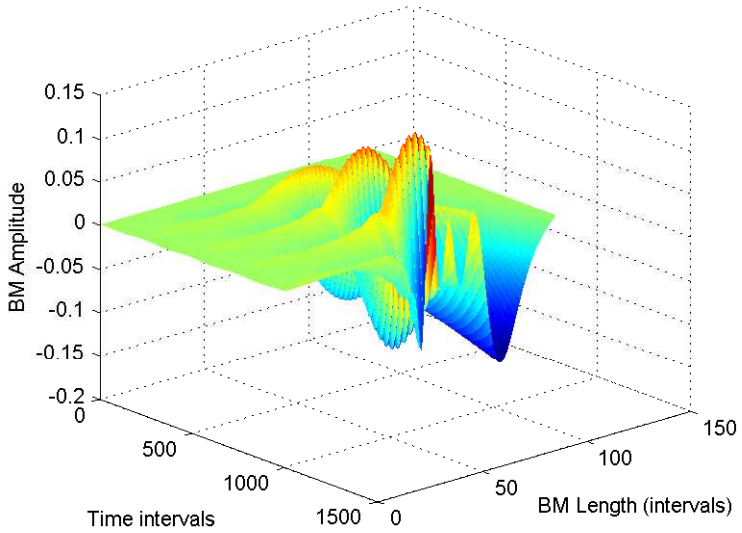

Figure 2.2: The time and space matrix output of the FDI model as the result of a 2 kHz input pure tone.

2.3.2 FDSE method in MATLAB

The second solution method, the Finite Difference Simultaneous Equations (FDSE) method, attempts to discretise the partial differential eq. 2.4 only spatially as opposed to space and time, obtaining a system of ordinary differential equations that can be solved by way of numerical integration in time. This system of equations is then simultaneously solved for one instance in time, making this method ideal for real time implementation.

2.3.2.1 Method

The difference for the pressure distribution is spatially discretised using the difference equation:

$$p''\left(x_i,t_j\right)=\frac{p\left(x_{i-1},t_j\right)-2p\left(x_i,t_j\right)+p\left(x_{i+1},t_j\right)}{\left(\Delta x\right)^2}. \tag{2.33}$$

When substituting this equation into the partial differential equation eq. 2.3,

$$\begin{aligned} -m(x_{i+1})\,\ddot{u}(x_{i+1},t) = m(x_{i-1})\,\ddot{u}(x_{i-1},t) \\ -\left[A\Delta x^2 + 2m(x_i)\right]\ddot{u}(x_i,t) + r(x_{i-1},t)\,\dot{u}(x_{i-1},t) \\ -2r(x_i,t)\,\dot{u}(x_i,t) + r(x_{i+1},t)\,\dot{u}(x_{i+1},t) \\ +s(x_{i-1},t)\,u(x_{i-1},t) - 2s(x_i,t)\,u(x_i,t) \\ +s(x_{i+1},t)\,u(x_{i+1},t) \end{aligned} \tag{2.34}$$

is obtained. From this one can write a set of equations in the form $\bar{A}\ddot{u} = \bar{b}$.

$$
\begin{pmatrix}
 & 0 & 1 & 2 & i-1 & i & i+1 & n-1 & n \\
0 & -m(x_0) & -m(x_1) & 0 & \dots & \dots & \dots & 0 & 0 \\
1 & -m(x_0) & A(\Delta x)^2+ & -m(x_2) & \dots & \dots & \dots & 0 & 0 \\
 & & 2m(x_1) & & & & & & \\
2 & 0 & -m(x_1) & A(\Delta x)^2+ & -m(x_{i-1}) & \dots & \dots & 0 & 0 \\
 & & & 2m(x_2) & & & & & \\
i-1 & 0 & 0 & -m(x_2) & A(\Delta x)^2+ & -m(x_i) & \dots & 0 & 0 \\
 & & & & 2m(x_{i-1}) & & & & \\
i & 0 & 0 & 0 & -m(x_{i-1}) & A(\Delta x)^2+ & -m(x_{i+1}) & 0 & 0 \\
 & & & & & 2m(x_i) & & & \\
i+1 & 0 & 0 & 0 & \dots & -m(x_i) & A(\Delta x)^2+ & -m(x_{n-1}) & 0 \\
 & & & & & & 2m(x_{i+1}) & & \\
n-1 & 0 & 0 & 0 & \dots & \dots & -m(x_{i+1}) & A(\Delta x)^2+ & 0 \\
 & & & & & & & 2m(x_{n-1}) & \\
n & 0 & 0 & 0 & \dots & \dots & \dots & 0 & -m(x_n)
\end{pmatrix}
\begin{pmatrix}
 \\
\ddot{u}(x_0,t) \\
\ddot{u}(x_1,t) \\
 \\
\ddot{u}(x_2,t) \\
 \\
\ddot{u}(x_{i-1},t) \\
 \\
\ddot{u}(x_i,t) \\
 \\
\ddot{u}(x_{i+1},t) \\
 \\
\ddot{u}(x_{n-1},t) \\
 \\
\ddot{u}(x_n,t)
\end{pmatrix}
= \bar{b} \tag{2.35}
$$

where

$$\bar{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_{i-1} \\ b_i \\ b_{i+1} \\ b_{n-1} \\ b_n \end{pmatrix} = \begin{pmatrix} g(x_0,t) - g(x_1,t) + f(t)\Delta x \\ g(x_0,t) - 2g(x_1,t) + g(x_2,t) \\ g(x_1,t) - 2g(x_2,t) + g(x_{i-1},t) \\ g(x_2,t) - 2g(x_{i-1},t) + g(x_i,t) \\ g(x_{i-1},t) - 2g(x_i,t) + g(x_{i+1},t) \\ g(x_i,t) - 2g(x_{i+1},t) + g(x_{n-1},t) \\ g(x_{i+1},t) - 2g(x_{n-1},t) \\ g(x_n,t) \end{pmatrix}. \tag{2.36}$$

This set of equations can be solved by either dividing $\bar{b}$ by $\bar{A}$, or by taking the inverse of $\bar{b}$ and multiplying it with $\bar{A}$. The following two equations describe the solution for the set of simultaneous equations:

$$\ddot{u}\left(x_{0\ldots n}, t_j\right) = \frac{\bar{b}}{\bar{A}}, \tag{2.37}$$

$$\ddot{u}\left(x_{0\ldots n}, t_j\right) = \bar{A}^{-1} \bullet \bar{b}. \tag{2.38}$$

The BM displacement and velocity can be solved by integrating the BM acceleration using numerical integration methods. Heun's method and the fourth-order Runga-Kutta method are described.

Heun:

$$\left.\begin{aligned} u^* &= u(t) + \Delta t v(t) \\ v^* &= v(t) + \Delta t w(t, u(t), v(t)) \end{aligned}\right\}, \tag{2.39}$$

$$\left.\begin{aligned} u(t+\Delta t) &= u(t) + \tfrac{\Delta t}{2}(v(t) + v^*) \\ v(t+\Delta t) &= v(t) + \tfrac{\Delta t}{2}(w(t, u(t), v(t)) + w(t+\Delta t, u^*, v^*)) \end{aligned}\right\}. \tag{2.40}$$

Fourth-order Runge-Kutta:

$$\begin{aligned} u^{(1)} &= u(t) + \tfrac{1}{2}\Delta t v(t) \\ v^{(1)} &= v(t) + \tfrac{1}{2}\Delta t w(t, u(t), v(t)) \\ u^{(2)} &= u(t) + \tfrac{1}{2}\Delta t v^{(1)} \\ v^{(2)} &= v(t) + \tfrac{1}{2}\Delta t w\left(t + \tfrac{\Delta t}{2}, u^{(1)}, v^{(1)}\right) \\ u^{(3)} &= u(t) + \Delta t v^{(2)} \\ v^{(3)} &= v(t) + \Delta t w\left(t + \tfrac{\Delta t}{2}, u^{(2)}, v^{(2)}\right) \end{aligned} \tag{2.41}$$

$$\left.\begin{aligned} u(t+\Delta t) &= u(t) + \tfrac{\Delta t}{6}\left[v(t) + 2v^{(1)} + 2v^{(2)} + v^{(3)}\right] \\ v(t+\Delta t) &= v(t) + \tfrac{\Delta t}{6}\left[w(t, u(t), v(t)) + 2w\left(t + \tfrac{\Delta t}{2}, u^{(1)}, v^{(1)}\right) + \right. \\ &\quad \left. 2w\left(t + \tfrac{\Delta t}{2}, u^{(2)}, v^{(2)}\right) + w\left(t + \Delta t, u^{(3)}, v^{(3)}\right)\right] \end{aligned}\right\}. \tag{2.42}$$

2.3.3 FDSE vectorised method in MATLAB

A third method follows the same methodology as the FDSE method, but in order to utilise the computational power of MATLAB vectorising code, the FDSE was rewritten to incorporate matrix calculations. This method was proposed in Diependaal and Viergever (1983). The second term in the first line of the matrix, however, is a positive term as opposed to a negative term in Diependaal and Viergever's (1983) matrix.

2.3.3.1 Method

$$\bar{A}\bar{M}\bar{\ddot{u}} = \bar{R}\bar{\dot{u}} \bullet \bar{S}\bar{u} + \bar{k} \tag{2.43}$$

where $\bar{M}$ is the lefthand matrix of (2.35) and $\bar{S}$, having the same form as the equation for $\bar{R}$ (eq. 2.44), where $r(x,t)$ is replaced by $s(x,t)$.

$$\bar{R} = \begin{pmatrix} r(x_0,t) & r(x_1,t) & 0 & \dots & \dots & \dots & 0 & 0 \\ r(x_0,t) & -2r(x_1,t) & r(x_2,t) & \dots & \dots & \dots & 0 & 0 \\ 0 & r(x_1,t) & -2r(x_2,t) & r(x_{i-1},t) & \dots & \dots & 0 & 0 \\ 0 & 0 & r(x_2,t) & -2r(x_{i-1},t) & r(x_i,t) & \dots & 0 & 0 \\ 0 & 0 & 0 & r(x_{i-1},t) & -2r(x_i,t) & r(x_{i+1},t) & 0 & \\ 0 & 0 & 0 & \dots & r(x_i,t) & -2r(x_{i+1},t) & r(x_{n-1},t) & 0 \\ 0 & 0 & 0 & \dots & \dots & r(x_{i+1},t) & -2r(x_{n-1},t) & 0 \\ 0 & 0 & 0 & \dots & \dots & \dots & 0 & r(x_n,t) \end{pmatrix} \tag{2.44}$$

$\bar{k}$,$\bar{\dot{u}}$ and $\bar{u}$ are defined by

$$\bar{k} = \left[\begin{array}{cccccccc} k(t_j), & 0, & 0, & 0, & 0, & 0, & 0, & 0 \end{array} \right]^{-1}, \tag{2.45}$$

$$\bar{\dot{u}} = \left[\begin{array}{ccccccc} \dot{u}(x_0,t), & \dot{u}(x_1,t), & \dot{u}(x_2,t), & \ldots & ,\dot{u}(x_{n-1},t), & \dot{u}(x_n,t) \end{array} \right]^{-1}, \tag{2.46}$$

and

$$\bar{u} = \left[\begin{array}{ccccccc} u(x_0,t), & u(x_1,t), & u(x_2,t), & \ldots & ,u(x_{n-1},t), & u(x_n,t) \end{array} \right]^{-1}. \tag{2.47}$$

2.3.3.2 Processing time

The table shows a comparison of the processing times for the four methods. The first method's simulations were done on a Pentium 4, 1.79 GHz processor with 1 Gigabyte of memory. For processing time comparison purposes, the simulations for the other three methods shown in the table were also done with this processor.

2.4 MIDDLE EAR MODEL

The main function of the middle ear is to provide impedance transformation between the high impedance of the cochlear fluid and the low impedance of the air, improving sound transmission from the eardrum to the perilymph in the cochlea. The ratio in area of the eardrum and the stapes and to a lesser extent the lever ratio of the ossicles result in such an impedance transformation (Moller, 1963). In their one-dimensional model solution (Diependaal *et*

Table 2.1: The time processing of the different implementation methods

Algorithm	Input signal (kHz)	x_n	t_n	Time length (ms)	Time sample rate(ms)	Processing time
FDI [a]	2	129	1000	2	0.002	29 hr, 36 min, 36 s
FDSE-M [b]	2	128	100k	100	0.001	8 min 5 s 420 ms
FDSE-C [c]	2	128	500k	500	0.001	11 min 58 s
FDSE-V [d]	2	128	100k	1001	0.01	2 min, 42 s, 495 ms

[a] Finite difference iterative model
[b] Finite difference simultaneous equations implemented in MATLAB
[c] Finite difference simultaneous equations implemented in C
[d] Finite difference simultaneous equations implemented in MATLAB using vectorised code

al., 1987) considered two ways of describing the boundary condition at the base of the cochlea. The first is a prescribed stapes movement, which also serves as input to the basilar membrane model. They mentioned, however, that a more realistic condition could be derived if the middle ear dynamics were incorporated. This section discusses the equations for the simplified middle ear model that were proposed in Diependaal *et al.* (1987) following the analysis of Matthews (1980) and Neely (1981a).

Neely (1981b) developed a middle ear model based on a circuit model developed by Matthews (1980), which is based on experimental results obtained by Moller (1963) and Guinan and Peake (1967). This model has three degrees of freedom (DOF), with each DOF associated with a mass attached to the temporal bone by means of a certain viscous damping and stiffness. The equations for the motion of the model follow:

$$A_e p_e = M_e \ddot{u}_e + R_e \dot{u}_e + S_e u_e, \tag{2.48}$$

$$A_m p_e = M_m \ddot{u}_m + R_m \dot{u}_m + S_m u_m + G_m S_i (G_m u_m - u_s), \tag{2.49}$$

$$-A_s p_s = M_s \ddot{u}_s + R_s \dot{u}_s + S_s u_s + S_i (u_s - G_m u_m). \tag{2.50}$$

The middle ear function can be simplified to a one-DOF system. In the one-DOF model the eardrum leak is disregarded ($u_e \equiv 0$) and the incudo-malleolar joint is considered to be infinitely stiff ($G_m u_m \equiv u_s$) (Diependaal and Viergever, 1989). Taking these approximations into consideration one finds:

$$A_m p_e + A_s p_s = \frac{M_m}{G_m}\ddot{u}_s + \frac{R_m}{G_m}\dot{u}_s + \frac{S_m}{G_m} + M_s\ddot{u}_s + R_s\dot{u}_s + S_s u_s. \tag{2.51}$$

Dividing by G_m yields

$$\frac{A_m}{G_m}p_e + ap(0,t) = \left(\frac{M_m}{{G_m}^2} + M_s\right)\ddot{u}_s + \left(\frac{R_m}{{G_m}^2} + R_s\right)\dot{u}_s + \left(\frac{S_m}{{G_m}^2} + S_s\right)u_s \tag{2.52}$$

where $\frac{A_s}{G_m} = a$ and $p(0,t) = p_s$.

If $T_m = \frac{1}{G_m}$ then eq. A1 in (Diependaal *et al.*, 1987) could be obtained:

$$T_m A_m p_e + ap(0,t) = \left({T_m}^2 M_m + M_s\right)\ddot{u}_s + \left({T_m}^2 r_m\right)\dot{u}_s + s_m u_s \tag{2.53}$$

where ${T_m}^2 r_m = {T_m}^2 R_m + R_s$ and $s_m = {T_m}^2 S_m + S_s$.

This simplified middle ear model has been integrated with the vectorised FSDE method in MATLAB for solving the BM displacement with respect to time. The next section summarises the parameters used in these two models.

2.5 BASILAR MEMBRANE AND MIDDLE EAR MODEL PARAMETERS

2.5.1 Basilar membrane

The parameter set for both the basilar membrane and the middle ear model was derived from experimental cat data in Neely (1981b).

Table 2.2: The parameter set for the basilar mebrane model.

Fluid density	ρ	1	$mg.mm^{-3}$
Cross-sectional channel area	$a(x)$	1	mm^2
Width of BM	$\beta(x)$	1	mm
Mass	$m(x)$	$0.045\,exp(0.08x)$	$mg.mm^{-2}$
Stiffness	$s(x,t)$	$4600\,exp(-0.32x)$	$mg.mm^{-2}.ms^{-2}$
Resistance	$r(x,t)$	$0.5\,exp(0.05x)$	$mg.mm^{-2}.ms^{-1}$
Length of BM	l	22.5	mm

2.5.2 Middle ear

Matthews (1980) used experimental measurements of cat response data to determine middle ear parameters. The middle ear parameters in (Neely, 1981b) are based upon those parameters set by Matthews.

Pressure is measured in dyn, where 1 dyn is:

$$1dyn = \frac{1g.cm}{s^2}$$

.

Table 2.3: The parameter set for the middle ear model.

Gains	G_m	0.5	
	T_m	2	
Mass	M_s	4.27	gm
	M_m	4.82×10^{-3}	gm
	$T_m{}^2 M_m + M_s$	17.0848	gm
Areas	A_s	1.26	mm^2
	A_m	34.7	mm^2
	a	1	mm^2
Damping	R_s	31.8	$dyn.sec.cm^{-1}$
	R_m	6.65	$dyn.sec.cm^{-1}$
	$T_m{}^2 r_m$	58.4	$mg.ms^{-1}$
Fluid density	ρ	1	$mg.mm^{-3}$
Stiffness	S_s	4.41×10^5	$dyn.cm^{-1}$
(defined as		4410	$mg.ms^{-1}$
K in (Neely, 1981b))	S_m	3.44×10^5	$dyn.cm^{-1}$
		3440	$mg.ms^{-1}$
	s_m	1817	$mg.ms^{-1}$

These parameter sets were used in the predicted neural spike train simulations in chapter 5. The only parameter that was adapted for humans in these simulations was the basilar membrane length. In these simulations the basilar membrane length was normalised to 1 and transformed to a length of 35 mm.

The different solution methods for the one-dimensional cochlear dynamic equations have been described. A comparison of the processing times for the various methods showed that the FSDE method implemented in MATLAB was computationally faster than the other methods. This method has been implemented in the next chapter to solve the one-dimensional dynamic equations in the hydrodynamic BM model. The simplified middle ear model has also been integrated with the BM model. The next chapter gives a summary of the results of both the BM (Diependaal and Viergever, 1983) and the middle ear model (Diependaal *et al.*, 1987).

Chapter 3

Travelling wave model results

3.1 INTRODUCTION

This chapter gives a summary of the results for the BM model implementation of Diependaal and Viergever (1983) and Diependaal *et al.* (1987). The first section investigates the transfer functions for the BM displacement frequency response using the parameter set in section 2.5. It further compares these results with experimental measurements. Along with this model, the middle ear model (Neely, 1981b) and the integration between these two models are also shown. The second part of the chapter investigates the model's time response for a set of pure tones for the parameter set in section 2.5. In the third section the frequency place mapping is discussed and compared to Greenwood's (1990) equation for approximating cat measurements. The last section discusses the travelling wave delay, with reference to experimental measurements. All the results in the first and last section were determined using the model implementation of Diependaal and Viergever (1983), integrated with the middle ear model proposed in (Neely, 1981b; Diependaal *et al.*, 1987). This two-model integration is referred to as the TW model in the rest of the document.

3.2 FREQUENCY RESPONSE

Flanagan (1962) developed a computational model of the cochlea which incorporates rational-function approximations for middle ear transmission and membrane movement as a result of stapes displacement. His BM model is based on the physiological data measured by von Békésy (1960). He also investigated the combined response of the middle ear and the BM response (fig. 3.1).

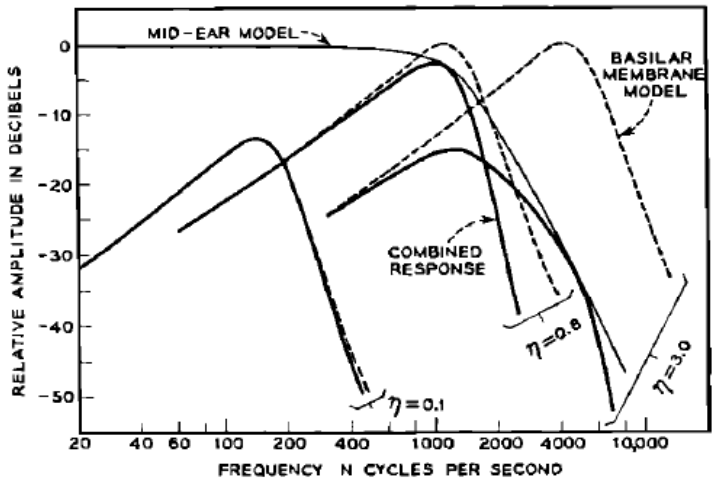

Figure 3.1: Amplitude versus frequency responses for the combined middle ear and BM models. Points on the membrane whose frequency of maximum response is less than the critical frequency of the middle ear are indicated by values $\eta < 1$ (Flanagan, 1962), (reprinted with permission from Flanagan (1962)).

His results showed that the transfer function for the middle ear (fig. 3.1) is relatively constant for frequencies less than the critical frequency around 1 kHz. However, at frequencies higher than the critical frequency of the middle ear, the BM displacement is strongly influenced by the transfer function of the middle ear. Flanagan's combined response for the middle ear and the basilar membrane movement (fig. 3.1) shows this influence of the middle ear. Results obtained in the model show a similar influence on the high frequencies by the middle ear

(fig. 3.2).

The magnitude of the frequency response of the model has been compared with data measured by Rhode (1971). Rhode measured the amplitude and the phase for basilar membrane vibration in the cochlea of living squirrel monkeys using the Moëssbauer technique. These data were measured at the cochlear place corresponding to 8 kHz. A constant magnitude of 40 dB was subtracted from Rhode's data in order to compare his magnitude with the magnitude from the TW model simulations. From his measurements he derived response data between the maleus and BM displacement. In order to compare these data with those from the BM model, the transfer function between the stapes and maleus needs to be determined. This displacement relationship was measured for cat data (Guinan and Peake, 1967). The magnitude and phase ratios for BM-malleus (fig. 3.3) and BM-stapes (fig. 3.4) displacement were compared with Rhode's frequency response data.

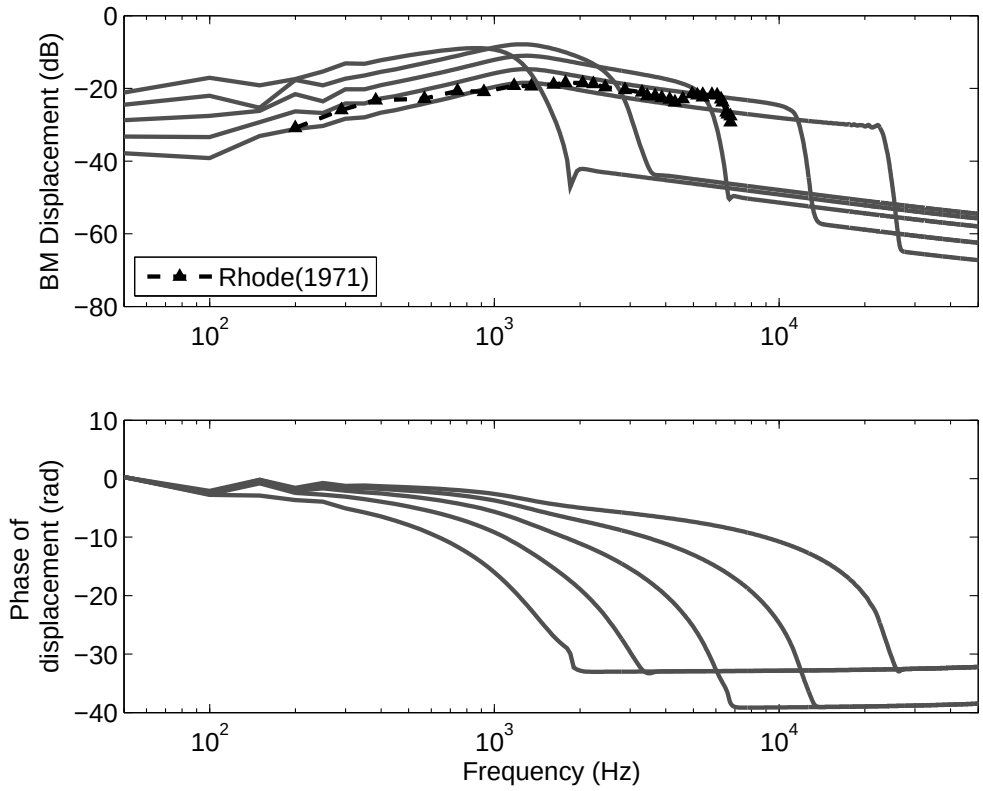

Figure 3.2: The combined BM and middle ear response compared to measured squirrel monkey data (data points replotted from Rhode (1971)).

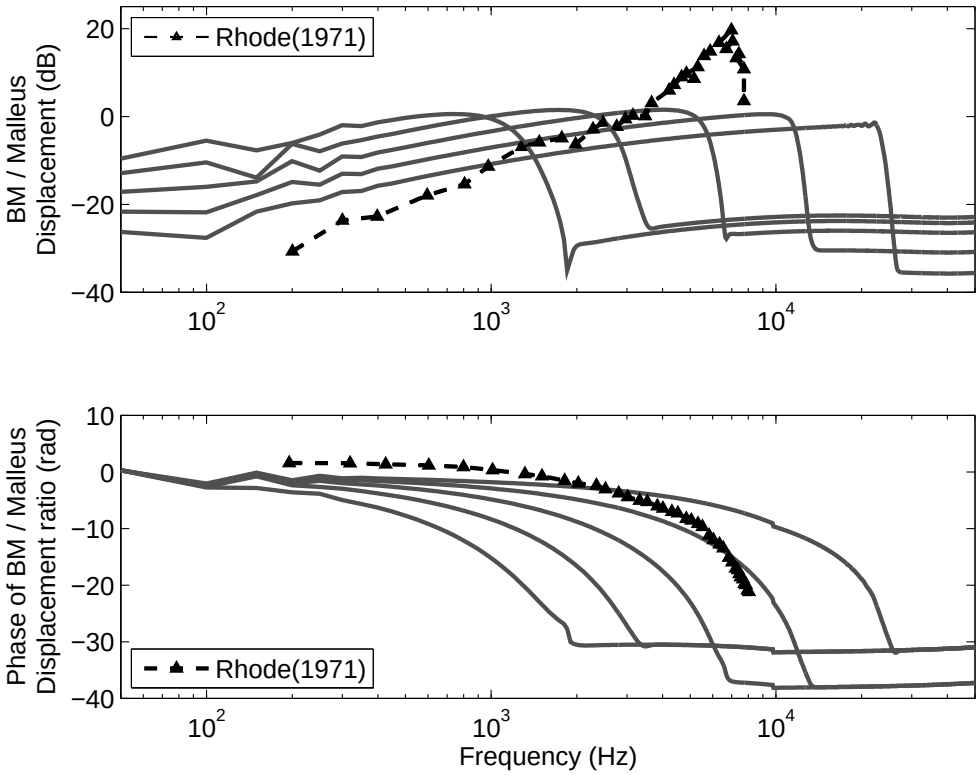

Figure 3.3: The ratio between the BM and malleus displacement compared to measured results (data points replotted from Rhode (1971)).

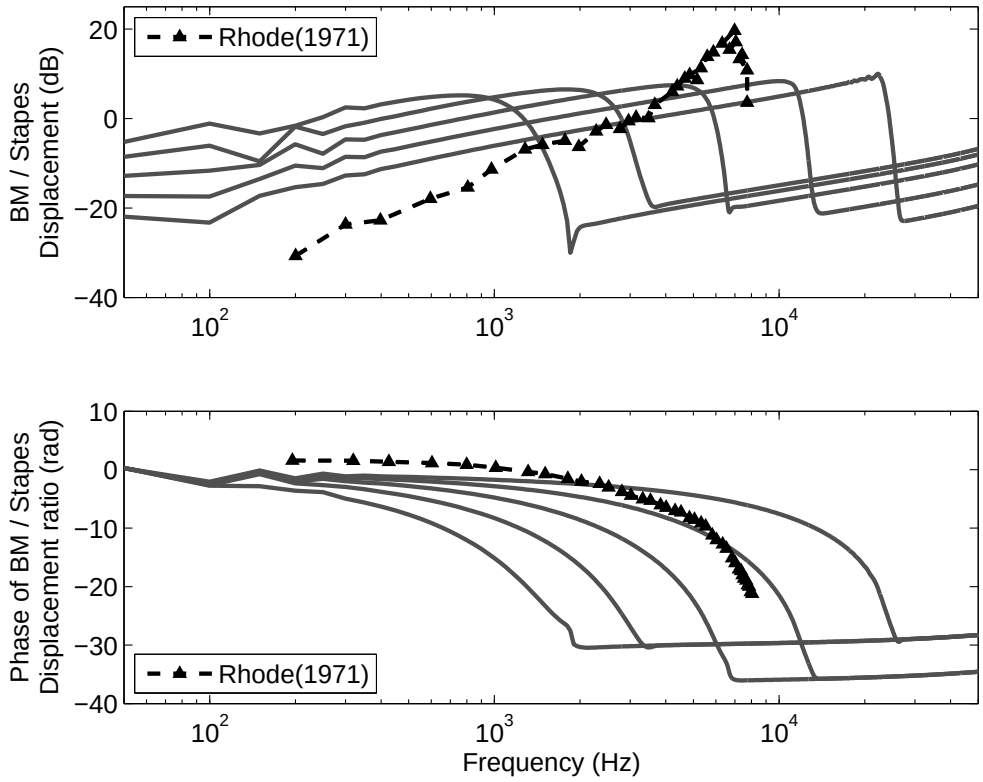

Figure 3.4: The ratio between the BM and stapes displacement compared to the measured BM-maleus displacement ratio (data points replotted from Rhode (1971)).

Another method can be used to calculate the magnitude and the phase response as a function of the BM length. The FFT for the impulse response is calculated at *n* number of places along the BM in the same way as it is calculated for the frequency response. At each specific place the magnitude and phase for a specific frequency (in this case 800 Hz) is determined from the FFT. Fig. 3.5 shows the results for this method for an expected pure tone input of 800 Hz.

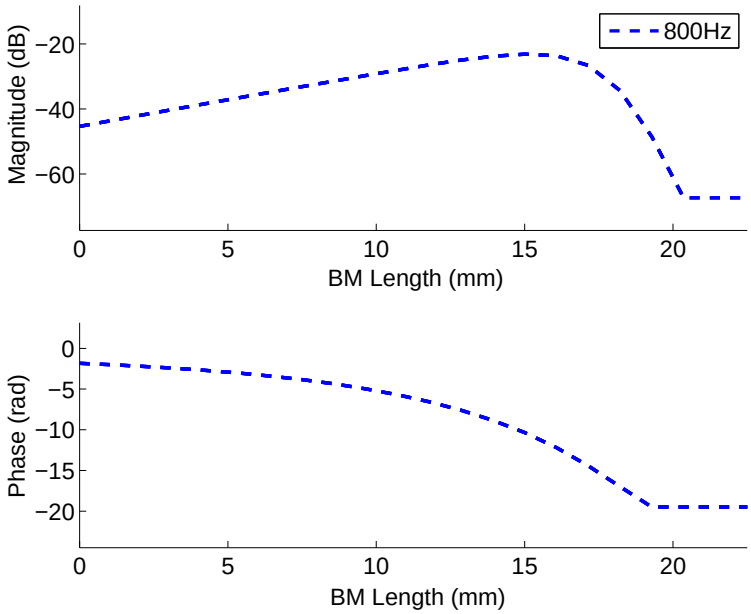

Figure 3.5: The magnitude and phase response in the space domain.

The magnitude and phase response results along the BM length for a 1 kHz input (Diependaal *et al.*, 1987) show a much narrower frequency response. This is a result of a different set of parameters with a difference in the BM length.

3.3 MIDDLE EAR FUNCTION

The middle ear function implemented in the model was proposed by Diependaal *et al.* (1987) following the analysis of Matthews (1980) and Neely (1981b). Moller (1963) investigated the transfer function of the middle ear in anesthetised cats. He measured the amplitude and phase angle of the vibration of the malleus, incus and round window for a constant eardrum sound-pressure using the capacitive probe method. Fig. 3.6 shows the transfer function for the middle ear, taken to be the ratio between the stapes vibration and the eardrum pressure. The impulse response for the stapes was calculated as the response resulting from an impulse eardrum pressure input. This transfer function was determined by calculating the FFT for the stapes impulse response.

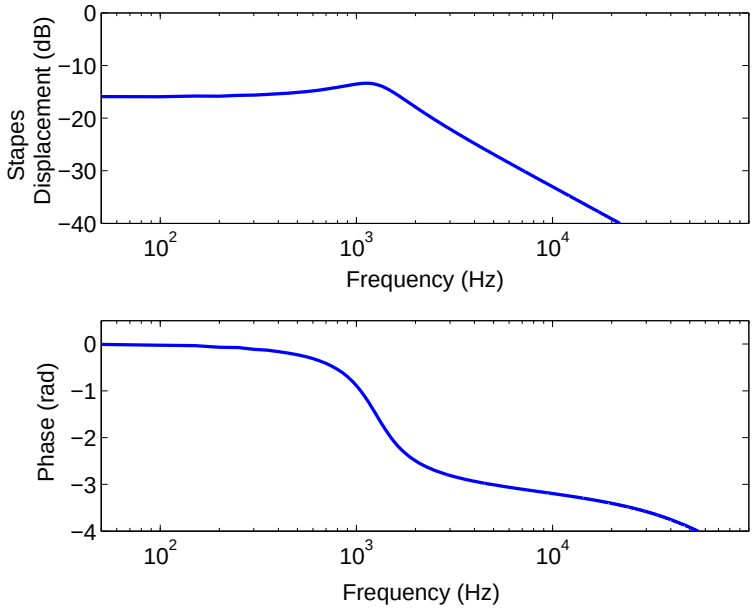

Figure 3.6: Transfer function of the middle ear model between the eardrum and the stapes.

In order to obtain the transfer function for the malleus, the relationship between the stapes and the malleus needs to be determined. Moller (1963) measured the amplitude of the incus

and the round window. These measurements show that the round window is tightly coupled to the incus for frequencies smaller that 2.5 kHz. Moller concluded that the displacement of the stapes is very well represented by the amplitude at the round window. Guinan and Peake (1967) found no significant difference between the amplitude and phase of the incus and the stapes. For the phase angle between the stapes and the eardrum pressure, they argued that if all the points on the eardrum and the ossicles moved in phase (as observed for low frequencies), it could be expected that the motion for high frequencies would be controlled by the inertia of the system and the phase angle would be $-180°$. As they concluded on the basis of measurements that this is not the case, they argued that a phase lag was introduced between the maleus and the incus-stapes (fig. 3.7). However, their measurements only included frequencies below 20 kHz. The amplitude and the phase magnitudes for frequencies higher than 20 kHz have been approximated to give a good fit (fig. 3.7) to the measured results for the malleus (fig. 3.8) .

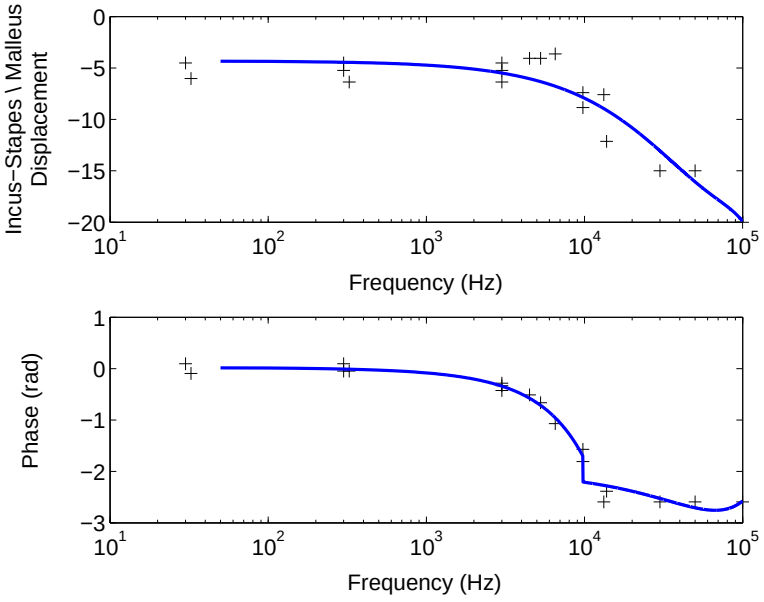

Figure 3.7: The ratio between the average incus-stapes displacement and the malleus displacement (data points replotted from Guinan and Peake (1967)).

The malleus-pressure transfer function for the model needs to be determined from these measured (fig. 3.7, approximated for frequencies above 20 kHz) and simulated (fig. 3.6) transfer functions. This stapes-pressure transfer function (eq. 3.1) is the cascaded result of the three functional blocks, namely the malleus-pressure, the incus-malleus and the stapes-incus. This can be rewritten to describe the malleus-pressure transfer function as a function of the other transfer functions, whereby results for the amplitude and phase response can be calculated (eq. 3.2). The equations which describe these relationships are:

$$\begin{aligned} H_{\frac{Stapes}{Eardrum\ pressure}} &= H_{\frac{Malleus}{Eardrum\ pressure}} \cdot H_{\frac{Incus}{Malleus}} \cdot H_{\frac{Stapes}{Incus}} \\ &= H_{\frac{Malleus}{Eardrum\ pressure}} \cdot H_{\frac{Incus-Stapes}{Malleus}}, \end{aligned} \tag{3.1}$$

from eq. 3.1

$$\begin{aligned} H_{\frac{Malleus}{Eardrum\ pressure}} &= \frac{H_{\frac{Stapes}{Eardrum\ pressure}}}{H_{\frac{Incus-Stapes}{Malleus}}} \\ &= \frac{\left|H_{\frac{Stapes}{Eardrum\ pressure}}\right| \angle H_{\frac{Stapes}{Eardrum\ pressure}}}{\left|H_{\frac{Incus-Stapes}{Malleus}}\right| \angle H_{\frac{Incus-Stapes}{Malleus}}} \\ &= \frac{\left|H_{\frac{Stapes}{Eardrum\ pressure}}\right|}{\left|H_{\frac{Incus-Stapes}{Malleus}}\right|} \left(\angle H_{\frac{Stapes}{Eardrum\ pressure}} - \angle H_{\frac{Incus-Stapes}{Malleus}} \right). \end{aligned} \tag{3.2}$$

The results for the transfer function of the malleus movement and the pressure at the eardrum for the model are found in fig. 3.8. This result is compared to the malleus displacement

results measured by Moller (1963).

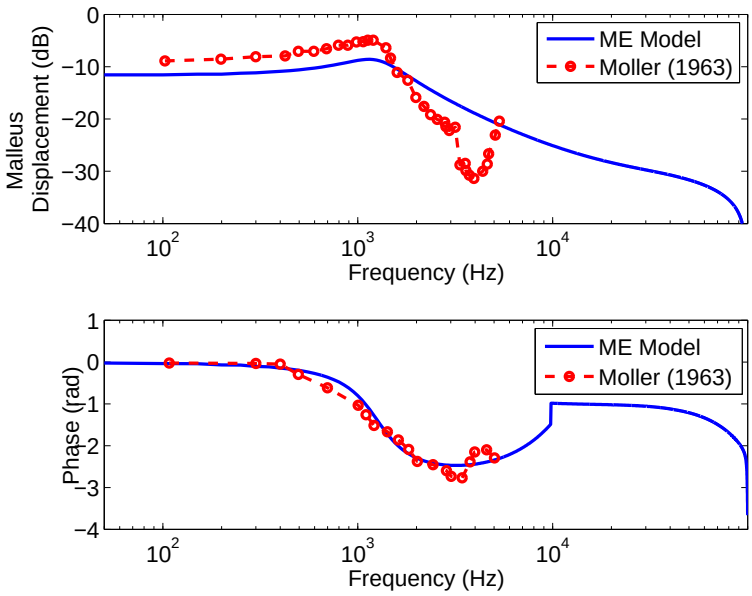

Figure 3.8: The maleus displacement of the model compared to the measured results (data points replotted from Moller (1963)).

The results obtained for the middle ear model gives a relatively good approximation of the measured results (Moller, 1963; Guinan and Peake, 1967) for frequencies up to the low kHz range. The measured results for the maleus displacement shows a higher cut-off rate compared to the simulated results for the model, resulting in more high-frequency suppression. A smoother phase response for the maleus displacement might be obtained with a better approximating function for the incus-stapes displacement ratio. These differences must be taken into account when considering the results for the BM model. The next section shows this middle ear low pass filter effect for an unramped signal input.

3.4 TIME RESPONSE

3.4.1 Input tones

Sinusoidal tone inputs have been used for the time response analysis of the model. Unramped signals (fig. 3.9 b) introduce a transient in the time response of the BM owing to its sudden onset. The integral solution method causes this transient to increase as the wave propagates towards the apex (fig. 3.10). Onset ramps have therefore been introduced to prevent these onset transients. The ramping function (fig. 3.11) multiplies a sinus function with a range of $[0, \frac{\pi}{2}]$, with the tone samples at the start of the unramped tone, resulting in a ramped signal (fig. 3.9 a).

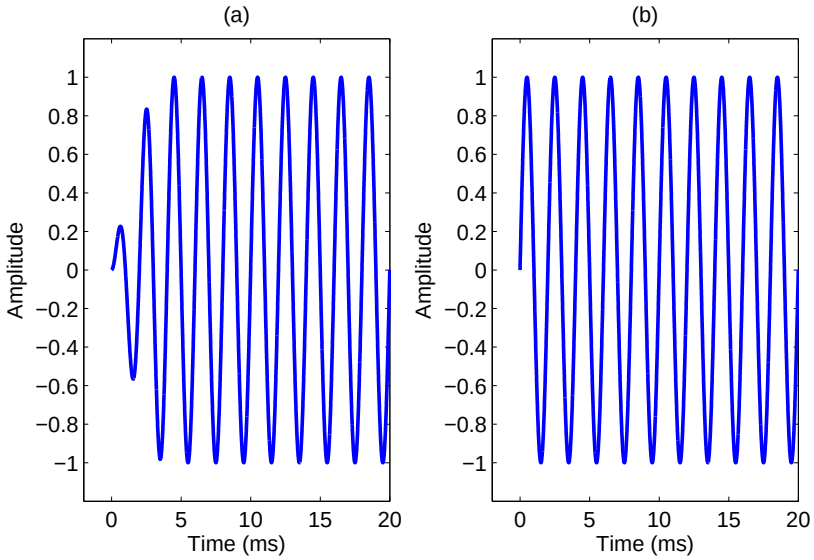

Figure 3.9: (a) A ramped 500 Hz pure tone. (b) An unramped 500 Hz pure tone.

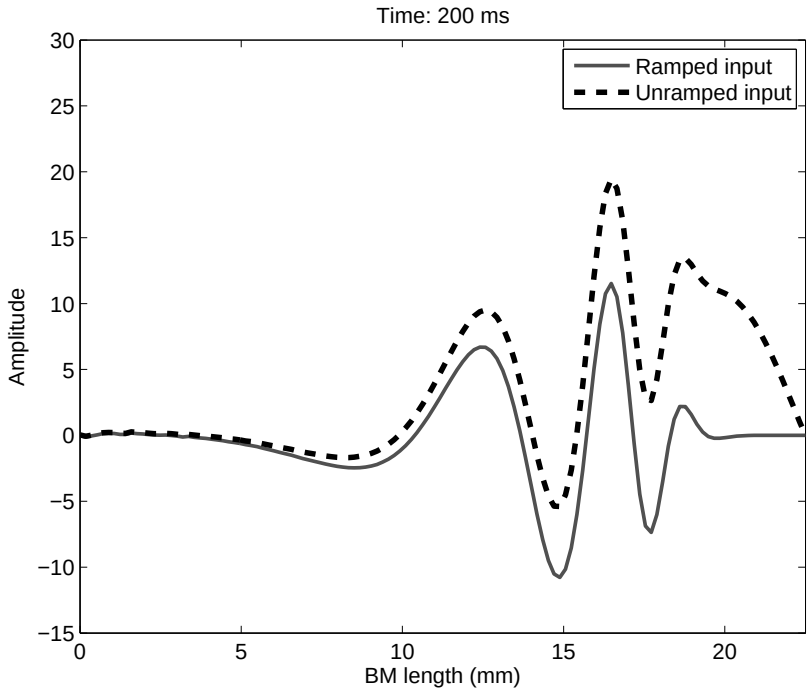

Figure 3.10: The displacement output of the basilar membrane model without the middle ear model.

