UNIVERSITY OF SOUTHAMPTON

Adaptation in Multisensory Neurons

by

Xutao Kuang

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Engineering, Science and Mathematics
School of Electronics and Computer Science

December 2008

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Xutao Kuang

The most studied region in the mammalian brain for multisensory integration is the deep superior colliculus (DSC). Neurophysiological experiments have revealed many response properties of DSC neurons, such as cross-modal enhancement (CME) and sub-additive/additive/super-additive operational modes. CME occurs when the response of a multisensory neuron to stimulation in one sensory modality is enhanced, often non-linearly, by temporally and spatially coincident stimulation of a second sensory modality. Response enhancement is frequently larger for weaker input stimuli than for stronger stimuli, a phenomenon known as inverse effectiveness. It is believed that a non-linear, saturating response function may underlie CME associated with inverse effectiveness. We explore this idea in more detail, showing that apart from CME, many other response properties of DSC neurons, including the different dynamic ranges of responses to unimodal and multimodal stimuli and the diverse operational modes, also emerge as a direct consequence of a saturating response function such as a sigmoidal function.

We then consider the question of how the exact form of a candidate, saturating sigmoidal function could be determined in a DSC neuron. In particular, we suggest that adaptation may determine its exact form. Adaptation to input statistics is a ubiquitous property of sensory neurons. Defining the operating point as the output probability density function, we argue that a neuron maintains an invariant operating point by adapting to the lowest-order moments of the input probability distribution. Based on this notion, we propose a novel adaptation rule that permits unisensory neurons to adapt to the lowest-order statistics of their inputs, and then extend this rule to allow adaptation in multisensory neurons, of which DSC neurons are an example. Adaptation in DSC neurons is expected to change the responses of a neuron to a fixed, probe or test stimulus. Such a neuron would therefore exhibit different CME when presented with the same stimulus drawn from different statistical ensembles. We demonstrate that, for suitable selections of test stimuli, adaptation to an increase in the mean, the variance or the correlation coefficient induce consistent changes in CME. By virtue of the robustness of the results, the underlying adaptation notion can be tested in neurophysiological experiments. Finally, it is known that descending cortical projections from the anterior ectosylvian sulcus and the rostral aspect of the lateral suprasylvian sulcus are indispensable for DSC neurons to exhibit CME. The structure of our proposed adaptation rule for multisensory neurons therefore permits us to speculate that the descending cortical inputs to multisensory DSC neurons facilitate the computation of the correlation coefficient between different sensory channels' activities.

Contents

List of Figures

List of Tables

Nomenclature

AES	Anterior Ectosylvian Sulcus
AEV	Visual Region of Anterior Ectosylvian Sulcus
AP5	2-amino-5-phosphonopentanoate
CDF	Cumulative Density Function
CME	Cross-Modal Enhancement
DR	Dynamic Range
DSC	Deep Layers of Superior Colliculus
EEG	Electroencephalograph
FAES	Field Anterior Ectosylvian Sulcus
fMRI	functional Magnetic Resonance Imaging
IC	Inferior Colliculus
ICA	Independent Component Analysis
ICc	Central Nucleus of the Inferior Colliculus
ICx	External Nucleus of the Inferior Colliculus
IE	Inverse Effectiveness
ITD	Interaural Time Difference
LGN	Lateral Geniculate Nucleus
MSS	Modality-Specific Suppression
NL	Nucleus Laminaris
NMDA	N-methyl-D-aspartate
OT	Optic Tectum
PCA	Principal Component Analysis
PDF	Probability Density Function
PET	Positron Emission Tomography
RF	Receptive Field
rLS	Rostral Aspect of Lateral Suprasylvian Area
SC	Superior Colliculus
SIV	Somatosensory Region of Anterior Ectosylvian Sulcus
SSC	Superficial Layers of Superior Colliculus
STS	Superior Temporal Sulcus
VIP	Ventral Intraparietal Area
V1	Primary Visual Cortex

Acknowledgements

I would like to express my sincere gratitude to my supervisors Terry Elliott, Klaus-Peter Zauner and Nigel R. Shadbolt, without whom this work would not have been possible. I would like to thank Terry for his inspiration, patience and guidance. I not only learned technical skills from him, but was also greatly inspired by his enthusiasm and philosophy in research. I want to thank Klaus-Peter, who has shown great patience in answering questions and provided continuous guidance. I would also like to thank Nigel, who has led me into the research area of biological multisensory integration and has been quite supportive throughout my Ph.D..

The process of this Ph.D. was like a long journey, in which I have not felt lonely because of my friends in Southampton. I would like to thank them for making my life colorful. A special thanks goes to Rong Zhang, Ke Li, Bo Hu, Zheng Chen, Yongjian Huang and Pei Zhang for proofreading. I would also like to thank my colleagues in the IAM group and the SENSe group for inspirational discussions and help.

Last, but not least, I wish to thank my family - my parents, Jianzheng Kuang and Qinglan Mo, my grandma, my aunt and my uncle - for their unconditional love and support.

To my parents.

Chapter 1

Introduction

1.1 Why Multisensory Integration?

We are born with multiple senses, each of which provides us with a unique experience of the world. The eyes, for example, receive ambient light signals, which project to the retina of the eyes and form an "image" of the environment. The ears, on the other hand, detect sound waves through air vibration in the eardrums. The skin is sensitive to the external pressure on its surface. Light, sound and pressure differ significantly in their physical properties, and provide distinct kinds of information of the environment. Presumably, the existence of multiple senses is a result of evolutionary development under the functional requirements of an animal in a physical niche.

Sometimes, the perception is based on a blend of multiple senses. Imagine a savoury dish with an attractive appearance and spicy smell. The combination of the aroma, color and taste of the dish produces a particular experience. When we hear a sudden sound, our attention is attracted and we will unconsciously orientate to the sound source to see what it is. These normal phenomena exhibit the interaction between different senses. Although cross-modal interaction seems natural, the underlying mechanisms remain unclear since multisensory processing in the brain is complicated.

Sensory information from different sensory modalities are all encoded in electric signals in the brain. Light signals are converted into electric signals by the photoreceptors in the retina and are passed through the subsequent neural pathway. Sound waves vibrate the basilar membrane in the cochlea where the hair cells attached to the basilar membrane produce the neural signals. Cutaneous receptors respond to the pressure on the skin to generate somatosensory signals. Accordingly, the information encoded in these signals corresponds to the traits of light, sound and pressure and varies significantly. For visual signals, information such as the color of an object is encoded, leading to a perception of the appearance; for auditory signals, the tone and the magnitude of the sound are

encoded; for somatosensory signals, qualities such as the roughness and vibration are recorded. Different sensory signals carry disparate information content.

The detection ranges of the sensors are also diverse, dependent on the mode of signal transmission in the medium and the structure of the sensors. Due to the rectilinear propagation of the light, the detectable areas are those in front of the eyes. Sound propagates as a wave in the air, therefore the ears can receive omnidirectional sound signals surrounding them. For the skin, the detection range is constrained to body surfaces. Moreover, spatial coding in individual sensory systems is also different. The retinas receive spatially ordered light signals from the environment and maintain this spatial representation in subsequent transmission. Similarly, the skin also directly inherits the spatial information from the touched surfaces. However, for the ears, the spatial information of a sound source is not explicitly encoded in the ear. The representations of auditory spatial locations require later neural processing to compare the arrival time of the sound signals between the two ears (interaural time difference).