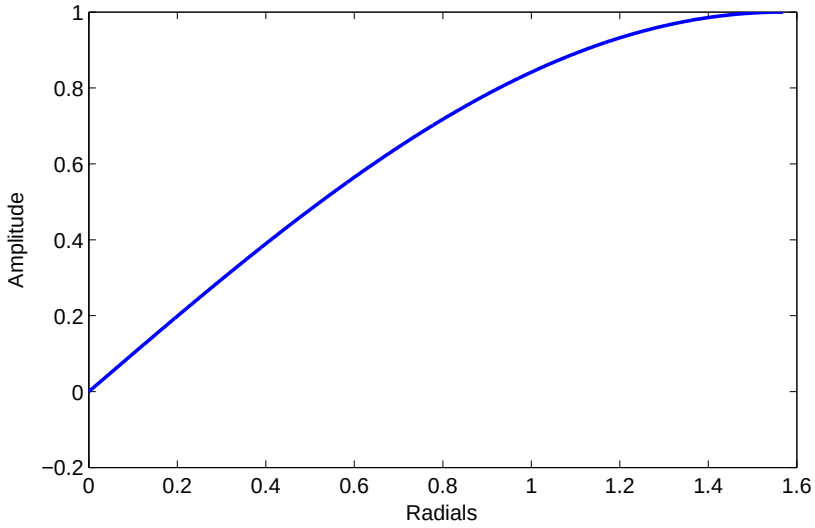

Figure 3.11: The sinus ramp function multiplied with the onset of the unramped signal to obtain a ramped onset.

Results from the BM model with the integrated middle ear model show that these onset

transients have been filtered out, resulting in the same displacement output as the output from that of a ramped input (fig. 3.12). In the rest of the study, however, ramped signals were used as input pure tones to the TW model.

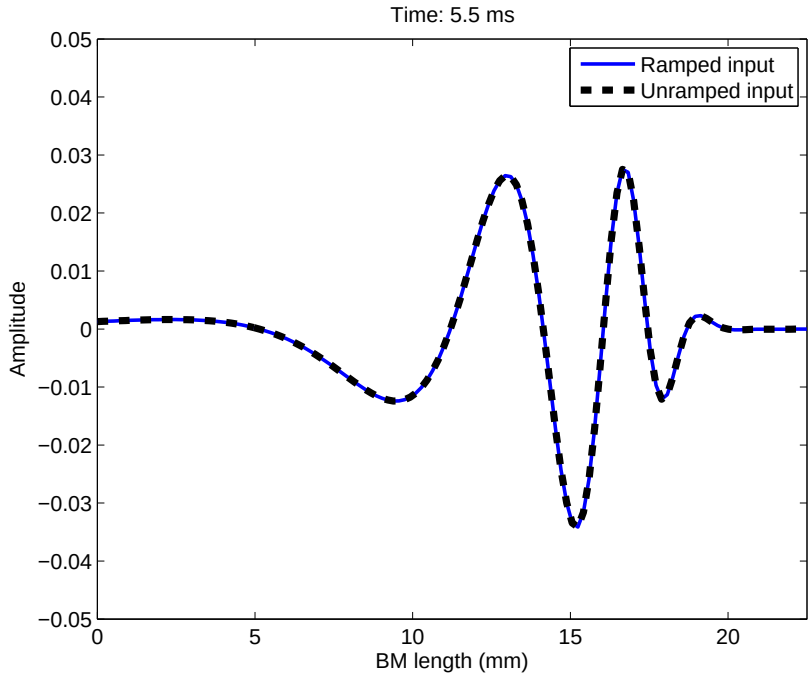

Figure 3.12: The displacement output of the basilar membrane model with the middle ear model.

3.4.2 Time domain pure tone responses

From the literature study it was concluded that a travelling wave (von Békésy, 1960) propagated along the length of the cochlea reaches a maximum peak at the CF. With this travelling wave the wavelength shortens as it reaches this point, whereafter the wave quickly dies out as it passes beyond this point. This section investigates this propagation in time along the length of the cochlea for different pure tone frequencies.

Fig. 3.13 - 3.16 shows the propagation of the travelling wave with respect to time for a 200

Hz, 500 Hz, 1 kHz and 6 kHz pure tone input. In Fig. 3.17 (a) - (d) the waveforms for specific times show that the propagation of the travelling wave reaches a maximum peak of deflection at different places along the cochlea for different frequencies.

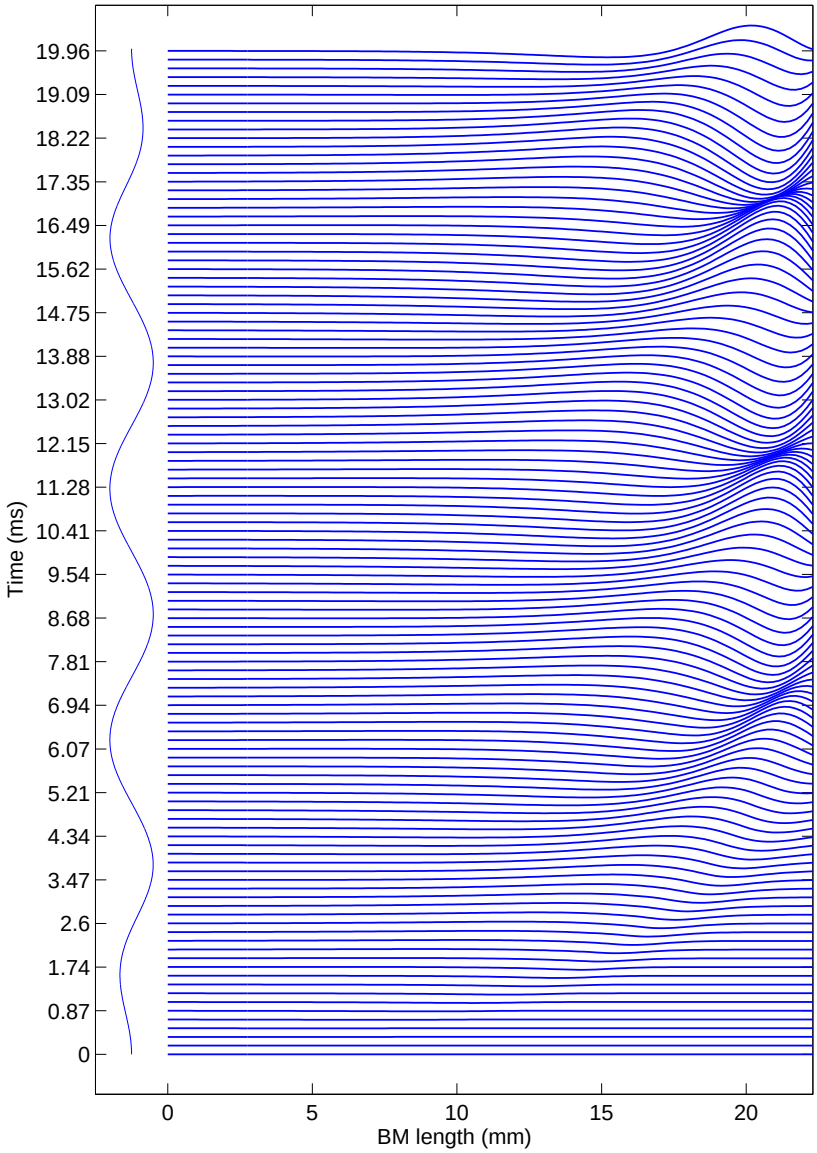

Figure 3.13: Space-time plot for a 200 Hz pure tone. The sinusoidal eardrum pressure input is shown in the graph on the left, relative to the incremental time steps.

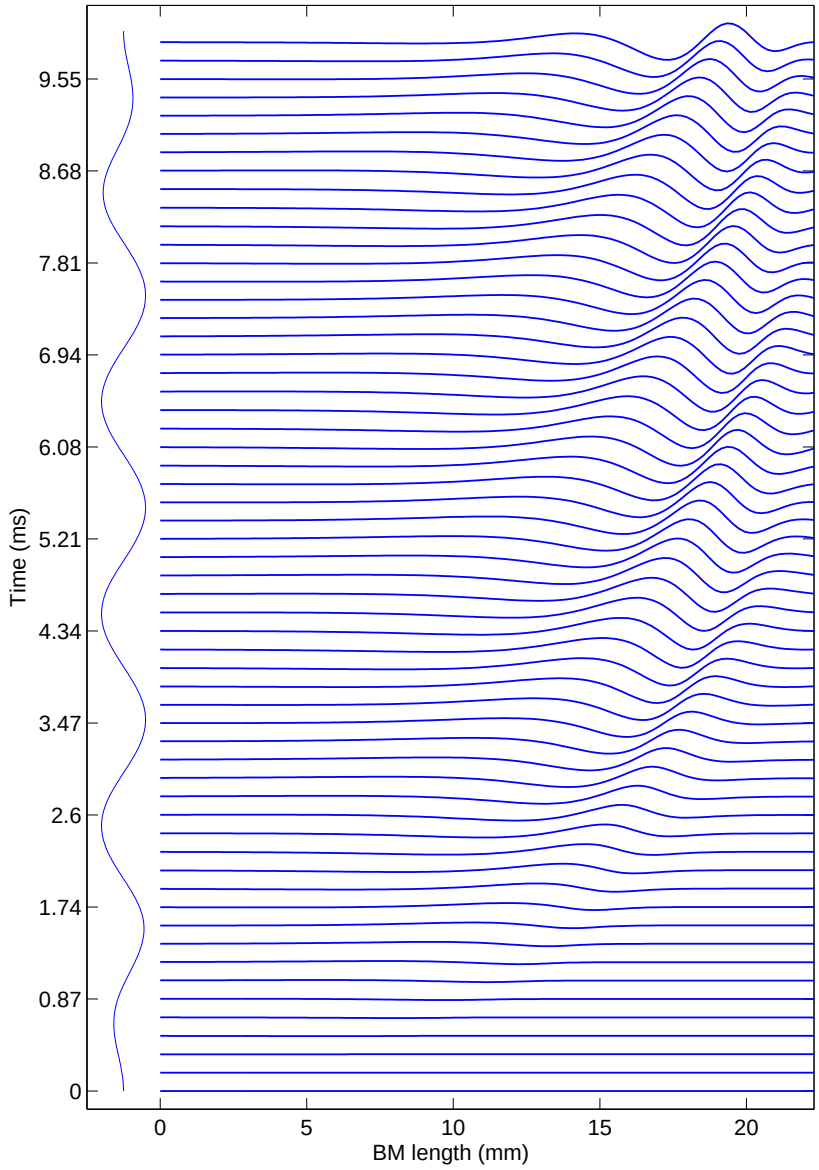

Figure 3.14: Space-time plot for a 500 Hz pure tone. The sinusoidal eardrum pressure input is shown in the graph on the left, relative to the incremental time steps.

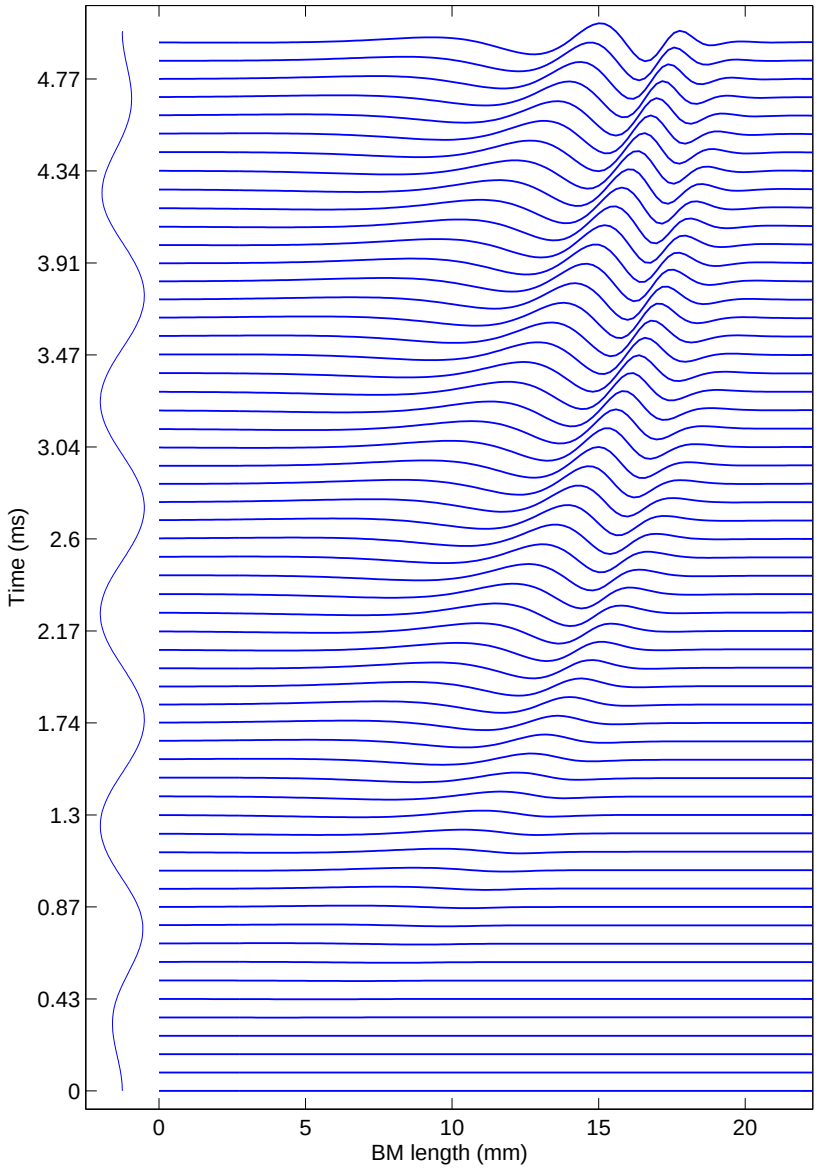

Figure 3.15: Space-time plot for a 1 kHz pure tone. The sinusoidal eardrum pressure input is shown in the graph on the left, relative to the incremental time steps.

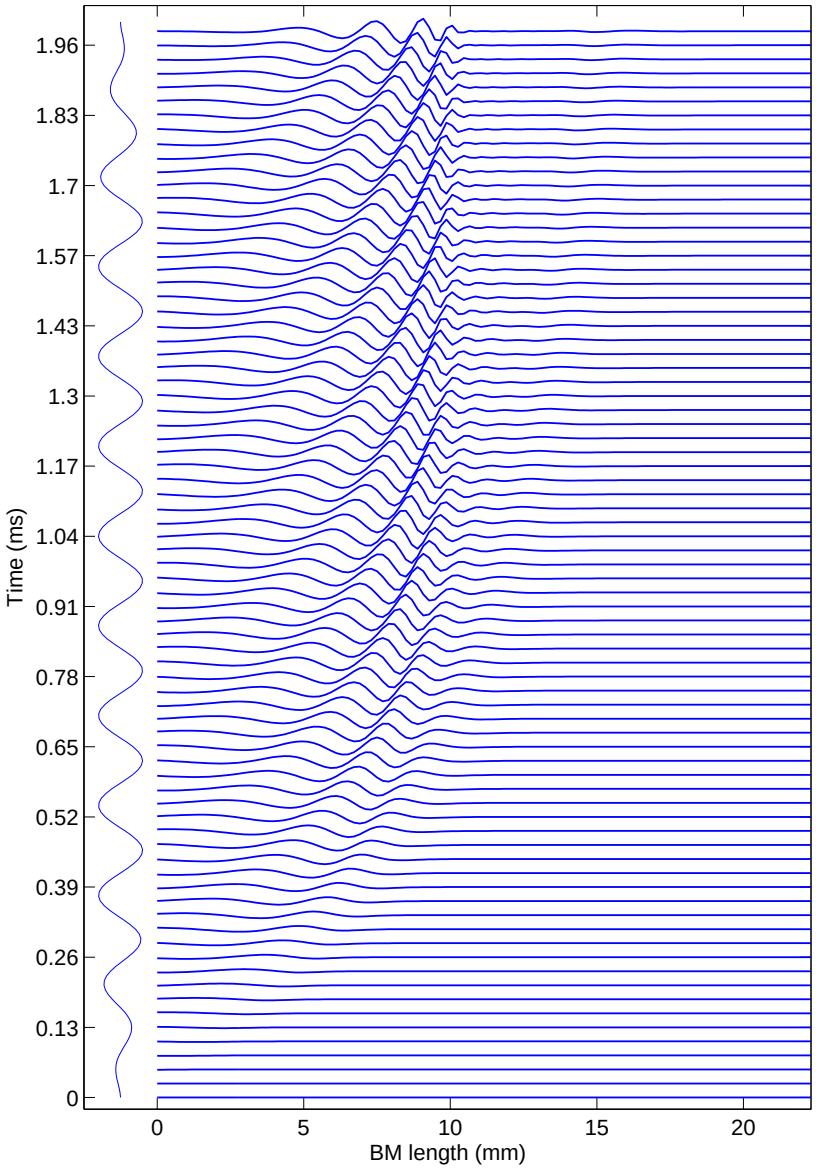

Figure 3.16: Space-time plot for a 6 kHz pure tone. The sinusoidal eardrum pressure input is shown in the graph on the left, relative to the incremental time steps.

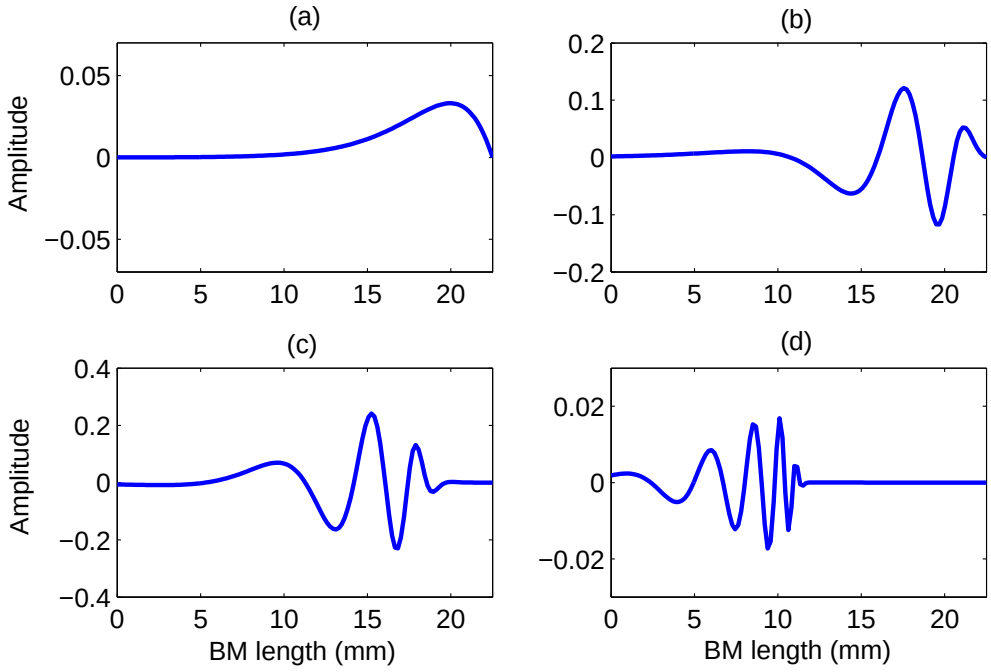

Figure 3.17: Time responses for pure tones. (a) 100 Hz (b) 500 Hz (c) 1 kHz (d) 5 kHz.

From these results it can be concluded that high frequency inputs cause the BM to reach a maximum peak toward the basal side. For low frequency inputs the peak moves toward the apical side. A further observation reveals a slower propagation velocity for low frequencies, with the propagation velocity increasing as the input frequency increases.

3.5 COCHLEAR FREQUENCY MAPS

The cochlea is tonotopicly spaced where specific frequencies stimulate specific regions of hair cells in the cochlea. Zwislocki (1950) derived equations for the maximum peak of vibration from the dynamic differential equation (eq. 2.1) describing the fluid flow in the cochlea. These two equations provide the maximum vibration at a specific place (eq. 3.3) corresponding to a specific frequency (eq. 3.5):

$$x = \frac{2}{3\beta}\left[ln\left(\frac{\beta {F_0}^{\frac{1}{2}}}{8\pi^2 R_2 C_0 (2\rho C_0)^{\frac{1}{2}}}\right) - 2ln\upsilon\right] mm \tag{3.3}$$

where

$$\upsilon = \frac{1}{2\pi(\sigma C)^{\frac{1}{2}}} \tag{3.4}$$

and

$$f = A \cdot exp\left[\frac{3}{4}\beta x - \frac{1}{\beta}\left(\frac{2\rho C_0}{F_0}\right)^{\frac{1}{2}} \omega^2 R_2 C_0 \cdot exp\left(\frac{3}{2}\beta x\right)\right] Hz, \tag{3.5}$$

with σ being the effective mass of the cochlear partition, β a constant (1.5 per cm), x the distance from the stapes, F_0 the average cross section of the cochlear canal, ρ the fluid density, R_2 the coefficient of friction and C_0 the compliance at the base of the cochlea. Zwislocki (1950), fig. 4, shows this relationship between place and frequency.

Greenwood (1961) measured critical bandwidths and how they relate to frequency and position on the basilar membrane. He approximated the place-frequency relationship in the cochlea with an exponential function (eq. 3.6),

$$f = A(10^{ax} - 1) \tag{3.6}$$

where f is the frequency related to the position x, A and a are two constants dependent upon the specific species, and x is the distance along the cochlea from the apex. Liberman (1982) fitted Greenwood's function to new frequency-position data for cat and a revised estimate for the basilar membrane length. Greenwood (1990) reviewed Liberman's fitting data and obtained the eq. 3.7 for the new cat data, using a basilar membrane length of 25 mm instead of 22 mm (fig. 2 of Greenwood (1990)). He also introduced an integration constant k, where $k = 0.8$. The equation follows:

$$f = 456\left(10^{2.1x} - 0.8\right) \tag{3.7}$$

with x being a normalised length from the apex between 0 and 1. Fig. 3.18 compares the two equations from Greenwood and Zwislocki to the results obtained from the travelling wave model. The model's frequency map underestimates the Greenwood equation and overestimates Neely's model results for the same parameter set. This might be the result of implementing a one-dimensional model as opposed to the two-dimensional implementation from Neely (1981b).

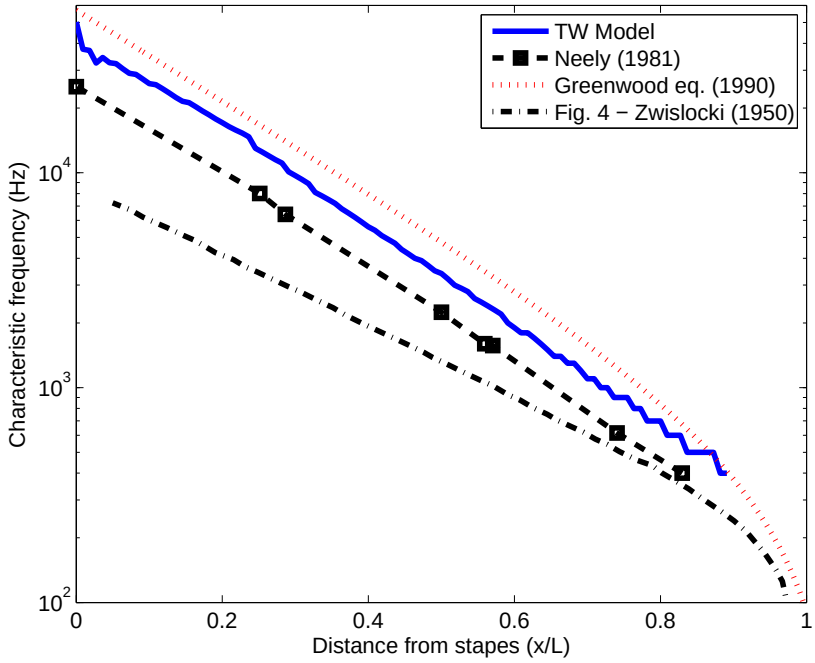

Figure 3.18: A comparison of the frequency map of the model, with Neely's (1981b) model and Greenwood's (1990) and Zwislocki's (1950) equations. Greenwood's equation approximates cat data. Data points replotted from Neely (1981b).

The figure further shows no results for the low frequencies. At these low frequencies the travelling wave model fails at predicting a propagation delay. This might be the result of the parameter set used. The results from Neely (1981b) also show no results at these frequencies, which might suggest that similar results were obtained at low frequencies. As a result the low frequency results for the travelling wave model have been omitted from the graph (fig. 3.18). It is important to take this into account when considering the results in the next chapters. The next section shows the travelling wave delay results for the travelling wave model in comparison to measured travelling wave delay results.

3.6 TRAVELLING WAVE PROPAGATION DELAY

As the travelling wave travels towards the apex, a travelling wave delay is observed, which increases exponentialy towards the basal region, up to a few milliseconds. Fig. 3.19 shows this time delay between the stapes and the place of maximum displacement for a 2 kHz pure tone input. The figure shows this delay to be approximately 1 ms. The figure also shows the time delay between the stapes and the eardrum pressure, where these two reponses are almost 180° out of phase. The phase response of the eardrum pressure-stapes transfer function (fig. 3.6) shows a similar phase shift of -180°.

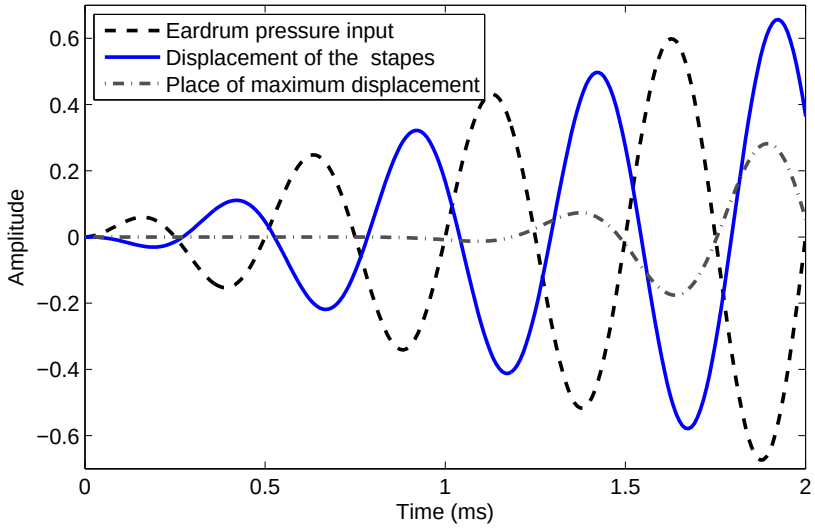

Figure 3.19: The figure shows the travelling wave delay of a 2 kHz tone input (as eardrum pressure) at the place of maximum displacement. The displacement of the stapes due to this input is also shown.

In the analysis of linear systems, Papoulis (1962:134-136) mentions three delays, namely phase, signal-front and group delays. Ruggero and Temchin (2007) show that these delays could also be calculated from the phase response. Apart from these delays, another delay measurement seems to be found in the literature. Kim and Molnar (1979) measured the la-

tency of responses of single cochlear nerve fibres in cat, as a response to clicks (rectangular electrical pulses). This latency is defined by Kim and Molnar (1979) as the onset of rise of the eardrum pressure to the beginning of the first peak of the post-stimulus time (PST) histogram. These delay estimates were used by Neely (1981b) to compare his model's latency data; Neely defines latency as the onset of a low-pass filtered impulse to the first zero crossing of the impulse response. These resulting delays have been compared in fig. 3.20 to latency of the travelling wave model. The latency for the travelling model was calculated using the calculation method of Neely (1981b). The results for the travelling wave model show that it fails in predicting a travelling wave delay for frequencies smaller than 400 Hz.

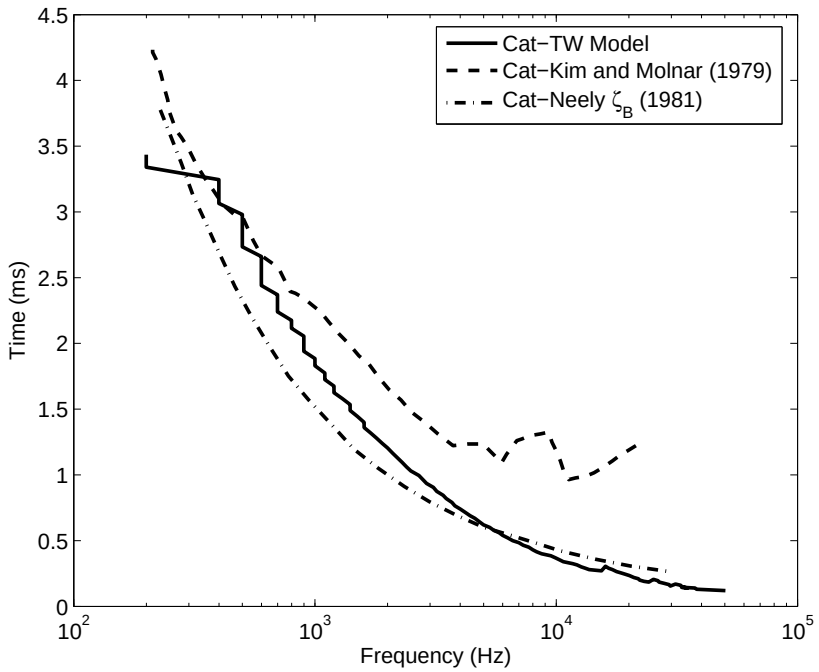

Figure 3.20: The click latency response as defined by Neely (1981b) compared to the results of the model and the PST histogram data (Kim and Molnar, 1979). Datapoints replotted from Neely (1981b) and Kim and Molnar (1979)).

The signal-front delays (fig. 3.21), also referred to as the latency of the impulse response (Ruggero and Temchin, 2007; Papoulis, 1962) shows a difference between different species

(cat (Siegel, Cerka, Recio-Spinoso, Temchin, Van Dijk and Ruggero, 2005), squirrel monkey (Robles, Rhode and Geisler, 1976) and cadavers (von Békésy, 1949)). Stimulus-frequency otoacoustic emissions (SFOAE) measurements taken by Shera and Guinan (2003) show that the SFOAE delays in cat, chinchilla and guinea pig are much smaller than in human subjects. Partly on the basis of these measurements they stated that the sharpening of the cochlear frequency filtering is greater in humans and monkeys than in other mammals. These differences in delays between species need to be taken into account when considering the neural response results.

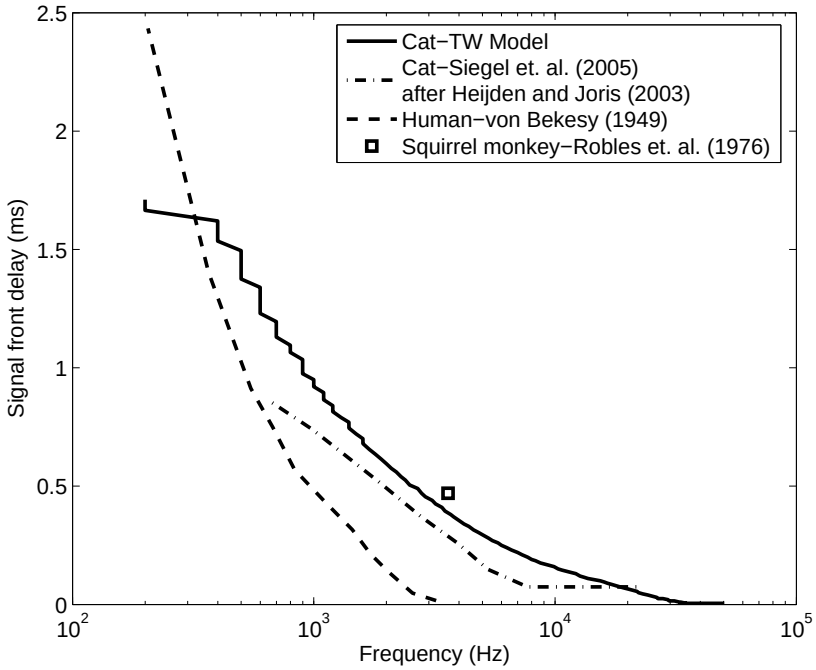

Figure 3.21: The model's signal front delay is compared to delay measurements for cat (Siegel *et al.*, 2005), squirrel monkey (Robles *et al.*, 1976) and cadavers (von Békésy, 1949). (Datapoints replotted from Siegel *et al.* (2005), Robles *et al.* (1976) and von Békésy (1949)).

After consideration of the results for the travelling model in this section, the next chapter discusses the implementation of this model in a speech processing strategy. The chapter

considers the functional analysis of such a processing strategy with regard to the trade-off between the stimulation rate and the number of electrode channels.

Chapter 4

Consideration of a travelling wave processing strategy in a CI

4.1 INTRODUCTION

A sound is received from the microphone of the cochlear implant, from where it is processed and stimulated at a number of electrodes, which are inserted surgically into the cochlear ducts. Each of these electrodes is stimulated at a specific stimulation pulse rate that differs with each device. The number of electrodes in current electrode arrays ranges from 12 electrodes for the Med-El device (PULSARci^{100}) to 22 electrodes for the Cochlear device (Nucleus 24) (Loizou, 2006). These devices support overall maximum stimulation rates of 50 704 pps for the PULSARci^{100} and 14 400 pps for the Nucleus (Loizou, 2006), with electrode spacings of 2.4 mm (Loizou, 2006) and 0.75 mm (Clark, 2003). Along with these the Advanced Bionics devices, HiRes90K and Clarion CII have an aggregate stimulation rate of 82 000 pps (Buechner, Frohne-Büchner, Gaertner, Stoever, Battmer and Lenarz, 2010). The stimulation rate per channel is calculated to be the overall stimulation rate divided by

the number of active channels. Therefore a trade-off remains between the stimulation rate (temporal resolution) and the number of electrode channels (spectral resolution). Spectral-maxima strategies (such as ACE) determine a user-dependent set of maxima channels. With this strategy electrodes could be stimulated at higher channel stimulation rates. For the above-mentioned Advanced Bionics devices an eight-channel strategy could be used, stimulating at a rate of 10 312 pps per channel. However, with smaller stimulated pulse widths, larger stimulation current is necessary to achieve sufficient loudness growth (Buechner *et al.*, 2010).

This chapter will investigate the effects of stimulation rate and the number of electrode channels in current processing strategies further and attempt to determine the effect of this channel-number stimulation-rate considerations on the implementation of the proposed travelling wave processing strategy. The next section gives an overview of the different functional blocks for this travelling wave processing strategy. These functional blocks will be discussed in further detail in the rest of the chapter in the light of processing strategies in literature.

4.2 A PROPOSED TRAVELLING WAVE PROCESSING STRATEGY

Fig. 4.1 shows the block diagram with the functional processing blocks for a proposed travelling wave processing strategy. The first block is the travelling wave model that has been discussed in detail in the previous chapters (chapter 3 and 2). This model is firstly sampled at a very high sampling rate, whereafter it is resampled in time to the Nucleus 24 device's stimulation rate. Further spatial sampling samples the high spatial resolution to 22 places, corresponding to the Nucleus 24 electrode array. After sampling, the travelling wave model's displacement amplitudes are mapped to the threshold-comfort current level (CL) units with a logarithmic loudness growth function (LGF) (Nogueira, Büchner, Lenarz and Edler, 2005).

This function compensates for the logarithmic increase in perceived loudness as a result of the increase in stimulation current. The Nucleus device uses sequential stimulation, therefore it is necessary to collate the CL units into a sequence. In a practical implementation this CL sequence would be transmitted via an RF link through the skin, from where it is demodulated, and mapped to physical stimulating currents. However, in this study the CL units are mapped to specific current values, which then serves as input to the current spread model (Strydom and Hanekom, 2011; Smith, 2011). The electrical model (Bruce *et al.*, 1999a; Bruce *et al.*, 1999b; Smith, 2011) approximates the electrical stimulation in the cochlea, from where spatio-temporal neural spike patterns could be generated.

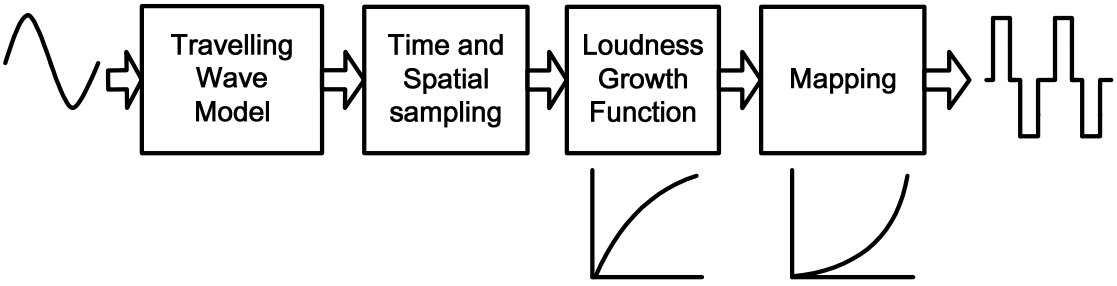

Figure 4.1: The functional block diagram of the proposed TW processing strategy.

The following sections give a more detailed analysis of these functional blocks. Each functional block for the travelling wave processing strategy is considered in the light of processing strategies in literature. The second block will be broken up in two sections and will be discussed in the next two sections. In the rest of the document the proposed travelling wave strategy will be referred to as the TW processing strategy.

4.3 SPATIAL SAMPLING CONSIDERATIONS IN A TRAVELING WAVE PROCESSING STRATEGY

In processing strategies a trade-off remains between the spatial and the temporal sampling. In this section the spatial sampling along the length of the cochlea is discussed with reference to what has been seen in processing strategies in literature. Given this backdrop, sampling is discussed related to the proposed travelling wave strategy.

A study (Friesen, Shannon and Slattery III, 2000) investigating the effects of electrode configurations, which were shifted along the apical, medial and basal regions of the cochlea, showed that electrode position is an important factor in perception. In most processing strategies each electrode (representing a certain place along the length of the cochlea) is associated with a specific frequency filter band. These filter bands, which span the frequency range 188 Hz to 8 kHz (Loizou, 2006), are either linearly or logarithmically spaced. The Nucleus device combines these two types of filter spacing, implementing linear spacing for frequencies below 1 kHz and logarithmic spacing for frequencies higher than 1 kHz. Fu and Shannon (2002) found the best phoneme recognition performance when filters were arranged logarithmically or near-logarithmically across the frequency range. This frequency-place relationship has been approximated by Greenwood's (1990) equation (eq. 3.7). Baskent and Shannon (2004) found the best speech recognition when information was delivered to the matching tonotopic place with the least frequency-place distortion.

Tonotopic arrangement, which filter band strategies use as their basis, corresponds roughly to the peak amplitudes of the frequency filter responses of the travelling waves. However, travelling waves, especially their temporal aspects, have been ignored in most commercial speech processing strategies. The travelling wave model in the proposed travelling wave processing strategy describes how the BM reacts to a sound input. Therefore a subset of

filter bands does not have to be implemented when designing a travelling wave-based speech processor. The travelling wave amplitude is associated with a specific place along the BM, obtained from the travelling wave model in chapter 3, as a measure of the BM displacement for that place. These places must relate to the number of electrodes. Therefore for a limited number of electrodes (for instance 22), these electrodes relate to a small number of places along the cochlea for which BM displacement is determined with the TW model. Shannon, Cruz and Galvin (2011) and Friesen, Shannon and Cruz (2005) found that speech recognition increased significantly from four to eight electrodes, but stayed similar for eight, 12 and 16 electrodes. These results suggest that implant users do not make full use of all the electrodes. Current devices such as Nucleus 24 implements 22 electrodes, whereas Clarion II and Med-El stimulate 16 and 12 electrodes respectively. With Nucleus 24, 22 electrodes are inserted inside the cochlea, with two external electrodes, for monopolar stimulation. The proposed travelling wave processing strategy divides the BM into 23 equal lengths, corresponding to the Nucleus array of 22 electrodes. Taking the BM length to be normalised to 1, this results in an electrode placement of electrode 1 at 0.044 and electrode 22 at 0.957. It was therefore necessary to resample the travelling wave model output spatially to correspond to these 22 places along the length of the BM. This maximum number of 22 electrodes is a tremendous decrease from the normal hearing cochlea, which contains in the region of 30 000 independent nerves.

Considering that the travelling waves generated by the TW model stimulate a relatively wide area, it seems necessary to keep electrode resolution high to keep the envelope and relative phase distortion between the spatial samples as small as possible. The effects of limited electrode resolution can be seen in fig. 4.2 for travelling wave envelopes at a low and high frequency. These effects are even more significant for higher frequencies; fig. 4.2 c) shows a greatly distorted travelling wave envelope. The spatial samples related to the electrode channels have been evenly distributed along the length of the cochlea. One possible solution

for an improved representation of the travelling wave's envelope is to shift these spatial samples towards the peaks of the travelling wave. Fig. 4.3 shows the shift in spatial samples for a 5 kHz signal. This shift, however, will cause more distortion in the specific place-amplitude of the travelling wave as a result of the uneven spread of samples, which might cause other unwanted effects.

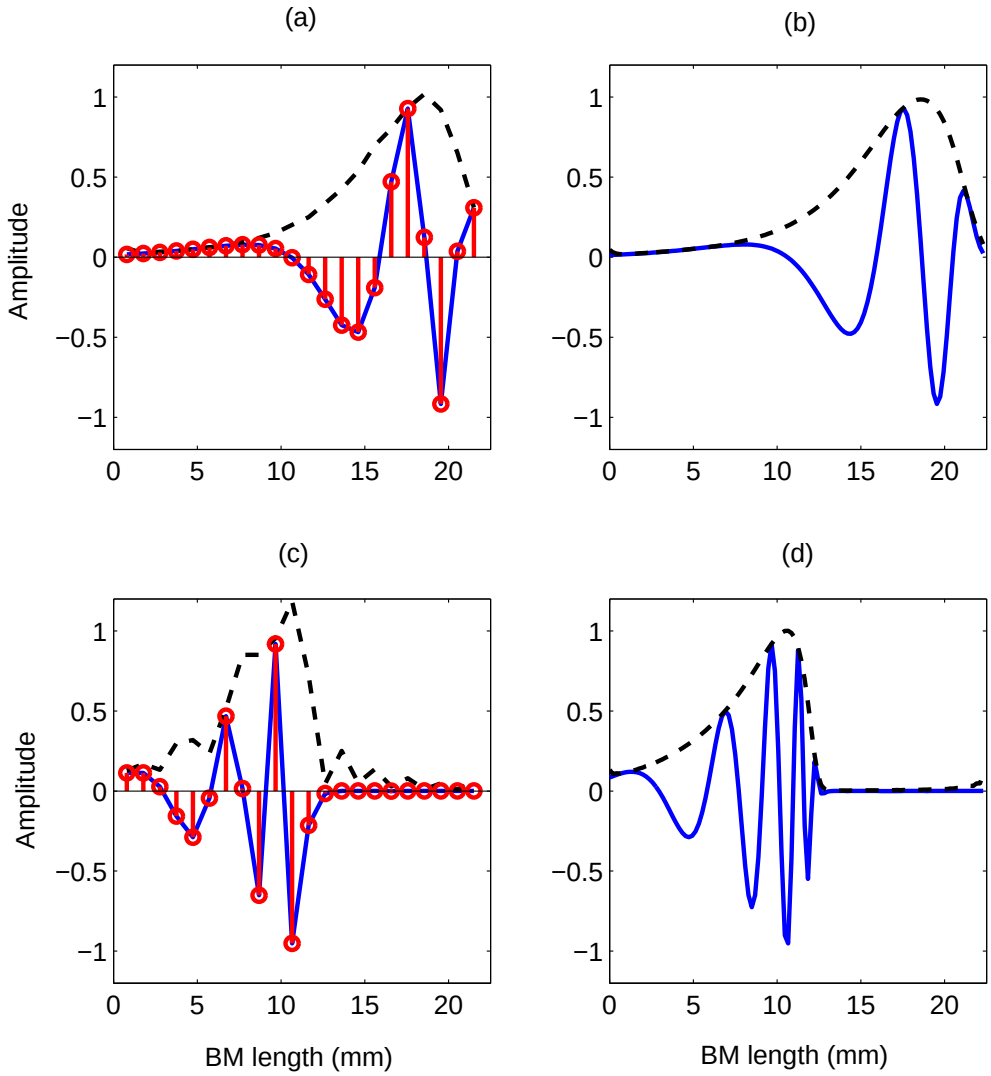

Figure 4.2: The effect of sampling the displacement waveform with only 22 spatial samples, at different input tones; (a) shows a 500 Hz pure tone input and (c) a 4 kHz input. The displacement waveforms b) for a 500 Hz input and (d) for a 4 kHz input were simulated for a high spatial resolution.

Another aspect that has to be taken into account when considering the implementation of travelling waves in a processing strategy is the insertion depth of the electrode array. The modelled electrode array consists of 22 electrodes, with an electrode spacing of 0.75 mm apart. This is the average electrode spacing used in the Nucleus 24 electrode array (Clark, 2003). Full insertion depth for the electrode array was assumed (Smith, 2011), with electrode 22 implanted 10 mm from the apex, corresponding to a distance of 25 mm from the base. This means that electrodes are compressed into 15.75 mm, rather than the full length of 35 mm of the cochlea, leading to a considerable distortion of the place-specific amplitude for the travelling waves. Possible solutions might be to have a deeper insertion depth, resulting in a wider spread of electrodes. The Med-El COMBI-40+ implant uses a soft electrode carrier, which is specifically designed for a deeper insertion depth of approximately 31 mm. This enables the electrode array to cover a longer distance of 26.4 mm, with electrodes spaced 2.4 mm apart (Loizou, 2006).

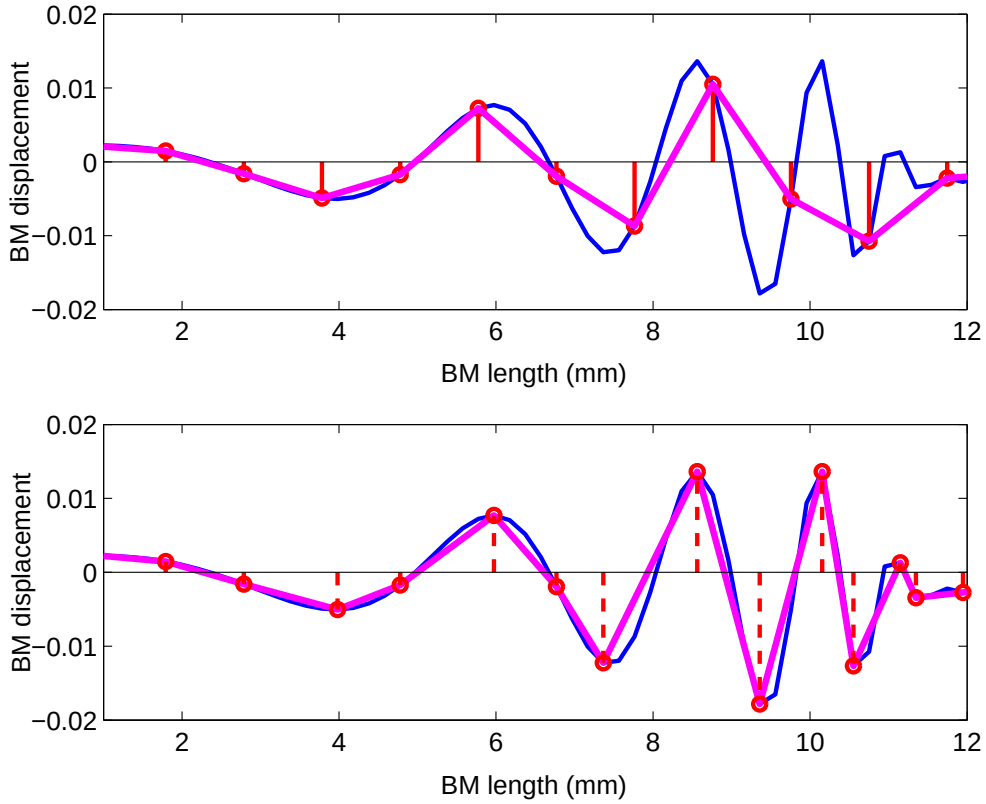

Figure 4.3: Shifting the spatial samples related to the electrode channels to the peaks of the travelling wave envelope. In a) the travelling wave model samples associated with the electrode channels are evenly spaced along the length of the cochlea. In b) the spatial samples related to the electrode channels are shifted towards the envelope peaks of the travelling wave.

4.4 TEMPORAL SAMPLING CONSIDERATIONS IN A TRAVELLING WAVE PROCESSING STRATEGY

This section investigates the effects of temporal sampling in the second functional block in fig. 4.1, with reference to processing strategies in literature. The proposed travelling wave processing strategy implements the same stimulation rate as the stimulation rate used in the Nucleus 24 implant from Cochlear, with an overall sampling rate of 14 400 pulses per second (pps) for the active channels.

The effects of stimulation rate on speech perception have been investigated through the years for a number of processing strategies. Although higher stimulation rates provide much better temporal sampling, the results obtained by most studies for speech perception are not significantly different from those obtained at lower stimulation rates. Fu and Shannon's (2000) results with the Nucleus 22 showed that implantees found no significant benefit in pulse rates between 150 - 500 pps. Users did, however, show significantly poorer scores for pulse rates lower than 150 pps. Another set of speech comprehension experiments conducted with the Nucleus 24 (Vandali, Whitford, Plant and Clark, 2000) also showed no significant increase in scores with higher stimulation rates. Although Holden, Skinner, Holden and Demorest (2002) found group mean scores across time periods to be significantly higher for a higher stimulation rate, they concluded that better scores at a higher stimulation rate were very subject-specific. Measured phoneme and speech recognition in a number of processors, which included the Clarion C1, C2 and the Nucleus 24, was conducted by Friesen *et al.* (2005). With channel stimulation rate being varied from 200 pps to more than 5000 pps, no significant change in the performance as a function of stimulation rate could be found. More recent experiments conducted by Shannon *et al.* (2011) measured the effect of varying the channel stimulation rate between 600 and 4800 pps and found little or no advantage in noise and quiet for the speech test used.

However, a study (Loizou, Poroy and Dorman, 2000) with Med-El/CIS-link implantees showed significantly higher scores for consonant and word recognition at higher stimulation rates. Loizou (2006) argued that these benefits reported at higher stimulation rates, might be the result of fewer channel interactions as a result of wider electrode spacings. The electrode spacing for the Nucleus 24 device is an average of 0.75 mm, whereas the Clarion II has a spacing of 1.1 mm. Med-El introduces the widest spread with a spacing of 2.4 mm. Loizou (2006) stated that it is reasonable to assume that a wider electrode spread would cause less interaction between neural populations excited by stimultion pulses from adjacent

electrodes. He further argued that this might be the reason why Nucleus users found no benefits in higher stimulation rates.

Although higher stimulation rates would incorporate the fine temporal resolution of the travelling wave, in sequential stimulation a trade-off remains between the stimulation rate and the number of channels. The total maximum stimulation rate with the Nucleus 24 device is 14 400 pps, which is divided between the different active channels. For instance, activating 14 channels would produce a stimulation rate of 1 028 pps on each channel. Therefore an increase in channels would result in a decrease in channel stimulation rate. The PULSARci^{100} and the HiRes90K and Clarion CII, with overall maximum stimulation rates of 50 704 pps and 82 000 pps, would be able to stimulate active channels at a higher stimulation rate. As a result of these limitations a trade-off remains between the stimulation rate and the number of electrode channels. The results of this trade-off will be discussed in chapter 6 for different configurations of stimulation rate and electrode channels.

This stimulation-rate channel-number trade-off has to be accounted for when sampling the simulated BM response. The numerical analysis for the travelling wave model (chapter 2) required a very high sampling rate to ensure a stable response. Therefore the TW model was sampled at a frequency of 115 274 Hz. This high sampling frequency was down-sampled to a frequency of 14 400 pps, with only every 8th sample retained. The travelling wave model amplitude is consequently sampled for every time interval of 69.4 μs. The Nucleus device uses sequential stimulation to keep the channel interaction to a minimum, and consequently excitation spread, whereas the Clarion II and the Med-El devices allow for sequential or simultaneous stimulation (Loizou, 2006). Sequential stimulation stimulates each active channel sequentially, therefore introducing a time delay between the stimulation pulses across active channels. A time delay of 69.4 μs between each active channel is introduced when the travelling wave model output samples related to a specific time are stimulated sequentially

across active channels. Fig. 4.4 shows the difference between simultaneous and sequential stimulation with reference to the TW processing strategy. The TW processing strategy compensates for this effect by sampling the TW model output of the next adjacent channel at the next time sample (fig. 4.5).

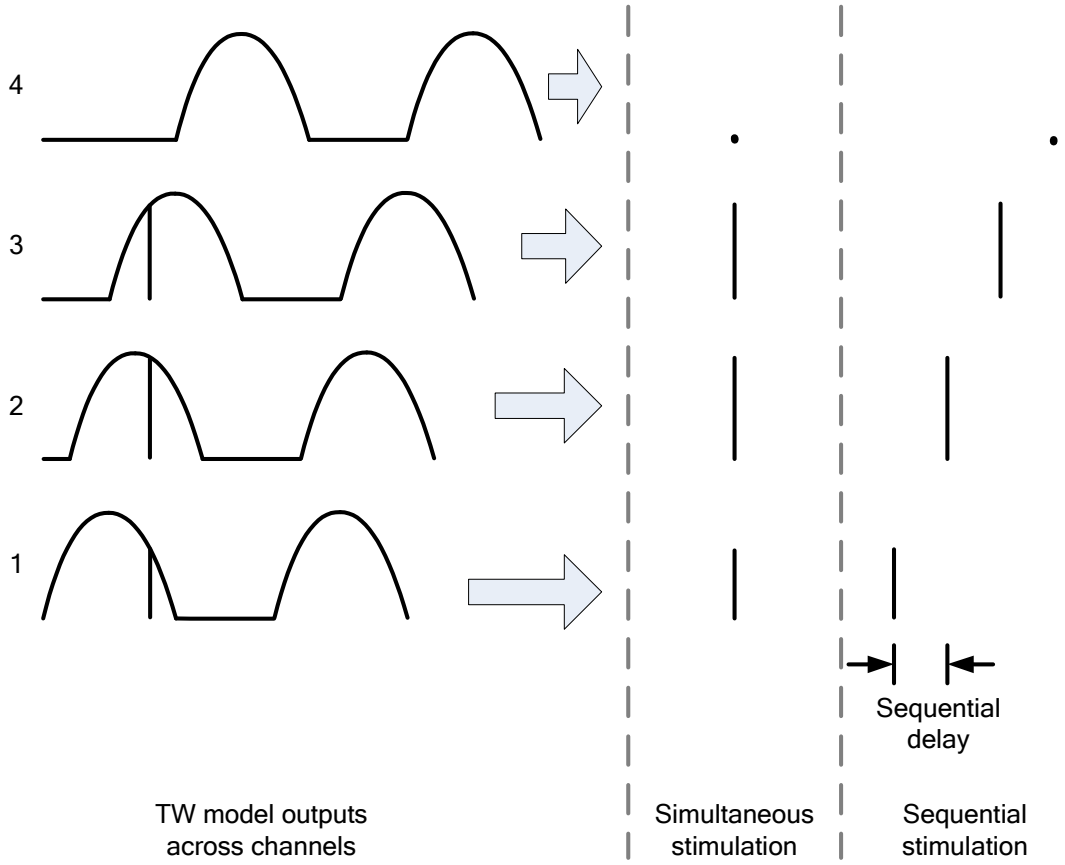

Figure 4.4: The difference between simultaneous stimulation and sequential stimulation. Sequential stimulation introduces a time delay across channels, which leads to phase distortion between adjacent channels.

Considering the trade-off between the stimulation rate and the number of electrodes, the resulting sampling rate for two channels would be half the total sampling rate. This implies an example where two channels are stimulated, at half the total maximum stimulation rate. Therefore every second sample is allocated to a channel. Samples 1, 3, 5, 7 ... of (a) would be allocated to the first channel (b) and samples 2, 4, 6, 8 ... (c) would be allocated

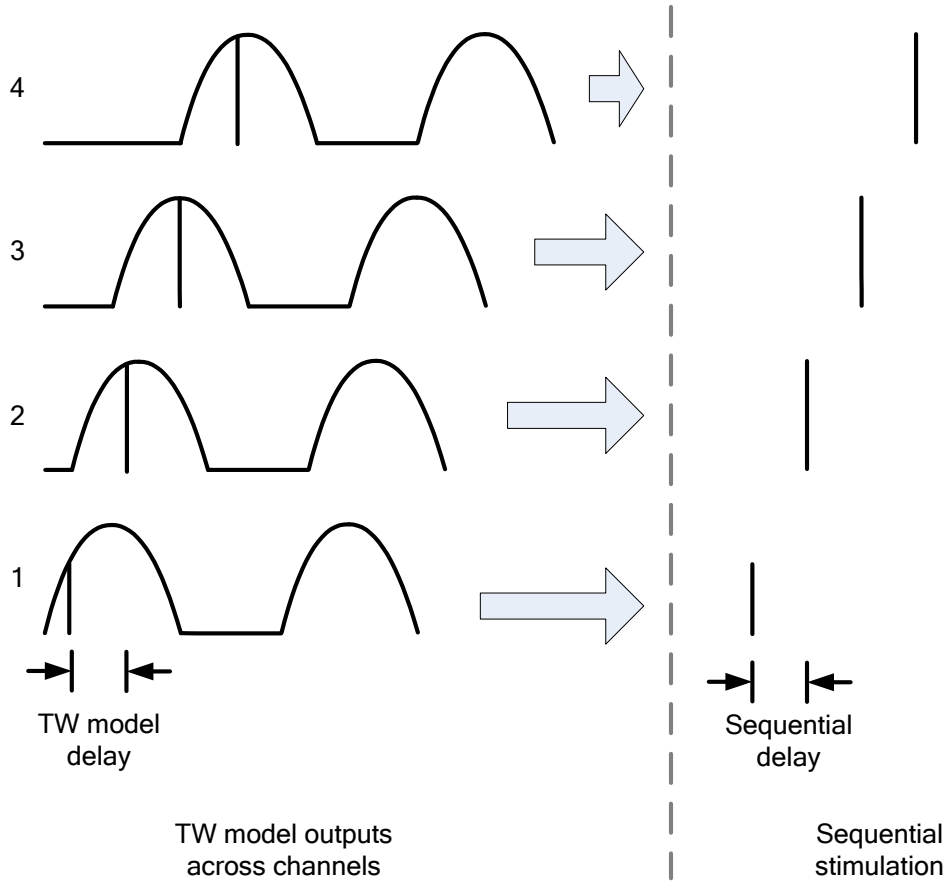

Figure 4.5: The TW model outputs for the adjacent channel are sampled at the next time sample. Stimulated samples are stimulated at more or less the same time as the time of the associated sample of the TW model output.

to the second channel (d). This would result in a stimulating pattern which compensates for the sequential time delay between the different channels. This time delay increases as the channels of stimulation are increased, resulting in six stimulated channels having a time delay of 277.6 μs between the first and the last channel. This division of the sampled BM response is explained in fig.4.6, where (a) and (c) show the sampled BM response for two different places along the cochlea and, (b) and (d) show how the samples are allocated to the two channels. This method allocates the samples to the right time slots associated with the specific places.

Fig. 4.7 and fig. 4.8 show the interleaved sequential resampling for a 4 kHz pure tone input simulated BM response with 22 channels activated. Each of these samples (bold lines) correspond to a specific place and time in the travelling wave model (thin lines). This sequence of sample amplitudes (fig. 4.9) is compressed into the dynamic range of the implant user, whereafter this dynamic range is mapped to the different CL units. The resulting CL unit pulses for electrode stimulation are shown in fig. 4.10.

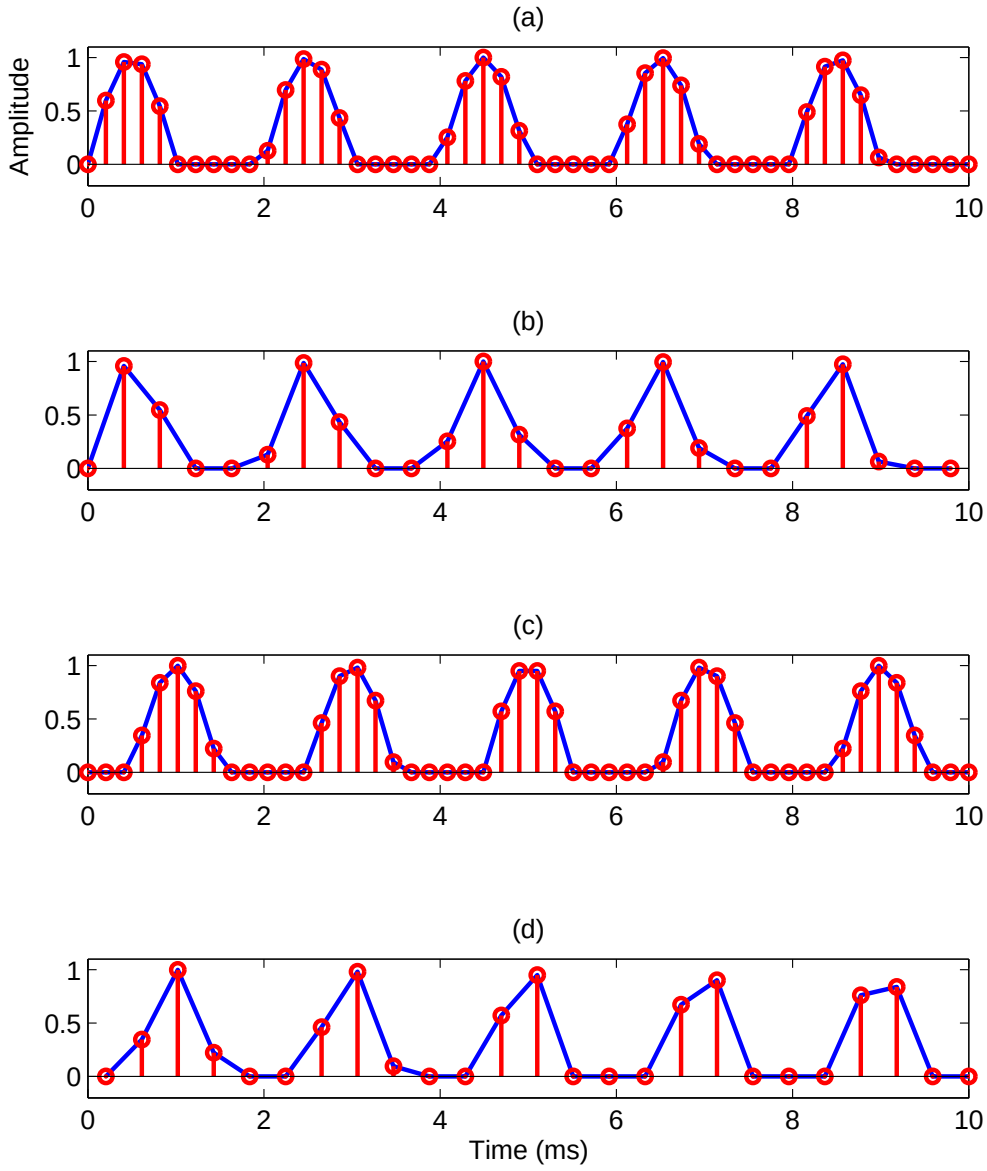

Figure 4.6: Illustrating the resampling of the TW model output into the active channels. (a) and (c) show the sampled simulated BM response on two different places along the cochlea (places 1 and 2) and (b) and (d) show how the samples are allocated between the two channels (channels 1 and 2). Sequential stimulation would lead to a stimulating sequence of [channel 1(t_1), channel 2(t_2), channel 1(t_3), channel 2(t_4) ...] where t_n is the time steps of stimulation. Therefore the places along the cochlea for the simulated BM amplitude response are assigned to the stimulating sequence in the order [place 1(t_{p1}) place 2(t_{p2}) place 1(t_{p3}) place 2(t_{p4}) ...], where t_{pn} refers to the simulated BM amplitude response time steps generated by the travelling wave model. The resulting sequence for channel 1 is [place 1(t_{p1}) place 1(t_{p3}) place 1(t_{p5}) place 1(t_p7) ...] and for channel 2 in [place 2(t_{p2}) place 2(t_{p4}) place 2(t_{p6}) place 2(t_{p8}) ...] for every stimulation time step $t_1 \ldots t_n$.

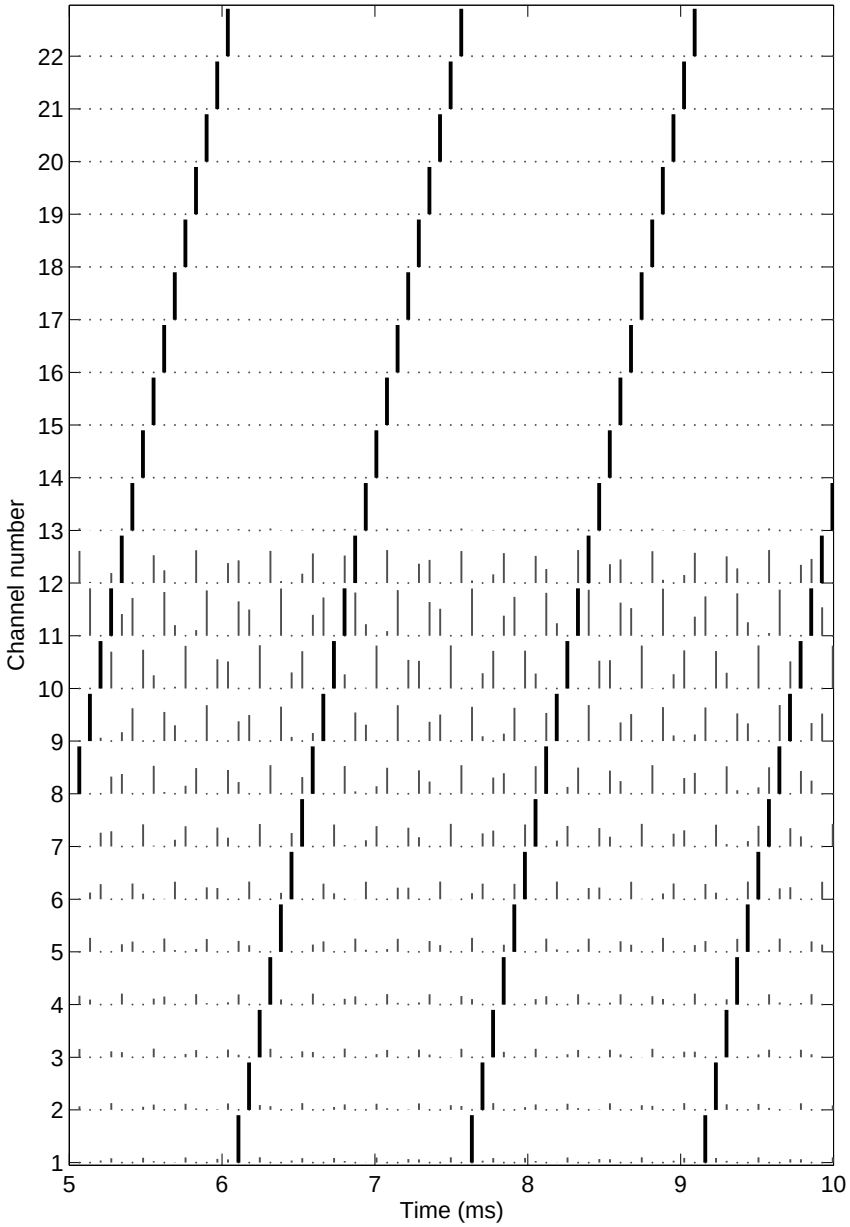

Figure 4.7: The travelling wave model samples (thin lines) which are sampled at 14 400 Hz are sequentially resampled (bold lines) for a 22 channel configuration. Each of the sequential samples correspond to a specific time and place in the travelling wave model. The small dots show a travelling wave model sample with either a negative or zero value.

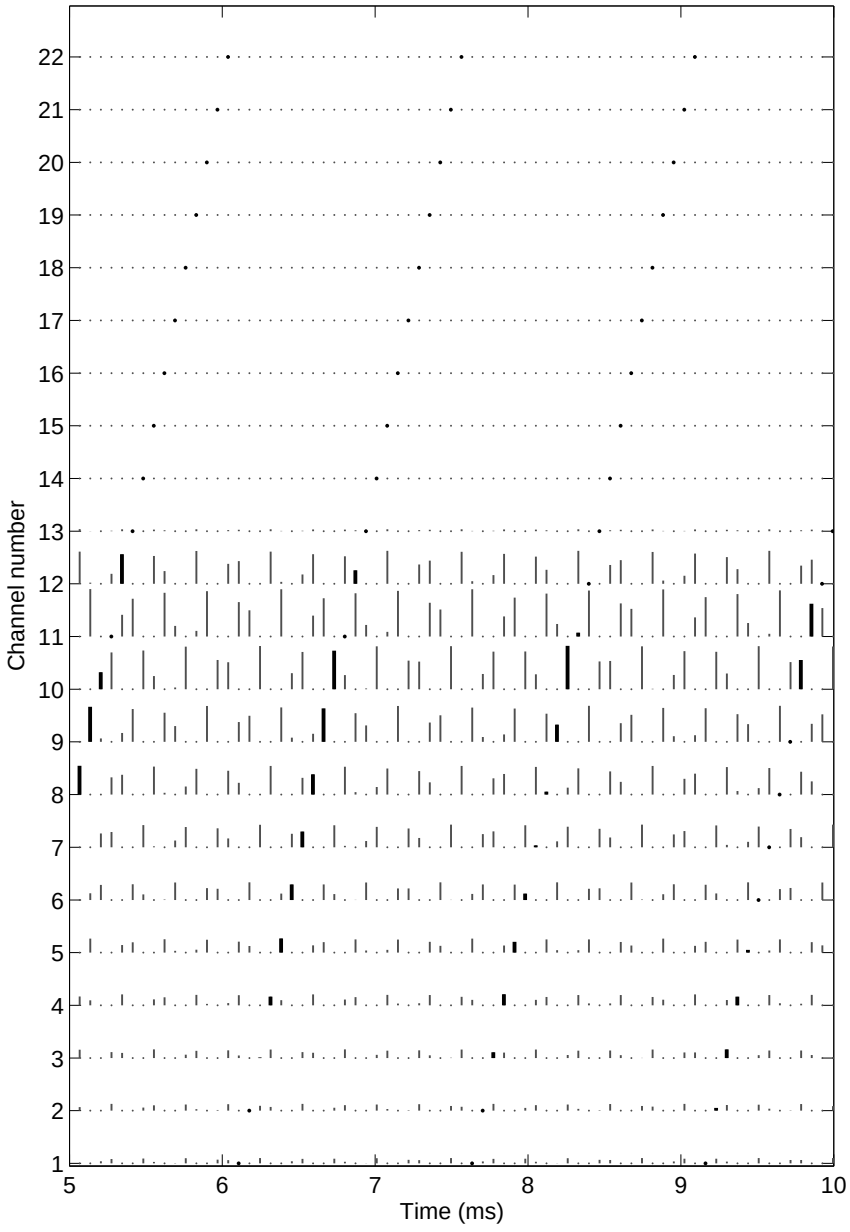

Figure 4.8: The sequential sampled values sampled at a sampling rate of 14 400/22 = 655 pps.

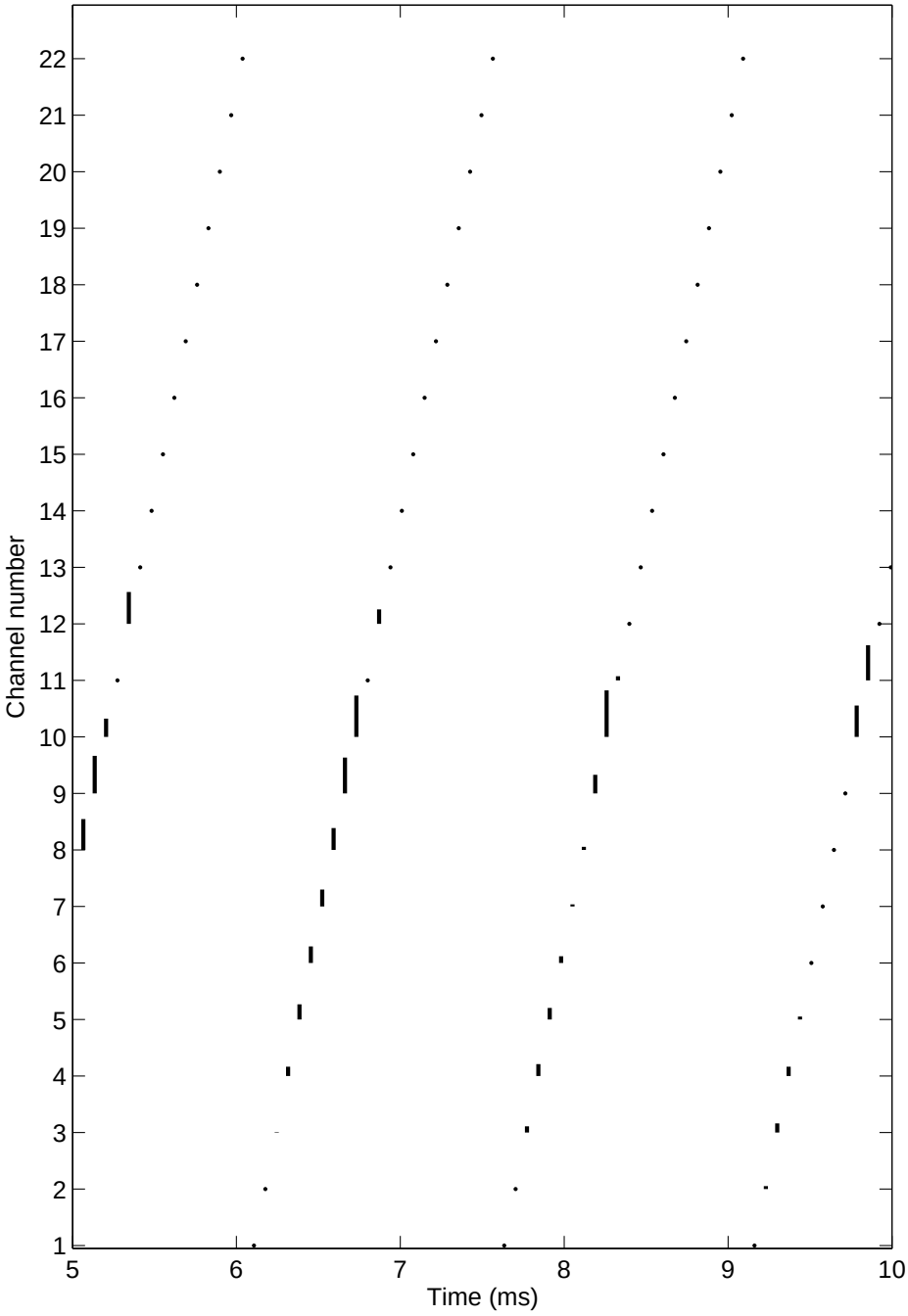

Figure 4.9: The electrode diagram of the sequential sampled values for the 22 channels, therefore maximum spatial resolution and minimum temporal resolution. The bold dots show zero sample values.

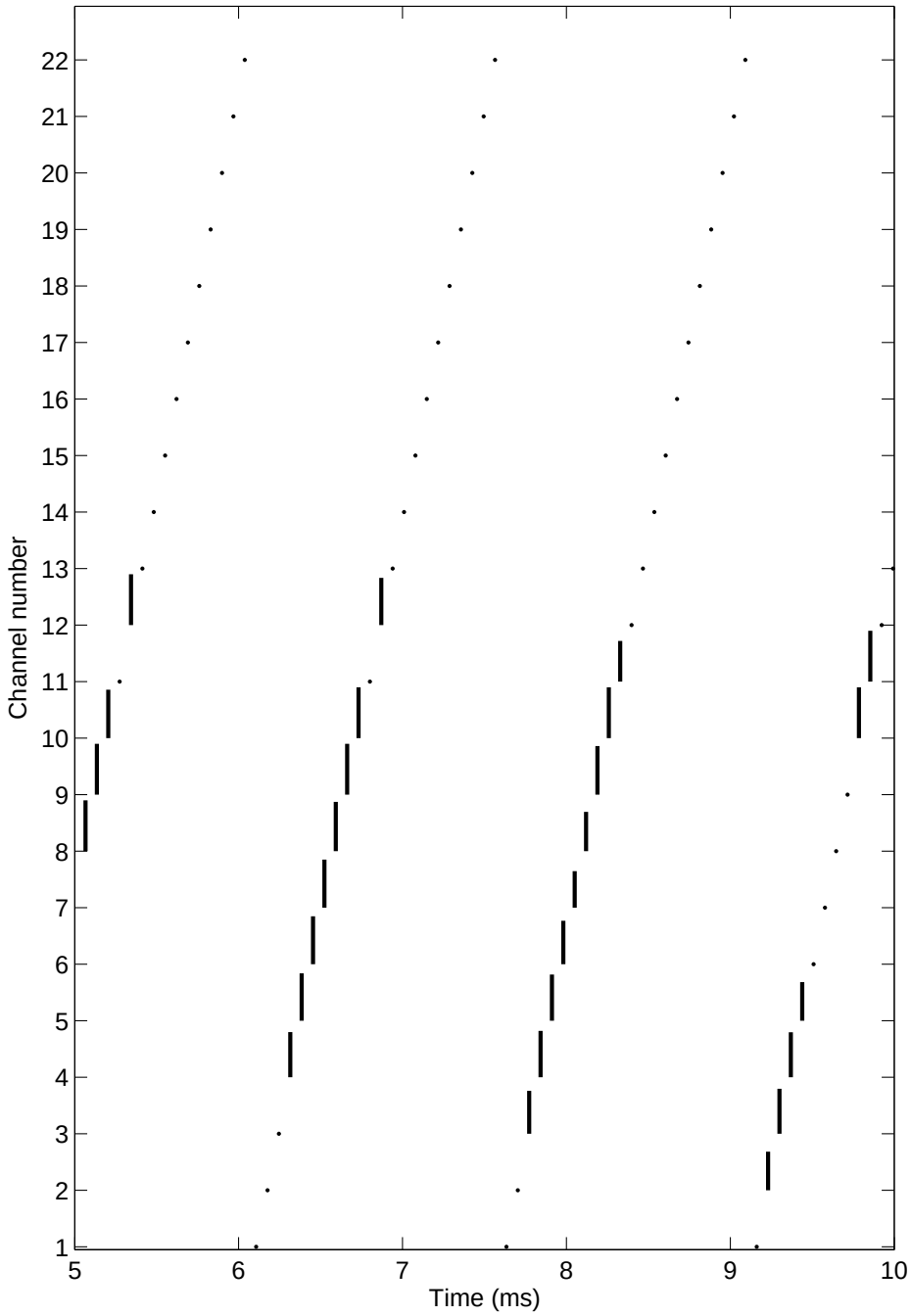

Figure 4.10: Pulses after the compression and CL mapping functional blocks. The bold dots show zero pulse values.

4.5 REAL TIME PROCESSING CONSIDERATIONS

The previous two chapters discussed the implementation and results of a hydodynamical model. This model solves the differential fluid equations and their corresponding boundary conditions, using a numerical method. For numerical stability, the numerical method requires a very small time interval, and therefore a high time resolution input signal. The previous section mentioned that the high time resolution BM displacement solution was resampled in time (at 14.4 kHz) to obtain a solution with a lower time resolution. If the BM displacement could be solved for a lower time resolution input (for instance sampled at 16 kHz) the number of calculations per second would be minimised, saving on computational time. This, however, is not possible with the existing numerical method, which would become numerically unstable for an input signal sampled at 16 kHz.

4.6 COMPRESSION FUNCTION

This non-linear function compensates for the non-linear increase in loudness perceived in electrical stimulation as a result of an increase in stimulation current. The LGF maps the simulated BM response amplitude to the dynamic range of the recipient. This dynamic range is defined as the range between the electrical threshold level (T) and the maximum comfortable level (C). This acoustic-electrical relationship is described (Loizou, 1998) using a logarithmic compression function of the form

$$Y = A log(x) + B, \tag{4.1}$$

where constants A and B are the acoustic and electrical amplitudes respectively. Another

type of compression function described in (Loizou, 1998) is the power law function:

$$Y = Ax^p + B, \qquad p < 1. \tag{4.2}$$

The advantage of the power law function is that the steepness can be controlled, simply by changing the value of p. The A and B constants are defined by

$$A = \frac{C - T}{x^p{}_{max} - x^p{}_{min}} \tag{4.3}$$

and

$$A = T - Ax^p{}_{min} \tag{4.4}$$

where the acoustic range is defined by $[x_{min}, x_{max}]$ and the electrical range by [T,C].

In the ACE strategy the energy in a specific channel is mapped to the current levels with a logarithmic-shaped function of the following form (Nogueira *et al.*, 2005):

$$q = \begin{pmatrix} \dfrac{log\left(1 + z\left(\dfrac{a-s}{m-s}\right)\right)}{log\,(1+z)} & , \quad s \leq a \leq m, \\ 0 & , \quad a < s, \\ 1 & , \quad a \geq m. \end{pmatrix} \tag{4.5}$$

where q is the output in the range between 0 and 1, corresponding to the T level and the C level, and a is the input simulated BM response amplitude. Inputs lower than the base level

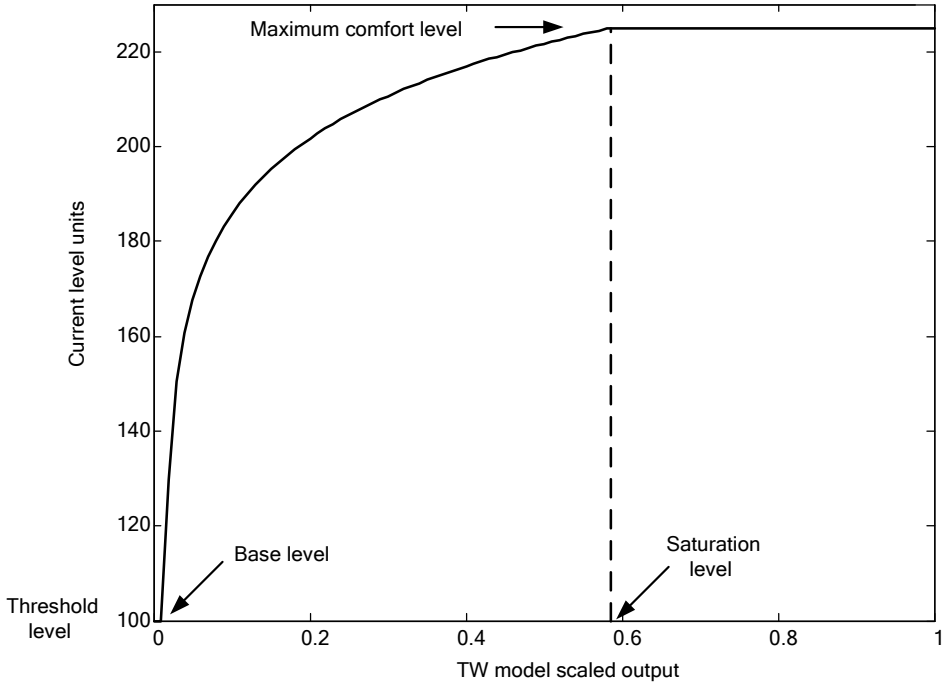

Figure 4.11: The relationship between the input amplitude and the output current levels in the LGF.

s are mapped to the T level and inputs higher than the saturation level *m* are mapped to the C level. The parameter *z* is the steepness factor, which controls the steepness of the curve. This function was also implemented in the TW processing strategy. Fig. 4.11 shows the relationship between the amplitude and the T-C levels for the LGF curve.

4.7 MAPPING

The LGF outputs a value between 0 and 1, corresponding to the T and C levels. Each electrode's T and C levels are subject-specific and adjusted separately. This LGF output range is mapped to the 255 current level (CL) units (Clark, 2003:488) available in the Nucleus 24 processor. This linear increase in CL units is mapped to a logarithmic increase in current in μA. Each CL unit is associated with a specific current, with CL unit 1 associated with a mini-

mum stimulus current of 10 μA and CL unit 255 with a maximum of 1.75 mA. An increase in current with electrical stimulation is perceived as an increase in loudness (Clark, 2003:488). The loudness growth function compensates for this logarithmic increase in loudness. The current value can be calculated with the following equation (Clark, 2003:488):

$$I_{CL} = 10 \times 175^{\frac{CL}{255}}, \qquad CL = [0, 255]. \tag{4.6}$$

This equation is implemented in the 3rd generation Cochlear implants. A similar equation to this one is implemented in the 4th generation Cochlear implants. The last equation was used in the travelling wave processing strategy. A T level of 100 and a C level of 225, which are typical values seen in CI implant subjects, were used in this study for all the active electrode channels.

4.8 ELECTRODE DIAGRAMS

The electrode diagrams (fig. 4.12 - 4.14) show the comparison between the travelling wave model samples and the sequential sampled values for a low frequency and high frequency input. The travelling wave model samples are the sampled TW model output (thin lines) at a sampling rate of 14 400 Hz and the sequential sampled values (bold lines) are the resampled TW model samples across the active channels. The sampling rate at which the travelling wave model samples are resampled is dependent on the number of active channels. A configuration where 22 channels are activated will therefore be sampled at a sampling rate of 655 pps, which is calculated by dividing the sampling rate of the travelling wave model samples of 14 400 Hz by 22 active channels. These sequential sampled values are the input to the LGF functional block.

Fig. 4.12 shows the result for a 200 Hz input signal, when all the electrodes' channels (1-22) are activated. Each of these channels is stimulated at a rate of 655 pps. The channel-number stimulation-rate trade-off can be seen when this figure is compared with fig. 4.13, where only six channels (17-22) in the area of excitation are activated. Fewer activated channels result in an increased stimulated rate of 2 400 pps. This results in 12 electrical pulses for one TW model sinusoidal cycle of 200 Hz, as opposed to three with 22 active channels. This shows a much better approximation of the output envelope of the travelling wave model. For a higher frequency pure tone input, the channel-stimulation rate trade-off becomes even much more clear.

Higher input frequencies require a higher sampling rate for a good approximation of the input envelope. A low sampling rate would result in under-sampling. This effect can be seen in fig. 4.14, where 13 channels are activated, with a stimulation rate of 1 108 pps each. This results in a very poor approximation of the TW model envelope for a 1 kHz input. Increasing the stimulation rate (3 600 pps) and decreasing the electrode channels (fig. 4.15) results in a narrower activation spread but a better approximation of the TW model envelope. Results from the electrode diagrams confirm the importance of the electrode channel-stimulation rate trade-off that was discussed in section 4.4.

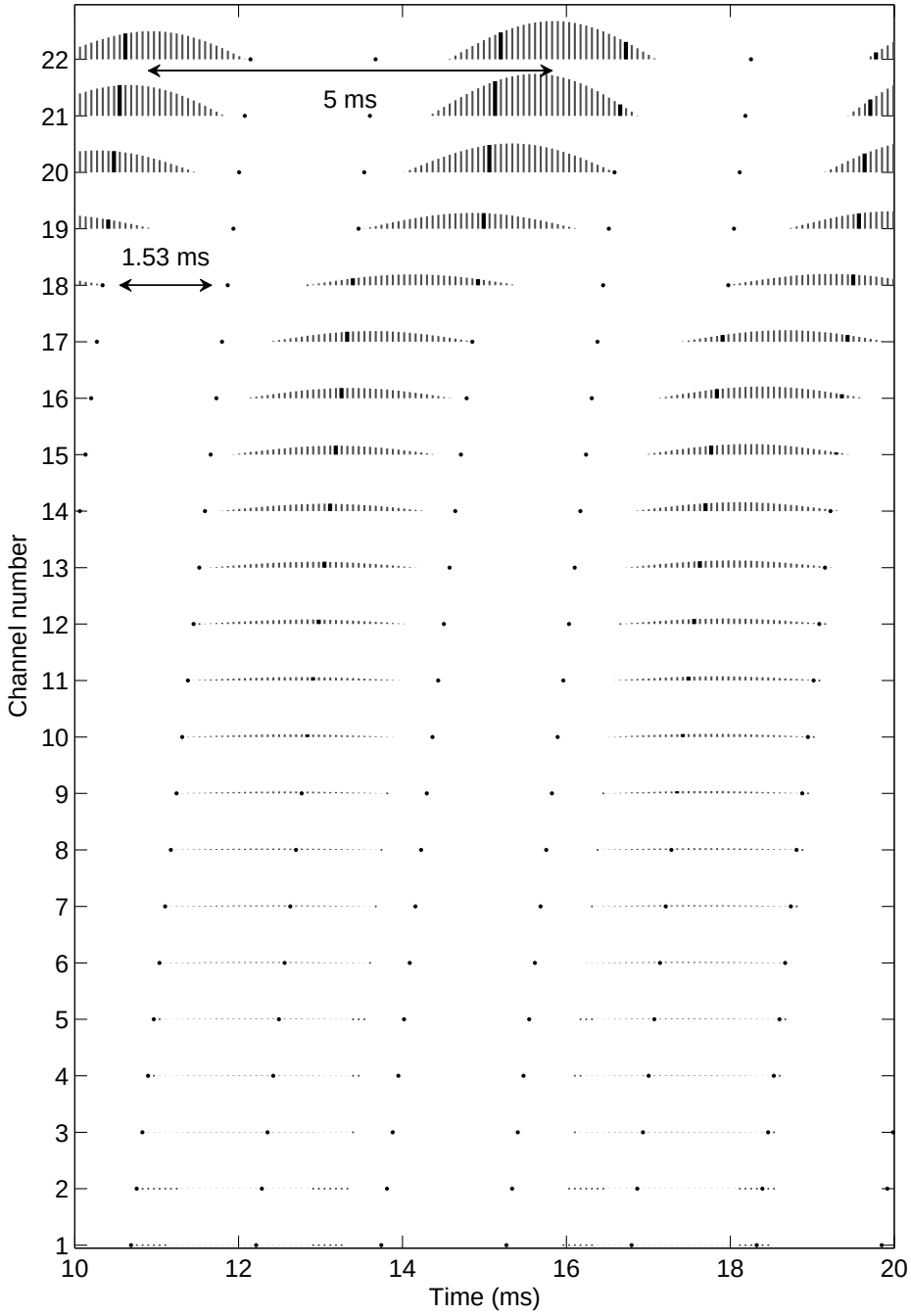

Figure 4.12: Illustrating the time samples for the TW model (thin lines) versus the sequential resampling samples (bold lines) for a 200 Hz pure tone, with electrode channels 1 to 22 activated. The dots represent resampling values with an amplitude of zero.