Despite the great disparity between different sensory modalities, the advantages of integrating cross-modal sensor information in biological systems are straightforward. A single sensory system only processes a limited portion of the environmental information, which is sometimes uncertain and ambiguous. For example, when the environment is dark, the visual information becomes ambiguous and the performance of visual sensors degrades. Moreover, sensory information in the environment is always coupled with noise, and will be further deteriorated by the inherent noise in the nervous system. Combining multiple sensory information not only enriches the information, but also tends to reduce uncertainty and noise through the interaction between disparate sensor information. To this end, it is beneficial to represent the diverse sensory information in some common areas where this information is comparable and where they can enhance or complement each other. For example, a common area for the integration of the auditory and visual spatial information could realize an effective combination where the omnidirectional auditory field complements the visual field and where the high resolution of the visual field enhances that of the auditory field. Consequently, the resultant multisensory spatial representation has an omnidirectional range as well as a high resolution in the front area. In biological systems, the brain areas to receive convergent cross-modal inputs are the main focus of the research on multisensory integration.

1.2 Multisensory Integration in the Brain

The presence of interactions between different sensory modalities have long been indicated by perceptual phenomena such as the synesthesia (Critchley, 1977) and the ventriloquism effect (Bertelson, 1999). Traditional neurophysiological experiments, however, are constrained by experimental techniques and only focus on individual sensory chan-

nels. The shift of emphasis towards an integrative view of sensory processing was due to the developments in experimental techniques, such as positron emission tomography (PET), functional magnetic resonance imaging (fMRI) and electroencephalography (EEG), by which the function of the whole brain can be investigated. This, on the one hand, deepens our understanding of separate sensory systems, and, on the other hand, enables us to study how these sensory systems interact. As a result, multisensory integration was discovered in a large number of animals ranging from insects such as bees and ants (Gronenberg and López-Riquelme, 2004), non-mammalian vertebrates such as owls (Knudsen, 1982) and snakes (Newman and Hartline, 1981), to mammals such as cats (Meredith et al., 1992), monkeys (Wallace et al., 1996) and humans (Calvert and Thesen, 2004). Integrating sensory stimuli across different sensory channels appears to be a general strategy adopted by many biological systems, enabling an animal to speed up sensory processing and enhance the detection of weak signals, and also increasing its competence to live in a complex environment. However, multisensory processes in the brain are very complicated, occurring at various stages of sensory processing, playing different functional roles and exhibiting disparate phenomena. Currently, the understanding of the mechanisms underlying multisensory integration is still at an initial stage.

1.2.1 Behavioural Studies

Behavioural study, concentrating on the perceptual phenomena, is one of the most popular approaches to investigate multisensory integration. Many behavioural results have demonstrated the existence of multisensory integration in the brain. For example, two identical visual targets moving across each other can be perceived as bouncing off or passing through. If a brief sound is presented at the moment the two targets meet, however, our percept will be biased towards bouncing off (Sekuler et al., 1997). This suggests that our perception of an event is based on a combination of multimodal information. In the experiments of sonogenic synesthesia, music induces intense visual experiences or cutaneous paresthesias (Critchley, 1977), indicating the existence of correlations between the sensory systems. These all show that cross-modal sensory processing does interact with each other in the brain.

In addition, behavioural studies also give a hint on the principles and the mechanisms underlying multimodal perception. Several factors that influence multisensory integration have been revealed by behavioural studies. For example, when two or more cross-modal stimuli occur at the same time and at the same location, the reaction time of orientation to the cross-modal stimuli is much shorter than when a single stimulus is present. As the discrepancies in time and space increase, however, the reaction time becomes longer and orientation performance is degraded (Stein et al., 1989). In another experiment, the sensitivity to visual stimuli below response threshold is increased by the simultaneous

presence of an auditory stimulus at the same spatial location. Similarly, this effect is reduced or eliminated when the visual and auditory stimuli are further separated in space or time (Frassinetti et al., 2002). Temporal synchrony and spatial correspondence thus appear to be important factors in target localization. In some other tasks such as cross-modal speech recognition, however, semantic congruence of different sensor information is instead the most influential factor. An example is the McGurk effect, in which an audible syllable "BA" coupled with a videotape of a speaker's mouth showing "GA" is actually perceived as "DA" (McGurk and MacDonald, 1976). Although semantic discrepancy disrupts our perception, the visual image and the sound do not necessarily originate from the same location. It is therefore reasonable to infer that the most influential factors of cross-modal perception depend on the functional roles of the integration.

One of the main concerns about the principles of multisensory integration is how the brain weights information from different sensory channels. Does the integration process bias towards a certain sense? If it does, what is the underlying principle? In the ventriloquism effect, synchronous auditory and visual stimuli presented in slightly different spacial locations give the illusion that the sound is coming from the position of the visual stimulus (Bertelson, 1999). In this case, auditory perception is significantly biased by the visual signal. In the freezing phenomenon, when some rapidly changing visual display is presented, an abrupt sound could "freeze" a display synchronized with the sound so that this "frozen" visual display was perceived brighter and sustaining for a longer time (Vroomen and de Gelder, 2000). In this experiment, sound instead alters the visual perception. Therefore, it appears that the brain would dynamically weight cross-modal sensory information based on different situations. It is possible that multisensory integration weights a sensory modality according to its reliability in a particular context (Welch and Warren, 1980). As in the above examples, for localization, visual information predominates in the perception because it has a higher spatial resolution; in the freezing effect, auditory signals influence the visual perception because they have a higher temporal resolution. There is evidence showing that, in a task to discriminate the widths of bimodal visual-haptic bars, the nervous system actually weights visual and haptic information according to the variances of the sensory information (Ernst et al., 2002).

Cross-modal plasticity to sensory modification or loss has also been investigated in behavioural studies. It has been shown that the performance of sound localization task is better in cats that were visually deprived from birth than sighted cats (Rauschecker and Kniepert, 1994). Moreover, identical results have also been found in visually deprived ferrets (King and Parsons, 1999) and blind humans (Muchnik et al., 1991). This indicates the occurrence of some plastic processes to enhance the capability of the remaining senses that belong to a different modality from the deprived sensors. Furthermore, in contrast to congenitally blind people, the performance of late blind and sighted partic-

ipants are equally affected by crossing the hands in the task of judging the temporal order of two tactile stimuli applied to either hand (Roder and Neville, 2003). Accordingly, there appears to be a "critical period" for this adaptive re-arrangement to take place. The performance of partially blind people in sound localization tasks, however, is worse than normally sighted people, indicating that compensatory plasticity in partial sensory loss might induce a conflict between vision and hearing (Lessard et al., 1998). In contrast to cats, ferrets and humans, blind-reared barn owls perform worse than normal owls in sound localization behaviour (Knudsen et al., 1991). For barn owl, visual experience seems to be necessary for normal development of auditory processing. This implies that either the principles of cross-modal plasticity are not the same for different species or the neural structures that exert influence on a sound localization task are different between mammalian and non-mammalian vertebrates.

1.2.2 Neurophysiological Studies

With the help of experimental techniques such as electrophysiological techniques, fMRI and EEG, our knowledge of the neural processing underlying multisensory integration has greatly increased. Complementary to behavioural studies that investigate the brain at the perceptual level, neurophysiological studies provide substantial results on the reactions of the brain from the single cell level, to neural population, and to the whole brain level.

Neuroanatomy

Traditionally, sensory processing in the brain has been considered to conform to a hierarchical organization, starting from the primary sensory-specific cortices, through secondary sensory cortices, sometimes even to tertiary sensory cortices, and then converging to multimodal association areas (Pandya and Kuypers, 1969; Jones and Powell, 1970). In this organization, multisensory integration occurs at a relatively "late" stage. During that traditional period, research emphasized the role of multimodal areas as the final stage in sensory information processing that are responsible for high-level activities such as language and cognition. However, with the development of more powerful experimental techniques, later research has provided evidence for a parallel and divergent organization, showing that some interactions between different sensory channels occur at an early stage of processing. For example, there are projections from the auditory parabelt, a part of the auditory cortex, to the primary visual cortex (V1) (Rockland and Ojima, 2001; Falchier et al., 2002). Although these projections from auditory cortex to visual cortex are sparse and might not have a significant impact on information processing in V1, they might modulate visual responses (Felleman and Van Essen, 1991). Moreover, in visual cortex, neural responses to a visual stimulus can be enhanced by a tactile cue (Macaluso et al., 2000). These results suggest that either there are feedback influences from the multisensory areas to these unimodal areas or that these unimodal

areas are able to access each other directly. Multisensory integration in the brain is therefore very complex, occurring at various levels ("early" and "late") of the information processing pathway and is mediated by feedforward, feedback and reciprocal connections with other areas.

Neuronal Responses

Neurons are the basic units of the brain which generate perceptions and produces signals to execute behaviours. The response properties of a single neuron might therefore to a certain extent contribute to the features exhibited by behavioural studies. As we will find out, some characteristics of neuronal responses described below do correlate with experimental results at the behavioural level. For example, when two or more sensory cues occur in temporal and spatial proximity, the firing rate of superior colliculus (SC) multisensory neurons can greatly exceed the maximum of those evoked by individual sensor stimuli (Meredith and Stein, 1986), corresponding to the improvement of orientation behaviour towards cross-modal stimuli (Stein et al., 1989). Moreover, this response enhancement is often maximal when the unimodal responses are weakest, a phenomenon called inverse effectiveness. On the other hand, enhancement in the firing rate of the SC neurons is greatly reduced, even eliminated when the spatial or temporal discrepancy of the cross-modal stimuli increases (Meredith et al., 1987; Meredith and Stein, 1996). Neurons in the superior temporal sulcus (STS), however, are not so constrained by precise spatial coincidence as superior colliculus neurons, but are mainly influenced by semantic congruency as well as temporal synchrony. They exhibit stronger responses to matching audiovisual speech and weaker responses to conflicting audiovisual speech, compared with the sum of their unimodal responses (Ojanen, 2005). Some studies suggest that inverse effectiveness could also be shown by neurons in STS (Callan et al., 2003), that is, weaker bimodal stimuli result in larger response enhancement. Since STS has been identified in tasks such as object identification and speech recognition (Beauchamp et al., 2004; van Atteveldt et al., 2004), where the sematic matching between the signals is more important than their spatial coincidence, the principles of multisensory integration are presumably a direct outcome of the functional requirements.

Development of Multisensory Integration

Concerning the development of multisensory processing in the brain, there are traditionally two radically different views (Lewkowicz, 1994). One view suggests that initially different sensory channels are completely separate without interactions, and that multisensory neurons appear later during postnatal experience. The other view suggests that sensory neurons are initially multisensory and then develop to be modality-specific with the maturation of the nervous system. Neurophysiological experiments on the development of neural responses in SC for cats show that SC neurons in kittens exclusively respond to somatosensory stimuli for the first several days after birth, and SC neurons that respond to cross-modal stimuli appear later during postnatal development (Wallace

and Stein, 1997). Moreover, the integrative ability to induce multisensory enhancement is not concomitant with the appearance of multisensory neurons, but is acquired at a later stage. Experiments on SC of rhesus monkeys, however, show that multisensory neurons in SC can be observed at birth, although with a percentage much lower than adult (Wallace and Stein, 2001). Again, these multisensory neurons only exhibit enhanced responses to cross-modal stimuli at a later stage. One possible cause of the different development chronologies of multisensory neurons between cats and monkeys is the specific functional needs and behavioural requirements at birth. The developmental processes of multisensory integration vary across different species.

Cross-modal Plasticity

Connection patterns in multisensory processing pathways are able to undergo rearrangement to accommodate the changes in the sensors. For example, neurons in the optic tectum (OT) are arranged in a topographic organization in which neighbouring neurons respond to cross-modal stimuli in adjacent spatial locations, and accordingly form aligned cross-modal spatial representations (Knudsen and Brainard, 1995). Prism glasses that shift the visual field in newborn barn owls thus induce a conflict between the visual and auditory topographic representations in optic tectum. Electrophysiological experiments have shown that this conflict is compensated by rewiring projections from the central nuclei to the external nuclei of inferior colliculus along the auditory pathway, which re-align the visual and auditory spatial maps in the OT (Feldman and Knudsen, 1997). In addition, the finding that brain areas initially dominated by visual stimuli are activated when blind participants read Braille suggests that neurons that exclusively respond to visual stimulation can be stimulated by other sensory modalities when vision is deprived (Eimer, 2004). Therefore, when one sensory input is permanently lost, frequent usage of other intact sensory modalities triggers neural re-arrangements that result in the recruitment of the deprived sensory areas by the intact sensors. A recent study showing that multimodal brain regions in the congenitally blind exhibit stronger responses than normally sighted people in an auditory localization task also indicates the existence of cross-modal plasticity (Wessinger et al., 1999). For blind-reared barn owls, however, the auditory spatial representation in OT is degraded compared with normally sighted owls, resulting in much less precise sound localization (Knudsen et al., 1991). This performance is in contrast to mammals such as cats and monkeys (Rauschecker and Kniepert, 1994; King and Parsons, 1999). One possible reason for this difference might be that, for mammals, the changes in subcortical areas such as SC as a result of sensory loss can be tamed by their highly-plastic neocortex, whereas, for non-mammalian vertebrates such as the barn owls, disruption in the development of OT could not be properly compensated (Rauschecker, 2004).

1.3 Thesis Outline

As discussed in previous sections, both behavioural and neurophysiological studies indicate that multisensory processes in the brain are complicated, and the mechanisms underlying these processes remain unclear. This thesis aims to contribute to the understanding of multisensory integration based on theoretical studies inspired by existing experimental results. Among the various multisensory areas in the brain, we narrow our focus to the superior colliculus (SC), a subcortical area that receives visual, auditory and somatosensory inputs and assists in orientation behaviour, particularly its basic neuronal response properties. We choose SC for three reasons. Firstly, SC is the most studied multisensory area in the brain, and thus offers a rich source of information (Stein and Meredith, 1993). The extensive experimental results on SC can provide the most detailed information not only to inspire theoretical studies but also to validate the resulting model. Secondly, information processing in SC is relatively low-level and it thus appears to be a suitable choice with which to start. Moreover, it has been reported that the response properties of SC neurons correspond well to the performance of orientation behaviour (Stein et al., 1989). Investigating the neuronal responses of SC might therefore contribute to the understanding of multisensory integration at the behavioural level. Thirdly, multisensory integration in SC exhibits some properties that also apply to other multisensory processes in the brain. For example, in addition to SC, multisensory enhancement associated with inverse effectiveness has also been observed in STS for audiovisual speech stimuli (Ojanen, 2005) and in the integration of taste and smell (Dalton et al., 2000). The study of SC might therefore contribute to the exploration of other multisensory processes in the brain.

In Chapter 2, we review the experimental results on SC, in particular the response properties of neurons in the deep layers of SC (DSC), which receive convergent input from different sensory modalities. For example, DSC neurons exhibit response properties such as cross-modal enhancement (CME), multisensory depression and modality-specific suppression (MSS). Moreover, multisensory responses of DSC neurons are significantly influenced by afferents from cortical areas such as anterior ectosylvian sulcus (AES) and the rostral aspect of the lateral suprasylvian (rLS) area. We also review findings on the development of the response properties of DSC neurons.