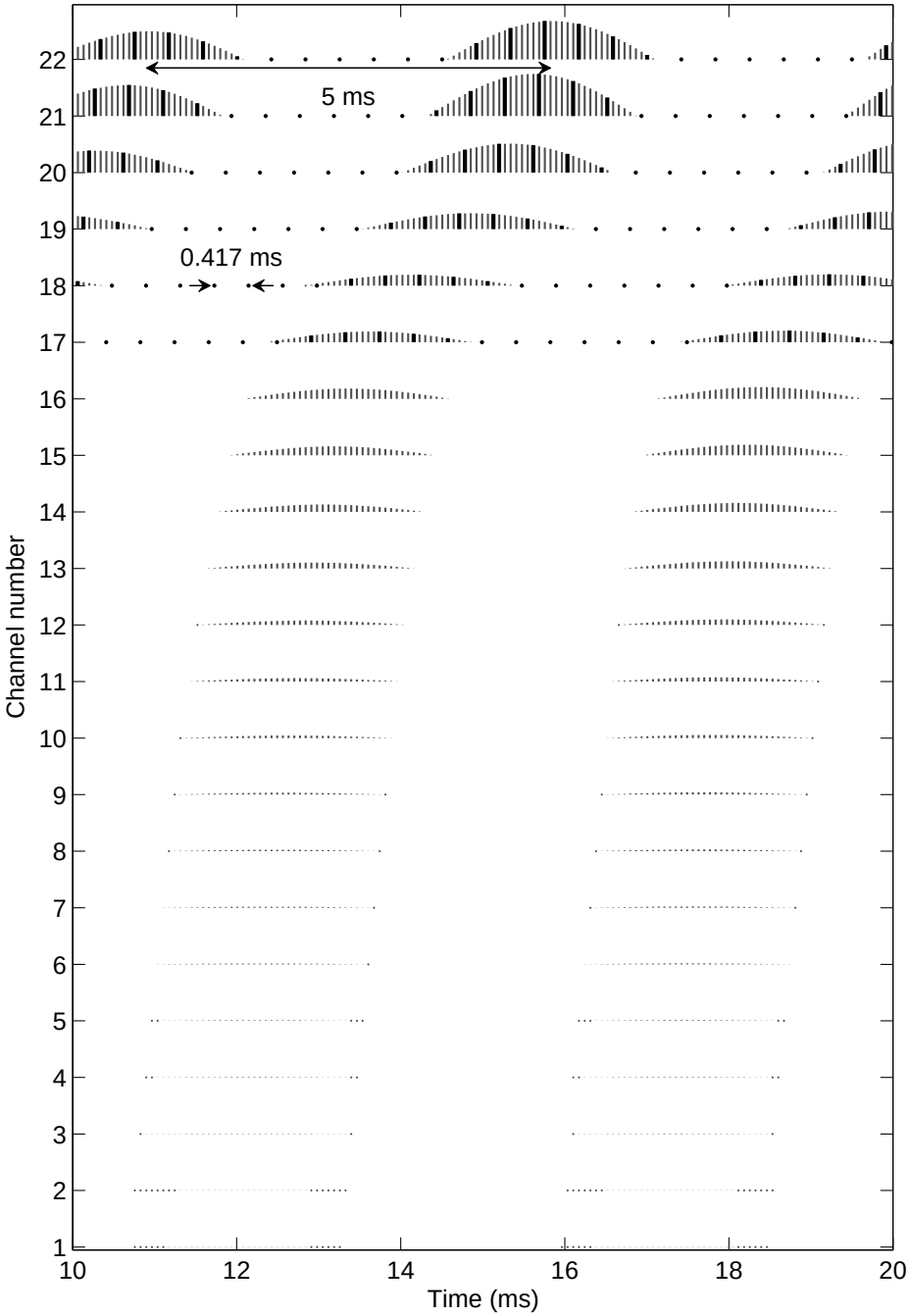

Figure 4.13: Minimising the the number of electrode channels (17-22) results in a higher stimulation rate and therefore a better approximation of the TW model envelope.

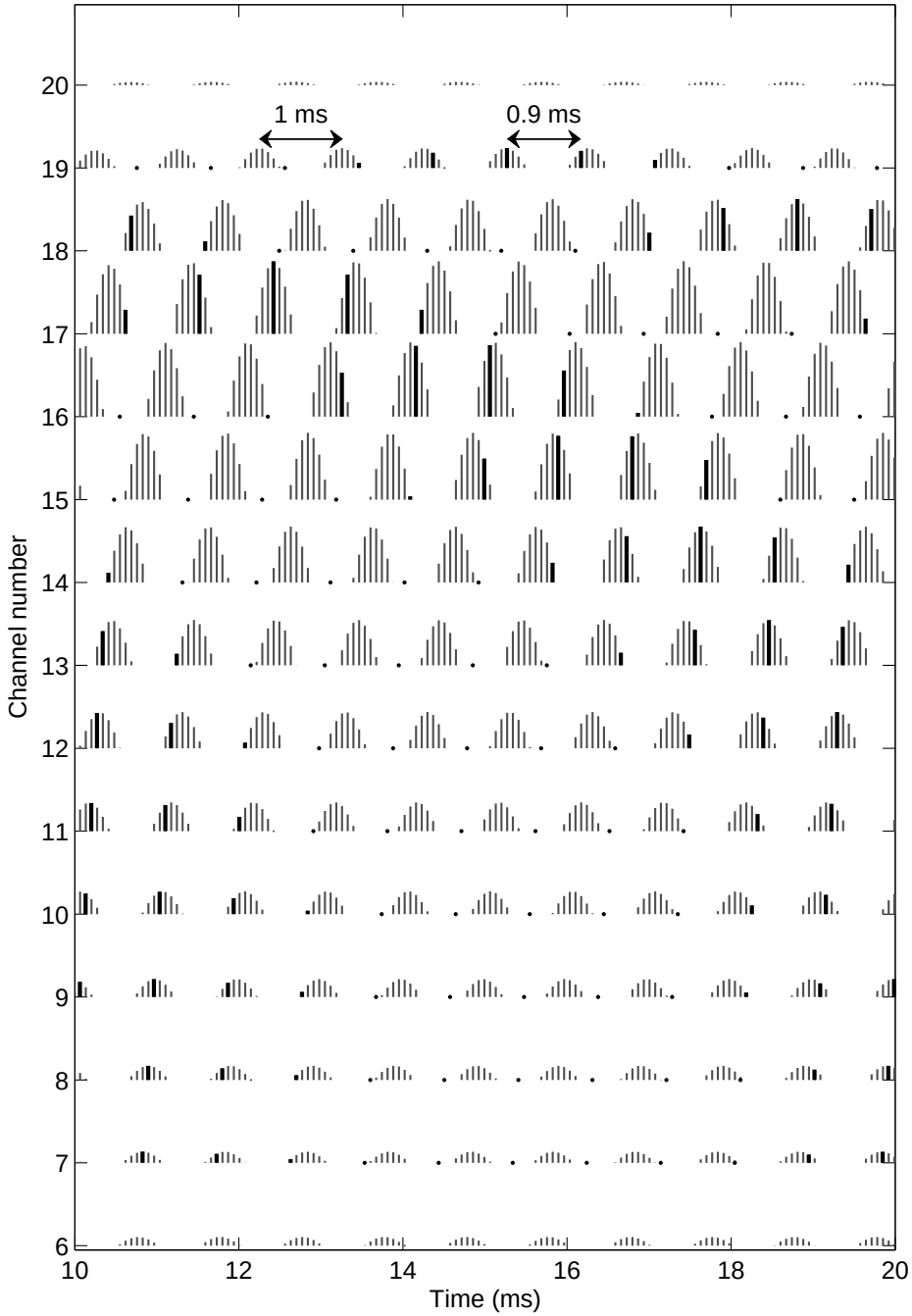

Figure 4.14: A higher frequency input requires a higher stimulation rate for the TW model envelope of 1 kHz to be adequately presented. A large activation area (electrodes channels 7-19) results in a smaller stimulation rate. The result of this trade-off is a large distortion between the TW model (thin lines) and the sequential resampled pulses (bold lines).

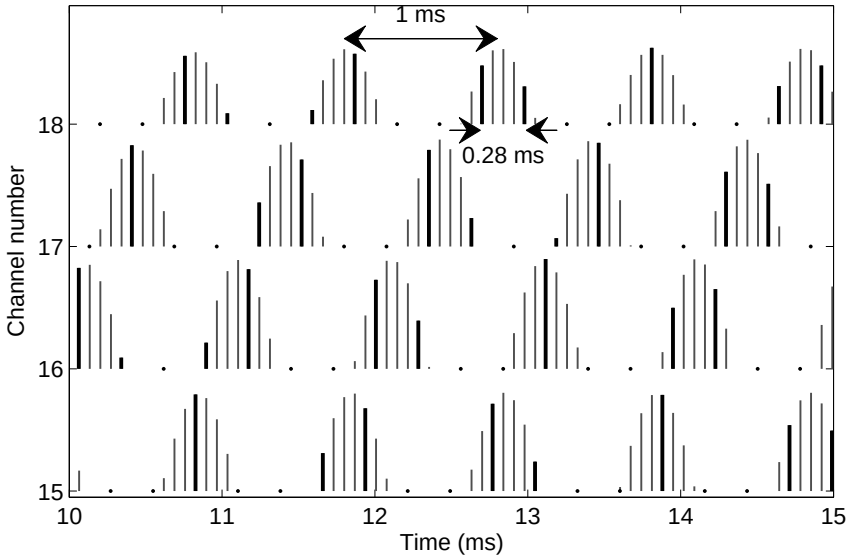

Figure 4.15: Minimising the number of electrode channels results in a higher stimulation rate and a less distorted representation of the model envelope. The results for a 1 kHz pure tone input, with electrode channels 15 to 18 activated, are shown.

4.9 CURRENT SPREAD AND ELECTRICAL STIMULATION MODELS

Fig. 4.16 shows the electrical stimulation model which approximates the neural excitation along the length of the cochlea. Current spread around the stimulation electrode was approximated with an exponential current decay away from the stimulation electrode. This current spread introduced a stimulating current across the length of the cochlea in the area of stimulation. The neural excitation in this area of stimulation was approximated with Bruce *et al.*'s (1999a) and Bruce *et al.*'s (1999b) electrical model, which was implemented by Smith (2011).

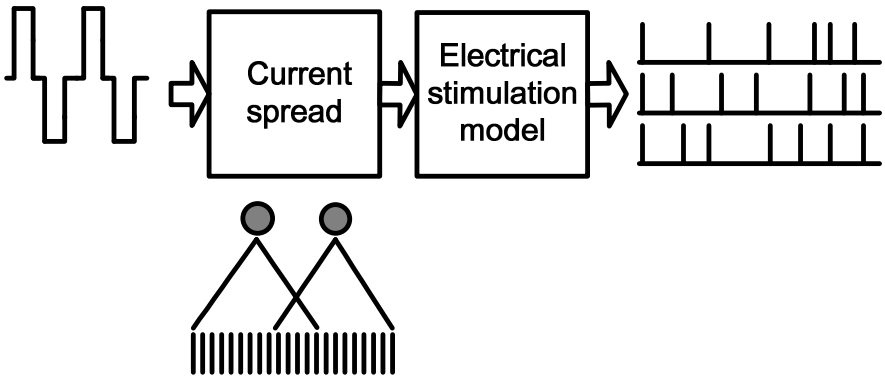

Figure 4.16: The functional block diagram of the current spread and electrical stimulation models.

4.10 CURRENT SPREAD MODEL

One important aspect of electrical stimulation is the degree to which current spreads from the stimulated electrode to the surrounding nerve fibres. The effect of current spread was therefore included to observe its effect on spike responses. The relation between the current and the stimulating electrode is approximated to be an exponential decay (eq. 4.7) over the length of the cochlea. This current decay rate of 2.8 dB/mm corresponds with a mean rate

calculated from the current decay rate estimates for five different current spread estimation methods (Bingabr, Espinoza-Varas and Loizou, 2008) for monopolar stimulation. The current as a result of current spread is described in the following equation (Smith, 2011; Strydom and Hanekom, 2011):

$$I_{stim} = I \cdot \Delta d \cdot \delta. \tag{4.7}$$

In eq. 4.7, I_{stim} is the current at a distance Δd from the stimulating electrode, producing a current I, with a decay rate of δ. Fig. 4.17 shows this interaction between the electrodes and the nerve plane. The stimulation current for each channel can be determined by summing the modified currents (eq. 4.7) of all active electrodes. The equation for the channel current is

$$I_{stim}(c) = \delta \sum_{z=1}^{Z} \Delta d_c(z) \cdot I(z) \tag{4.8}$$

where c is the channel index, and z is the electrode index. The equation is summed for all electrodes, Z (Smith, 2011; Strydom and Hanekom, 2011).

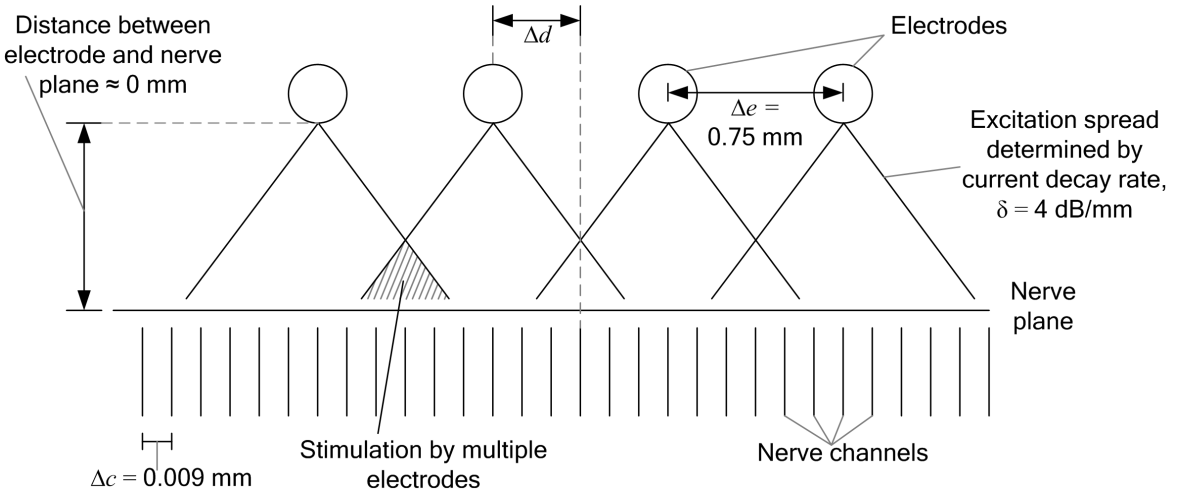

Figure 4.17: An illustration of the excitation spread from the electrode array (fig. 3.25 in Smith (2011)).

4.11 ELECTRICAL STIMULATION MODEL

A stochastic computational model implemented by Smith (2011) describes the nerve activation for stimulated electrical pulse-trains. This implementation is based upon the models of (Bruce *et al.*, 1999a; Bruce *et al.*, 1999b). The model compares the stimulus current I_{stim}, represented by the input pulse train, with the summation of the threshold, noise and refractory currents ($I_{thr} + I_{noise} + I_{refr}$). If the stimulus current is greater than or equal to the added threshold, noise and refractory currents, neural firing occurs (fig. 4.18). This comparison is performed at every time sample in the cathodic phase.

Spiking occurrences, however, are dependent on the refractory effect. After a spike occurs, there is a certain time period where the neuron is unable or less able to fire. This time period is divided into two time periods known as the absolute refractory period and relative refractory period. For the first period, taken to be 0.7 ms (Bruce *et al.*, 1999b), I_{refr} has an infinitely high value, resulting in no spikes. Following this period, I_{refr} follows an exponential decay with a time constant of 1.32 ms, where the probability of firing increases as the value of I_{refr}

decreases. A constant time length of 20 ms has been assigned to this exponential function of I_{refr} (Bruce *et al.*, 1999b). The stochastic nature of the noise current I_{noise} introduces a further variation in spike times.

It must be added that neural firing only occurs in the cathodic phase, based on the assumption that spikes are more likely to occur in this phase. This assumption is based on the fact that the cathodic phase has a lower threshold than the anodic phase (Bruce *et al.*, 1999a).

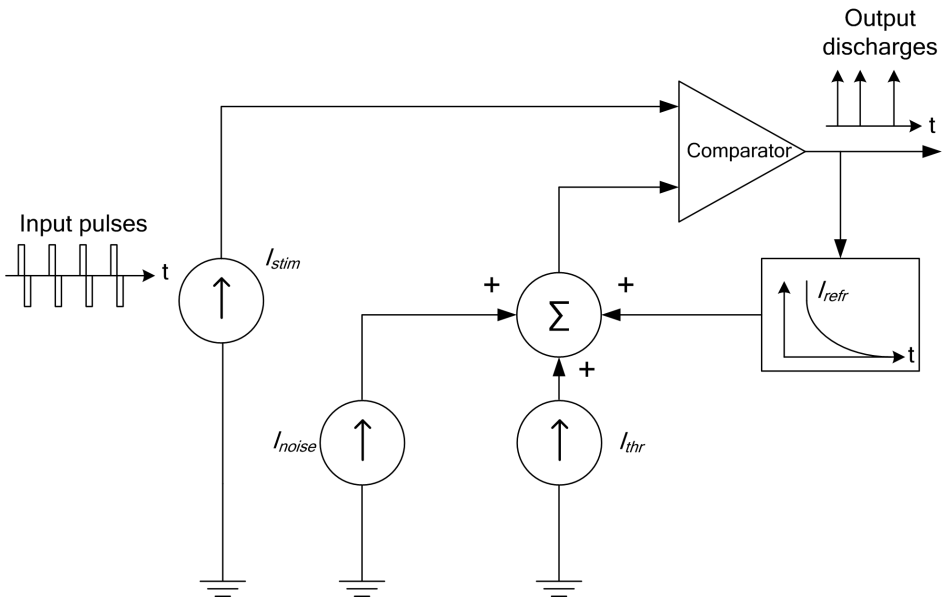

Figure 4.18: The circuit model of the electrical stimultion model implemented by Smith (2011), which was adapted from Bruce *et al.* (1999a). (Fig. 3.28 in Smith (2011)).

4.12 A PROPOSED NORMAL HEARING MODEL

To investigate the difference between electrical stimulation and normal hearing, a normal hearing neural model, which incorporates the BM movement, is proposed in this section.

The travelling wave movement of the BM causes a hair cell shearing movement, which in turn results in neural spike activation for upward displacement (Russell, Richardson and

Cody, 1986). For a pure tone input this displacement reaches a peak at a specific place along the cochlea. The model assumes the most probable place for neuron-firing, $p(spike)$, to be at this peak displacement. Arbitrary probability values of 0.9, 0.02 and 0.005 were chosen. It follows that for every time instant $(t_1 \ldots t_n)$ a uniformly distributed random variable between [0,1] is generated using the *rand* function in Matlab. If the random variable is greater than $1 - p(spike)$, neural firing occurs at that time instant. When the variable generated is smaller than $1 - p(spike)$, there is no firing. If this assumption is correct, an increase in the BM displacement would cause an increase in the firing probability. Therefore with this model a close relation remains between the BM displacement amplitude and the firing probability. De Jongh (1977) found similar correspondence between a linear filtered stimulus waveform and the firing probability. The absolute refractory period 0.7 ms, for which no firing occurs, was also incorporated in the model, corresponding to the value used in Bruce *et al.*'s (1999a) model.

This distribution of neural spikes is repeated for every channel along the length of the cochlea. Although the TW model only generates displacement values for 114 channels, these values were linearly interpolated to obtain 3 889 channels, spaced 9 μm apart. This results in a firing pattern that follows travelling wave characteristics (see chapter 5).

The next section discusses the neural pattern results for different channel-number stimulation-rate configurations, with respect to the channel-number stimulation-rate trade-off discussed in this chapter. These results are compared to those of the proposed normal hearing model to show the differences for approximated neural patterns in normal hearing and electrical stimulation, with respect to travelling waves.

Chapter 5

Predictions of neural spike train patterns for travelling wave speech processors

5.1 INTRODUCTION

In order to investigate the effect that the travelling wave processor has on electrical neural activation, the travelling wave model has been integrated with the excitation spread and electrical stimulation models implemented in Smith (2011). A normal hearing model with temporal travelling wave characteristics was proposed in chapter 4. The first section of the present chapter discusses the simulated spatio-temporal neural fibre response patterns for this normal hearing model. These results, along with ISI histogram results, are compared to measured response and neural firing data. The second section discusses the spatio-temporal neural responses resulting from electrical stimulation. For comparing the TW processing strategy to present processing strategies, neural response data were predicted for pure tone inputs with the ACE speech processing strategy as implemented by Smith (2011). In the main part of this chapter the trade-off between the number of activated electrode channels

and stimulation rate, with respect to electrical stimulation, is investigated. Results for spatio-temporal neural patterns with different channel-number stimulation-rate configurations are shown. Using ISI histograms, the differences in predicted ISIs could be investigated for various channel-number stimulation-rate configurations. Another type of histogram further investigated these channel-number stimulation-rate configurations. The post-stimulus time (PST) histogram calculates the number of spikes within a time bin for a band of neural channels. These neural channel bands are a result of neural channel summation, where a number of neural channels are collapsed onto one another. The last section investigates the effect of deeper insertion depth and wider electrode spacing.

5.2 MODELLING OF NEURAL SPIKE TRAINS FOR NORMAL HEARING

In the normal cochlea, nerve fibres are excited by a unidirectional movement (towards the scala vestibule) in the basilar membrane (Brugge, Anderson, Hind and Rose, 1969). This results in firing times corresponding to the elevations of the stimulating BM waveform (Russell *et al.*, 1986). BM displacement for a pure tone input at a certain place along the cochlea (chapter 3) shows a sinusoidal movement of the BM at the input frequency. This sinusoidal movement is conveyed by the spike patterns of the auditory nerve fibre (fig. 17.13 in Javel (1990)). The proposed model also shows these temporal characteristics in the predicted neural spike train patterns. Fig. 5.1 shows the predicted spatio-temporal neural spike train patterns for the parameter set (section 2.5.1), with a BM length of 22.5 mm. These parameter values have been derived from measured cat data (Neely, 1981b). These spatio-temporal neural patterns show the times of the predicted spikes of the nerve fibres along the length of the cochlea, measured from the base. The nerve fibres have been placed at a distance of 9 μm from each other, resulting in 3 889 modelled fibres along the length of the cochlea. Each

spike is represented by a dot at the specific time of firing on the graphs. The left-hand y-axis represents the length along the cochlea, measured from the base, whereas the right-hand y-axis represents Greenwood's (1990) approximated frequency associated with that specific place.

These results show spike occurrence on the upward displacement of the BM, corresponding to a half cycle of the travelling wave model's sinusoidal input. Results at the lower frequencies show a wider band of neural firing (fig. 5.1, a) and b), due to longer cycle time lengths, rather than higher frequencies (fig. 5.1, c) and d), with narrower, more frequent bands of neural firing. To investigate the effect of the firing probablity on the ISI histogram results of the normal hearing model, three arbitrary firing probabilities (0.9, 0.2 and 0.05) were chosen and compared to ISI histogram results from single nerve fibres (Rose *et al.*, 1967; Javel, 1990). Fig. 5.1 shows the spatio-temporal neural spike patterns for the model with a firing probability of 0.9 and fig. 5.2, shows the calculated ISI histograms for the normal hearing model's neural response with this firing probability.

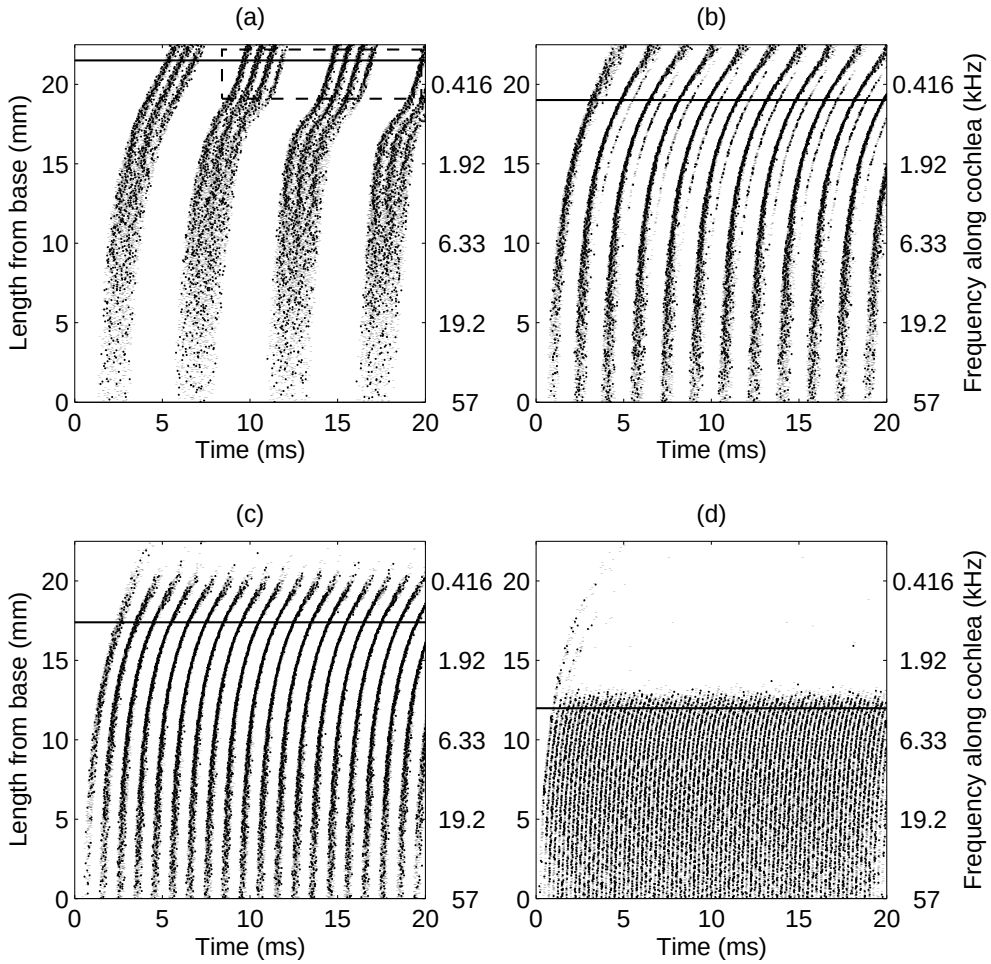

Figure 5.1: Spatio-temporal neural spike train patterns for the normal hearing model (firing probability of 0.9) with an input frequency of a) 200 Hz, b) 600 Hz, c) 1 kHz and d) 4 kHz. Each dot represents a firing at that specific time. The horizontal line in each graph is the predicted place related to the input frequency according to Greenwood's (1990) equation.

These histograms are calculated by counting the number of times a given interval occurs. An interval is defined as the time duration between two successive spikes for a nerve fibre. All the intervals within a certain time duration (time bin) are added together. The ISI histogram indicates the number of intervals within a certain time bin. A time bin duration of 0.1 ms was used in this section, unless specified otherwise.

Rose *et al.* (1967) recorded neural response for single auditory nerve fibres in the cochlear

nerve. They constructed ISI histograms for 10 tonal stimuli with a 1 second duration. Their results (fig. 5.3) showed histograms that had clusters of intervals at multiples of the stimulation pulse rate period. The number of intervals in each cluster decreased exponentially, as the interval length increased. For 1 000 cycles per second (cps) stimuli, there were almost no interval lengths longer than 15 ms. Javel's (1990) histogram results (fig. 5.4) for an auditory nerve fibre showed a similar exponential decaying function. It must be noted that the comparison results (Rose *et al.*, 1967; Javel, 1990) were recorded for relatively long stimuli in a single auditory nerve fibre in anesthetised squirrel monkeys (Rose *et al.*, 1967) and in cats (Javel, 1990) (The authors do not specify whether their results were conducted on deaf or hearing cats.) However, as a result of a processing time limitation, the results obtained from this model were calculated for short stimuli (20 ms) across a modelled population of 3 889 nerve fibres. This leads to the following assumption: Observing that spikes do not occur on every cycle of the input pure tone (Rose *et al.*, 1967), whether calculating the ISI histogram for a population of fibres for a short duration, or the ISI histogram for a single fibre for a long duration, it is assumed that the normal hearing ISI histogram results for both these cases will follow a similar trend. This assumption was followed in this section.

Comparing their results with the ISI histograms results obtained from the model with a firing probability of 0.9 (fig. 5.2 a) - d), shows clusters at multiples of the input tone period in their results, as opposed to only one cluster at the tone period for the model's ISI results. As the probability of firing in the proposed normal hearing model is directly related to the displacement of the BM, a high probability ($p(spike) = 0.9$) would result in spikes occurring on almost every displacement cycle. As a result, most firing will occur on the first cycle, with very little on the second and third cycles (fig. 5.2).

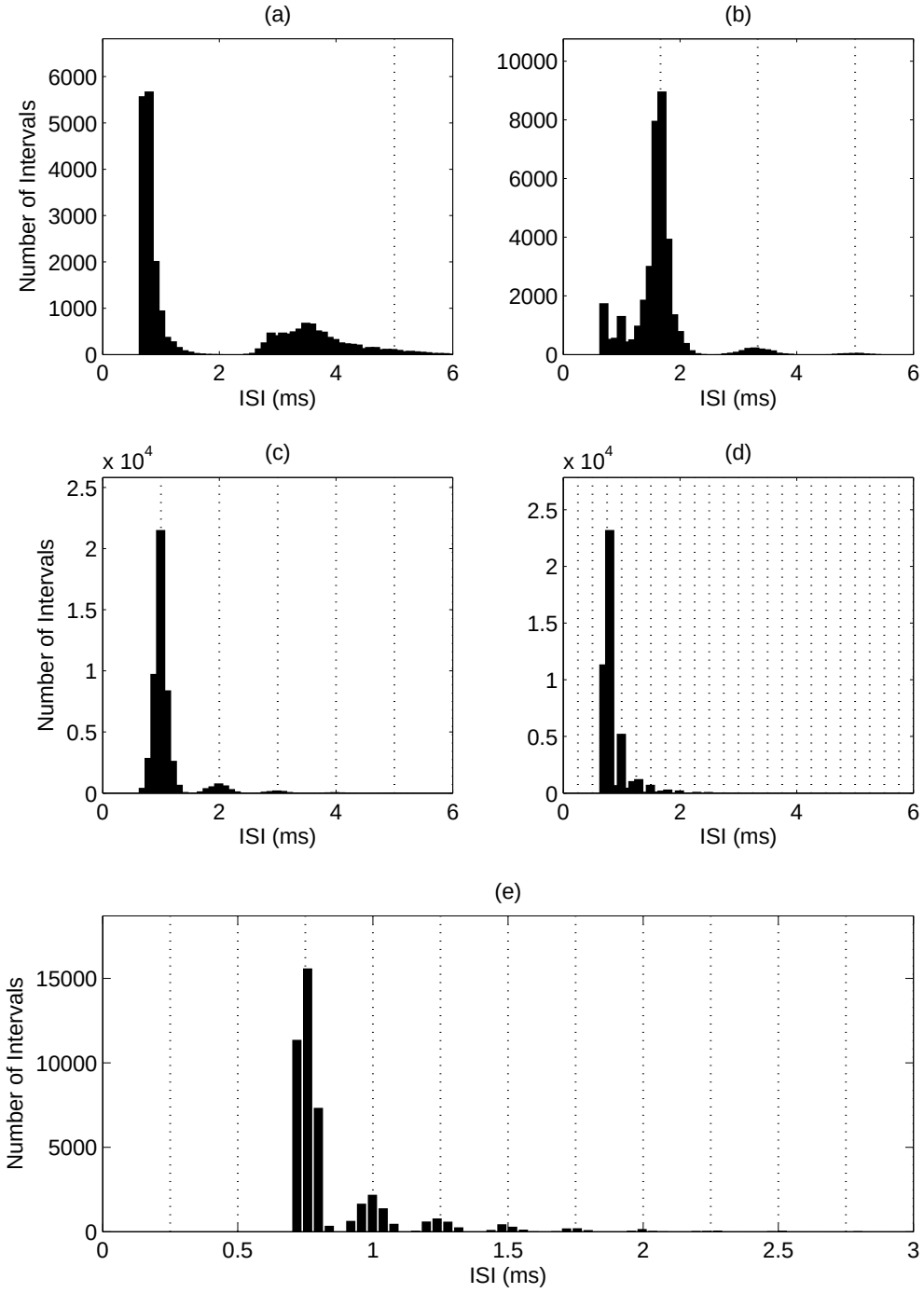

Figure 5.2: ISI histograms for the normal hearing model with a high firing probability of *p(spike)* = 0.9. These results show a definite clustering of spike intervals around the tone period, but almost no spikes at multiples of the tone period. The respective input frequencies are a) 200 Hz, b) 600 Hz, c) 1 kHz and d) 4 kHz. In the enlarged version for the 4 kHz input (e), the number of time bins was increased, with a time bin length of 0.04 ms.

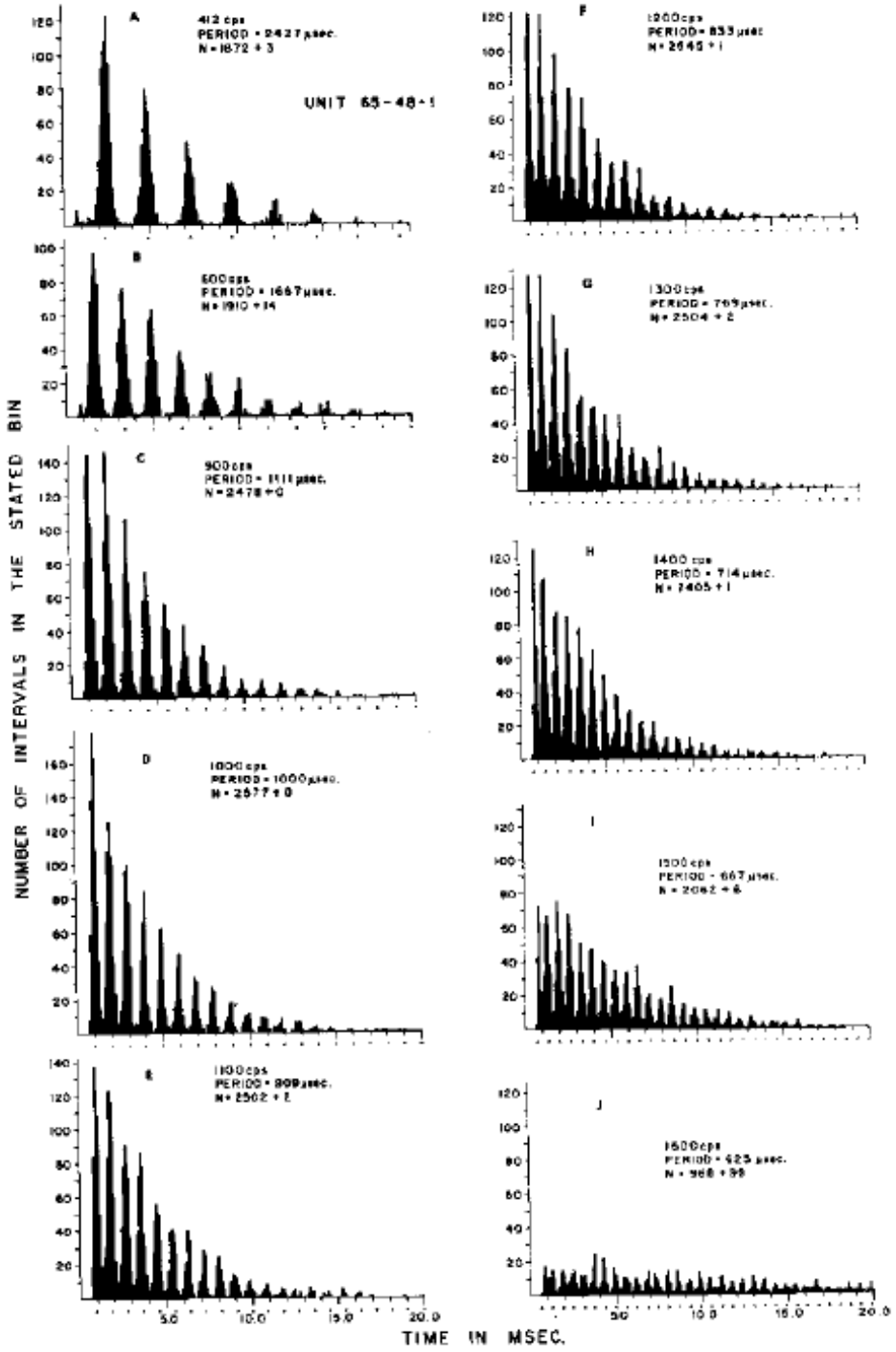

Figure 5.3: ISI histograms for recorded neural responses at different pure tone frequencies (reprinted with permission from Rose *et al.* (1967)).

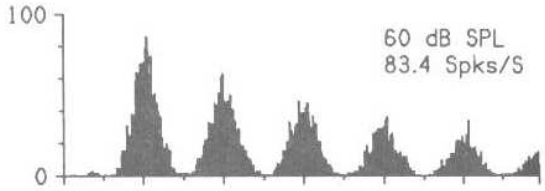

Figure 5.4: ISI histograms for neural response to a 200 Hz acoustic tone (reprinted from Javel (1990), fig 17.20, with kind permission of Springer Science and Business Media).

Results for the lower probabilities of 0.02 and 0.005 (fig.5.5 and 5.6 respectively) show definite peaks at multiples of the tone period. These peaks can also be seen in the calculated ISI histograms for the measured responses of Rose *et al.* (1967) and Javel (1990). Fig. 5.5, a) and d), shows large clusters around the absolute refractory period of 0.7 ms (marked by the arrows). For the lower frequency case (fig. 5.5 a)), having a long time cycle, the probability of firing remains high for a longer time, leading to multiple firings within one cycle. Rose *et al.* (1967) also found a similar, much smaller peak at 1 ms (fig. 4 in Rose *et al.* (1967)) and concluded that this must be the result of multiple firings rather than a single spike. The model's predicted histogram results, however, show this peak to be larger than the peak at 5 ms. For the higher frequency case ((fig. 5.5 d)), with a period smaller than 0.7 ms, firing will not occur faster than the absolute refractory time period, explaining the reason for the large cluster of spikes at 0.7 ms.

A further observation (fig. 5.5 a)) shows only four peaks at multiples of 5 ms as opposed to at least six peaks in Javel's (1990) histogram data. The proposed model only introduces the constant absolute refractory period of 0.7 ms. However, in the nerve firings, this period is followed by the relative refractory period, where the probability of firing increases as time increases. Implementing the relative refractory period in the model will increase the time intervals between spikes. This might result in fewer firing intervals at 0.7 ms and more firing

intervals at multiples of the input tone's period.

Rose *et al.* (1967) calculated the probability that the next firing will occur, multiples of tone periods further in time, given that a nerve fibre has fired. The probabilities they calculated ranged from an average of 0.502 for 412 cps, to 0.0608 for 1 600 cps, decreasing for higher frequency input tones. Although the firing probabilities of 0.02 and 0.005 are much smaller (more than an order) compared to those calculated in Rose *et al.* (1967), the predicted results also seem to suggest that the firing probability might decrease with an increase in tone frequency. This is seen in the ISI histogram (fig. 5.5 (a) - (c) for 200 Hz, 600 Hz and 1 kHz input tones, which shows a similar exponential decrease in peaks to the peaks seen in the ISI histogram data (Rose *et al.*, 1967; Javel, 1990) for input tones of up to 1 500 cps. ISI histograms (Rose *et al.*, 1967) however, show this exponential decrease in peaks to disappear at 1 600 cps. In fig. 5.6 this decrease also disappears with the higher frequency results (e), as a result of the low firing probability. This can also be seen in the spatio-temporal neural patterns in fig. 5.7 (f), where synchronous firing for populations of fibres is less evident compared to the spatio-temporal neural patterns seen in figs. 5.1 (d) and 5.7 (e). Therefore introducing a decreasing firing probability as input frequency increase might result in a better approximation of the ISI histogram data from measured neural responses.

Another observation from figs 5.1 and 5.7 is the definite travelling wave delay seen as the wave approaches the place of peak BM deflection. This travelling wave delay is much clearer in the lower frequencies, a) and b), in the apical region. Fig. 5.1 a) however, shows an increase in speed as the wave travels towards the most apical part. It was mentioned in section 3.6 that the model fails to approximate the travelling wave delay beyond the 400 Hz place (indicated by the dashed line block). This needs to be taken into account when considering the predicted spatio-temporal neural patterns in the rest of this chapter. Beyond (towards the apex) the place of maximum deflection (shown by the horizontal line) the predicted spatio-

temporal neural patterns in fig.s 5.1 and 5.7 show less synchronous firing owing to a decrease in travelling wave BM displacement.

Shamma (1985a) constructed spatio-temporal response patterns from large populations of auditory nerve fibre measurements in cat. These neural patterns show a synchronised response to the movement of the BM for a 1 kHz pure tone, where fibres are activated at a 1 kHz half cycle period (fig. 5.8). Another aspect that he mentions is the rapid accumulation of phase lag near the CF place equal to 1 kHz. This rapid phase lag results in the firing pattern of adjacent nerve fibres being out of phase close to the CF region of 1 kHz. This phase shift between the two nerve fibres near the 1 kHz CF region is clearly shown in the spatio-temporal response auditory nerve patterns in fig. 3 of (Shamma, 1985a)(fig. 5.8).

A similar phase shift is seen in the results for the model, where firing PST histograms were calculated for populations of 300 nerve fibres. The population firing times were taken to be the combined response for 300 fibres. The combined response was calculated by collapsing all the fibres' spike times onto one another. PST histograms results were then calculated for these combined responses by adding the spikes within a time bin period of 0.1 ms. Each plotted line on the graph in fig. 5.9 represents the combined firing times of 300 fibres. The x-axis shows the the firing time period, whereas the y-axis shows the length of that specific population from the base.

A definite phase shift can be seen as the place of maximum deflection at approximately 17.4 mm is reached. Observation of the phase shift between the two populations in this region shows the one population of discharges to be out of phase with the other. However, when comparing these results with the recorded responses of Shamma (1985a), the model results show a more gradual phase shift, as opposed to a very rapid phase shift between the response of two nerve fibres in Shamma's (1985a) results.