As we will propose an adaptation rule for DSC neurons, we also review in Chapter 3 experimental studies on the adaptation of unimodal sensory neurons to their input statistics. We discuss statistical properties of natural signals, such as natural images and natural sounds, followed by a discussion of the adaptation of sensory neurons to the input mean, variance and other statistical properties. We finally review the proposed information principles underlying these adaptations.

To account for multisensory processes in DSC, computational and mathematical models have been proposed. We introduce several representative models in this area in Chap-

ter 4, and briefly discuss their advantages and disadvantages. We suggest that these theoretical studies on DSC neurons are questionable.

As an example, we analyse the model proposed by Patton and Anastasio (2003) to explain the CME and MSS properties of DSC neurons (Chapter 5). This model is constructed based on Bayes' theorem, and it exhibits MSS if the covariance between two sensory inputs approaches the variances when firing spontaneously. However, our numerical and mathematical studies of the model reveal that the MSS property is in fact produced by particular choices of parameters and inputs. This suggests that the model is not an appropriate one to account for the mechanisms underlying all DSC neuronal responses.

We discuss in Chapter 6 that, in fact, many response properties naturally emerge if we model the responses of DSC neurons under a non-linear, saturating function, such as the sigmoidal function, extending the notion of Stanford et al. (2005). For example, the sigmoidal transfer function contributes to a simple interpretation of the CME property associated with inverse effectiveness, the sub-additive/additive/super-additive responses and the four different operational modes in DSC neurons.

We then propose an adaptation rule in Chapter 7 for a sensory neuron, both unimodal and multimodal, responding under a sigmoidal transfer function. Defining the probability density function of a neuronal response as the operating point, we provide mathematical analysis showing that a sensory neuron maintains an exactly or approximately invariant operating point under our adaptation rule when the input statistics change. We also present several examples of adaptation to input distributions for both unimodal and bimodal neurons.

Based on the adaptation rule for multimodal inputs, we derive predictions about the impact of adaptation to different statistical parameters, including the mean, the variance and the correlation coefficient, on the CME property of DSC neurons (Chapter 8). Our results show that, with suitable choices of bimodal test stimuli, consistent changes in the CME measure could be observed when the statistical parameters increase. According to these predictions, neurophysiological experiments could be conducted to verify the adaptation notion.

Finally, we discuss in Chapter 9 several issues surrounding our adaptation model. We first give a brief summary of the work we have done and suggest extensions and applications of our model. We then discuss issues such as the interpretation of multisensory responses in DSC neurons, the implication of our model for neuronal computation and model predictions for experimental tests. We also reflect on the current status of experimental and theoretical studies.

Chapter 2

Superior Colliculus

2.1 General Introduction

Superior colliculus (SC) is a subcortical area in the mammalian brain that, together with inferior colliculus, comprises the tectum (Fig. 2.1). It receives convergent inputs from the visual, auditory and somatosensory channels and is considered to integrate these sensory inputs. Based on cross-modal information, SC assists the orientation behaviours of an animal such as saccades, head and limb movements. In non-mammalian vertebrates, there is a comparable structure called optic tectum (OT) that subserves the same functional role as SC. Compared with other multisensory areas in the brain, SC/OT has received the most attention, and a large number of experiments, both behavioural and neurophysiological, have been carried out to study them.

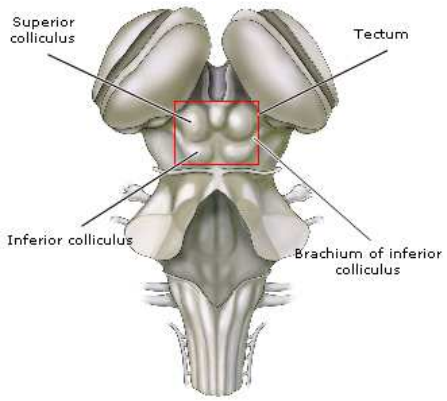


FIGURE 2.1: The tectum. From Purves et al. (2004).

SC is divided into seven layers, with the top three layers generally referred to as the superficial layers and the bottom four layers referred to as the deep layers (Kanaseki and Sprague, 1974). This conventional division differentiates the layers according to their neuronal properties and physiological characteristics. The superficial layers of SC

(SSC) receive projections from the retina and the visual cortex, and neurons in these layers respond exclusively to visual stimuli. In contrast, neurons in the deep layers of SC (DSC) respond to visual, auditory and somatosensory stimuli (Stein and Meredith, 1993). Our focus is therefore on DSC, which is considered as the region in SC where multisensory integration occurs. Although multisensory neurons in SC are located in its deep layers, there are also unisensory neurons in DSC. For example, electrophysiological experiments in cats have investigated the composition of DSC based on the sensory inputs to which DSC neurons respond (Kadunce et al., 1997). The results show that about 42% of DSC sensory neurons are unisensory, responding exclusively to one of the three sensory modalities, and the other 58% could be triggered by cross-modal stimuli (Fig. 2.2). Among the multisensory neurons, the largest portion of neurons respond to bimodal visuoauditory stimuli, with a percentage up to about 35%. This is intuitive since visual and auditory sensors in cats receive a large amount of stimuli. There are also trimodal neurons, occupying about 8% of DSC neurons. Although a similar distribution of multisensory and unisensory neurons could probably be observed in other animals, the actual modalities that DSC neurons respond to certainly vary across different species. Rattlesnake, for example, extensively utilizes its heat sensor to localize a prey, and the majority of its multisensory neurons are visual-infrared neurons (Stein and Meredith, 1993).

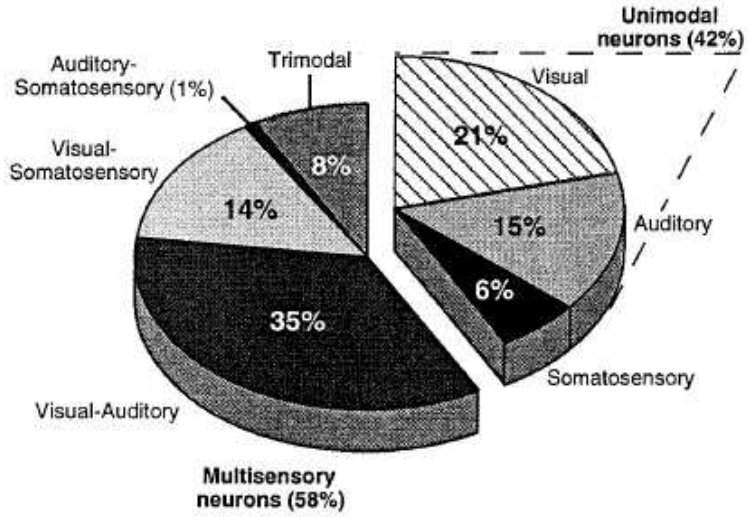


FIGURE 2.2: The composition of DSC neurons of a cat. From Kadunce et al. (1997).

2.2 Anatomical Organization

DSC is connected with a large range of brain structures, receiving ascending afferents from sub-cortical areas and sensory organs, descending afferents from cerebral cortex, and sending efferents to other regions of the brain. Varying in function and structure, these regions form a complex connection pattern together with DSC (see Table 2.1).