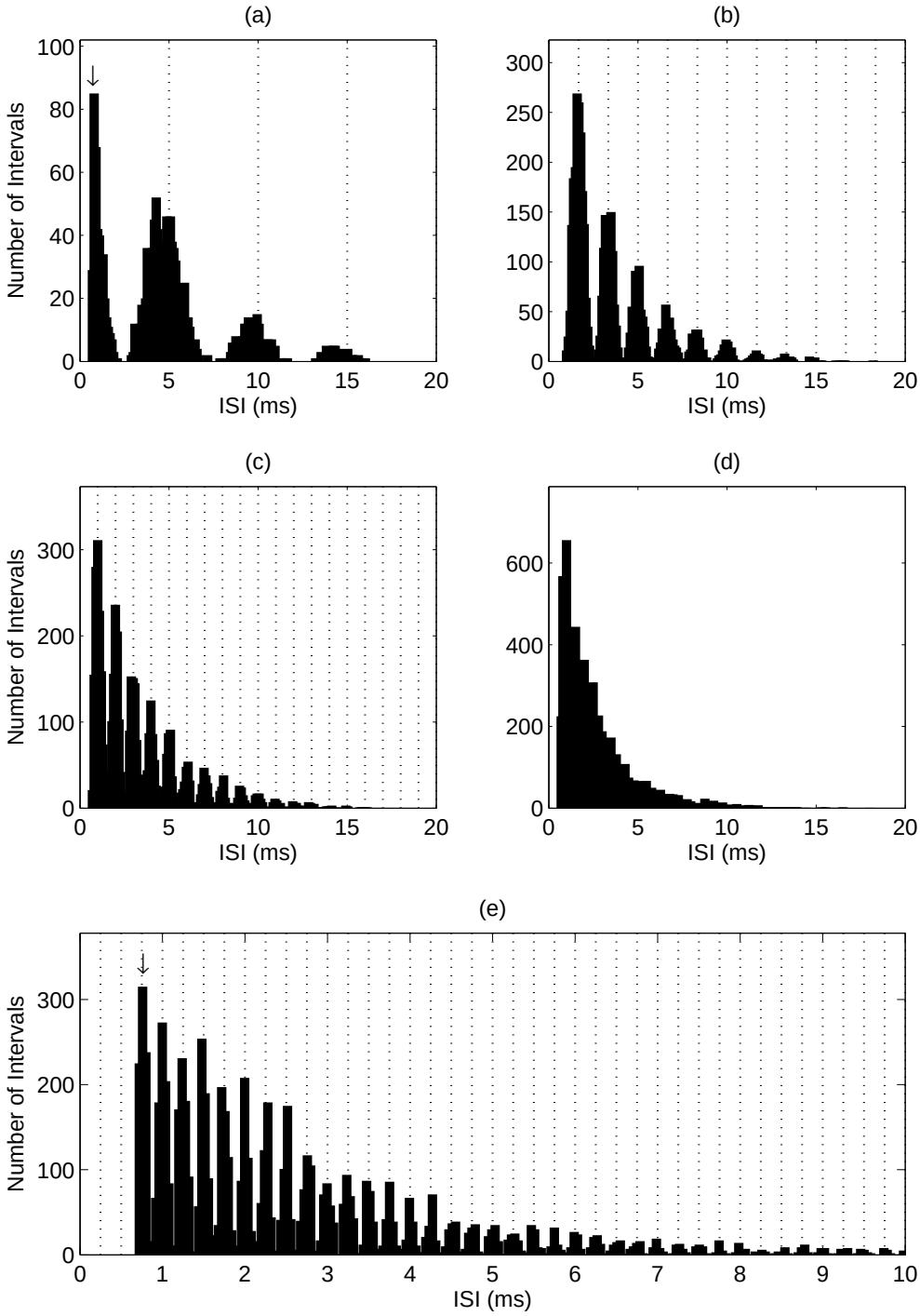

Figure 5.5: ISI histograms for the normal hearing model with a much lower firing probability of *p(spike)* = 0.02. These results show a definite clustering of spike intervals around multiples of the tone periods. The respective input frequencies are a) 200 Hz, b) 600 Hz, c) 1 kHz and d) 4 kHz. In the enlarged version for the 4 kHz input (e), the number of time bins was increased, with a time bin length of 0.04 ms.

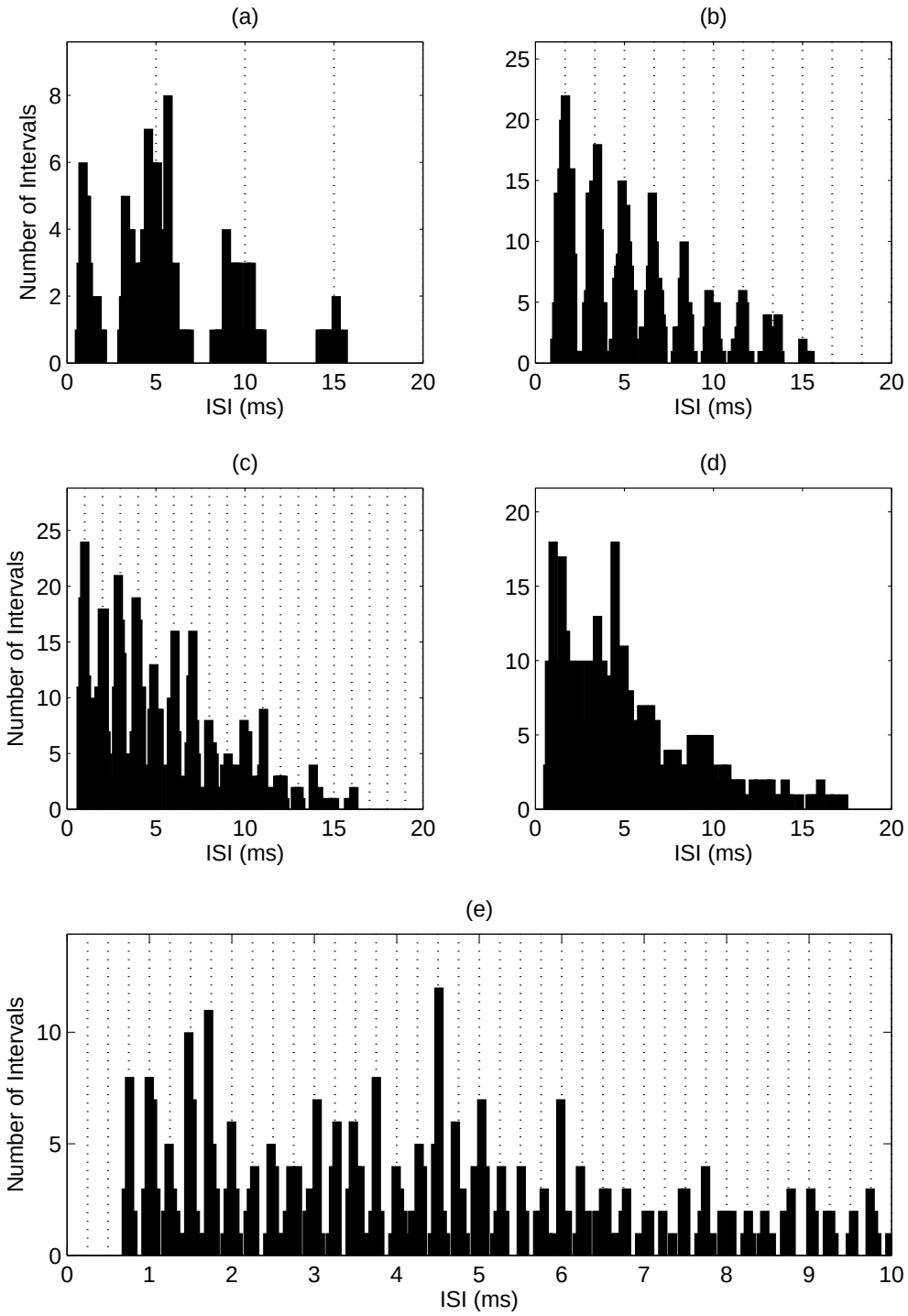

Figure 5.6: ISI histograms for the normal hearing cat model with a very low probability of *p(spike)* = 0.005. The respective input frequencies are a) 200 Hz, b) 600 Hz, c) 1 kHz and d) 4 kHz. In the enlarged version for the 4 kHz input (e), the number of time bins was increased, with a time bin length of 0.04 ms.

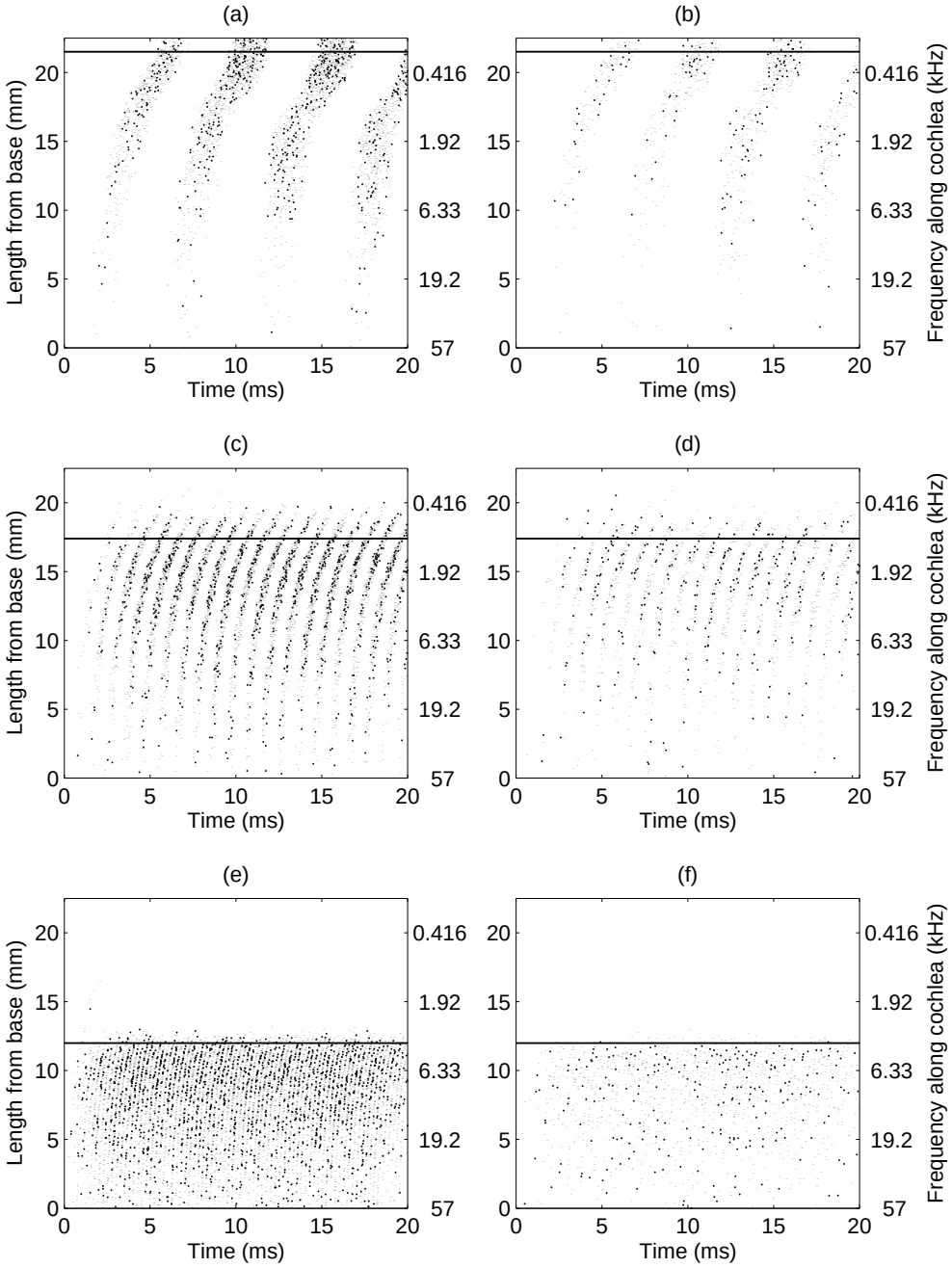

Figure 5.7: Spatio-temporal neural spike train patterns for the normal hearing model with a firing probability of *p(spike)* = 0.02, for a) 200 Hz, c) 1 kHz and e) 4 kHz frequency inputs. Response patterns for a probability of *p(spike)* = 0.005 is shown in b) - f) for 200 Hz, 1 kHz and 4 kHz frequency inputs. The horizontal line in each graph is the predicted place related to the input frequency according to Greenwood's (1990) equation.

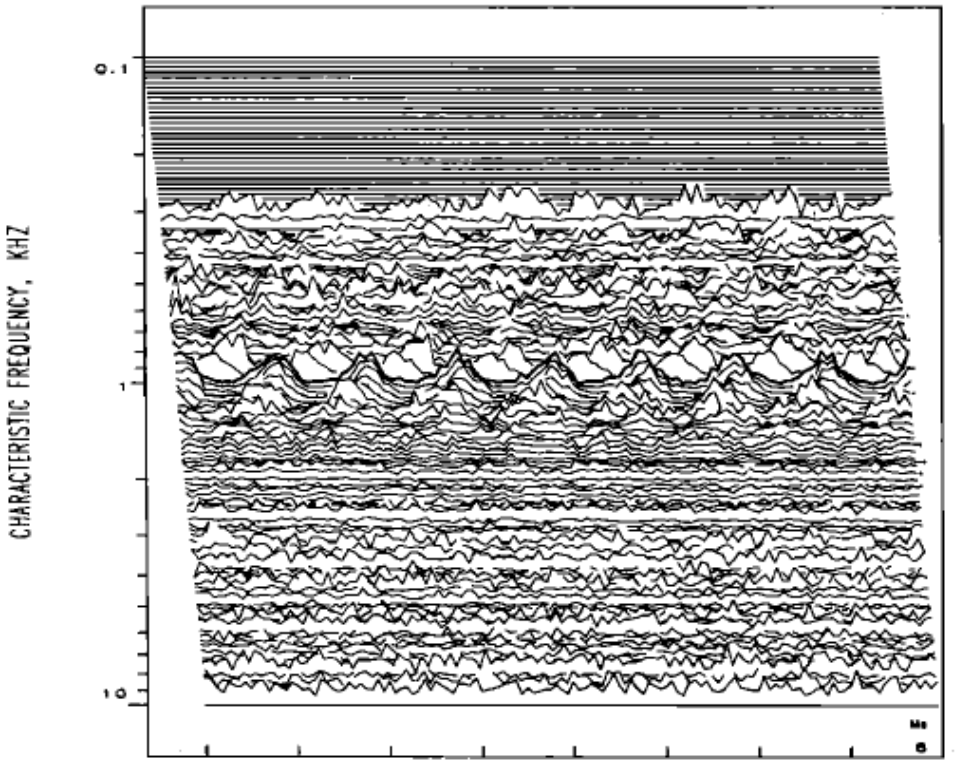

Figure 5.8: Spatio-temporal response auditory nerve patterns for a 1 kHz-noise stimulus input (reprinted with permission from Shamma (1985a)).

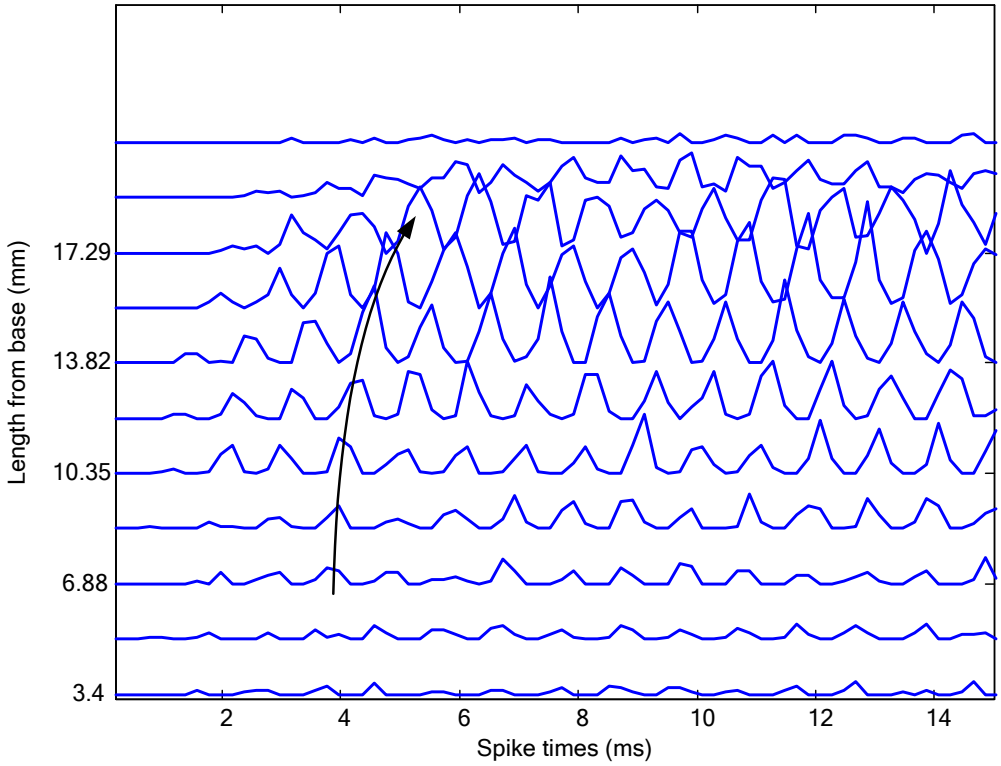

Figure 5.9: PST histogram results for populations of 300 nerve fibres with a 1 kHz pure tone input. The arrow indicates the increase of the phase shift for the populations of nerve fibres.

The results confirm the validity of the normal hearing model for its intended use, i.e. to predict neural spike train patterns in normal hearing that could be used in comparison with spike trains for electrical stimulation. In order to investigate the effect that the channel-number stimulation-rate trade-off has on predicted neural spike train patterns for a travelling wave processor, it is important to have predicted normal hearing spike train data to which these electrical stimulation spike train data could be compared.

The model parameters used in the travelling wave model have been approximated from cat measurements (sections 2.4 and 2.5) with a cochlear length of 22.5 mm. However, in the rest of this chapter the insertion depth and electrode spacing have been based upon the Nucleus device, which is implanted in the human cochlea with a BM length of 35 mm. In the rest of this chapter the travelling wave model output length of 22.5 mm has been scaled to a value

between 0 and 1 and transposed onto the length of 35 mm. This results in travelling wave output amplitudes being spread across 35 mm, with a different frequency-place mapping according to Greenwood's (1990) equation. When considering the results of the electrical stimulation in the rest of this chapter, this cochlear length scaling must be taken into account.

5.3 MODELLING OF NEURAL SPIKE TRAINS ELICITED BY ELECTRICAL STIMULATION

5.3.1 Spike trains predicted for electrical stimulation

The results of the previous section were for normal hearing. The researcher now proceeds to the predicted neural spike train patterns for electrical stimulation, with reference to a travelling wave processor.

Before the results for the electrical stimulation can be summarised, it is important to understand how the neural spike train patterns correspond to the electrical stimulation and the travelling wave model pulses. Fig. 5.10 shows the sampled travelling wave model displacement (thin lines) for a 1 kHz pure tone input in comparison with the CL unit stimulating pulses (bold lines) for a six-channel activation configuration (channels - 2, 6, 10, 14, 18, 22) with a stimulating rate of 2 400 pps per channel. The sampling process for n channel activation has been discussed in detail in chapter 4. The electrical stimulating pulses are represented in CL units, ranging between a threshold level of 100 and a comfort level of 225 (section 4.7). This relationship between the CL units and the current values is described by a similar equation to eq. 4.6. Fig. 5.11 shows this corresponding current sampling values in μA (bold lines). These values serve as input to the electrical stimulation model. The electrical model approximates the neural spike times across the channels along the cochlea. Each

of these predicted spikes is represented by a dot across time and space. Fig. 5.11 shows that these spikes appear as lines around the place and time of stimulus pulses. These lines are in fact a population of nerve fibres, each generating a spike at the same time instance. One of the assumptions made in Bruce *et al.*'s (1999a) model is that spikes can only occur during the cathodic phase of the current pulses. This assumption is clear from the spatio-temporal neural spike patterns (fig. 5.11 and 5.12), where spikes only occur during the cathodic phase of the stimulation pulses. Therefore no spikes are generated between stimulation pulses.

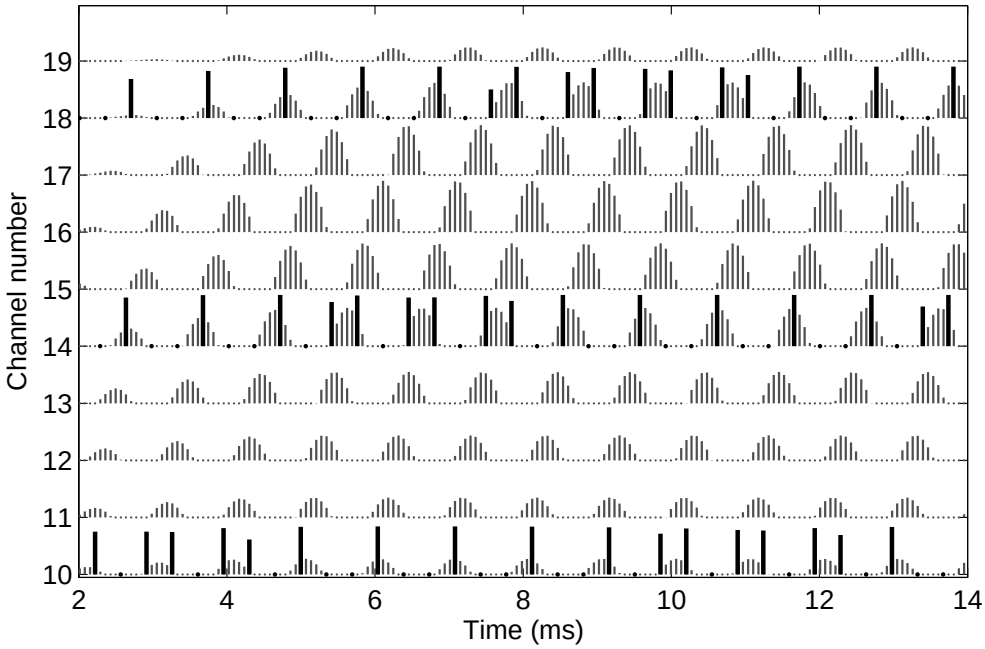

Figure 5.10: The thin lines represent the sampling displacement values of the TW model for a 1 kHz pure tone input. The values for electrode channels 10 - 18 are shown. The bold lines represent the stimulating pulses (in CL units) for a six-channel (2 ,6 10, 14, 18, 22) activation implementation. The reason for the stimulating pulses having similar amplitude values is the LGF function.

The lines of neural spikes also show a relatively large excitation spread, up to 4 mm for some pulses, for the chosen decay rate of 2.8 dB/mm (see section 4.10). For bipolar stimulation with a decay rate of around 7 - 8 dB/mm (Bingabr *et al.*, 2008; Bruce *et al.*, 1999a) the current spread will be much smaller. Because of this large current spread, the same population of nerve fibres is stimulated by adjacent electrodes. As a result of this current spread, the refractory period, which is the period where no or very few spikes will be generated after a firing has occurred, has a definite limiting effect on the firing patterns of adjacent neural channels. This can be seen in the neural spikes generated as a result of the third stimulation pulses for electrodes at 19 (marked as A in fig. 5.11) and 22 mm (marked as B). For the pulse (22 mm) occurring just after the pulse at 19 mm, there is almost no basal neural activation.

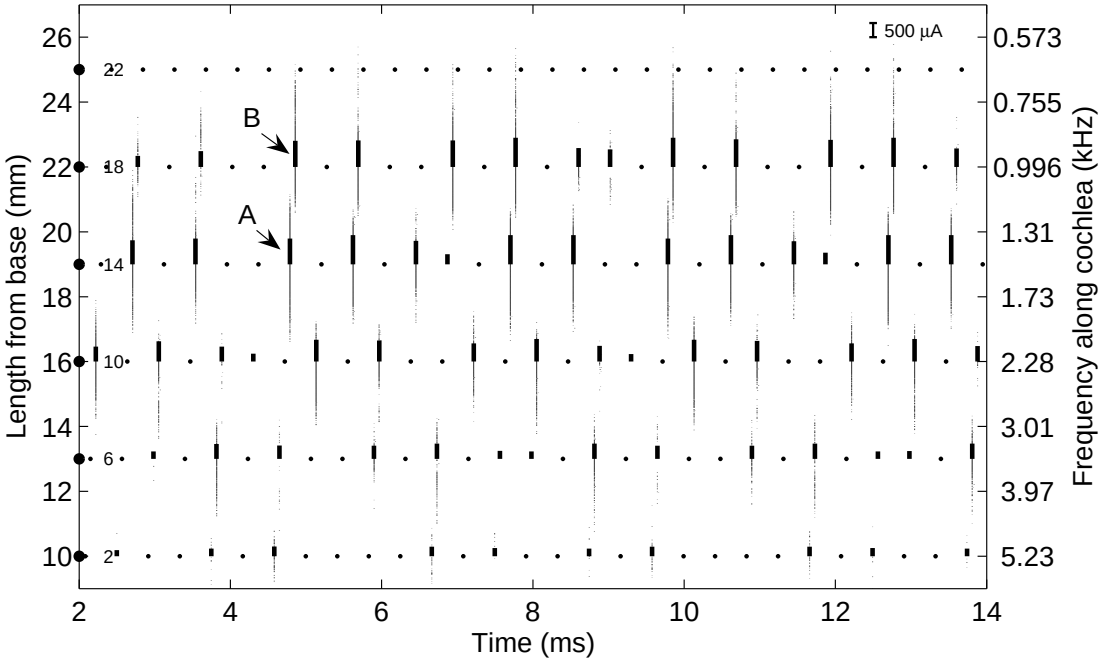

Figure 5.11: The bold lines represent the stimulation pulses from channels 2, 6, 10, 14 and 18 at a stimulation rate of 2 400 pps. The bold dots represent sample times for stimulation pulses with a CL of zero. The small dots represent the spikes generated along the cochlea as a result of these stimulating pulses. A and B are explained in the text.

Observation of a higher stimulation rate of 3 600 pps, with a period of 0.27 ms (fig. 5.12), shows that adjacent stimulating pulses (marked as A and B) spaced closer than the absolute refractory period, result in much less neural activation for the duration of the second stimulating pulse (marked as B).

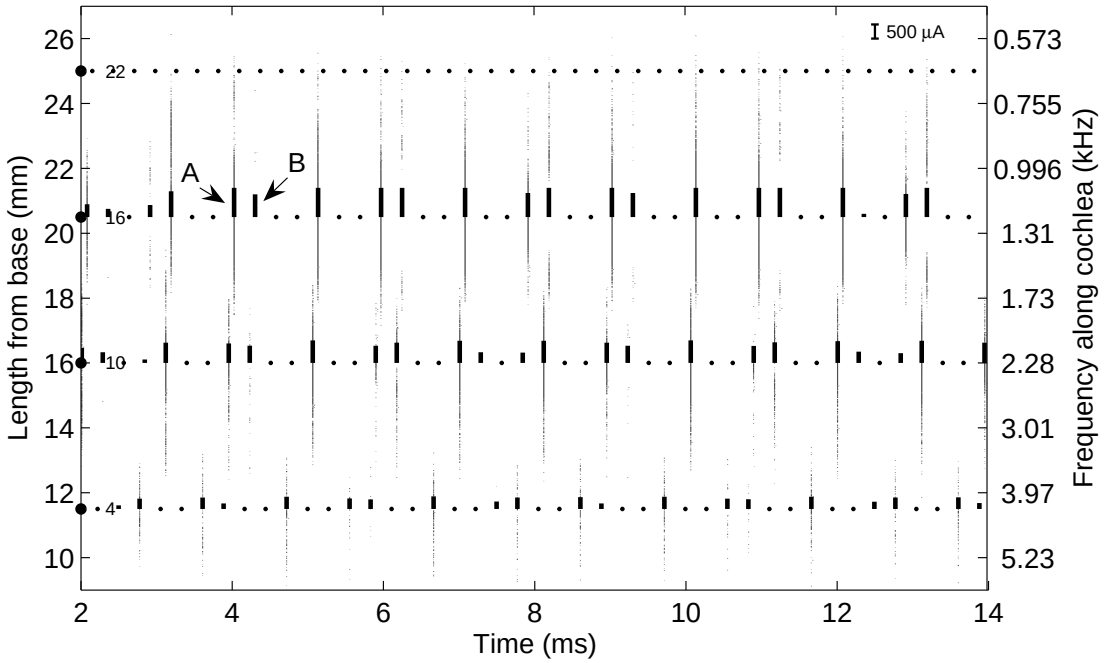

Figure 5.12: The higher stimulation rate (3600 pps) results in adjacent stimulating pulses (A and B) spaced more closely than the absolute refractory period. This results in much less neural activation for the duration of the second stimulating pulse (B).

In the next section the predicted neural spike train patterns for the ACE strategy is investigated.

5.3.2 Predicted spike trains for ACE processing

Next, simulated responses from the *n-m* processing strategy, ACE (Vandali *et al.*, 2000) are used as an example of a present processing strategy. This strategy was chosen for the

reason that a large number of implantees have been implanted with implants using it. These simulated responses are compared to the response patterns for the TW processing strategy (section 5.3.3). ACE introduces a set of FFT bins to produce filter bands, which are spaced across the frequency range. The power of adjacent bins is added together to estimate a total of *n* envelopes from which *m* envelope amplitudes are selected in each stimulation time frame (Loizou, 2006). The ACE strategy gives the user the ability to decrease the number of activated channels (*m*), employing a higher channel stimulation rate. The *m* channels are stimulated at the same rate. Predictions for a 200 Hz pure tone input fig. 5.13 (a) show populations of nerve fibres firing at the stimulation frequency of 600 pps. The ISI histogram (b) shows that ISIs occur at multiples of the *stimulation rate*. The neural spike train pattern for a low pure tone input of 200 Hz (a) and those for a higher frequency input of 1 kHz (c) show a similar temporal pattern, with firings occurring at the rate of channel stimulation (600 pps). This can also be seen in the calculated ISI histograms((b) and (d)), which show peaks at the multiples of the stimulation rate period of 1.67 ms. Spatially the 200 Hz frequency (a) is stimulated much more basally than the 1 kHz frequency, resulting in the stimulation of different neural populations along the length of the cochlea.

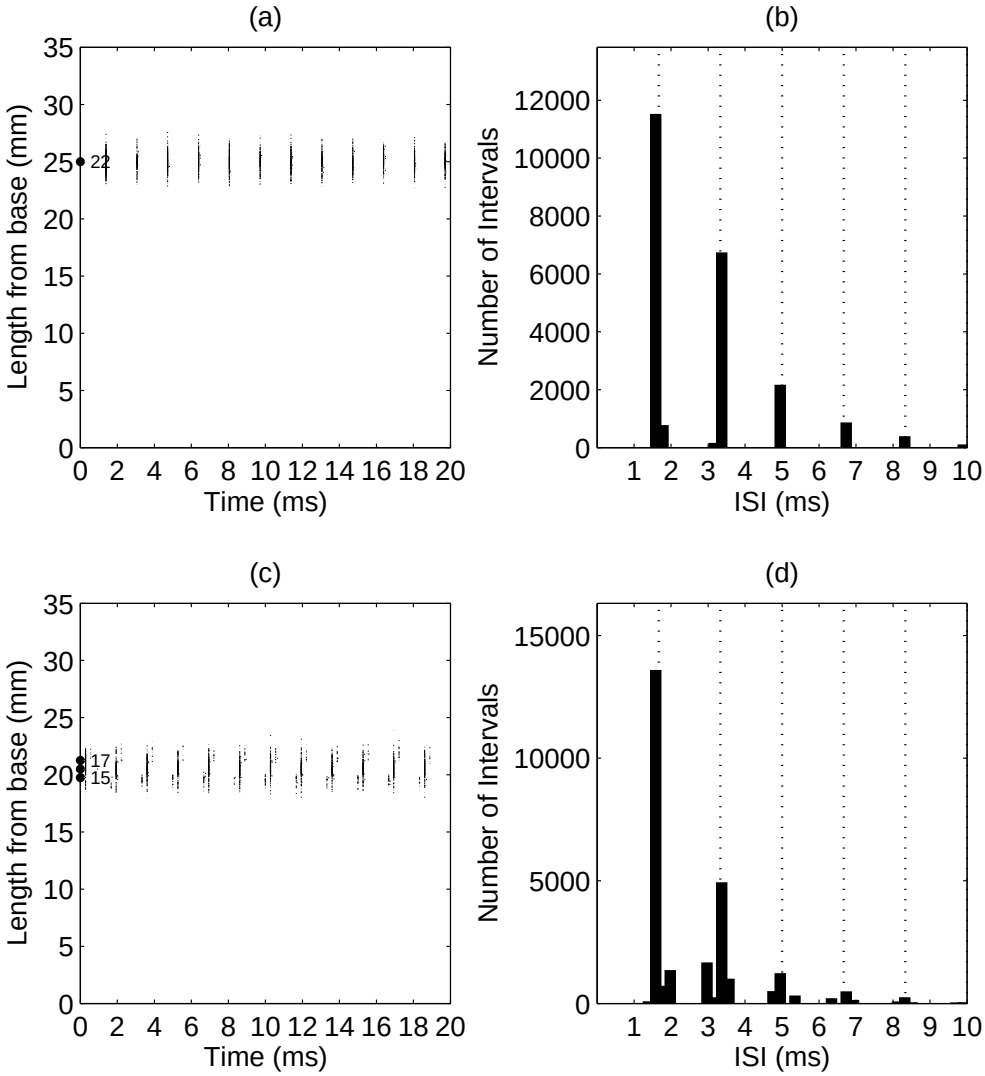

Figure 5.13: The predicted spike pattern results for the an ACE-like strategy implemented by Smith (2011). (a) Spatio-temporal pattern results for a 200 Hz pure tone. The activated electrode is shown on the left-hand y-axis with the electrode's corresponding electrode number, numbered in the order: electrode 1 is nearest to the base and electrode 22 is nearest to the apex. (b) ISI histogram results for a 200 Hz tone. (c) Spatio-temporal neural pattern results for a 1 000 Hz pure tone. Predicted stimulation pulses show activation on three electrodes, namely 15, 16 and 17. Only the first and the last activated electrode numbers are shown. (d) ISI histogram results for a 1 000 Hz tone.

5.3.3 Trade-off between number of electrodes and stimulation rate

5.3.3.1 Predicted spatio-temporal neural spike train patterns

When one implements the travelling wave in a processor algorithm, a trade-off remains between the channel stimulation rate and the number of activated channels. These trade-off limitations have been discussed in chapter 4. This section investigates this trade-off further, with reference to an electrical stimulation model.

Spatio-temporal neural spike train patterns for a 200 Hz pure tone input for various channel-number stimulation-rate configurations compared with spatio-temporal neural pattern results from the normal hearing model in section 4.12 are shown in fig. 5.14. A firing probability of 0.04 was used in the normal hearing model in this section. Electrode 22 has been placed 10 mm from the apex, corresponding to an electrode carrier insertion depth of 10 mm. With an electrode spacing of 0.75 mm, the first electrode's placement is calculated at 25.75 mm from the apex, or 9.25 mm from the base, for a cochlear length of 35 mm. The normal hearing model has been compressed within this range, corresponding to these electrode placements, resulting in a compressed neural spike pattern (a). In comparison to this result, (b) - (f) shows the results for the electrical stimulation where each spatio-temporal neural spike train's pattern is associated with a different channel-number stimulation-rate configuration. The channel-number stimulation-rate configurations are: (b) 22 activated channels (1-22) at 655 pps per channel , (c) 11 channels (2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22) at 1 309 pps per channel , (d) 8 channels (2, 5, 8, 11, 14, 17, 20, 22) at 1 800 pps, (e) 6 channels (2, 6, 10, 14, 18, 22) at 2 400 pps and (f) 4 channels (4, 10, 16, 22) at 3 600 pps . These spatio-temporal neural spike train patterns for these configurations were further simulated for tones 600 Hz, 1 and 4 kHz (figs 5.15, 5.16 and 5.17) to investigate the effect across a range of frequencies.

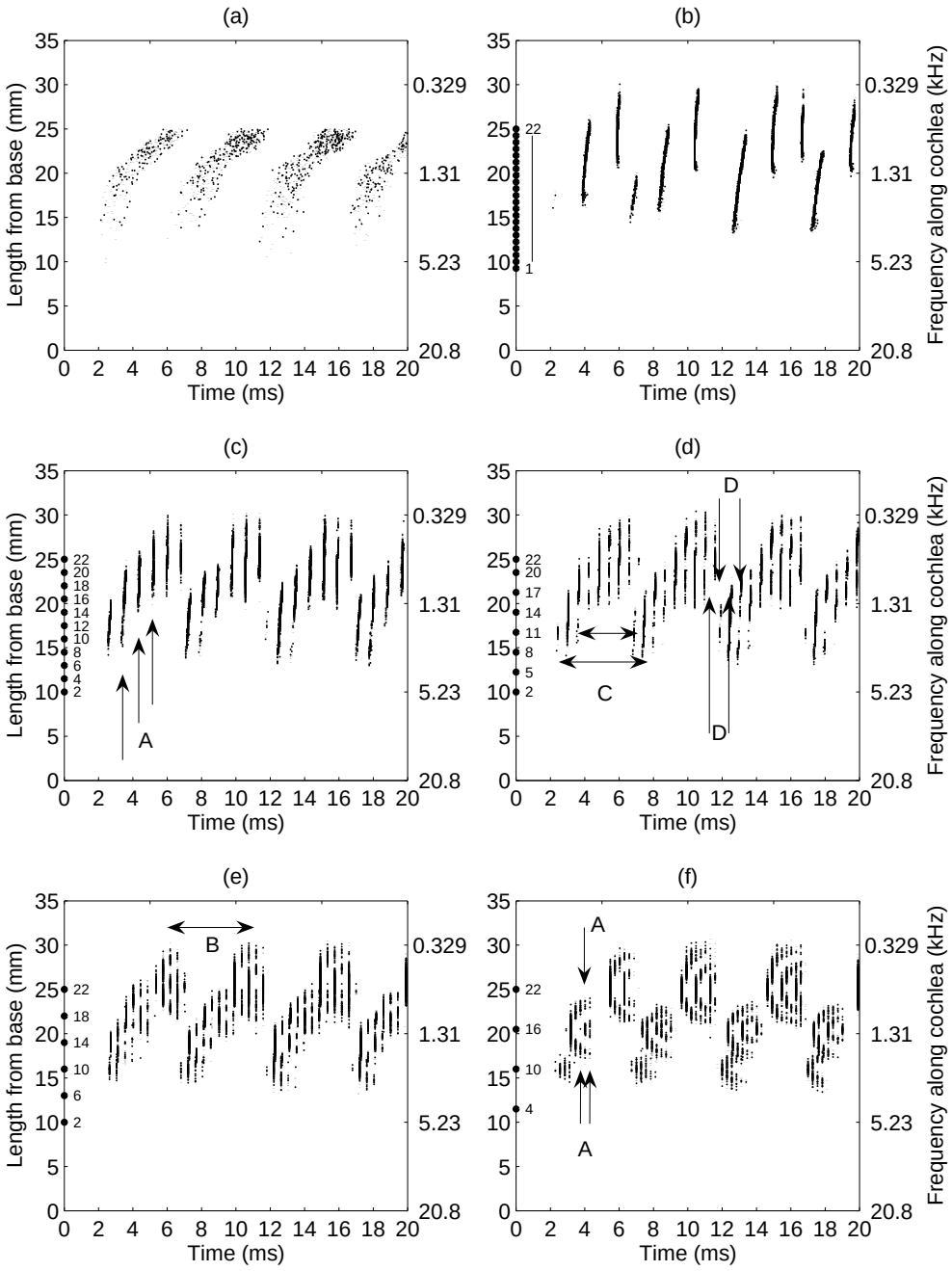

Figure 5.14: Predicted spatio-temporal neural spike train patterns for an input pure tone frequency of 200 Hz. The results show that of (a) the normal hearing model and channel-number stimulation-rate configurations: (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps. The activated electrodes along the length of the cochlea are shown as big dots on the left-hand y-axis with their corresponding electrode number. Where adjacent electrodes were activated (b), only their specific places are indicated, with the first and last electrode number in the electrode array indicated. A, B, C and D are explained in the text.

A first observation shows a clear difference between predicted electrical (fig. 5.14 (b) - (f)) and normal hearing neural spike train patterns (a). Predicted electrical neural spike train patterns show that these patterns occur closer toward the apex, compared to the 22nd electrode placement. This is as a result of current spread. Further observation shows that neural population activation occurs at the stimulation rate of the different configurations (shown by A, (c)). Even for stimulation rate periods ((d) - (f)) smaller than the absolute refractory period of 0.7 ms, populations of nerve fibres still seem to fire synchronously (shown by A, (f)). The results suggest that although the firing rate within a single nerve fibre is limited by the refractory period, the firing rate for a population of nerve fibres is not limited by this period.

Observation of a time window of 20 ms shows that clusters of neural channels seem to be activated at multiples of 5 ms (shown by B, (e)). Although this is less clear for neural pattern results at a low stimulation rate (b), multiple repetitions of neural activation remain around 5 ms intervals. Further inspection shows that as neural population positions move toward the apical region, the response of adjacent populations is progressively delayed. This delay becomes more evident with the 11-channel activation (c). A comparison between the electrical simulation for various configurations (apart from the 22-channel configuration) and the normal hearing simulation shows similarity between the spike phase delay patterns.

Spatio-temporal neural spike train patterns were also simulated for a tone frequency of 600 Hz (fig. 5.15). Predicted normal hearing neural patterns (a) show a travelling wave delay of a few milliseconds. This travelling wave delay, however, is not evident in the 22- (b) and 11-channel (c) configurations. Responses in (d) - (f) with less channel activation shows a neural population travelling wave phase shift toward the apical region (shown by the arrows). This is the result of increased electrode spacing and therefore less neural population interaction as a result of the current spread.

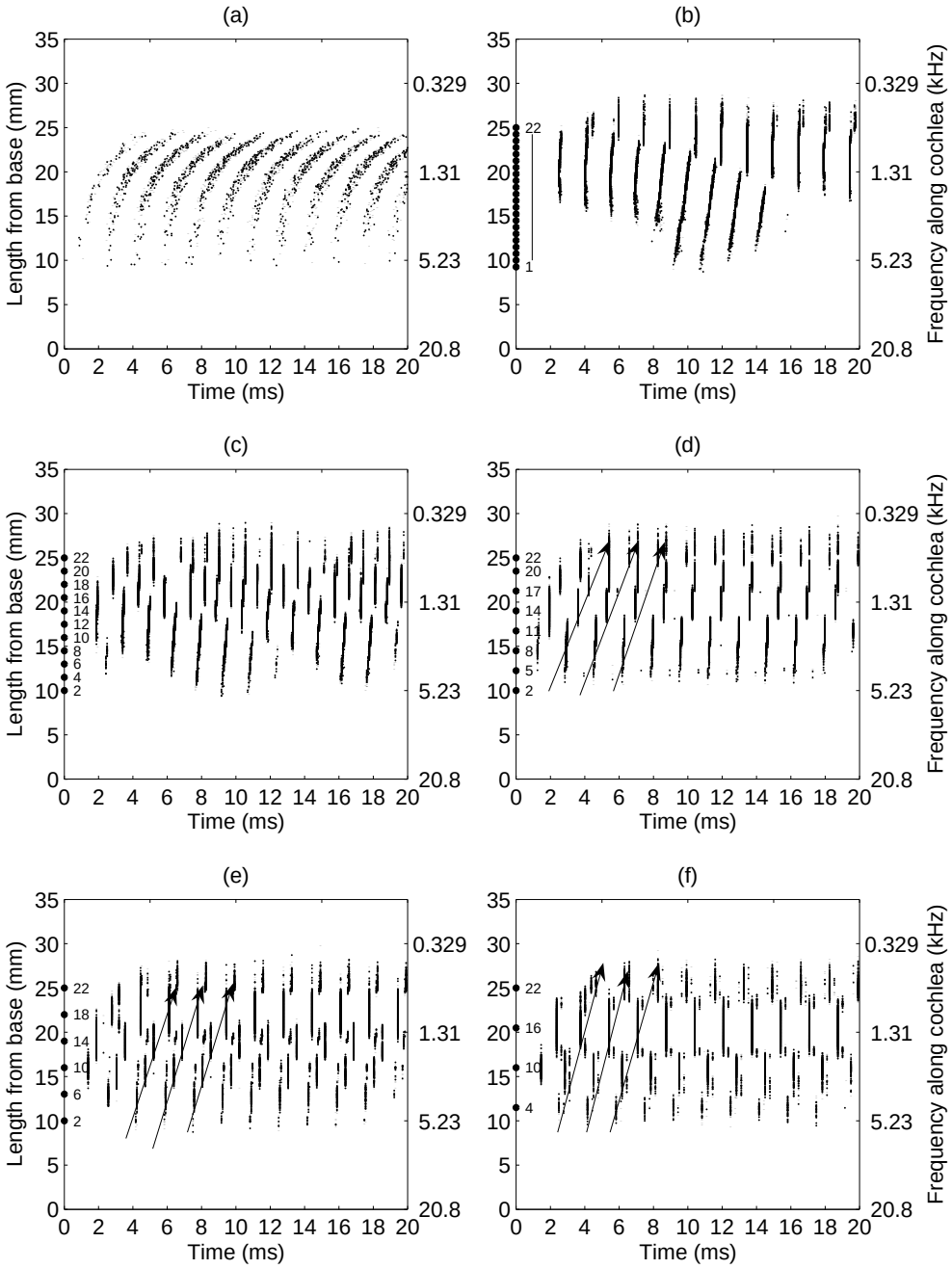

Figure 5.15: Predicted spatio-temporal neural spike train patterns for an input pure tone frequency of 600 Hz. The results show those of the normal hearing model(a) and channel-number stimulation-rate configurations: (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps. The activated electrodes along the length of the cochlea are shown as dots on the left hand y-axis with the corresponding electrode number.