	Ascending	Descending
Visual	Retina Pretectum Ventral lateral geniculate n.	Posterior & lateral suprasylvian cortex Anterior ectosylvian sulcus Orbital cortex
Auditory	Exterior n. inferior colliculus N. brachium inferior colliculus N. sagulum Dorsomedial periolivary n. N. trapezoid body	Field AES (anterior ectosylvian sulcus)
Somatosensory	Sensory trigeminal complex Dorsal column n. Lateral cervical n. Spinal cord, dorsal horn	SIV (anterior ectosylvian sulcus) Para SIV (anterior ectosylvian sulcus) Rostral lateral suprasylvian cortex
Motor	Substantia nigra Entopeduncular n. Medial, interposed, and lateral n. of deep cerebellum Perihypoglossal n. Zona incerta	Frontal eye fields

TABLE 2.1: Afferents to the deep layers of the cat superior colliculus. From Stein and Meredith (1993).

In comparison to SSC, DSC receives relatively sparse projections from the retina. The majority of visual projections to DSC originates from cerebral cortex such as the visual area of anterior ectosylvian sulcus (AES) (Mucke et al., 1982) and lateral suprasylvian cortex (Baleydier et al., 1983). Although there are also connections between SSC and DSC (Behan and Appell, 1992; Moschovakis and Karabelas, 1985), the function of these connections is still unclear. Descending auditory projections to DSC comes exclusively from the auditory area in AES (Meredith and Clemo, 1989), while the ascending afferents originate from several sources, including the external nucleus of inferior colliculus (ICx). Since IC neurons are topographically organized to form a map of auditory space (Knudsen and Konishi, 1978), the projections from ICx to DSC are considered to contribute to the formation of the auditory receptive field of DSC neurons. Somatosensory corticotectal inputs to DSC mainly arise from the somatosensory area in AES (McHaffie et al., 1988). Other descending somatosensory projections include the rostral lateral suprasylvian (rLS) area (Stein and Gaither, 1983). There are also a number of ascending somatosensory afferents from areas such as sensory trigeminal complex and dorsal column nuclei (Edwards et al., 1979). In addition to sensory afferents, motor afferents to DSC also exist. For example, the frontal eye fields are the most well-known motor projection to DSC in cat and primate (Kawamura and Konno, 1979; Leichnetz et al., 1981). Other oculomotor-related areas that project to DSC include the basal ganglia (Moschovakis and Karabelas, 1985), zona incerta, thalamic reticular nucleus and nucleus of the posterior commissure (Stein and Meredith, 1993).

DSC efferents are conventionally considered to project in four pathways: ascending, commissural, crossed and uncrossed descending pathways. The ascending outputs project to thalamus and presumably alert higher-level processing to the changes in the functional condition of DSC neurons; the commissural projection connects with the opposite

SC and coordinates the operations between them; both the crossed and uncrossed descending efferents project to brain stem and spinal cord, and are generally involved in initializing behaviours such as eye, head and limb movement (Stein and Meredith, 1993).

2.3 Spatial Maps

2.3.1 Sensory Maps

Each DSC neuron responds to stimuli in a specific spatial region called the "receptive field" (RF) of the neuron. Since DSC neurons are organized in a way that neighbouring neurons have adjacent RF for visual stimuli, a population of DSC neurons represent a continuous visual spatial map. In addition, DSC neurons also form an auditory spatial map, although each with a larger RF and thus lower "resolution" compared to the visual spatial map. The somatosensory RF of a DSC neuron refers to the region of body surface on which a stimulus induces the neuron's response. Analogously, neighbouring DSC neurons have adjacent somatosensory RF, resulting in a somatosensory map of the body surface. Although not in exact alignment, these sensory spatial maps align with each other to form a coherent multisensory map (see Fig. 2.3). Accordingly, a stimulus in a specific spatial location, no matter whether it is visual, auditory or somatosensory, triggers the same DSC neuron.

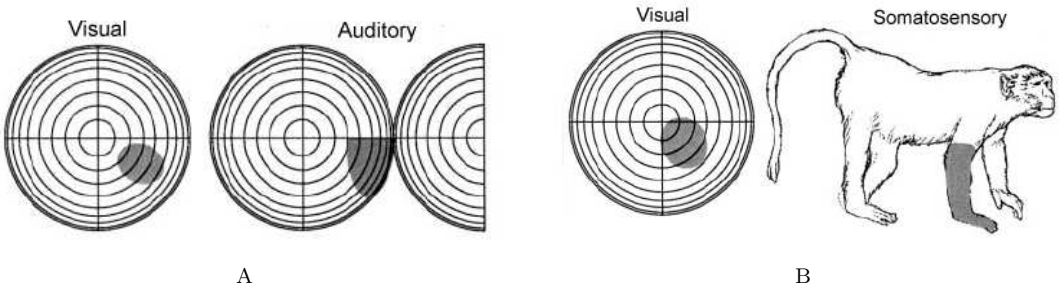


FIGURE 2.3: Receptive filed of DSC neurons in monkey. The shaded area represents the receptive field for (A) a visual-auditory neuron and (B) a visual-somatosensory neuron. From Wallace and Stein (2001).

Research on spatial maps has studied issues such as the development of spatial maps, compensation to sensor modification and dynamic mapping between different modalities. During the development of spatial maps, a rough map in OT firstly appears through the establishment of afferent projections under genetic mechanisms, after which the spatial maps are refined and aligned under activity-dependent mechanisms (Knudsen and Brainard, 1995). This developmental process in DSC has been reported to be highly plastic and the way cross-modal spatial maps align depends on the correlation between cross-modal stimuli in the environment (Wallace and Stein, 2007). When the cross-modal spatial maps are disrupted by, for example, sensor modification, re-alignment processes

occur to compensate for this. For example, shifting the visual field of juvenile barn owls by a prism results in a misalignment between the visual and auditory map in OT. This is later compensated by a re-arrangement of connections from the central nucleus to the external nucleus of the inferior colliculus in the auditory pathway (Brainard and Knudsen, 1993). Furthermore, since different sensors appear to have different reference frames, for example, visual information is eye-centred while auditory information is head-centred, conflict in the multisensory map induced by movement of the eye or the head needs to be reconciled dynamically. Knudsen and Brainard (1995) suggested that the reference transformation from eye-centred to head-centred is accomplished within the OT through an intermediate transformation.

2.3.2 Motor Maps

The primary functional role of SC is to translate the sensory signals into motor signals to orientate the sensory organs towards the stimulus. Some DSC neurons have been reported to be involved in the generation of these motor movements. For example, there are oculomotor DSC neurons that send signals to trigger saccadic eye movement. Activation of oculomotor DSC neurons determines how the eye should move: the frequency of spikes is related to the velocity of the movement (Berthoz et al., 1986), and the activated neuron determines a particular direction and amplitude. Analogous to the sensory spatial representations in DSC, experiments in cat SC reveal an eye movement map represented by these neurons (McIlwain, 1990), in which neighbouring neurons induce eye movements to foveate adjacent spatial locations. In addition to oculomotor neurons, there are neurons in DSC that evoke ear movements towards a sound via polysynaptic connections to the pinna muscles. A well-ordered ear movement map has also been revealed in cat (Stein and Clamann, 1981).

In analogy to the alignment of cross-modal sensory spatial maps, some studies have shown that motor maps for different sensory organs are aligned with each other. For example, the eye movement map and the ear movement map are in register so that microstimulation at one point of DSC evokes both eye and ear movements towards approximately the same spatial location (Stein and Clamann, 1981). This multimotor map in the DSC neurons could induce coordinated movements of the eyes and ears. Furthermore, according to the study by McIlwain (1990), the reported eye movement map is in general register with the visual sensory map, indicating that sensory and motor representations might comply with a uniform topographical scheme. The formation of an integrated multisensory-multimotor map is presumably the simplest and most economical way to translate cross-modal sensory signals into motor signals to induce appropriate motor responses (Stein and Clamann, 1981).

2.4 Neuronal Response Properties

There has been a large amount of work to study the response properties of DSC neurons. Research has found that DSC neuronal responses exhibit some properties distinct from those of unimodal neurons. In the following sections, we introduce these reported response properties of DSC multisensory neurons.

2.4.1 Enhancement and Depression

Compared with unimodal neurons, multisensory neurons in DSC receive convergent input from different sensory modalities. The responses of DSC multisensory neurons therefore exhibit some properties different from unimodal neurons when there are cross-modal stimuli (Stein and Meredith, 1993; Stein et al., 2004).

One important response property of DSC neurons is multisensory enhancement, or cross-modal enhancement (CME), in which the response of a DSC neuron to multimodal stimuli is much larger than its response to unimodal stimuli (Figs. 2.4B and 2.4C). Furthermore, it has been demonstrated that a large enhancement is achieved with weak stimuli, a property called inverse effectiveness (Fig. 2.4D). CME is robustly observed and prevails in the majority of multisensory neurons. This enhancement phenomenon is advantageous in the sense that the probability of detecting a weak stimulus can be enhanced by the presentation of another stimulus from a different sensory modality. Behavioural studies in cat have shown that the performance of orientation behaviour to a weak visual stimulus is indeed improved with the presentation of an auditory stimulus (Stein et al., 1989).

To characterize the properties of multisensory integration, several metrics have been proposed. One commonly used metric is the interactive index, or the cross-modal enhancement (CME) index as we will call it, generally defined as the difference between the response of a DSC neuron evoked by cross-modal stimuli and its response evoked by the most effective unimodal stimulus (Meredith and Stein, 1983). Consider the case of a DSC bimodal neuron receiving visual and auditory inputs. The CME index is then defined by

$$\%\text{CME} = \frac{Bi - \max\{V, A\}}{\max\{V, A\}} \times 100\%, \tag{2.1}$$

where Bi represents the bimodal response, V is the visual response and A is the auditory response. The CME index characterizes whether a bimodal response exhibits enhancement (CME$>$ 0) or depression (CME$<$ 0).

Another popular metric is the additivity index, which is calculated by the difference between multisensory response and the sum of unimodal responses (King and Palmer, 1985). In this definition, however, the sum of unimodal responses V and A double the

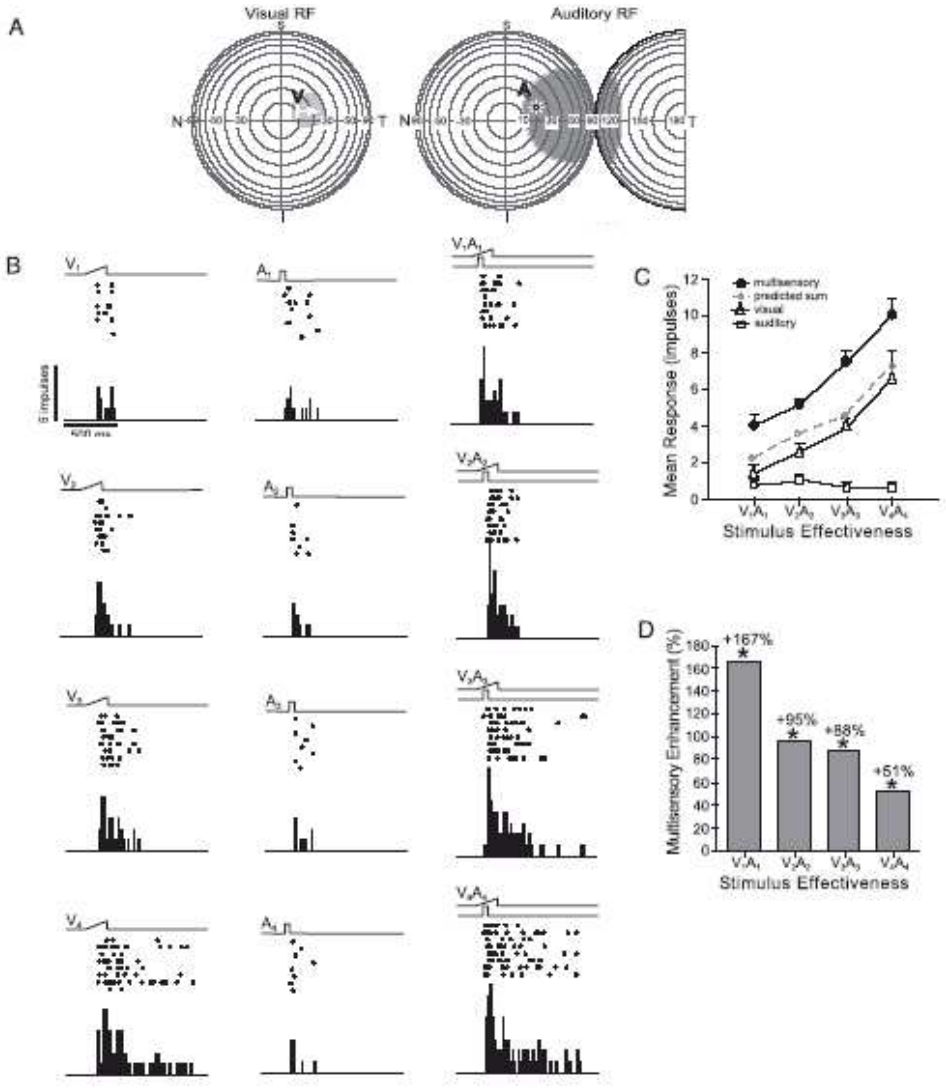


FIGURE 2.4: Multisensory enhancement associated with inverse effectiveness for spatially coincident stimuli. Each of the concentric circles in the figures represents 10^o, N represents nasal and T represents temporal. (A) Visual and auditory receptive fields (shaded areas) of the neuron. The icons labeled V and A represent the locations of visual and auditory stimuli. (B) Ramp wave for visual stimulus and square wave for auditory stimulus are above rasters and peristimulus time histograms. In the four sets of stimuli, the intensity of visual stimuli is increased (V_1-V_4). Responses to bimodal stimuli (For example, V_1A_1) are also shown in the third column. (C) Response curves for the four sets of unimodal visual stimuli, unimodal auditory stimuli, the sum of visual and auditory stimuli and bimodal stimuli. (D) The CME index calculated under Eq. (2.1) for the four sets of stimuli. From Jiang et al. (2006).

influence from spontaneous activities. To overcome this problem, the effect of spontaneous activities should be eliminated from the stimulus-evoked activity (Stanford et al., 2005). A modified version of the additivity index can thus be defined as

$$\%\text{ADD} = \frac{|Bi| - |V| - |A|}{|V| + |A|} \times 100\%, \tag{2.2}$$

where the bars represents the neuronal responses from which the spontaneous activities have been eliminated. According to the additivity index, a bimodal response can be characterized as super-additive (ADD>0), additive (ADD≈ 0), or sub-additive (ADD<0).