With a higher input frequency of 1 kHz (fig. 5.16) neural pattern responses with lower stimulation rates, such as the 22-, 11- and 8-channel configurations, do not show a traveling wave delay, as seen in the predicted normal hearing neural patterns. This is as a result of spectral overlapping (shown by the arrows, (c)). According to Nyquist's rule, the sampling rate should be at least twice the frequency of the highest frequency component in the sampled signal, in this case 1 kHz. This is not the case for channel-number stimulation rate configurations shown in ((b) - (d)), where the stimulation rates are 655 pps, 1 309 pps and 1 800 pps. The neural response for a 6-channel activation (shown by the arrows,(e)) shows some travelling wave delay information contained in the responses. This, however, is very distorted owing to wide current spread. Therefore with this configuration the auditory system might only be able to extract some of the travelling wave delay information. The 4-channel configuration results reveal synchronous neural firing with a period of 1 ms (shown by arrows at A, (f)), with very little (shown by the arrows, (f)) travelling wave delay information contained in the neural pattern responses.

At high frequencies the travelling wave delay is much smaller and the spread across the length of the cochlea narrower, as the peak for maximum deflection is spaced closer to the base. The predicted spatio-temporal neural spike train patterns for a high input frequency tone of 4 kHz, fig. 5.17 (b) - (d), do not show any synchronous firing at the tone frequency because of the sampling rates (655 pps, 1 800 pps, 2 400 pps and 3 600 pps), which are too small, according to the Nyquist criteria, to sample the input tone frequency of 4 kHz adequately. This will result in spectral overlapping.

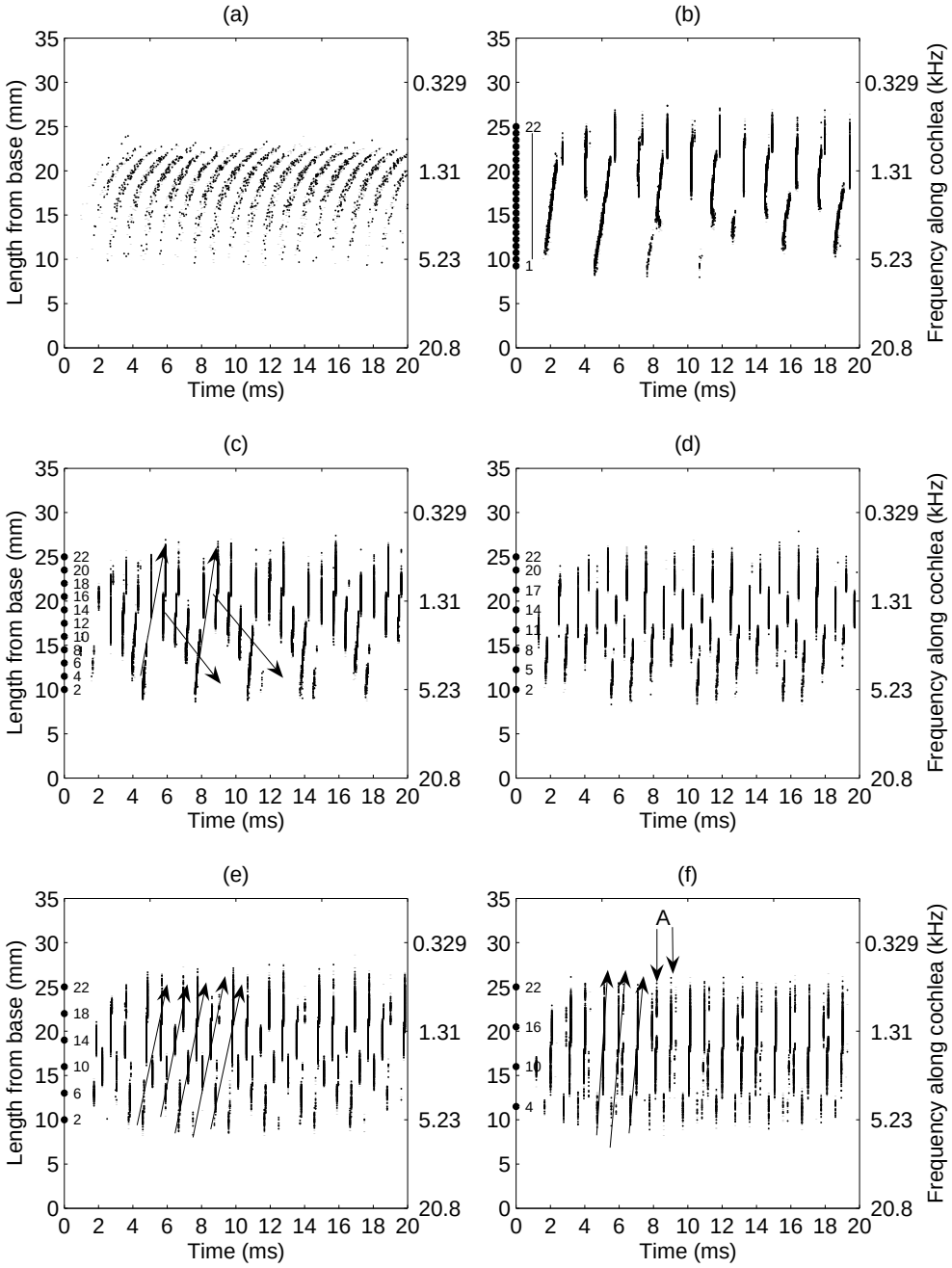

Figure 5.16: Predicted spatio-temporal neural spike train patterns for an input pure tone frequency of 1 kHz. The results show those of the normal hearing model(a) and channel-number stimulation-rate configurations: (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps. The activated electrodes along the length of the cochlea are shown as dots on the left-hand y-axis with the corresponding electrode number. A is explained in the text.

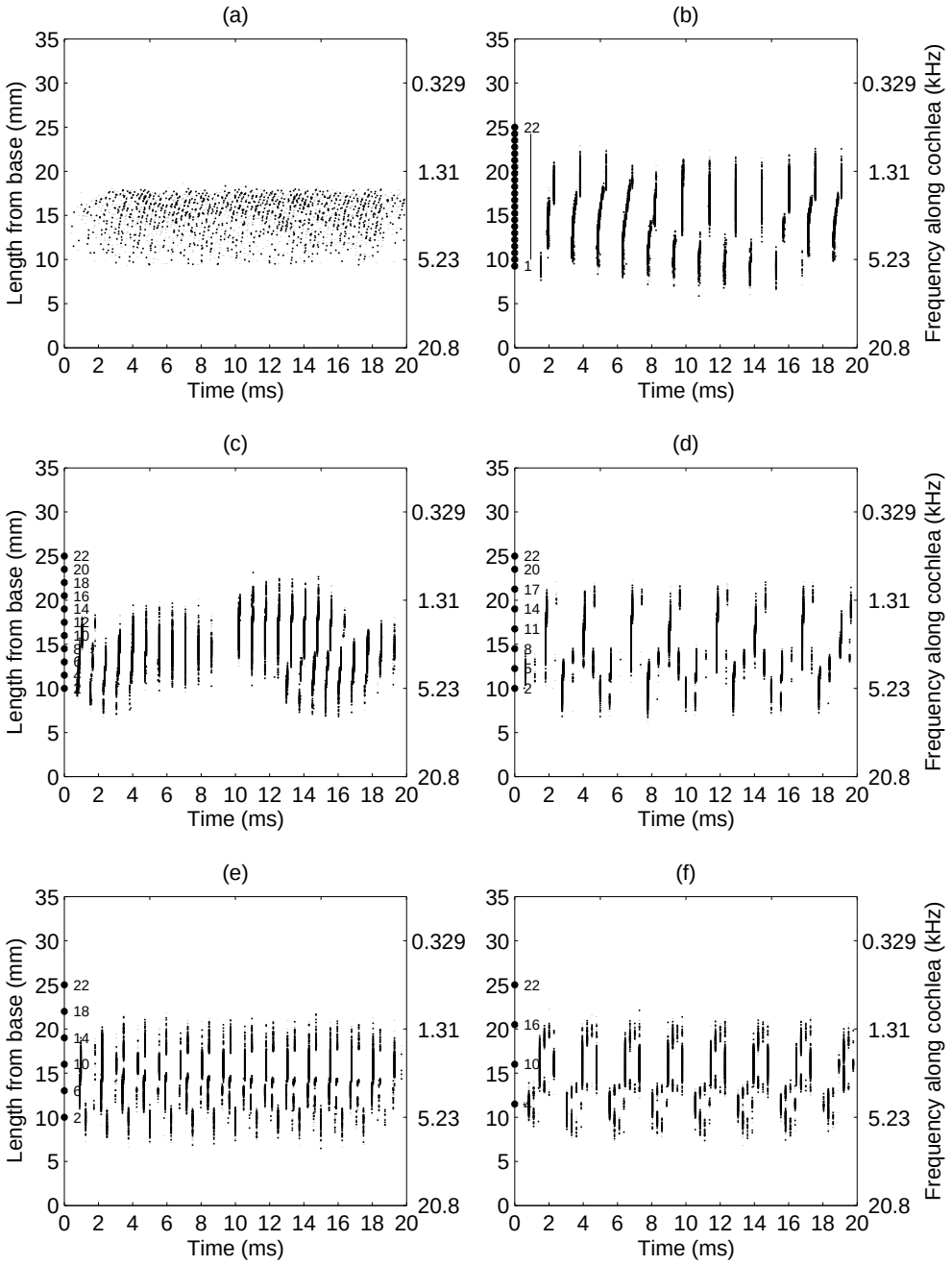

Figure 5.17: Predicted spatio-temporal neural spike train patterns for an input pure tone frequency of 4 kHz. The results show those of the normal hearing model(a) and channel-number, stimulation-rate configurations: (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps. The activated electrodes along the length of the cochlea are shown as dots on the left-hand y-axis with the corresponding electrode number.

5.3.3.2 ISI histograms for predicted neural spike train patterns

Observation of the predicted ISI histogram results (fig. 5.18) for the 200 Hz pure tone reveals rather different results for the electrical stimulation compared to the predicted ISI histograms for normal hearing. It is significant that the normal hearing ISI histogram shows clusters of intervals *around the input tone period*. These peaks are not seen in (b). Rather, peaks at the stimulation rate period of 1.5 ms (shown by A, (b)) and multiples thereof can be seen (shown by B and C, (b)). However, a peak remains around 5 ms (shown by B, (b)), corresponding to the cycle period of the 200 Hz signal. The 11-channel configuration (c) shows a cluster of peaks around 0.76 ms (shown by D), which corresponds to the stimulation rate period. This relation to a multiple of the stimulation rate can also be seen in (d), with a peak at 1.1 ms = 0.55 ms $\times$ 2 (shown by D), (e) with a peak at 0.83 = 0.416 ms $\times$2 (shown by D) and (f) at 0.83 ms = 0.27 ms $\times$ 3 (shown by D). These peaks occur at multiples of the stimulation rate because the input tone periods are smaller than the refractory period. Therefore these peaks relate to spikes firing after the refractory period. Smaller peaks also remain, in (c) - (f) between 1 and 2.5 ms, which correspond mainly to the intervals between upward BM movement (every 5 ms) (fig 17.9 in Javel (1990)), where the travelling phase shift rapidly increases (shown by the distance between the arrows at D, fig. 5.14). The peaks between 2.5 ms and 5 ms mainly constitute the intervals corresponding to the upward movement of the BM when the travelling wave phase slowly changes (shown in C, fig. 5.14). These clusters of peaks evident around 5 ms can be seen for the 8-, 6- and 4-channel configurations, and even more so for the 4 - 3 600 pps channel-number stimulation-rate configuration. This is consistent with the normal hearing model, where clusters of peaks occur at multiples of the tone cycle period.

ISI histogram results for a 600 Hz pure tone (fig. 5.19), show one main peak corresponding to the tone period for all the configurations, except (b). The 22-channel configuration with

a stimulation rate of 655 pps shows a peak corresponding to that stimulation rate (shown by A, (b)). Distinctive peaks for the 8-, 6- and 4-channel configurations, with *stimulation rates* between 1 800 pps and 3 600 pps ((d)- (f)), show that spikes are generated at intervals of the *tone period* of 1.66 ms (shown by B). A further observation also shows a smaller peak at 3.3 ms, a *multiple of the tone period* (shown by C). This shows that spikes also occur two tone cycles apart, which is consistent with normal hearing simulation results (fig. 5.19 (a)) where spikes occur multiple tone cycles apart.

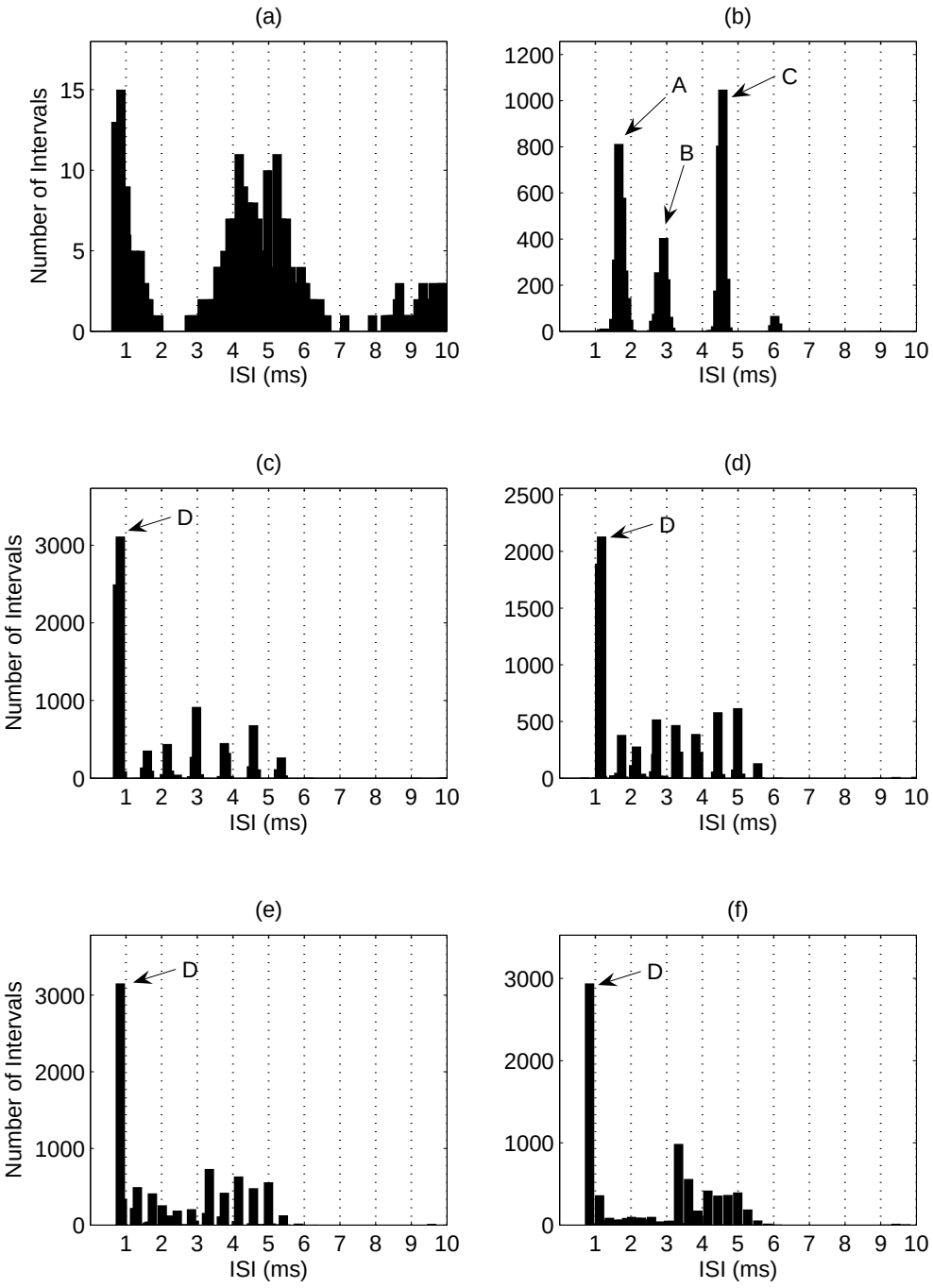

Figure 5.18: ISI histogram results that correspond to the predicted spatio-temporal neural patterns for a pure tone frequency of 200 Hz. Results are shown for the normal hearing model (a) and channel-number, stimulation-rate configurations: (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps. A, B, C and D are explained in the text.

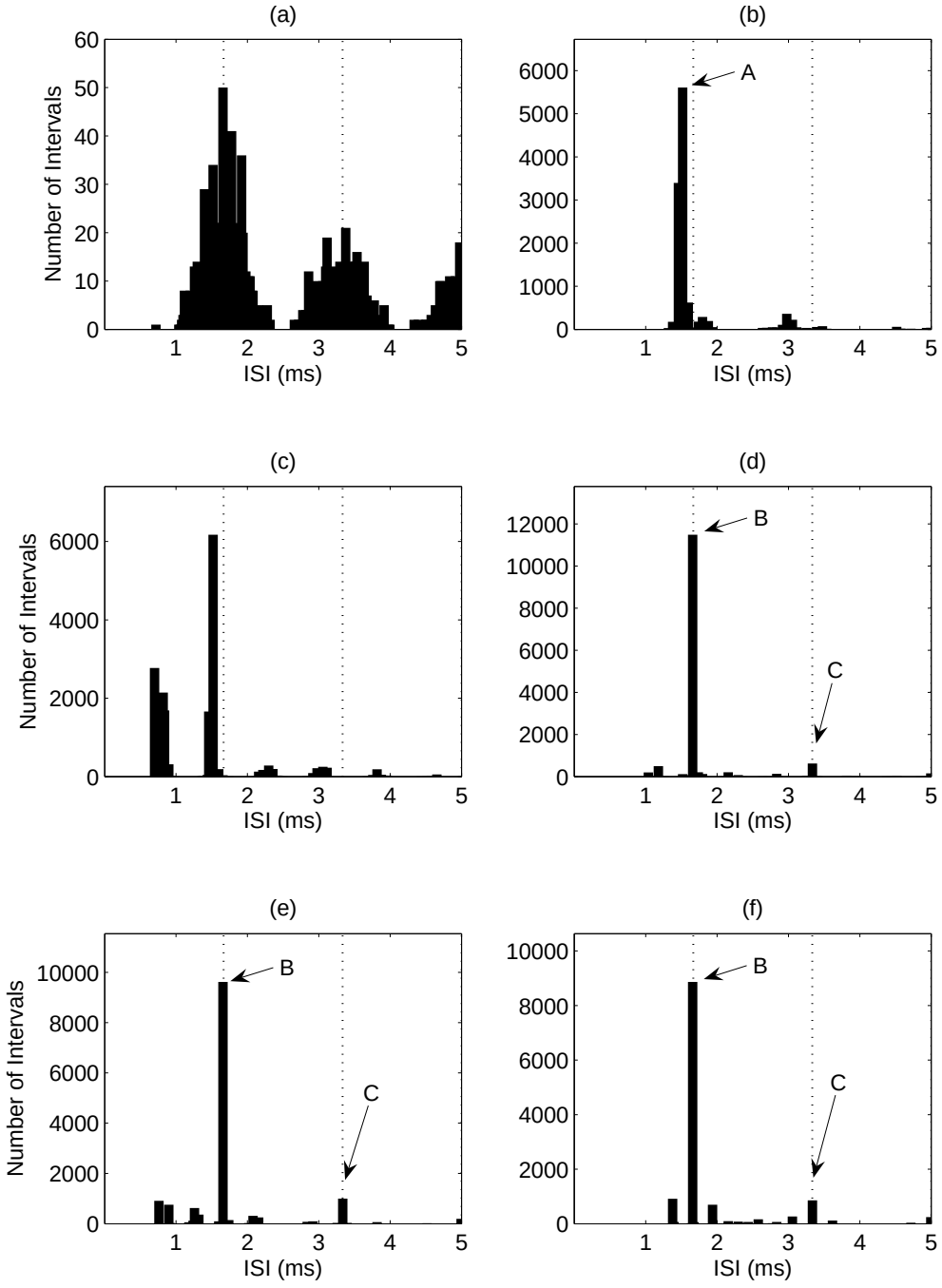

Figure 5.19: ISI histogram results that correspond to the predicted spatio-temporal neural patterns for a pure tone frequency of 600 Hz. Results are shown for the normal hearing model (a) and channel-number, stimulation-rate configurations: (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps. A shows a peak around the stimulation rate of 655 pps, whereas B and C show peaks around the input tone period (B) and multiples thereof (C).

Fig. 5.20 and 5.21 show the ISI histogram results of the higher tone frequencies, 1 kHz and 4 kHz, respectively. Electrical stimulation results for the 1 kHz tone (fig. 5.20) show that spikes are generated at multiples of the various stimulation rates. The stimulation rates: 665, 1 309 for the 22 (b), 11 (c) channel configurations are too low to sample the tone frequency of 1 kHz accurately. This results in histogram peaks, at multiples of the two configurations' stimulation rates. Results in (d) for a stimulation rate of 1 800 pps show a main peak at 1.1 ms (shown by A), double the stimulation period. This peak corresponds to the tone period of 1 ms. Therefore most of the spikes are generated on the first tone cycle. However, a stimulation rate of 1 800 pps, according to Nyquist, is not adequate to sample the tone period of 1 kHz properly. Higher stimulation rates generate more pulse samples per tone cycle, owing to better temporal representation of the tone frequency. This is evident in the histogram results for the 6- and 4-channel configurations, (e) and (f), with stimulation rates at 2 400 and 3 600 pps. Both of these configurations show two peaks around 1 ms (shown by A) corresponding to multiples of their various stimulation rates. These results also show peaks around 2 and 3 ms (shown by B), corresponding to predicted normal hearing ISI histogram (a).

Observation of the predicted normal hearing ISI histogram for the tone frequency of 4 kHz (fig. 5.21 (a)) reveals that there are clusters of spike intervals at the tone period which are much smaller than the absolute refractory period. Predicted electrical stimulation ISI histogram results (fig. 5.21 (b) - (f)) are quite different from the predicted normal hearing ISI histogram. One of the reasons for this is that the stimulation rates for the 22-, 11-, 8-, 6-, 4-channel configuration ((b) - (f)) under-sample the tone frequency of 4 kHz. The ISI histogram results show that spike intervals occur at multiples of the various stimulation rates (shown by A,(b) - (f)). The 4-channel configuration, with stimulation rate 3 600 pps (f), shows spike intervals occurring at multiples of 0.27 ms. These spike peaks of 0.27 ms correspond to the stimulation period of 3 600 pps, which is close to the input frequency of 4 kHz. These spikes, however, do not occur at every stimulation period, but at multiples of

this period because the refractory period is longer than the stimulation period.

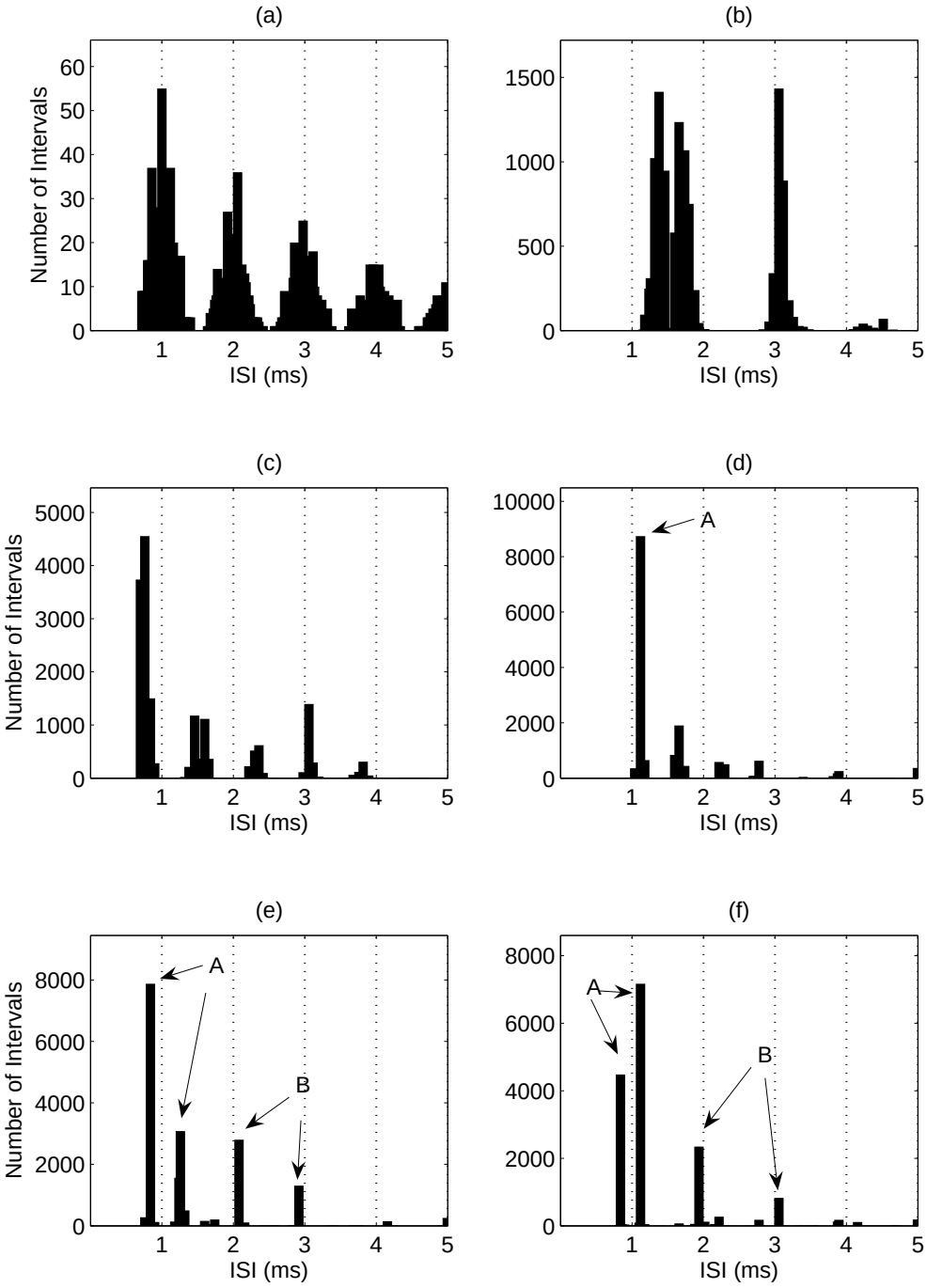

Figure 5.20: ISI histogram results that correspond to the predicted spatio-temporal neural patterns for a pure tone frequency of 1 kHz. Results are shown for the normal hearing model (a) and channel-number, stimulation-rate configurations: (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps. A shows a peak around the input tone period of 1 ms and B shows peaks around multiples of the input tone period (2 and 3 ms).

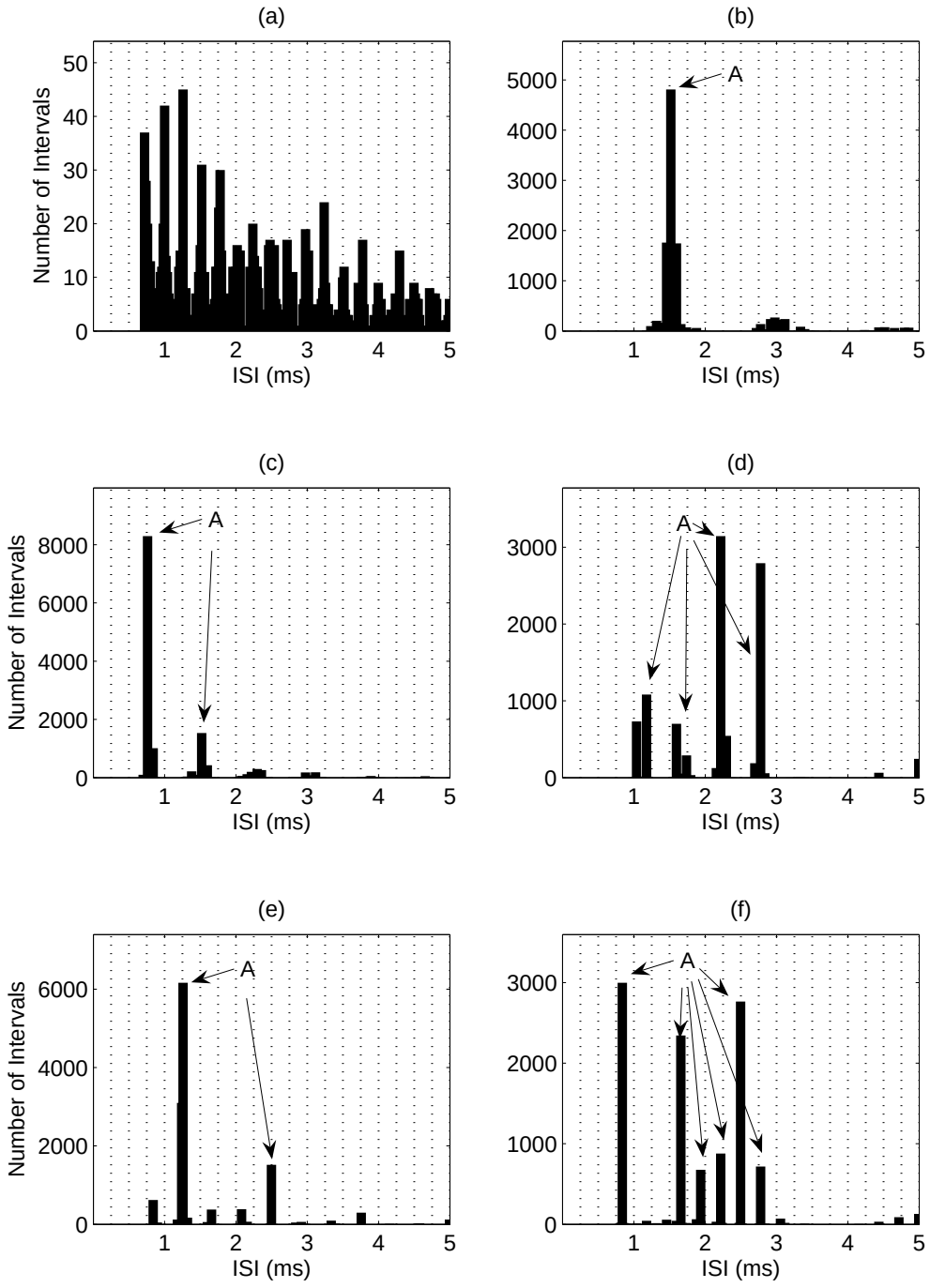

Figure 5.21: ISI histogram results that correspond to the spatio-temporal results for a pure tone frequency of 4 kHz. Results are shown for the (a) normal hearing model and the (b) 22 - 655 pps, (c) 11 - 1 309 pps, (d) 8 - 1 800 pps, (e) 6 - 2 400 pps and (f) 4 - 3 600 pps electrode-channel stimulation-rate configurations. A shows peaks at the various stimulation rates and multiples thereof.

Therefore to investigate the ISI histogram for the 4 kHz input tone, stimulated at an adequate stimulation rate, it was necessary to investigate a 2-channel configuration (fig. 5.22), with a channel stimulation rate of 7 200 pps. This configuration shows populations of nerve fibres (a) firing at a period much closer to the tone period of 0.25 ms. Fig. 5.22 (e) shows the PST histogram with a time bin of 0.04 ms. Results from the PST histogram show that the population spike rate is much closer to the tone period than results obtained with a 4-channel configuration (b). ISI results show peaks close to multiples of the tone period of 0.25 ms, indicated by the gridlines.

The predicted spatio-temporal neural spike train pattern and the corresponding ISI histograms suggest that when a travelling wave processor is implemented, fewer active channels with a higher stimulation rate might produce neural spike train patterns that more closely approximate those seen in normal hearing. In a travelling wave type processor, which samples the envelope of the pure tone, spectral overlapping might be avoided by sampling channels at least twice the stimulation rate.

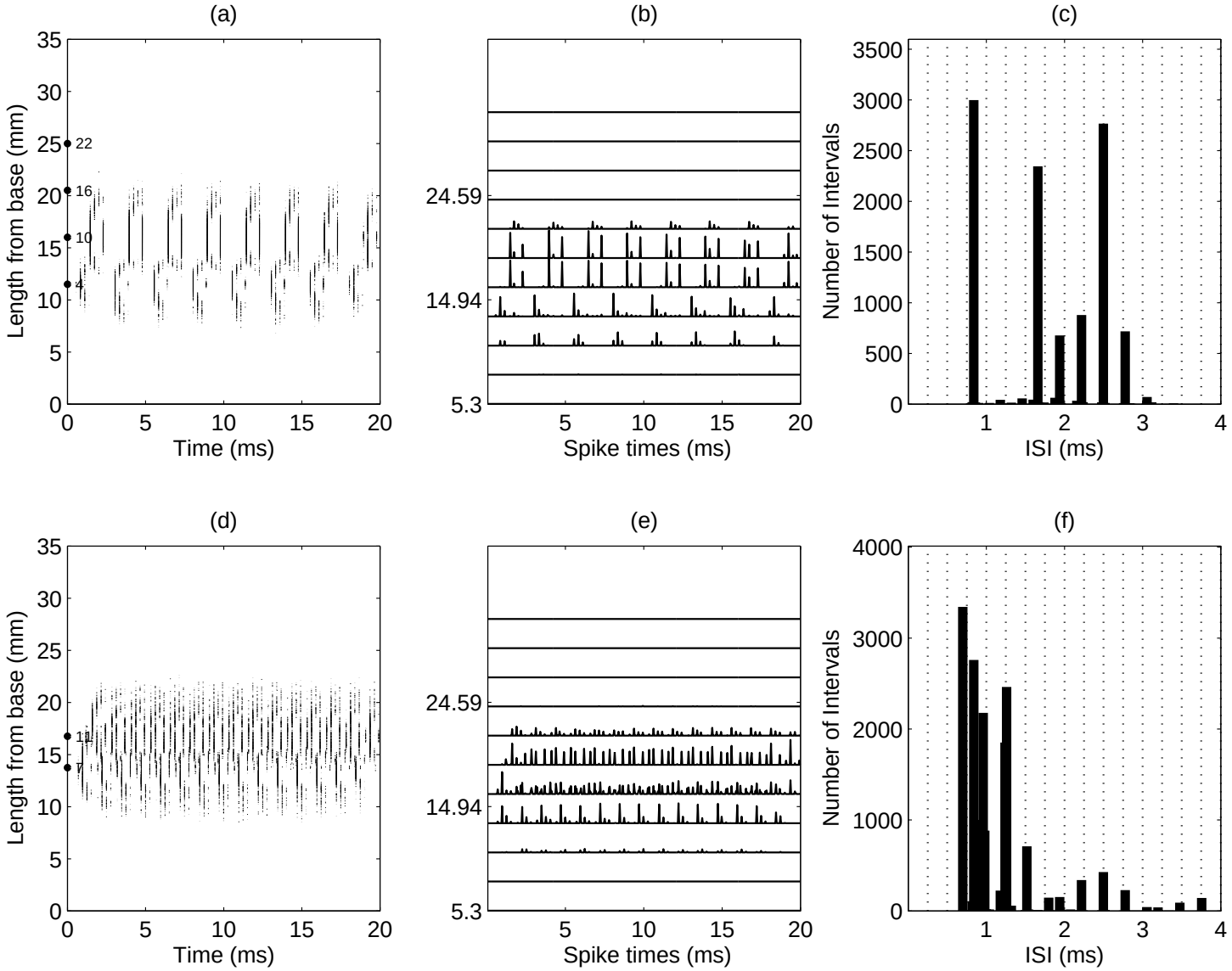

Figure 5.22: Results for two configurations, (top) 4 and (bottom) 2 channels (electrode channel 7 and 11) at 7 200 pps for a 4 kHz pure tone input. The results show the comparison between the spatio-temporal patterns (a) and (d), the PST histogram (b) and (e), and the ISI histogram (c) and (f).

5.3.3.3 ISI histograms for critical bands of predicted neural spike train patterns

As the auditory system views the basilar membrane input through a critical band window, it is insightful to consider temporal and spatial information regarding the travelling wave, available in different critical bands. This section investigates predicted ISI histograms for critical bands of the predicted neural spike patterns for the input tone frequency of 200 Hz. The critical bands are considered along the length of the cochlea for the area of neural activation. (Zwicker and Fastl, 1999:162) determined the relationship between critical-band rate and the distance along the BM as 1 bark being equal to 1.3 mm.

Figs. 5.23 - 5.26 show the calculated ISI histogram data for the predicted spatio-temporal patterns for the channel-number stimulation-rate configurations: 8 - 1 800 pps (a,b), 6 - 2 400 pps (c,d) and 4 - 3 600 pps (e,f) at a tone input of 200 Hz. The same electrode channel configurations as those in the previous sections were used. Fig. 5.23 shows the calculated ISI histogram data for the predicted spatio-temporal patterns in the critical band between 15 and 16.3 mm. The results for the three channel-number stimulation-rate configurations show a large peak at the tone period of 5 ms. These peaks are much larger than the peaks seen in the ISI histograms, calculated for the whole BM length (section 5.3.3.2). These show a much closer approximation to the ISI histogram results for predicted normal hearing responses. Similar peaks at the tone period are observed for the predicted ISI histograms (fig. 5.24) for four critical bands between 14.8 and 20 mm. Although much smaller, these peaks can also be seen in the predicted ISI histograms (fig. 5.26) for four critical bands between 25.2 and 30.4 mm. Predicted spatio-temporal neural spike patterns (fig. 5.25, (a), (c) and (e)) show a distortion in the area around the place of maximum deflection where the travelling wave delay increases rapidly. Predicted ISI histograms for the four critical bands in this region between 20 and 25.2 mm (fig. 5.25) show smaller peaks with peaks shifted to the right of 5 ms, as a result of this distortion. This temporal and spatial distortion of the travelling

wave delay in this region might be attributed to the significant current spread of adjacent electrodes that decrease the threshold of adjacent populations of nerve fibres, resulting in increased firing.

The ISI histograms results suggest that predicted neural spike train patterns in certain critical bands contain more temporal information than those in other critical bands. The predicted ISI histograms for a low frequency input suggest that in the region around the place of BM maximum deflection, there is more distortion of the input tone temporal pattern. If the auditory system considers separate critical bands, critical bands that are further away from the place of maximum BM deflection might contain more temporal information about the low frequency pure tone.

The next section investigates the effect on predicted neural spike trains, when electrodes are spaced more closely to one another around the area of maximum deflection. Along with this, wider electrode spacing with a deeper insertion depth will also be investigated.

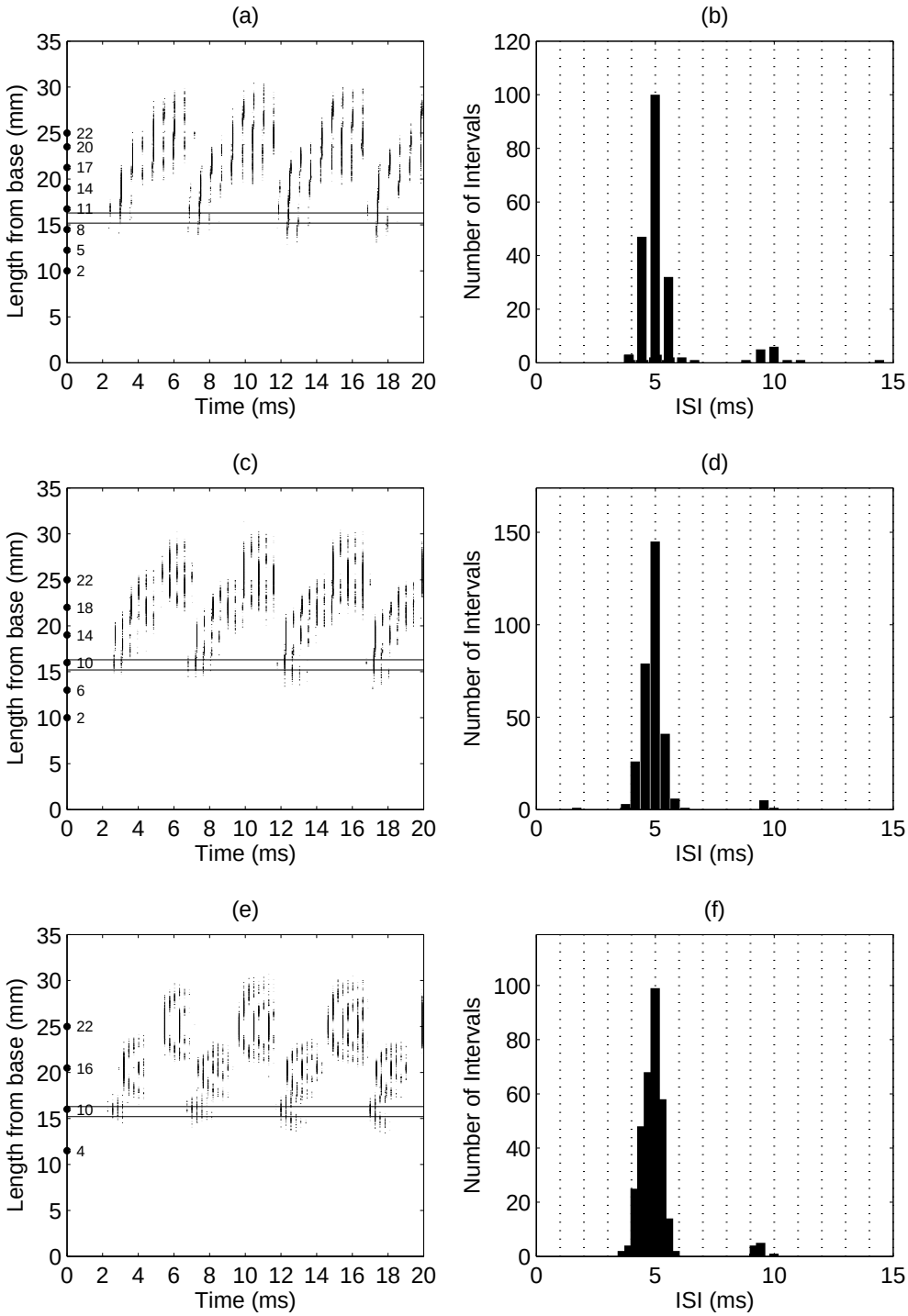

Figure 5.23: The predicted spatio-temporal neural spike train patterns for the channel-number, stimulation-rate configurations: 8 - 1 800 pps (a), 6 - 2 400 pps (c) and 4 - 3 600 pps (e), for an input tone of 200 Hz are shown. Their corresponding ISI histograms for the critical band between 15 and 16.3 mm (indicated by the two lines in (a), (c) and (e)) are shown in (b), (d) and (f).