From Eq. (2.1), we can see that the minimum value of the CME index is -100%, whereas the maximum CME could theoretically go to infinity. To have a symmetric index describing CME and also eliminate the influences of spontaneous actitivies, a normalized form of the CME index is proposed (Avillac et al., 2007),

$$\%\text{CME}_n = \frac{|Bi| - \max\{|V|, |A|\}}{|Bi| + \max\{|V|, |A|\}} \times 100\%, \tag{2.3}$$

which is within the range from -100% to 100%. A normalized form of the additivity index is also proposed (Avillac et al., 2007),

$$\%\text{ADD}_n = \frac{|Bi| - |V| - |A|}{|Bi| + |V| + |A|} \times 100\%. \tag{2.4}$$

An experiment to study multisensory integration in the ventral intraparietal (VIP) area shows that, calculated from the responses of VIP neurons, normalized forms of the CME index and the additivity index are closely related in a way that a depression is always associated with sub-additive interactions (Avillac et al., 2007). However, as we will discuss in Chapter 6, these two indices are in fact inherently correlated such that they always exhibit the reported association, even for random choices of Bi, V and A.

One of the requirements for DSC neurons to exhibit enhancement is spatial proximity (Stein and Meredith, 1993; Meredith and Stein, 1996). As introduced in the previous sections, DSC neurons are topographically organized, forming a spatial map for each sensory modality. Moreover, cross-modal spatial maps are in register with each other. Therefore, spatially coincident cross-modal stimuli trigger the same DSC neuron and induce enhancement. When the spatial discrepancy between these two stimuli increases, however, there is no enhancement. Further separation of the cross-modal stimuli induces multisensory depression (Kadunce et al., 1997), a phenomenon in which the response of a DSC neuron to multisensory stimuli is weaker than its unimodal responses (Fig. 2.5). The transition from enhancement to depression for large spatial discrepancy is intuitive in the sense that the closer two stimuli are, the more likely that they originate from the same object or event. On the other hand, spatially separated sensory stimuli are

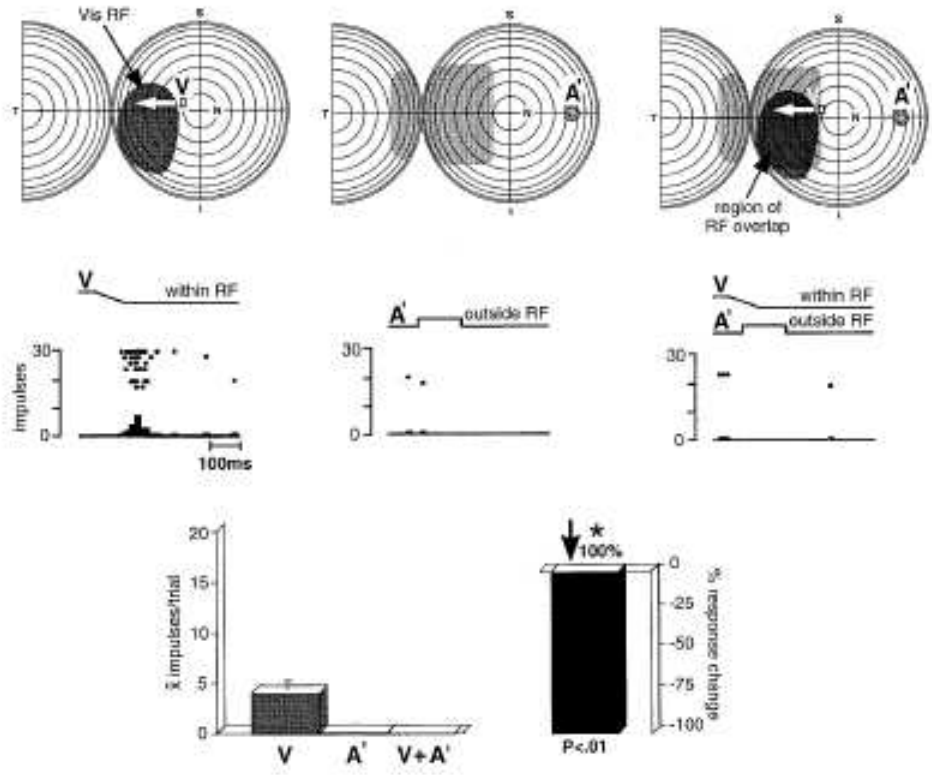


FIGURE 2.5: Multisensory depression for spatially disparate stimuli. The conventions are the same as in Fig. 2.4. From Kadunce et al. (1997).

more likely to originate from different objects, therefore there is no need to integrate them. The neural mechanism behind multisensory enhancement and depression can be explained in terms of the excitatory and inhibitory regions of DSC neurons (Stein and Meredith, 1993). When cross-modal stimuli occur in spatial proximity, they are located in the RF of the same DSC neuron and coactivate the DSC neuron to produce CME. However, if one stimulus is in the RF of a DSC neuron while the other falls outside, there would be no interaction, and thus no enhancement. Furthermore, if one stimulus is in the inhibitory region of the DSC neuron, it suppresses the neuronal response to another stimulus within the RF, leading to response depression. Neurophysiologically, multisensory depression, or the inhibitory regions, might be mediated by lateral inhibitions between DSC neurons.

In addition, temporal proximity between cross-modal stimuli also affects the integration of these sensory stimuli (Meredith et al., 1987; Stein and Meredith, 1993). The occurrence of cross-modal stimuli needs to be close in time to induce interaction. When the onset time between two stimuli is larger than the "temporal window" of a multisensory neuron, which denotes the longest interval for cross-modal stimuli to interact, a DSC neuron would consider them as two separate events. The temporal window is important because the speeds of signal transmission are not the same for different sensory channels. For example, the time for a light signal to reach our eyes is much shorter than the time

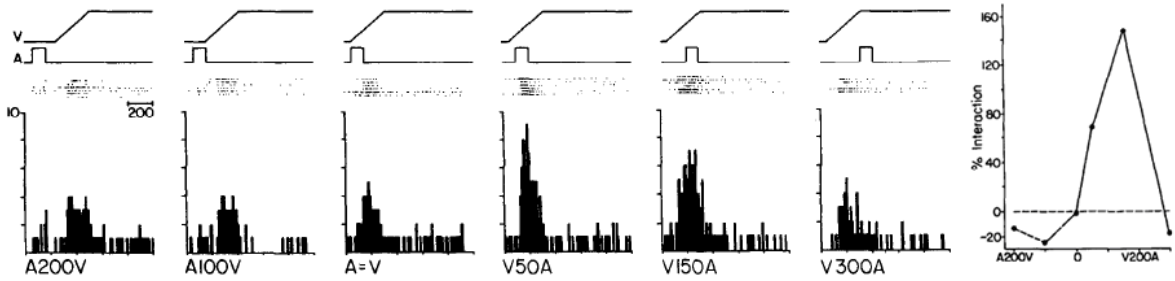


FIGURE 2.6: Multisensory enhancement is converted to multisensory depression by varying the temporal interval between stimuli. As the convention in Fig. 2.4, the waves for bimodal stimuli are shown above the rasters and histograms. The text under the histogram represents the time interval between the onset of two stimuli. For example, A200V represents that the auditory stimulus was presented 200 ms before the visual stimulus. The rightmost figure shows the enhancement curve for the stimuli with varying time interval. From Meredith et al. (1987).

for a sound signal to reach our ears, resulting in a temporal latency. Moreover, the relay time for cross-modal sensory signals varies because they are transmitted via different pathways to DSC neurons. For example, an auditory stimulus near the ear takes about 13 ms to reach a DSC neuron, whereas a visual stimulus near the eye takes 65-100 ms to reach the same neuron (Stein and Meredith, 1993). It is therefore also natural to observe that the maximum levels of multisensory interactions do not necessarily require exact matching between the onset of cross-modal stimuli, but occur for an overlapping of their peak activity periods (Meredith et al., 1987). On the other hand, analogous to the spatial inhibitory region, experiments have also revealed the presence of postexcitatory inhibitory periods in some DSC neurons, so that varying the temporal interval between two stimuli would convert an enhancement to a depression (Fig. 2.6). However, changing the temporal interval does not convert a depression induced by spatially disparate stimuli to an enhancement.