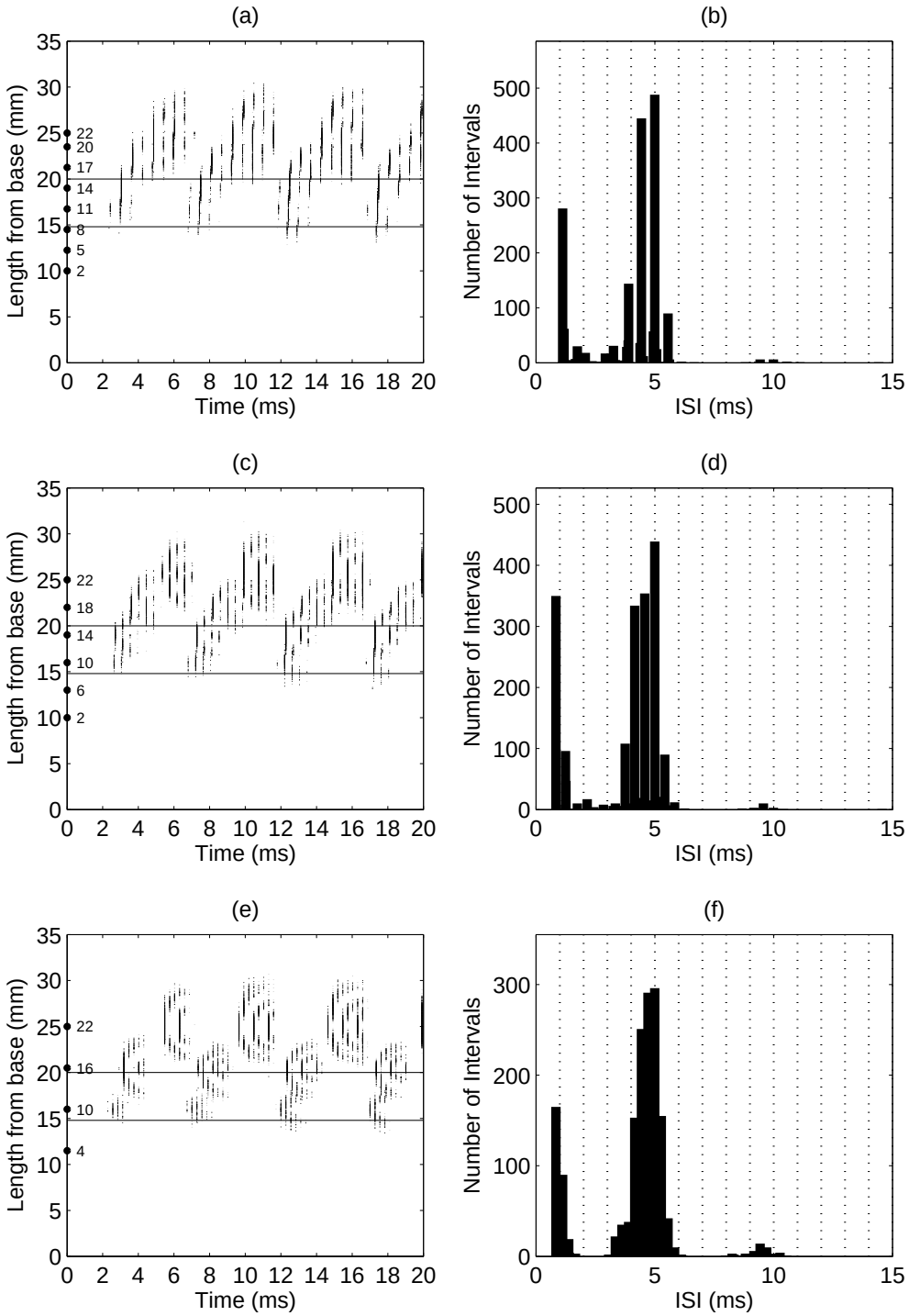

Figure 5.24: The predicted spatio-temporal neural spike train patterns for the channel-number stimulation-rate configurations: 8 - 1 800 pps (a), 6 - 2 400 pps (c) and 4 - 3 600 pps (e), for an input tone of 200 Hz are shown. Their corresponding ISI histograms for the four critical bands between 14.8 and 20 mm (indicated by the two lines in (a), (c) and (e)) are shown in (b), (d) and (f).

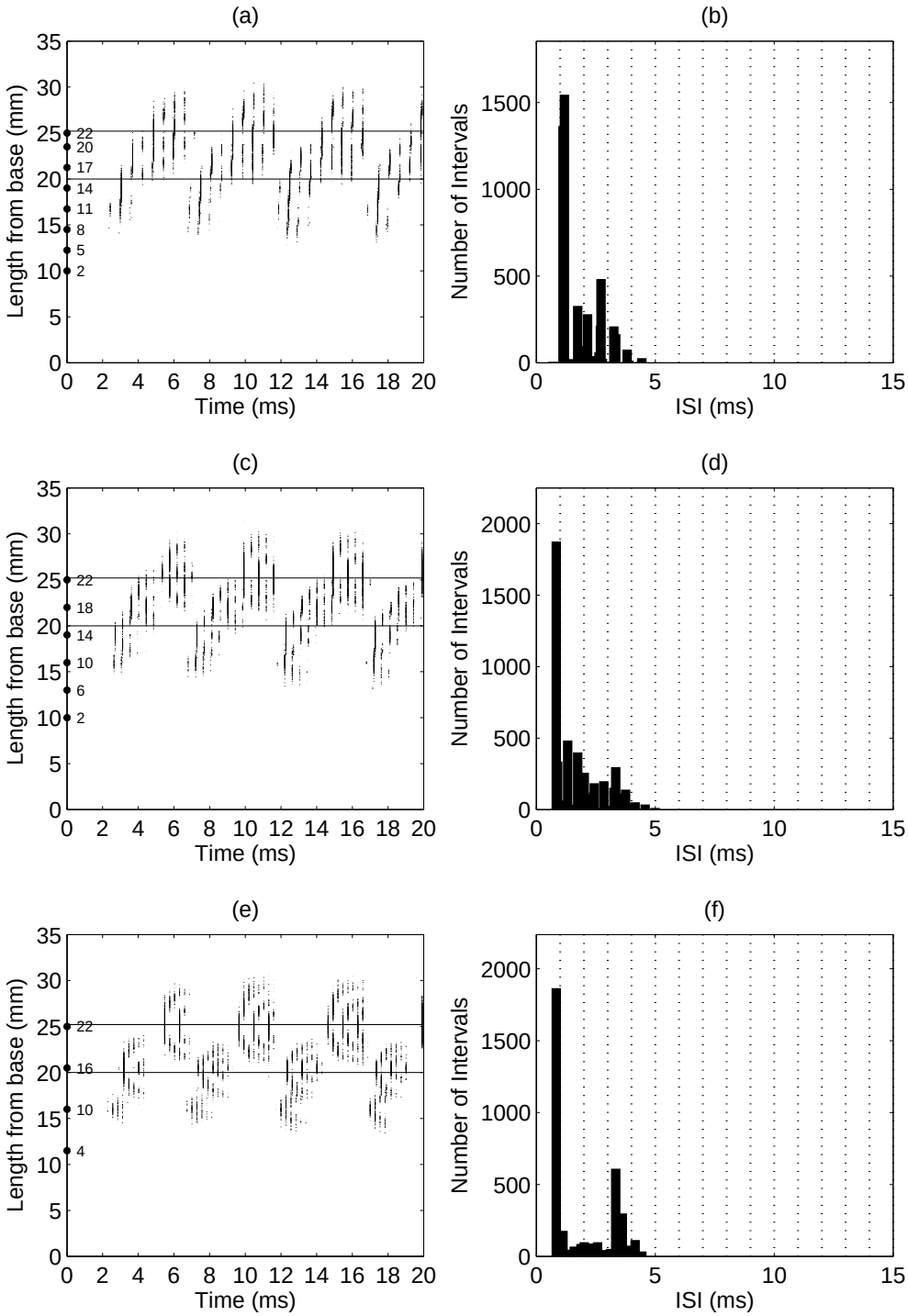

Figure 5.25: The predicted spatio-temporal neural spike train patterns for the channel-number stimulation-rate configurations: 8 - 1 800 pps (a), 6 - 2 400 pps (c) and 4 - 3 600 pps (e), for an input tone of 200 Hz are shown. Their corresponding ISI histograms for the four critical bands between 20 and 25.2 mm (indicated by the two lines in (a), (c) and (e)) are shown in (b), (d) and (f).

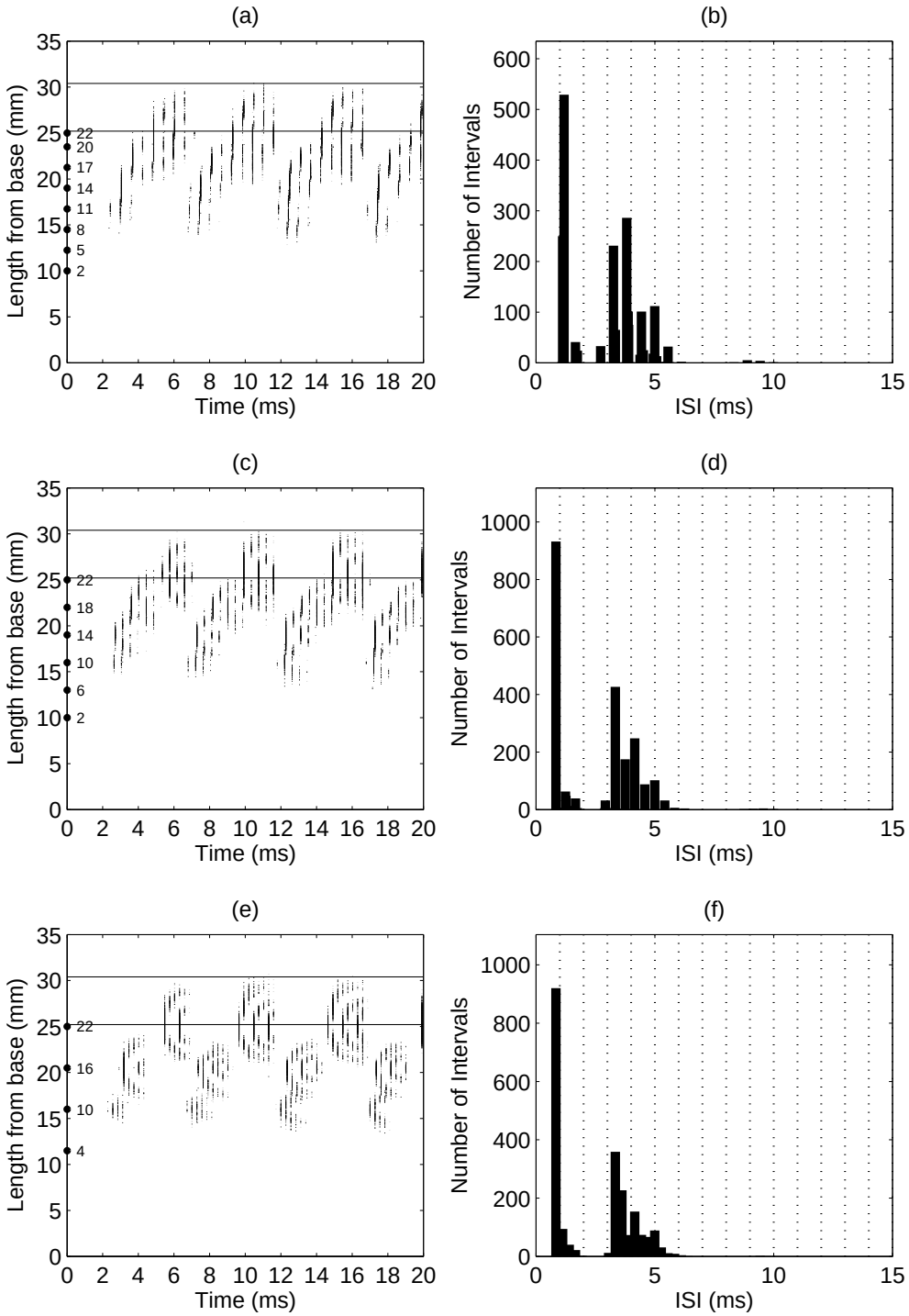

Figure 5.26: The predicted spatio-temporal neural spike train patterns for the channel-number stimulation-rate configurations: 8 - 1 800 pps (a), 6 - 2 400 pps (c) and 4 - 3 600 pps (e), for an input tone of 200 Hz are shown. Their corresponding ISI histograms for the four critical bands between 25.2 and 30.4 mm (indicated by the two lines in (a), (c) and (e)) are shown in (b), (d) and (f).

5.3.4 Predictions for closely spaced electrodes

To study the effect of channel interaction on neural responses for the travelling wave processor, six electrodes adjacent to one another were activated around the place of maximum BM activation, for pure tone inputs of 200 Hz and 1 kHz. These adjacent electrodes had an electrode spacing of 0.75 mm.

The closer electrode configuration, with activated electrode channels 17, 18, 19, 20, 21 and 22, was compared with the six-channel configuration investigated in the previous two sections, where electrode channels 2, 6, 10, 14, 18 and 22 were activated and stimulated at a channel stimulation rate of 2 400 pps. Fig. 5.27 shows the comparison between the six-channel configuration ((a) - (c)) and the configuration with more closely spaced electrodes ((d) -(f)) for a 200 Hz input tone. The spatio-temporal neural spike train patterns show the increased effect of current spread for the configuration with more closely spaced electrodes on the populations of nerve fibres. In the place of maximum BM activation the wavelength of the travelling wave decreases rapidly, leading to an increase in the phase. Therefore the BM response pattern between two adjacent areas in this area shows a large phase shift i.e. time delay. However, when electrodes are more closely spaced, adjacent populations are stimulated by more than one electrode. This results in more temporal (electrodes stimulated sequentially) and spatial distortion (current spread decreasing nerve fibres' firing threshold, leading to increased firing probability). This distortion can also be seen when observing the periods (shown by the distance between arrows at A, B and C, (a) and (d)) between clusters of spikes corresponding to the upward BM movement. The predicted results suggest these periods decrease leading to more spike intervals at periods smaller (shown by A, (f)) than the tone period of 5 ms (f). A comparison of the PST histograms (with a time bin length of 0.02 ms) (b) and (d) does not seem to show this distortion, observed in the spatio-temporal neural patterns ((a) and (d)), that clearly.

Predicted spatio-temporal neural spike train patterns for the 1 kHz tone fig. 5.28 (d) shows that the closer spacing of electrodes results in a smaller period of activation for populations of nerve fibres. The PST histogram (with a time bin length of 0.01 ms) (d), shows that the interval period between population firings is much smaller than the 1 ms tone period. The ISI histogram (f) shows fewer interval lengths at multiples of the tone period, 2 and 3 ms. Fig. 5.28 (c) shows more spikes at multiples of the tone period, which is more consistent with normal hearing where a large percentage of spikes occur at multiples of the tone period.

The results suggest that current spread has a quite significant effect on the travelling wave delay, especially in the place of maximum deflection, with neural spike trains pattern containing less temporal information about the tone period. To gain better understanding of the influence of current spread, it is necessary to observe the effect when electrodes are spread across the length of the cochlea as well. The next section investigates the effect for the placement of 22 electrodes across a length of 32 mm.

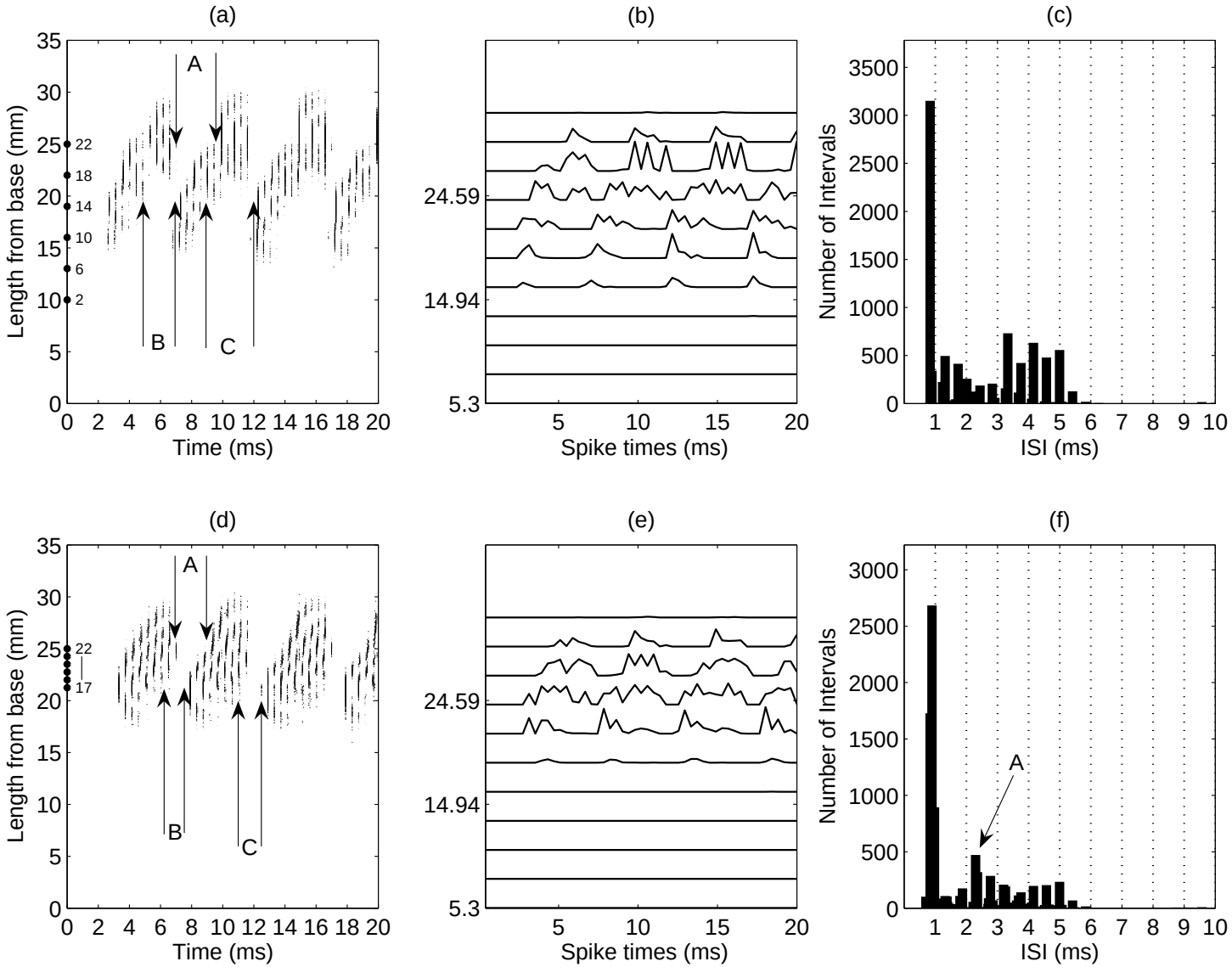

Figure 5.27: Results for two configurations, (top) 6 - normal spread and (bottom) 6 - narrow spread for a 200 Hz pure tone input. The results show the comparison between the spatio-temporal patterns (a) and (d), the PST histogram (b) and (e), and the ISI histogram (c) and (f). A, B and C are explained in the text.

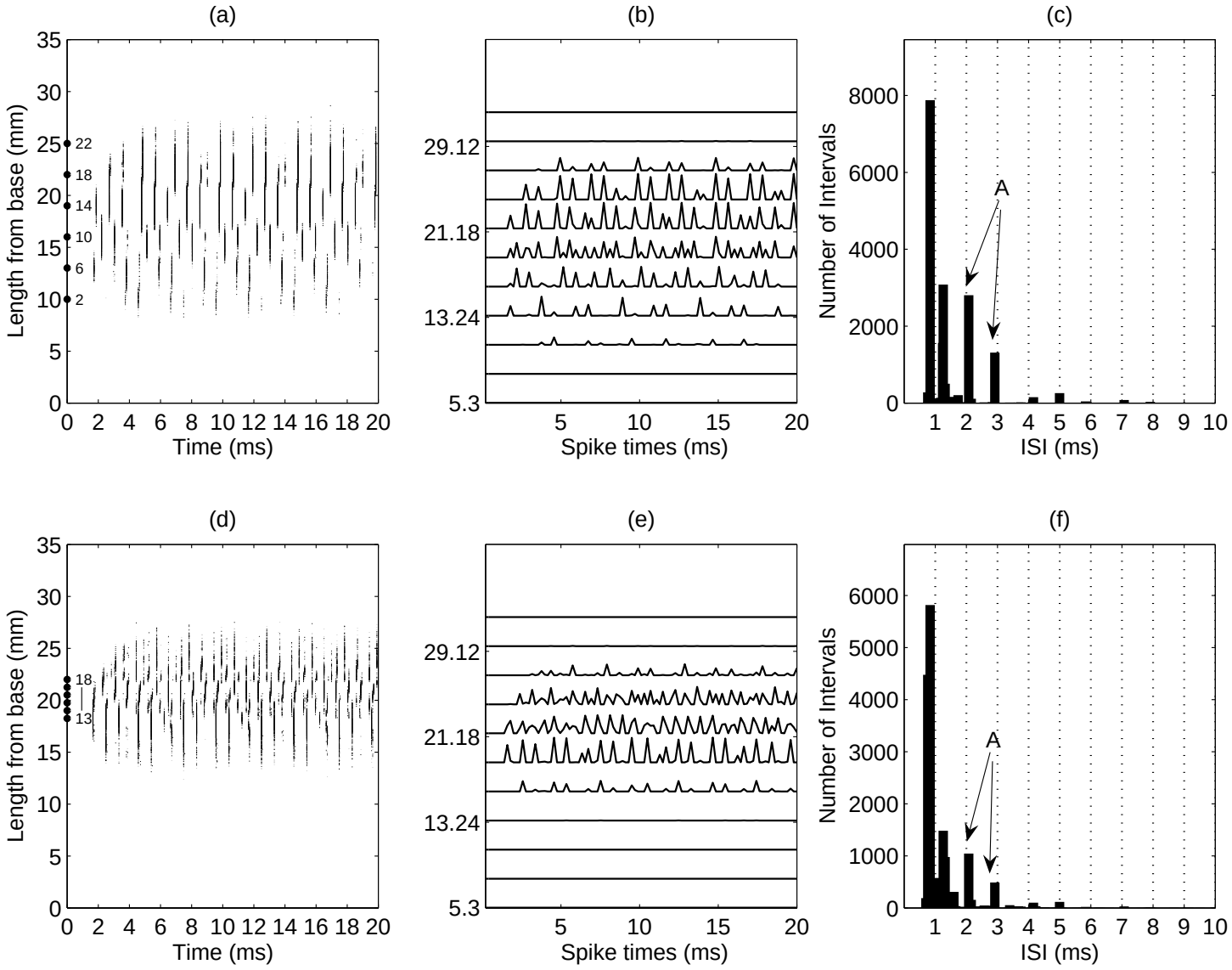

Figure 5.28: Results for two configurations, (top) 6 electrodes normal spread and (bottom) 6 electrodes narrow spread for a 1 kHz pure tone input. The results show the comparison between the spatio-temporal patterns (a) and (d), the PST histogram (b) and (e), and the ISI histogram (c) and (f). A is explained in the text.

5.3.5 Predictions for widely spaced electrodes

The COMBI-40+ Med-El cochlear implant uses a soft electrode carrier designed to facilitate deeper electrode insertion into the cochlea. Because of this deeper insertion into the cochlea, the electrodes can be spaced 2.4 mm apart, resulting in a considerably larger electrode span (26.4 mm) (Loizou, 2006). To investigate wider electrode spacing, 22 electrodes were spaced across the length of the cochlea covering a length of 32 mm, with an electrode spacing of 1.52 mm. A configuration with six activated electrodes: 2, 6, 10, 14, 18 and 22, was stimulated at a channel stimulation rate of 2 400 pps. With this configuration every 4th electrode is stimulated, resulting in a spacing between activated electrodes of 6.08 mm. This configuration was compared to the six-channel configuration investigated in the previous sections.

Fig's. 5.29 and 5.30 show the comparison between the two six-channel configurations for a 200 Hz and 1 kHz input tone. The spatio-temporal neural patterns (figs. 5.29 and 5.30, (d)) and PST histogram results (e), show more defined clusters of spikes around multiples of the tone periods of 5 ms and 1 ms. The spatio-temporal neural patterns for the 200 Hz tone (fig. 5.29 (d)) showed significantly less travelling wave distortion. This is the result of stimulating discrete neural populations, which fire synchronously at the input tone period. The ISI histograms (fig.5.29 (f) and fig.5.30 (f)) confirm this synchronous firing with a large cluster of intervals around the tone periods of 5 ms and 1 ms (shown by A). A further comparison between the two ISI histograms (fig.5.30 (c) and (f)) for the tone input of 1 kHz shows spike intervals increasing around multiples of the tone periods (shown by B, (f)). Large clusters of spikes at multiples of the tone period are also seen in predicted normal hearing ISI histogram results (fig. 5.20).

The results suggest that there might be less travelling wave distortion at low frequency inputs when electrodes are spaced at a distance where discrete populations of nerves could be stimulated. For both low and high frequency inputs, stimulating discrete populations across the length of the cochlea might provide the auditory system with better temporal information. Further, each electrode channel in the TW processing strategy is associated with a specific place along the basilar membrane. Placing these electrodes across the cochlea would place them closer to that associated basilar membrane place. Therefore, spacing the electrode array across the length of the cochlea may potentially decrease the frequency-place mismatch that exists with the electrode array simulated in the rest of this study.

The next chapter discusses the results seen in this chapter with reference to what has been observed in literature.

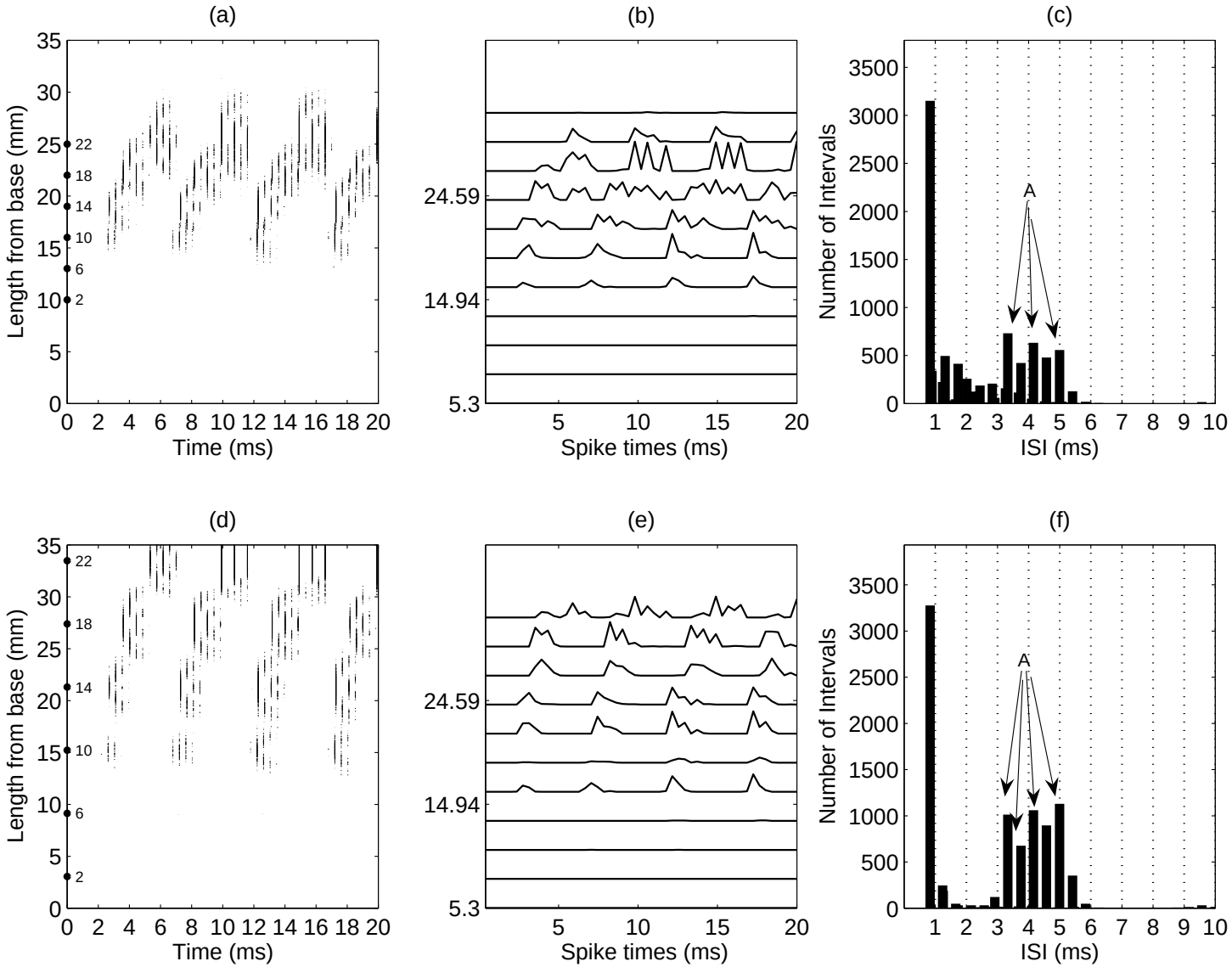

Figure 5.29: Results for two configurations, (top) 6 - normal spread and (bottom) 6 - wide spread for a 200 Hz pure tone input. The results show the comparison between the spatio-temporal patterns (a) and (d), the PST histogram (with a time bin length of 0.02 ms) (b) and (e), and the ISI histogram (c) and (f). A is explained in the text.

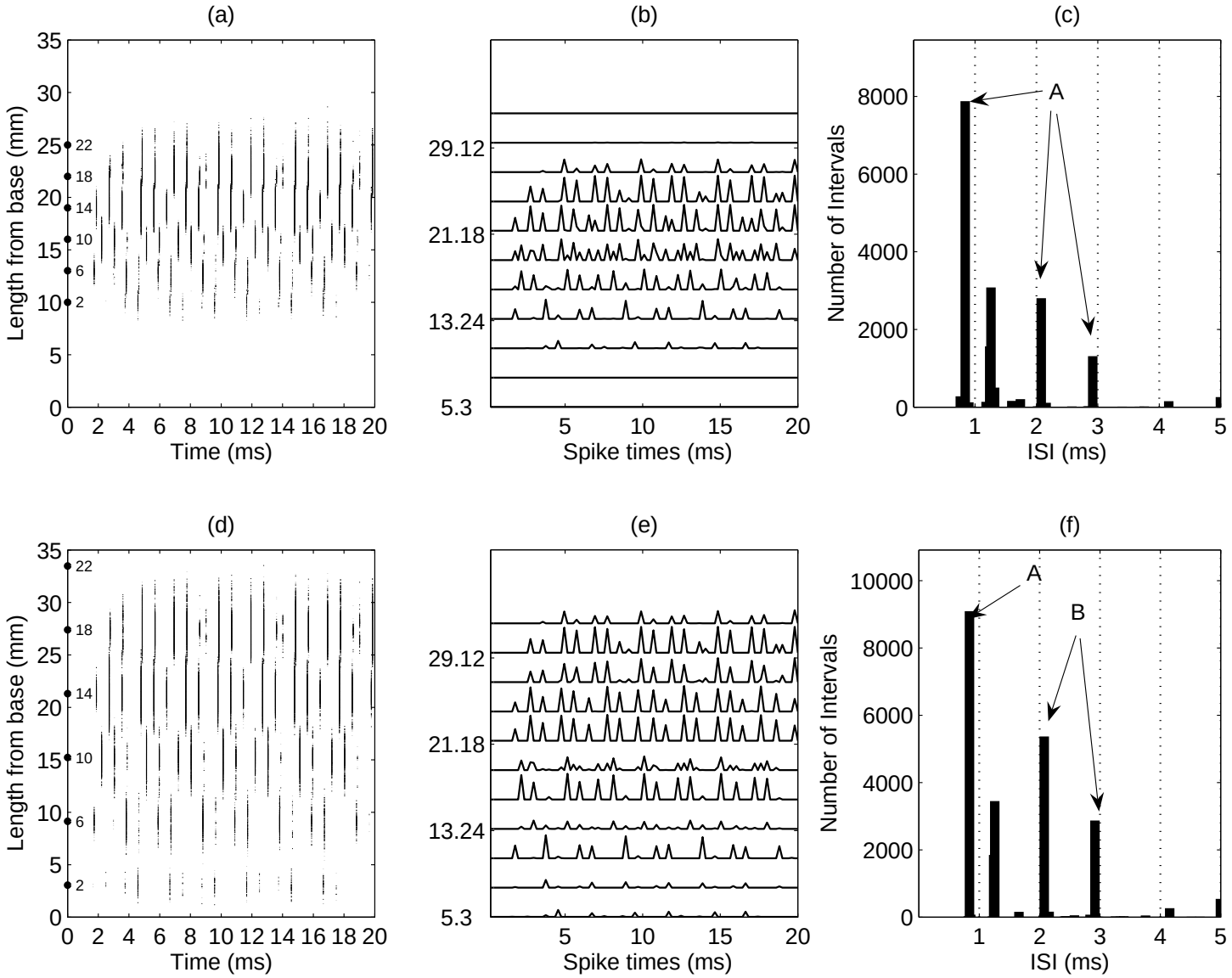

Figure 5.30: Results for two configurations,(top) 6 - normal spread and (bottom) 6 - wide spread for a 1 kHz pure tone input. The results show the comparison between the spatio-temporal patterns (a) and (d), the PST histogram (with a time bin length of 0.01 ms) (b) and (e), and the ISI histogram (c) and (f). A and B are explained in the text.

Chapter 6

Discussion

This chapter discusses the neural response results for a travelling wave CI processor strategy, expanding on what has been considered in chapters 4 and 5. The findings explore the implications of these predicted results with reference to previous findings in literature.

Traveling waves might provide a way to get better pitch perception by providing the auditory system with better salience pitch information. The reason for this speculation is that traveling waves stimulate a wide area of neural populations, which might make pitch more robust to noise. In cochlear implants, traveling waves have mostly been ignored, except for a few studies (Clark, 2004; Blamey *et al.*, 2006; Taft *et al.*, 2009; Taft *et al.*, 2010). This study has considered these traveling waves in a processing strategy.

As a preliminary investigative study the travelling wave model has been useful in providing a way to investigate the predicted spike train patterns resulting from a travelling wave strategy. The travelling wave strategy provided a way to investigate the temporal and spatial neural encoding of stimulation pulses from a travelling wave processor, for different channel-number stimulation-rate configurations. These configurations were chosen to range

from high spatial resolution (22 activated electrode channels) stimulated at a low stimulation rate of 655 pps per channel, to a low spatial resolution (four activated electrode channels) stimulated at a high stimulation rate of 3 600 pps per channel. This range was chosen to investigate the trade-off effect between spatial and temporal resolution on predicted neural spike train patterns for a travelling wave type processor, in order to determine which choices of channel number and stimulation rate will better approximate normal hearing neural spike train patterns. The neural spike train patterns were simulated for the various configurations at four tone frequencies across the frequency range, namely 200 Hz, 600 Hz, 1 kHz and 4 kHz. This study therefore serves as a stepping stone for further studies that will consider the implementation of travelling waves in speech processors.

6.1 DISCUSSION OF MODEL ASSUMPTIONS AND THEIR IMPLICATIONS

6.1.1 Model dimensionality

Before discussing the implications of the study, a number of assumptions have to be considered. As this is an investigative study, it was considered adequate to implement a passive linear one-dimensional model describing the cochlear dynamics in the travelling wave processing strategy. Although this one-dimensional model fails at approximating the observed sharp neural tuning seen in measured BM vibration results such as those presented in Rhode (1971), it presents the time domain representation of the travelling wave that could be implemented in the TW processing strategy fairly well.

The frequency domain results (chapter 3) showed that the travelling wave model fails at approximating the travelling wave delay in the most apical part (below 400 Hz). This dis-

crepancy in the travelling wave delay in lower frequencies is discussed shown in chapter 5. Neely's (1981b:74) model's frequency map has no results calculated for the apical part, which might suggest that they observed the same discrepancy in their data. This needs to be taken into account when considering the predicted neural spike train patterns for 200 Hz. These predicted neural patterns, however, were adequate for investigating whether the neural patterns contained spatial and temporal information. The parameters used in the travelling wave model were derived from experimental cat data (Neely, 1981b). These parameters were not adapted for humans. The cat cochlea, with a length of 22.5 mm, is shorter than cochlear lengths measured in human subjects at 35 mm. To incorporate electrode distances from currently available electrode array carriers, this length for the travelling wave model was normalised to 35 mm. Another consideration is that the frequency range of the cat, according to the equation derived by Greenwood, is much wider (20 Hz - 50 kHz) than the human auditory range of 20 Hz - 20 kHz. Comparing the cochlear delay measured in cats to those in human cadavers shows an apparent difference between species (von Békésy, 1949; Van der Heijden and Joris, 2003; Siegel *et al.*, 2005).

6.1.2 Modelling of current spread

A simplifying assumption was made regarding the placement of the electrode array, with electrodes spaced equidistant, 0 mm from the nerve fibres. Cohen, Richardson, Saunders and Cowan (2003) show how the non-uniform placement of electrodes affected psychophysical performance. Introducing non-uniform electrode spacing from the nerve fibres will cause larger populations of nerve fibres to be stimulated. This might introduce more spatial and temporal distortion in the predicted neural spike trains for the travelling wave processor, resulting in less neural activation at the tone period. Furthermore, the current decay rate used in the model was 2.8 dB/mm, corresponding to a mean rate calculated from the current decay

rate estimates for five different current spread estimation methods (Bingabr *et al.*, 2008) for monopolar stimulation. For bipolar stimulation with a decay rate of 7.4 (Bingabr *et al.*, 2008) and 8 dB/ mm (Bruce *et al.*, 1999a) a smaller population of nerve fibres will be excited. This might result in less travelling wave delay distortion around the CF place, with more neural activation at the tone period.

6.1.3 Modelling of neural firing

The normal hearing model proposed in (section 6.4.1) predicts neural spike patterns in normal hearing with respect to the travelling wave BM movement. The model assumes a constant firing probability for all nerves fibres across the length of the cochlea. Rose *et al.* (1967) calculated the probability of the next firing occurring, multiples of the tone cycle further in time, given that a nerve fibre has fired. This probability decreased relative to an increase in tone frequency. Including an increasing firing probability in the model could better predict ISI histograms, calculated from measured neural responses. Furthermore, the model's refractory period has been limited to the absolute refractory period. Including a relative refractory period could result in more interval clusters at multiples of the tone period in the predicted ISI histograms. However, with only the absolute refractory period included, the predicted ISI histogram results show peaks at multiples of the tone period, similar to those seen in calculated ISI histograms for measured responses (Rose *et al.*, 1967; Javel, 1990). The predicted ISI histograms further show an exponential decay for lower and midrange input frequencies, with non-uniform peaks seen at higher frequencies when the firing probability decreases. This has also been observed in (Rose *et al.*, 1967).

6.1.4 Modelling of electrical stimulation

Predicted spatio-temporal neural patterns (chapter 5) show that electrical neural stimulation patterns differ considerably compared to the predicted normal hearing model patterns. Electrical stimulation (Javel, 1990:282) results show that lower current stimulation pulses result in lower spike rates, with ISIs being extended to multiples of the stimulation period. This is similar to what is observed in normal hearing ISI histogram results (Rose *et al.*, 1967; Javel, 1990), where intervals cluster around multiples of the acoustic tone period. Higher stimulation currents (> 50 dB μA or 316.22 μA, current level of 170), on the other hand, lead to firings occurring on every stimulation cycle, with their ISI results showing one peak at the stimulation period. In Bruce *et al.*'s (1999a) model, neural firing only occurs in the cathodic phase, based on the assumption that spikes are more likely to occur during this phase. This assumption is based on the fact that the cathodic phase has a lower threshold than the anodic phase (Bruce *et al.*, 1999a). Observation of the predicted spatio-temporal neural patterns for the 200 Hz input tone reveals that clusters of spikes occur at the tone period. However, within these clusters there are separate clusters that occur at the stimulation rate. Predicted ISI histograms for low stimulation rates show large peaks at the stimulation rate. The auditory system might extract this temporal information about the stimulation rate, rather than the temporal information about the tone period.

6.2 DISCUSSION OF REAL TIME PROCESSING CONSIDERATIONS

This study implements a hydrodynamical model which solves the cochlear differential fluid equations and their corresponding boundary conditions in time, using a numerical method. This hydrodynamical model predicts the BM movement across time, which provided a way to implement a travelling wave processor. The numerical method, however, requires a very

high time resolution input signal, for numerical stability. This increases the computational time as the signal must be resampled to 14.4 kHz. Therefore sampling at a low stimulation rate of 16 kHz is not possible with the existing numerical method. However, with faster processors, faster analogue-to-digital converters, and larger memory it would be possible to implement this method in real time. However, longer computational time and a higher sampling rate mean higher clock speed and higher power consumption, resulting in decreased battery life. Therefore, when implementing a traveling wave processor in real time, it might be more beneficial to consider a less complex method, which does not have to solve the cochlear fluid dynamics using a numerical method. Zwislocki (1948) introduced a transmission line model that incorporated travelling wave experimental results. Blamey *et al.* (2006) developed a real time travelling wave sound processor based on psychophysical tuning curves in normal hearing individuals. These methods do not solve the fluid dynamics in time, saving on computational time.

6.3 DISCUSSION OF CHANNEL STIMULATION RATE, ELECTRODE AND PURE TONE CONSIDERATIONS

As mentioned in the first section of this chapter, the channel-number stimulation-rate configurations were chosen to span the range of high spatial resolution (22 activated electrode channels) with a channel stimulation rate of 655 pps to a low spatial resolution (four activated electrode channels) with a channel stimulation rate of 3 600 pps. The chosen channel-number stimulation-rate configurations between these boundaries were 11 - 1 309 pps, 8 - 1 800 pps and 6 - 2 400 pps. Four pure tone frequencies were chosen to cover a range of frequencies, from a low input pure tone of 200 Hz up to a very high frequency of 4 kHz, with a middle frequency tone of 600 Hz and a high frequency of 1 kHz added to the pure tone set. According to the Nyquist rule, none of the sampling rates could adequately sample the

input frequency of 4 kHz, therefore the sampling rate of 7 200 pps was chosen, which is the highest frequency for a multi-channel configuration. This frequency, however, also fails at adequately sampling the pure tone of 4 kHz (according to Nyquist), but it was used to investigate whether the predicted neural spike trains patterns contain any information regarding the input tone. The predicted neural patterns suggest that there is temporal information about the tone that could be extracted by the auditory system.

The electrode array and insertion depth were chosen to be consistent with those used in currently available processing strategies. Modelled insertion depths of 22 and 23 mm from the base showed significant degradation in performance, while an insertion depth of 25 mm showed normal performance (Dorman, Loizou and Rainey, 1997). Insertion depths closer than 19 mm caused significant loss of speech information (Faulkner, Rosen and Stanton, 2003). An insertion depth of 25 mm from the base was used in the simulations. The simulated electrode array was based upon the Nucleus 24 electrode array with 22 electrodes inserted in the cochlea having an electrode spacing of 0.75 mm. The last section in chapter 5 investigates a more widely spaced electrode array. This array was chosen to span the length of the BM with an electrode spacing of 1.52 mm. This electrode array does not correspond to electrode arrays used in current commercially available processing strategies. The commercially available electrode array with the deepest insertion depth is the soft 12-electrode carrier used in the COMBI-40+ Med-El cochlear implant with an approximate insertion depth of 31 mm, and the electrode array spanning a distance of 26.4 mm (Loizou, 2006). The simulated electrode array spanning the length of the cochlea was used to investigate the predicted spike train patterns when discrete neural populations are stimulated across the cochlea.

6.4 DISCUSSION OF PREDICTED NEURAL SPIKE TRAIN PATTERNS

With present implants, such as Nucleus, a trade-off remains between the number of electrodes and the stimulation rate. The maximum stimulation rate per channel is dependent upon the number of activated channels and is calculated by dividing the overall maximum stimulation rate by the number of activated channels (Loizou, 2006). There are a number of recent processing strategies that introduce different ways of stimulating an electrode array, namely HiRes 120 (Buechner, Brendel, Krüeger, Frohne-Büchner, Nogueira, Edler and Lenarz, 2008), which introduces current steering, and Fine Hearing (Hochmair, Nopp, Jolly, Schmidt, SchöBer, Garnham and Anderson, 2006; Arnoldner, Riss, Brunner, Durisin, Baumgartner and Hamzavi, 2007), which introduces stimulation pulses at the zero crossings of bandpass-filtered low frequency channels. The ACE strategy employs a set of FFT bins to produce filter bands which are spaced across the frequency range. The power of adjacent bins is added together to estimate a total of *n* envelopes from which *m* envelope amplitudes are selected in each stimulation time frame (Loizou, 2006). This envelope extraction results in the loss of most or all of the fine structure (temporal fine structure) information (Wilson, Schatzer, Lopez-Poveda, Sun, Lawson and Wolford, 2005). Predicted neural spike train patterns (section 5.3.2) confirm this loss in fine structure information. These predicted neural patterns show hardly any temporal information regarding the pure tone input, with spike populations occurring at multiples of the stimulation rate for low and high frequency inputs. The predicted spatio-temporal neural patterns for the ACE strategy employs that the auditory system can use mainly the spatial information regarding the place of stimulation, to determine the pure tone frequency.

When electrical stimulation takes place, there is still a huge loss in spatial and temporal resolution compared to the resolution seen in the normal auditory system. Another effect

of electrical stimulation is that of current spread around the stimulating electrode. These might be some of the reasons why travelling waves have mostly been ignored in commercial speech processors. Recent studies however, have investigated the implementation of travelling wave type speech processors. A preliminary study with a travelling wave processor showed that subjects found it easier to distinguish high-pitched sounds than low-pitched sounds (Clark, 2004). This study showed a better pitch ranking between 150 Hz and 250 Hz, compared to ACE, with subjects finding it easier to recognise musical instruments. However, no published work could be found about the details of this study. Blamey *et al.* (2006) developed a real time travelling wave sound processor based on psychophysical tuning curves in normal hearing individuals. This travelling wave processor has been patented, but no publications were found in open literature that evaluate this strategy. Taft *et al.* (2009) incorporated a set of calibrated travelling wave delays in a speech processor, by shifting the band-pass filter output relative to the travelling wave delay. Their results showed significant improvements with sentence recognition in noise for all the subjects. A further study (Taft *et al.*, 2010) that tried to explain this improvement in speech perception showed that de-synchronising the filter bands with across-frequency delays is merely an improvement of the processing strategy rather than a cochlear wave mechanism. As far as the authors know, this is the only published work on travelling wave processors.

Observation of processing in the normal auditory system suggests that it probably extracts both spectral and temporal data about the specific input tone from the nerve fibre responses. This study implemented a hydrodynamical model in a processing strategy that predicts the travelling wave movement of the BM. Although this model can predict more complex tones (Diependaal and Viergever, 1989), the aim was not to consider these complex tones, but to observe the neural spike train patterns for pure tone inputs. There were two reasons for only investigating pure tones. The first was to determine whether the predicted neural spike train patterns showed any temporal information regarding the pure tone period that might be

extracted by the auditory system. Predicted ISI histograms for the electrical stimulation were compared to calculated ISI histograms from measured neural spike patterns for acoustic pure tones (Rose *et al.*, 1967; Javel, 1990). The travelling wave shows a very distinct travelling wave delay increasing toward the place of maximum deflection. Therefore the second reason for choosing pure tones was to determine whether this travelling wave delay and peak could be observed at all in the predicted neural spike train patterns. However, the investigation of only pure tones prevented an investigation of the effects of wider spectrum signals.

Predicted neural patterns for the TW processing strategy (chapter 5) suggest that these patterns could contain spectral and temporal information about the pure tone that the auditory system might be able to extract. The predicted neural results further suggest that the choice of channel number and stimulation rate has a significant effect on the predicted neural spike train patterns. One of the factors contributing to this significant difference in neural patterns remains the current spread around the stimulating electrodes. The predicted results showed that the current spread caused not only a spatial but also a temporal distortion as a result of the sequential stimulation, which is also used in the Nucleus implant. Therefore, although sequential stimulation probably minimises current spread (Loizou, 1998) across the length of the cochlea, the predicted results suggest that it leads to increased temporal distortion. In terms of the trade-off between channel number and stimulation rate, the predicted neural spike train patterns suggest that a low stimulation rate of 655 pps, with a channel-number resolution of 22 electrode channels, shows very little spatial and temporal information that could be useful to the auditory system. The predicted ISI histograms show peaks around the stimulation rate, with no temporal information about the input tone period. The predicted neural results for channel-number stimulation-rate configurations: 11 - 1 309 pps, 8 - 1 800 pps, 6 - 2 400 pps and 4 - 3 600 pps, for the input frequency of 200 Hz, showed that these neural patterns could contain both spatial and temporal information that could be used by the auditory system. Further results for a tone frequency of 600 Hz showed that

results for channel-number configurations 8 - 1 800 pps, 6 - 2 400 pps and 4- 3 600 pps, could contain useful spatial and temporal information. With a higher frequency of 1 kHz channel-number, stimulation-rate configurations: 11 - 1 309 pps and 8 - 1 800 pps do satisfy Nyquist's rule, which states that the sampling frequency should be at least twice the frequency of the highest frequency component in the sampled signal. This needs to be taken into account in channel-number stimulation-rate configurations in travelling wave processors, where the temporal pattern of the input tone needs to be sampled adequately to prevent spectral overlapping. This spectral overlapping might provide the auditory system with distorted information, which might influence the perception of the specific tone. For a processing strategy such as ACE, where the input signal's envelope is not sampled, but rather estimated (by the power of adjacent FFT bins being added together), this spectral overlapping as a result of lower stimulation rate might not need to be considered. Channel-number stimulation-rate configurations: 6 - 2 400 pps and 4 - 3 600 pps adequately sample the TW model output (for a frequency of 1 kHz) for temporal information to be contained in the predicted neural patterns. These predicted neural patterns, however, show hardly any travelling wave delay, which would make it very difficult for the auditory system to extract any spatial information regarding the pure tone. However, the input frequency of 4 kHz could not be adequately sampled by one of the above-mentioned configurations. In section (section 5.3.3.2) the channel-number, stimulation-rate configuration 2 - 7 200 pps was proposed with a higher stimulation rate to sample the travelling wave model output. According to Nyquist this sampling rate is still too low to sample the input adequately. The predicted neural patterns, however, suggest that there is temporal information that could be extracted, but show no travelling wave delay.

6.4.1 Possible temporal information in predicted neural patterns

The previous paragraph mentioned that there might be temporal information in the predicted neural spike train patterns that the auditory system might use. Studies have shown that cochlear implantees can perceive pitch with a single electrode up to a stimulation rate of 300 pps (Shannon, 1983; Zeng, 2002). Apart from perceiving a different pitch with a different stimulation rate, cochlear implantees can also extract pitch information from amplitude modulated stimuli below 300 Hz (Shannon, 1992; McKay, McDermott and Clark, 1994). With sufficiently large modulation depths, modulated pulse trains are perceived at a pitch close to that of unmodulated pulse trains (McKay, McDermott and Clark, 1995). F0 amplitude modulated strategies that used techniques to minimise the phase difference across channels showed improved pitch ranking results (Vandali, Sucher, Tsang, McKay, Chew and McDermott, 2005; Milczynski, Wouters and Van Wieringen, 2009). One of these strategies (Milczynski *et al.*, 2009) applied 100 percent amplitude modulation depth. The TW model simulation results (section 3.6) show that for a sinusoidal input, the BM movement at a specific place will be a sinusoidal movement at the input tone's frequency. The travelling wave model modulates this BM movement with a carrier rate (channel stimulation rate) at a 100 percent modulation depth. Predicted spatio-temporal neural spike train patterns (section 5.3.3.1) for the travelling wave processor for a low frequency pure tone input of 200 Hz at the different channel configurations show clusters of spikes appearing at the upward BM movement. Stimulating at higher stimulation rates (1 800, 2 400 and 3 600 pps) produces denser clustering of neural spikes around the modulated BM envelope. This is similar to that seen in the predicted normal hearing results and predicted PST histograms where spikes show synchronised activity at the phases of the stimulating tone. Single nerve fibres have been shown to phase lock to relatively high stimulus frequencies (Kiang *et al.*, 1965; Rose *et al.*, 1967; Liu *et al.*, 2006). Populations of nerve fibers have also been observed to show entrainment to input stimuli (Shamma, 1985a).

Although studies (Shannon, 1992; McKay *et al.*, 1994) show that cochlear implantees could only extract temporal information for low-frequency inputs it could still be beneficial to consider predicted neural spike train patterns for higher frequency inputs. However, an issue that has to be considered with high-frequency inputs is that when sampling rates are too low, aliasing has a significant effect on the phase information. Overall, predicted ISI histogram results for channel-number stimulation-rate configurations with higher channel stimulations rates, 8 - 1 800 pps, 6 - 2 400 pps and 4 - 3 600 pps, showed that clusters of spikes appear at intervals around the tone period (200 Hz tone) or multiples thereof (600 Hz and 1 kHz tones). The channel-number stimulation-rate configuration of 2 - 7 200 pps showed intervals at multiples of the tone period. ISI histograms calculated for predicted neural spike train patterns within a critical band show that this clustering of spikes becomes more evident for certain critical bands (section 5.3.3.3). Predicted ISI histograms for normal hearing and ISI histograms calculated from measured responses ((Rose *et al.*, 1967; Javel, 1990) also show similar peaks at multiples of the input tone period. These ISI lengths seen in the ISI histograms might be used by the auditory system to extract temporal information, which might be the reason why subjects could consistently rank modulated stimuli to modulation frequency for carrier rates higher than 800 pps (McKay *et al.*, 1994). McKay further suggested that it might be beneficial to stimulate at higher stimulation rates when modulating signals (McKay *et al.*, 1994:2672). Another study, where carrier rates ranged from 254 to 4069 pp, showed a decrease in modulation sensitivity with increased carrier rate (Middlebrooks, 2008). The predicted neural spike train patterns suggest that stimulating at higher stimulation rates could convey more temporal information with predicted neural patterns showing clusters of spikes at intervals around the tone period.

6.4.2 Traveling wave delay: Possible spatial and temporal information

As opposed to speech processors that minimise the phase across channels (Vandali *et al.*, 2005; Milczynski *et al.*, 2009), the travelling wave strategy introduces a travelling wave delay across populations of nerve fibres. This delay increases along the length of the cochlea as it approaches the place of maximum deflection. This travelling wave delay (phase shift), related to the different places along the cochlea, might contain temporal information regarding the pure tone. The auditory system might further also extract spatial information regarding these places as well as the CF place.

The low frequency (200 Hz) spatio-temporal neural spike patterns show this increase in the travelling wave delay clearly, as the wave approaches the place of maximum excitation. However, at the CF place where this travelling delay increases rapidly, the current spread shows a significant effect, distorting this rapid phase shift (section 5.3.4). Predicted ISI histograms for critical bands in this region (section 5.3.3.3) showed almost no peaks at the tone period for the 200 Hz input. The TW model's (chapter 3) simulated results showed that with an increase in pure tone frequency the travelling wave delay becomes smaller as the place of maximum deflection moves closer to the basal region. The predicted neural pattern results show this travelling wave delay becoming more distorted as input frequency increases (section 5.3.3.1). With an input frequency of 1 kHz there is hardly any travelling wave delay that could be observed from the predicted spatio-temporal neural patterns. From the predicted neural spike train patterns for the 4 kHz tone input, no travelling wave delay could be observed. This might also be the result of stimulating only two channels. Decreasing the electrode spacing to 0.75 mm rather than 3 mm (section 5.3.4) shows the increased effect of current spread. This increased current spread results in increased firings in the adjacent nerve fibre populations around the stimulating electrodes, as a result of the decreased nerve fibre threshold. Predicted spatio-temporal neural patterns show wider areas of spike activation

across time, resulting in the distortion of the tone period. This distortion of the tone period can also be seen from the predicted ISI results where the peak at the tone period decreases. The PST histograms for populations of nerve fibres further display this distortion of the tone period, with closer electrode spacing. These results suggest that the auditory system might perceive a different temporal pattern, other than that of the input tone. McKay and McDermott (1996) investigated the perception of temporal patterns on two different intra-cochlear sites, using phase-shifted modulated stimuli. Results showed that when pulses were alternated between two electrode positions, subjects perceived the combined temporal pattern for the two electrodes, when electrode spacing was smaller than 3-4 mm. However, when electrodes were spaced further apart, individual temporal patterns were perceived. These results suggest that when electrodes are spaced closely together the combined temporal pattern of the electrodes might be perceived. Predicted neural spike train patterns (section 5.3.5) further show less phase shift distortion between population regions, when the electrode array is spaced across the cochlea, with a wider active electrode spacing of 6.08 mm. Predicted ISI histograms show larger peaks at the tone period and multiples thereof (for a frequency of 1 kHz). Carlyon, Geurts and Wouters's (2000) results showed that implantees showed sensitivity to timing differences applied to discrete populations. Their results suggest that timing differences might be extracted by the auditory system when electrodes are spaced widely enough apart, in order to stimulate two discrete populations of nerve fibres. Loizou (2006) argued that the benefits of higher stimulation rates with Med-El implantees might be a result of less channel interaction as a result of wider electrode spacing of 2.4 mm in Med-El.

In conclusion, predicted neural spike train patterns suggest that travelling wave processors might contain the temporal and spatial information needed for better pure tone discrimination. Although only pure tones were investigated, we speculate that predicted neural spike train patterns for speech signals might also contain the spatial and temporal information that is needed for better speech perception. However, implementing a travelling wave in a

speech processor needs careful consideration. The predicted neural spike patterns show that decisions on the channel number and the stimulation rate have a significant effect on the predicted neural spike train patterns. From these predicted neural pattern results one can conclude that stimulating with fewer electrodes at a higher stimulation rate might be more beneficial in presenting the temporal and spatial information of the travelling wave. With wide electrode spacing between activated electrodes, predicted neural spike train patterns showed more synchronous firing at the tone period, with less travelling wave phase distortion. In the neural spike train patterns for the travelling wave processor, the relationship between temporal patterns at adjacent intra-cochlear sites might be of the utmost importance for the auditory system.

Chapter 7

Conclusion

This chapter gives a brief summary of the main findings of this study and concludes with suggestions for future work.

The primary goal of this study was to investigate the implementation of a travelling wave processor. An existing travelling wave model (Diependaal and Viergever, 1983) was implemented in a speech processor. The current spread model (Strydom and Hanekom, 2011; Smith, 2011) along with the electrical model (Bruce *et al.*, 1999a; Bruce *et al.*, 1999b) was used to compare the spatio-temporal neural patterns, ISI and PST histograms for various channel-number, stimulation-rate configurations for a travelling wave processor.

7.1 RESEARCH OVERVIEW

The main findings for this study are as follows:

- A current spread and electrical model was used to predict the spatio-temporal neural

spike patterns for various channel-number, stimulation-rate configurations. Predicted neural patterns for the travelling wave processor with four and six activated electrode channels, at channel stimulating rates of 2 400 pps and 3 600 pps, showed closer approximations to the predicted normal hearing spatio-temporal neural patterns (for input frequencies of 200 Hz, 600 Hz and 1 kHz), compared to the predicted spatio-temporal neural patterns for configurations with more activated electrodes and lower channel stimulation rates. These results showed a rather synchronous firing for populations of nerve fibres to the tone period. Predicted neural patterns contain travelling wave delay information, which can be seen in the predicted neural patterns for the 200 Hz input tone. These neural patterns further show this travelling wave delay to be rather distorted around the CF place, where the phase rapidly increases. For input frequencies of 600 Hz and 1 kHz, these travelling wave delays become even more distorted. In comparison, predicted spatio-temporal neural spike train patterns for the ACE strategy show no such travelling delay phase information. Spectral overlapping in predicted neural spike train patterns could be observed when the travelling wave model output was not sampled with a channel stimulation rate of at least twice that of the pure tone frequency. The predicted spatio-temporal neural spike train patterns for the channel-number, stimulation-rate configuration of 2 - 7 200 pps showed synchronous firing to the tone frequency of 4 kHz but no evidence of a travelling wave delay. The predicted neural spike train patterns for the travelling wave strategy suggest that these patterns might contain useful spatial and temporal information about the specific pure tone that the auditory system might be able to extract.

- Overall predicted ISI histogram results for channel-number, stimulation-rate configurations: 6 - 2 400 pps and 4 - 3 600 pps showed peaks at the tone period (200 Hz) or multiples thereof (for the input tone frequencies of 600 Hz, 1 kHz). The channel-number stimulation-rate configuration of 2- 7 200 pps showed peaks at multiples of the tone period for the input frequency of 4 kHz. The ISI histograms calculated from

predicted normal hearing results and those calculated from measured results (Rose *et al.*, 1967; Javel, 1990) showed similar peaks at multiples of the tone period. Therefore it can be concluded that the predicted neural spike train patterns for the travelling wave strategy contain temporal information regarding the tone period, which was not observed for predicted ISI histograms for the ACE strategy. Predicted neural spike train patterns for the ACE strategy showed peaks at the stimulation, which suggests that the ACE strategy contains very little temporal information about the pure tone. Pitch perception in CI subjects may possibly be improved if the auditory system could extract this temporal information seen in the predicted neural spike train patterns for the travelling wave strategy.

- The predicted neural spike train patterns show that current spread has a significant effect on the travelling wave delay and tone period distortion, especially around the place of maximum BM deflection. More closely spaced electrodes therefore show an increased distortion effect on the travelling wave phase, decreasing the tone period between the clusters of spikes. Predicted ISI histograms for a configuration where the electrode array was spaced across the cochlea show more distinct peaks at the tone period and multiples thereof, with less travelling wave delay distortion in the predicted spatio-temporal neural patterns. The predicted neural spike train patterns for more widely spaced active electrodes (6.08 mm between activated electrodes), spaced across this electrode array, suggest that increasing the electrode spacing between activated electrodes in a travelling wave processor could increase temporal and spatial information conveyed to the auditory system.

With reference to the research questions, it has been shown that:

- The current spread caused significant temporal and spatial distortion. Therefore the predicted neural spike train patterns for a travelling wave processor suggests that stim-

ulating fewer channels at a higher rate could increase temporal and spatial travelling wave information conveyed to the auditory system.

- Predicted neural spike train patterns and the corresponding ISI histograms suggest that temporal information about the travelling wave is contained in neural spike patterns. These patterns for low frequencies suggest that travelling wave delay information is contained in these neural spike patterns. However, as a result of the rapid travelling wave phase shift in the region of maximum BM deflection, travelling wave delay distortion increases significantly. Although travelling wave delays could be observed for higher frequencies, predicted neural spike results show them to be rather distorted as a result of the large current spread introduced by the stimulating electrodes. Much less travelling wave delay distortion was observed in the predicted neural spike patterns when the electrode spacing was increased for between electrodes, with the electrode array spanning the length of the cochlea.

7.2 FUTURE WORK

This study provides a basis for future work. Suggestions for further studies on the implementation of travelling wave speech processors are listed below.

In this study the cochlear mechanics were approximated with a linear passive one-dimensional model. The measured results from Rhode (1971) and Sellick *et al.* (1982) showed sharpened tuning responses, with active and nonlinear processes evident in the cochlear mechanics. For a better approximation of the cochlear mechanics, a non-linear active one-dimensional model could be implemented in the travelling wave strategy.

The hydrodynamical model solves the differential fluid equations and their corresponding boundary conditions, using a numerical method that requires a very high time resolution in-

put signal. This model was introduced in the TW processing strategy. However, implementing a travelling wave model in a real time processing strategy requires a fast computational model. The model used in this study might not be adequate for real time processing, therefore a computationally faster model, that solves a lower time resolution (16 kHz or 44 kHz) input signal, could be investigated. Blamey *et al.* (2006) developed a real time travelling wave sound processor based on psychophysical tuning curves in normal hearing individuals. Zwislocki (1948) introduced a transmission line model that describes cochlear mechanics. These methods could be helpful in investigating a computationally faster real time traveling wave processing strategy.

The normal hearing model suggested in this study includes the effect of the absolute refractory period, but does not include the effect of the relative refractory period. Bruce *et al.* (1999a) approximated this effect in their electrical stimulation model as an exponential decreasing current, with respect to time. This current was added to the threshold current, increasing the threshold current for which a spike would occur. In the normal hearing model this effect might be included, with an exponential increase in firing probability with respect to time. Including the relative refractory period might increase the ISIs between spikes, which might cause more spikes to occur at multiples of the tone period. Rose *et al.* (1967) calculated the probability of the next firing occurring, multiples of the tone cycle further in time, given that a nerve fibre has fired. This probability decreased relative to an increase in tone frequency. Including a decreasing firing probability (relative to an increasing frequency) in the model, could better predict ISI histograms, calculated from measured neural responses (Rose *et al.*, 1967).

This study only investigated the effect of pure tones at different channel-number stimulation-rate configurations. Predicting neural spike train patterns for speech signals with a travelling wave processor might provide better understanding of the considerations regarding speech

signals in travelling wave processors.

The current decay rate used in the current spread model was 2.8 dB/mm, corresponding to a calculated mean rate calculated from the current decay rate estimates for five different current spread estimation methods (Bingabr *et al.*, 2008) for monopolar stimulation. For bipolar stimulation with a decay rate of 7.4 (Bingabr *et al.*, 2008) and 8 dB/mm (Bruce *et al.*, 1999a) a smaller population of nerve fibres will be excited. Using a higher current decay rate would result in a smaller current spread. This might result in configurations with a higher number of electrode channels, showing better approximations to predicted normal hearing neural patterns.

The predicted neural spike train patterns suggest that a travelling wave processor might provide the auditory system with more temporal and spatial information regarding the pure tone. It further suggest that the choice of channel number and stimulation rate significantly influences the predicted neural spike train patterns and should be considered carefully before implementing a travelling wave type processor. The predicted neural patterns show that current spread has a significant effect, especially in the place where the travelling wave phase changes rapidly. Therefore the current spread should be minimised when implementing a travelling wave processor. A deeper insertion depth with more widely spaced active electrodes could significantly decrease travelling wave distortion, providing the auditory system with more travelling wave information. In view of all these conclusions, experiments need to be developed that could investigate them. A proposed frequency discrimination experiment, which determines the just noticeable frequency difference, could be used to compare the different channel-number stimulation-rate configurations. The effect of current spread on a travelling wave processor could be determined by investigating bipolar and monopolar stimulation. These experiments for the travelling wave processor could be compared to similar experiments done with a CI subject using the ACE strategy.

References

Allen, J. (1977). Two dimensional cochlear fluid model: New results, *Journal of the Acoustical Society of America* **61**(1): 110–119.

Allen, J. (1985). Cochlear modeling, *IEEE Acoustics, Speech, and Signal Processing Magazine* **2**(1): 3–29.

Allen, J. and Sondhi, M. (1979). Cochlear macromechanics, *Journal of the Acoustical Society of America* **66**(1): 123–132.

Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W. and Hamzavi, J. (2007). Speech and music perception with the new fine structure speech coding strategy: Preliminary results, *Acta Oto-Laryngologica* **127**(12): 1298–1303.

Baskent, D. and Shannon, R. (2004). Frequency-place compression and expansion in cochlear implant listeners, *Journal of the Acoustical Society of America* **116**(5): 3130–3140.

Bingabr, M., Espinoza-Varas, B. and Loizou, P. (2008). Simulating the effect of spread of excitation in cochlear implants, *Hearing Research* **241**(1-2): 73–79.

Blamey, P. J., Swanson, B. A., McDermott, H., Patrick, J. F. and Clark, G. (2006). Sound processor for a cochlear implant, us7082332b2.

Borsboom, M. (1979). *Linear and Nonlinear one-dimensional cochlear model*, PhD thesis, Department of Mathematics and Informatics, Delft University of Technology.

Bruce, L., Irlicht, L., White, M., O'leary, S., Dynes, S., Javel, E. and Clark, G. (1999b). A stochastic model of the electrically stimulated auditory nerve: Pulse-train response, *IEEE Transactions on Biomedical Engineering* **46**(5): 630–637.

Bruce, L., White, M., Irlicht, L., O'leary, S., Dynes, S., Javel, E. and Clark, G. (1999a). A stochastic model of the electrically stimulated auditory nerve: Single-pulse response, *IEEE Transactions on Biomedical Engineering* **46**(5): 617–629.

Brugge, J., Anderson, D., Hind, J. and Rose, J. (1969). Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds, *Journal of Neurophysiology* **32**(3): 386–401.

Buechner, A., Brendel, M., Krüeger, B., Frohne-Büchner, C., Nogueira, W., Edler, B. and Lenarz, T. (2008). Current steering and results from novel speech coding strategies, *Otology and Neurotology* **29**(2): 203–207.

Buechner, A., Frohne-Büchner, C., Gaertner, L., Stoever, T., Battmer, R. and Lenarz, T. (2010). The advanced bionics high resolution mode: Stimulation rates up to 5000 pps, *Acta Oto-Laryngologica* **130**(1): 114–123.

Carlyon, R., Geurts, L. and Wouters, J. (2000). Detection of small across-channel timing differences by cochlear implantees, *Hearing Research* **141**(1-2): 140–154.

Clark, G. (2003). *Cochlear Implants: Fundamentals and Applications*, Springer-Verlag, New York.

Clark, G. (2004). Research directions for future generations of cochlear implants, *Cochlear Implants International* **5**(SUPPL. 1): 2–8.

Cohen, L., Richardson, L., Saunders, E. and Cowan, R. (2003). Spatial spread of neural excitation in cochlear implant recipients: Comparison of improved ecap method and psychophysical forward masking, *Hearing Research* **179**(1-2): 72–87.

De Boer, E. (1980). Short and long waves in the cochlea, *Hearing Research* **2**(3-4): 465–473.

De Boer, E. (1997). Cochlear models and minimum phase, *Journal of the Acoustical Society of America* **102**(6): 3810–3813.

De Jongh, H. (1977). Which stimulus transform controls firing probability of an auditory nerve fibre, *in* E. Evans and J. Wilson (eds), *Psychophysics and Physiology of Hearing*, Academic Press, New York, pp. 149–152.

Diependaal, R. and Viergever, M. (1983). Nonlinear and active modelling of cochlear mechanics: A precarious affair, *in* E. De Boer and M. Viergever (eds), *Mechanics of hearing 1983*, Delft University Press, The Netherlands, pp. 153–160.

Diependaal, R. and Viergever, M. (1989). Nonlinear and active two-dimensional cochlear models: Time-domain solution, *Journal of the Acoustical Society of America* **85**(2): 803–812.

Diependaal, R., Duifhuis, H., Hoogstraten, H. and Viergever, M. (1987). Numerical methods for solving one-dimensional cochlear models in the time domain, *Journal of the Acoustical Society of America* **82**(5): 1655–1666.

Dorman, M., Loizou, P. and Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs, *Journal of the Acoustical Society of America* **102**(4): 2403–2411.

Duifhuis, H. (1988). Cochlear marcomechanics, *in* G. Edelman, W. Gall and W. Cowan (eds), *The Neurosciences Institute publication series*, number 6, Wiley, New York, pp. 189–211.

Duifhuis, H., Hoogstraten, H., Van Netten, S., Diependaal, R. and Bialek, W. (1986). Modelling the cochlear partition with coupled van der pol oscillators, *in* J. Allen, J. Hall, A. Hubbard, S. Neely and A. Tubis (eds), *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin Heidelberg, pp. 290–297.

Faulkner, A., Rosen, S. and Stanton, D. (2003). Simulations of tonotopically mapped speech processors for cochlear implant electrodes varying in insertion depth, *Journal of the Acoustical Society of America* **113**(2): 1073–1080.

Flanagan, J. (1962). Computational model for basilar-membrane displacement, *Journal of the Acoustical Society of America* **34**(8): 1371–1376.

Fletcher, H. (1951). On the dynamics of the cochlea, *Journal of the Acoustical Society of America* **23**(6): 637–645.

Friesen, L., Shannon, R. and Cruz, R. (2005). Effects of stimulation rate on speech recognition with cochlear implants, *Audiology and Neurotology* **10**(3): 169–184.

Friesen, L., Shannon, R. and Slattery III, W. (2000). Effects of electrode location on speech recognition with the nucleus-22 cochlear implant, *Journal of the American Academy of Audiology* **11**(8): 418–428.

Fu, Q. and Shannon, R. (2000). Effect of stimulation rate on phoneme recognition by nucleus-22 cochlear implant listeners, *Journal of the Acoustical Society of America* **107**(1): 589–597.

Fu, Q. and Shannon, R. (2002). Frequency mapping in cochlear implants, *Ear and Hearing* **23**(4): 339–348.

Greenberg, S. (1997). The significance of the cochlear traveling wave and the theories of frequency analysis and pitch, *in* E. Lewis, C. Steele and R. Lyon (eds), *Diversity in Auditory Mechanics*, World Scientific Publishing, Singapore.

Greenberg, S., Poepel, D. and Roberts, T. (1998). A space-time theory of pitch and timbre based on cortical expansion of the cochlear traveling wave delay, *in* A. Palmer, Q. Summerfield, A. Rees and R. Meddis (eds), *Psychophysical and Physiological Advances in Hearing*, Whurr Publishers, London, pp. 293–300.

Greenwood, D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane, *Journal of the Acoustical Society of America* **33**(10): 1344–1356.

Greenwood, D. (1990). A cochlear frequency-position function for several species - 29 years later, *Journal of the Acoustical Society of America* **87**(6): 2592–2605.

Guinan, J. and Peake, W. (1967). Middle-ear characteristics of anesthetized cats, *Journal of the Acoustical Society of America* **41**(5): 1237–1261.

Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., SchöBer, H., Garnham, C. and Anderson, I. (2006). Med-el cochlear implants: State of the art and a glimpse into the future, *Trends in Amplification* **10**(4): 201–220.

Holden, L., Skinner, M., Holden, T. and Demorest, M. (2002). Effects of stimulation rate with the nucleus 24 ace speech coding strategy, *Ear and Hearing* **23**(5): 463–476.

Hubbard, A. and Geisler, C. (1972). A hybrid-computer model of the cochlear partition, *Journal of the Acoustical Society of America* **51**(6): 1895–1903.

Javel, E. (1990). Acoustic and electrical encoding of temporal information, *in* J. Miller and F. Spelman (eds), *Cochlear Implants: Models of the Electrically Stimulated Ear*, number 17, Springer-Verlag, New York, pp. 247–295.

Johnson, D. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones, *J.Acoust.Soc.Am.* **68**(4): 1115–1122.

Kanis, L. and De Boer, E. (1993). Self-suppression in a locally active nonlinear model of the cochlea: A quasilinear approach, *Journal of the Acoustical Society of America* **94**(6): 3199–3206.

Kanis, L. and De Boer, E. (1996). Comparing frequency-domain with time-domain solutions for a locally active nonlinear model of the cochlea, *Journal of the Acoustical Society of America* **100**(4): 2543–2546.

Kates, J. (1991). A time-domain digital cochlear model, *IEEE Transactions on Signal Processing* **39**(12): 2573–2592.

Kiang, N., Watanabe, T., Thomas, E. and Clark, L. (1965). *Discharge patterns of Single Fibers in the Cat's Auditory Nerve*, Vol. 35, Technology Press, Cambridge, Mass.

Kim, D. (1986). A review of nonlinear and active cochlear models, *in* J. Allen, J. Hall, A. Hubbard, S. Neely and A. Tubis (eds), *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin Heidelberg, pp. 239–249.

Kim, D. and Molnar, C. (1979). A population study of cochlear nerve fibers: Comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones, *Journal of Neurophysiology* **42**(1): 16–30.

Kim, D., Molnar, C. and Pfeiffer, R. (1973). A system of nonlinear differential equations modeling basilar membrane motion, *Journal of the Acoustical Society of America* **54**(6): 1517–1529.

Kim, H., Shim, Y., Chung, M. and Lee, Y. (2000). Benefit of ace compared to cis and speak coding strategies, *Advances in Oto-Rhino-Laryngology* **57**: 408–411.

Lesser, M. and Berkley, D. (1972). Fluid mechanics of the cochlea, part i, *Journal of Fluid Mechanics* **51**(Part 3): 497–512.

Liberman, M. (1982). The cochlear frequency map for the cat: Labeling auditory-nerve fibers of known characteristic frequency, *Journal of the Acoustical Society of America* **72**(5): 1441–1449.

Lighthill, J. (1981). Energy flow in the cochlea, *Journal of Fluid Mechanics* **106**: 149–213.

Liu, L., Palmer, A. and Wallace, M. (2006). Phase-locked responses to pure tones in the inferior colliculus, *Journal of Neurophysiology* **95**(3): 1926–1935.

Loizou, P. (1997). Signal processing for cochlear prosthesis: A tutorial review, *IEEE 40th Midwest Symposium on circuits and sytems , Sacramento, CA*, pp. 200–204.

Loizou, P. (1998). Mimicking the human ear, *IEEE Signal Processing Magazine* **15**(5): 101–130.

Loizou, P. (2006). Speech processing in vocoder-centric cochlear implants, *Advances in Oto-Rhino-Laryngology* **64**: 109–143.

Loizou, P., Poroy, O. and Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding, *Journal of the Acoustical Society of America* **108**(2): 790–802.

Mammano, F. and Nobili, R. (1993). Biophysics of the cochlea: Linear approximation, *Journal of the Acoustical Society of America* **93**(6): 3320–3332.

Matthews, J. (1980). *Mechanical Modeling of Nonlinear Phenomena observed in the Peripheral Auditory system*, PhD thesis, Washington University, Saint Louis, Missouri.

McKay, C. and McDermott, H. (1996). The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites, *Journal of the Acoustical Society of America* **100**(2 I): 1081–1092.

McKay, C., McDermott, H. and Clark, G. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees, *Journal of the Acoustical Society of America* **96**(5 I): 2664–2673.

McKay, C., McDermott, H. and Clark, G. (1995). Pitch matching of amplitude-modulated current pulse trains by cochlear implantees: The effect of modulation depth, *Journal of the Acoustical Society of America* **97**(3): 1777–1785.

Middlebrooks, J. (2008). Cochlear-implant high pulse rate and narrow electrode configuration impair transmission of temporal information to the auditory cortex, *Journal of Neurophysiology* **100**(1): 92–107.

Milczynski, M., Wouters, J. and Van Wieringen, A. (2009). Improved fundamental frequency coding in cochlear implant signal processing, *Journal of the Acoustical Society of America* **125**(4): 2260–2271.

Moller, A. (1963). Transfer function of the middle ear, *Journal of the Acoustical Society of America* **35**(10): 1526–1534.

Neely, S. (1981a). Finite difference solution of a two-dimensional mathematical model of the cochlea, *Journal of the Acoustical Society of America* **69**(5): 1386–1393.

Neely, S. (1981b). *Fourth-Order Partition Dynamics for a Two-dimensional Model of the Cochlea*, PhD thesis, University of Washington, Saint Louis, Missouri.

Neely, S. (1993). A model of cochlear mechanics with outer hair cell motility, *Journal of the Acoustical Society of America* **94**(1): 137–146.

Neely, S. and Kim, D. (1983). An active cochlear model showing sharp tuning and high sensitivity, *Hearing Research* **9**(2): 123–130.

Neely, S. and Kim, D. (1986). A model for active elements in cochlear biomechanics, *Journal of the Acoustical Society of America* **79**(5): 1472–1480.

Nogueira, W., Büchner, A., Lenarz, T. and Edler, B. (2005). A psychoacoustic "nofm"-type speech coding strategy for cochlear implants, *Eurasip Journal on Applied Signal Processing* **2005**(18): 3044–3059.

Papoulis, A. (1962). *The Fourier Integral and its Applications*, McGraw-Hill, New York.

Peterson, L. and Bogert, B. (1950). A dynamical theory of the cochlea, *Journal of the Acoustical Society of America* **22**(3): 369–381.

Ranke, O. (1950). Theory of operation of the cochlea: A contribution to the hydrodynamics of the cochlea, *Journal of the Acoustical Society of America* **22**(6): 772–777.

Rhode, W. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the moëssbauer technique, *Journal of the Acoustical Society of America* **49**(4): 1218–1231.

Roberts, T., Ferrari, P. and Poeppel, D. (1998). Latency of evoked neuromagnetic m100 reflects perceptual and acoustic stimulus attributes, *NeuroReport* **9**(14): 3265–3269.

Robles, L., Rhode, W. and Geisler, C. (1976). Transient response of the basilar membrane measured in squirrel monkeys using the mossbauer effect, *Journal of the Acoustical Society of America* **59**(4): 926–939.

Rose, J., Brugge, J., Anderson, D. and Hind, J. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey, *Journal of Neurophysiology* **30**(4): 769–793.

Ruggero, M. (1994). Cochlear delays and traveling waves: Comments on 'experimental look at cochlear mechanics', *Audiology* **33**(3): 131–142.

Ruggero, M. and Temchin, A. (2007). Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods, *JARO - Journal of the Association for Research in Otolaryngology* **8**(2): 153–166.

Russell, I., Richardson, G. and Cody, A. (1986). Mechanosensitivity of mammalian auditory hair cells in vitro, *Nature* **321**(6069): 517–519.

Sellick, P., Patuzzi, R. and Johnstone, B. (1982). Measurement of basilar membrane motion in the guinea pig using the moëssbauer technique, *Journal of the Acoustical Society of America* **72**(1): 131–141.

Shamma, S. (1985a). Speech processing in the auditory system i: The representation of speech sounds in the responses of the auditory nerve, *Journal of the Acoustical Society of America* **78**(5): 1612–1621.

Shamma, S. (1985b). Speech processing in the auditory system ii: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve, *Journal of the Acoustical Society of America* **78**(5): 1622–1632.

Shannon, R. (1983). Multichannel electrical stimulation of the auditory nerve in man. i. basic psychophysics, *Hearing Research* **11**(2): 157–189.

Shannon, R. (1992). Temporal modulation transfer functions in patients with cochlear implants, *Journal of the Acoustical Society of America* **91**(4 I): 2156–2164.

Shannon, R., Cruz, R. and Galvin, J. (2011). Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise, *Audiology and Neuro-Otology* **16**(2): 113–123.

Shera, C. and Guinan, J. (2003). Stimulus-frequency-emission group delay: A test of coherent reflection filtering and a window on cochlear tuning, *Journal of the Acoustical Society of America* **113**(5): 2762–2772.

Siegel, J., Cerka, A., Recio-Spinoso, A., Temchin, A., Van Dijk, P. and Ruggero, M. (2005). Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering, *Journal of the Acoustical Society of America* **118**(4): 2434–2443.

Skrodzka, E. (2005). Mechanical passive and active models of the human basilar membrane, *Applied Acoustics* **66**(12): 1321–1338.

Smith, S. (2011). *Modelling psychoacoustic experiment outcomes from space-time neural patterns for acoustic and electrical hearing*, Master's thesis, University of Pretoria.

Sondhi, M. (1978). Method for computing motion in a two-dimensional cochlear model, *Journal of the Acoustical Society of America* **63**(5): 1468–1477.

Steele, C. and Taber, L. (1979). Comparison of wkb calculation and experimental results for three dimensional cochlear models, *Journal of the Acoustical Society of America* **65**(4): 1007–1018.

Strydom, T. and Hanekom, J. (2011). An analysis of the effects of electrical field interaction with an acoustic model of cochlear implants, *Journal of the Acoustical Society of America* **129**(4): 2213–2226.

Taft, D., Grayden, D. and Burkitt, A. (2009). Speech coding with traveling wave delays: Desynchronizing cochlear implant frequency bands with cochlea-like group delays, *Speech Commununication* **51**(11): 1114–1123.

Taft, D., Grayden, D. and Burkitt, A. (2010). Across-frequency delays based on the cochlear traveling wave: Enhanced speech presentation for cochlear implants, *IEEE Transactions on Biomedical Engineering* **57**(3): 596–606.

Van der Heijden, M. and Joris, P. (2003). Cochlear phase and amplitude retrieved from the auditory nerve at arbitrary frequencies, *Journal of Neuroscience* **23**(27): 9194–9198.

Van der Raadt, M. and Duifhuis, H. (1990). A generalized van der pol-oscillator cochlea model, *in* P. Dallos, C. Geisler, J. Matthews, M. Ruggero and C. Steele (eds), *Lecture Notes in Biomathematics*, Vol. 87, Springer-Verlag, Berlin, pp. 227–234.

Van Netten, S. and Duifhuis, H. (1983). Modelling an active, nonlinear cochlea, *in* E. De Boer and M. Viergever (eds), *Mechanics of hearing 1983*, Delft University Press, The Netherlands, pp. 143–151.

Vandali, A., Sucher, C., Tsang, D., McKay, C., Chew, J. and McDermott, H. (2005). Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies, *Journal of the Acoustical Society of America* **117**(5): 3126–3138.

Vandali, A., Whitford, L., Plant, K. and Clark, G. (2000). Speech perception as a function of electrical stimulation rate: Using the nucleus 24 cochlear implant system, *Ear and Hearing* **21**(6): 608–624.

Viergever, M. (1980). *Mechanics of the inner ear - a mathematical approach*, PhD thesis, Delft University of Technology, Delft UP.

Viergever, M. (1986). Cochlear marcomechanics - a review, *in* J. Allen, J. Hall, A. Hubbard, S. Neely and A. Tubis (eds), *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin Heidelberg, pp. 63–72.

Viergever, M. and Kalker, J. (1975). A two-dimensional model for the cochlea - i. the exact approach, *Journal of Engineering Mathematics* **9**(4): 353–365.

von Békésy, G. (1928). Zur theorie des hörens; die shwingungsform der basilarmembran, *Physikalische Zeitshrift* **29**: 793–810.

von Békésy, G. (1944). Über die frequenzauflösung in der menslechen schnecke, *Acta Oto-Laryngologica* **32**: 60–84.

von Békésy, G. (1947). The variation of phase along the basilar membrane with sinusoidal vibrations, *Journal of the Acoustical Society of America* **19**(3): 452–460.

von Békésy, G. (1949). On the resonance curve and the decay period at various points on the cochlear partition, *Journal of the Acoustical Society of America* **21**(3): 245–254.

von Békésy, G. (1960). *Experiments in Hearing*, McGraw Hill, New York.

Wickesberg, R. and Geisler, C. (1986). Longitudinal stiffness coupling in a 1-dimensional model of the peripheral ear, *in* J. Allen, J. Hall, A. Hubbard, S. Neely and A. Tubis (eds), *Lecture Notes in Biomathematics*, Springer-Verlag, Berlin Heidelberg, pp. 113–120.

Wilson, B., Schatzer, R., Lopez-Poveda, E., Sun, X., Lawson, D. and Wolford, R. (2005). Two new directions in speech processor design for cochlear implants, *Ear and Hearing* **26**(4): 73S–80S.

Xin, J., Qi, Y. and Deng, L. (2003). Time domain computation of a non-linear nonlocal cochlear model with applications to multitone interaction in hearing, *Communications in Mathematical Sciences* **1**(2): 211–227.

Zeng, F. (2002). Temporal pitch in electric hearing, *Hearing Research* **174**(1-2): 101–106.

Zweig, G. (1976). Basilar membrane motion, *Cold Spring Harbor Symposia on Quantitative Biology* **40**: 619–633.

Zwicker, E. and Fastl, H. (1999). Critical bands and excitation, *in* M. Schroeder (ed.), *Springer Series in Information Sciences*, number 6, Springer-Velag, Germany, pp. 149–173.

Zwislocki, J. (1948). Theorie der schneckenmechanik - qualitative and quatitative, *Acta Oto-Laryngologica* **Supplement 72**: 1–76.

Zwislocki, J. (1950). Theory of the acoustical action of the cochlea, *Journal of the Acoustical Society of America* **22**(6): 778–784.

Appendix A: Time delays

This section discusses the definitions of the three time delays mentioned in section 3.6.

In the analysis of linear systems, Papoulis (1962:134-136) describes three delays, namely phase, signal-front and group delays. If a system function is defined as

$$H(\omega) = A(\omega)e^{-j\theta(\omega)}, \tag{A.1}$$

the phase delay is defined as

$$\tau_{Phase} = \frac{\theta(\omega)}{\omega} rad. \tag{A.2}$$

The phase delay is calculated by taking the slope of the line through the origin and the point $(\omega, \theta(\omega))$, which is the place of best frequency on the phase response. Ruggero and Temchin (2007) define the best frequency (BF) as the frequency where the largest response is produced (not taking the stimulus level or cochlear health or development into account), and CF as the BF response to low-level stimuli in mature and healthy cochleae.

The group delay is a measure of the transit time of the signal through a linear system and

gives a measure of the phase distortion. The group delay is defined as the differentiation of the phase response:

$$\tau_{Group} = \frac{d\theta(\omega)}{d\omega} rad. \tag{A.3}$$

The average signal-transit time is determined by calculating the slope of the linear portion of the phase response. Deviations from this value are transformed into deviations from constant group delay, where these deviations in group delay cause signal distortion, just as deviations in linear phase cause distortions. The group delay is thus another measure of linear phase distortion and is determined by calculating the slope of the phase response at a specific frequency ω.

The signal-front delay is the latency of the impulse response (Ruggero and Temchin, 2007; Papoulis, 1962). The signal-front delay can also be calculated from the phase response by calculating the asymptotic slope of the phase response (eq. A.4). The signal-front-delay is calculated as

$$\tau_{Signal-front} = \lim_{\omega \to \infty} \frac{\theta(\omega)}{\omega} rad. \tag{A.4}$$