Neurophysiological experiments have reported within-modality suppression, or modality-specific suppression (MSS), in which the response of a neuron to a sensory stimulus within its RF is suppressed by the presence of another stimulus of the same sensory modality outside the RF. MSS has been observed more often than multisensory depression. Specifically, when multisensory depression is demonstrated, MSS could also be observed. However, when a neuron exhibit MSS, it does not necessarily exhibit multisensory depression (Kadunce et al., 1997). Therefore, the mechanisms underlying multisensory depression and MSS might be different. As suggested by Kadunce et al. (1997), it might be possible that MSS is mediated by the inhibitory region at the previous unisensory level, whereas multisensory depression, which requires the convergence of cross-modal inputs, depends on the internal circuitry of DSC. In this case, MSS exhibited by DSC neurons is inherited from modality-specific channels, and it thus can not influence multisensory depression.

2.4.2 Operations Underlying Multisensory Integration

Traditionally, the most popular metric to characterize the operations underlying multisensory interaction is the CME index (Meredith and Stein, 1983; Stein and Meredith, 1993; Stein et al., 2004). Although a majority of DSC neurons robustly exhibit response enhancement, the underlying operations for multisensory integration might not be sufficiently characterized by the CME index. In fact, operational rules to combine multiple inputs in unisensory areas that cannot be properly captured by the CME index have been reported. For example, a neuron in the inferior colliculus of owl multiplies its tuning function for the interaural time difference and interaural level difference (Pena and Konishi, 2001). Binaural neurons in the superior olivary complex, however, implement a linear operation (Yin and Chan, 1990). In consideration of this, experiments have been designed to investigate the underlying operations in DSC neurons.

To investigate the operations of DSC neurons, the additivity index is introduced. Perrault et al. (2005) adopted a mean multisensory contrast measure, which is essentially an average version of the additivity index in Eq. (2.2) without being divided by $|V|+|A|$, and showed that four apparently different operational modes, namely, super-additive, super-additive/sub-additive, sub-additive and additive/sub-additive operations, are involved in the responses of DSC neurons (Fig. 2.7). For example, neurons with the super-additive/sub-additive operation mode refer to those exhibiting a transition in their operations from super-additivity to sub-additivity when the effectiveness of the sensory stimulus is increased. However, since the division is based on the additivity index, the capacity of which to characterize multisensory integration has not been fully justified, it is not necessarily an appropriate categorization. In fact, an examination of Fig. 2.7 reveals a general principle in the four operation modes that super-additive operation occurs only for weak neuronal responses while sub-additive operation occurs only for strong responses. The proposed categorization fails to capture this relation. It might therefore be helpful to interpret the data in another way to account coherently for the different operations. In addition, Perrault et al. (2005) have also shown that the dynamic range (DR), defined as the difference between the neuron's response at threshold and saturation, of the bimodal response of a DSC neuron is larger than its unimodal visual or auditory response (Fig. 2.8). Intuitively, this result indicates that a DSC neuron dynamically enlarges its DR to accommodate bimodal inputs. However, it is also possible that, instead of dynamically adjusting the DR, prior saturation in the unimodal neurons that project to the DSC neurons contributes to the different DRs. In this case, the smaller DR of DSC neurons for unimodal stimuli is the result of saturated inputs they receive from those unimodal neurons. We will discuss this in more detail in Chapter 6.

A recent study has systematically explored the responses of a population of DSC neurons to sensory stimuli with different intensities. Stanford et al. (2005) recorded the responses

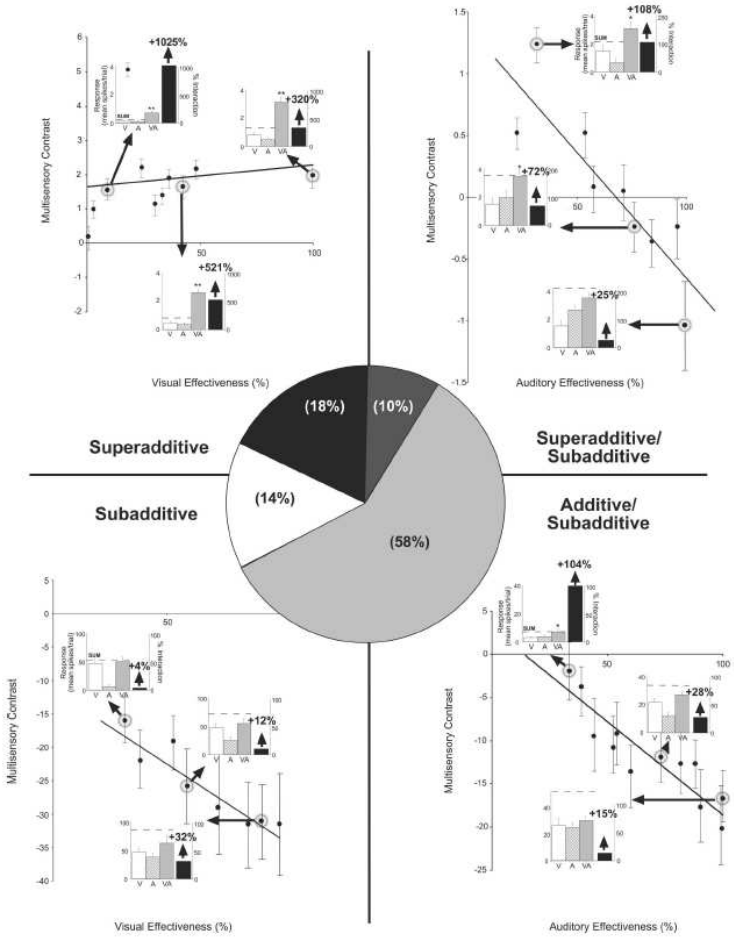


FIGURE 2.7: Multisensory neurons are divided into four categories based on their underlying operations. Multisensory contrast, which compares the bimodal response with the sum of the two unimodal responses and is comparable to the additivity index, is plotted as a function of the effectiveness of a unimodal stimulus over the neuron's dynamic range. In addition, the CME index is produced for three selected stimulus combinations in the insets. Top left: a super-additive neuron. Top right: a super-additive/sub-additive neuron. Its operation changes from super-additivity to sub-additivity as the auditory effectiveness increases. Bottom right: an additive/sub-additive neuron, the operation of which changes from additivity to sub-additivity. Bottom left: a sub-additive neuron. The central pie graph shows the percentage of the four categories in the whole multisensory neurons. Taken from Perrault et al. (2005).

of a DSC neuron to a fixed visual, auditory or visual-auditory stimuli 30 times and then produced a set containing the sum of the complete 900 combinations of the observed visual and auditory responses. To form a distribution of predicted neuronal responses, they then produced 10,000 datum points, each of which was an average of 30 samples randomly selected from the set. The z-score of the observed bimodal response with respect to this predicted response distribution was then calculated. For a DSC neuron, they calculated 36 z-scores corresponding to a complete combination of 3 visual stimuli (weak, medium and high), 3 auditory stimuli (weak, medium and high) under 4 inter-stimulus time configurations. In total, 41 neurons were tested, most of which subject to 36 different stimuli combinations, resulting in 1482 z scores. According to convention,

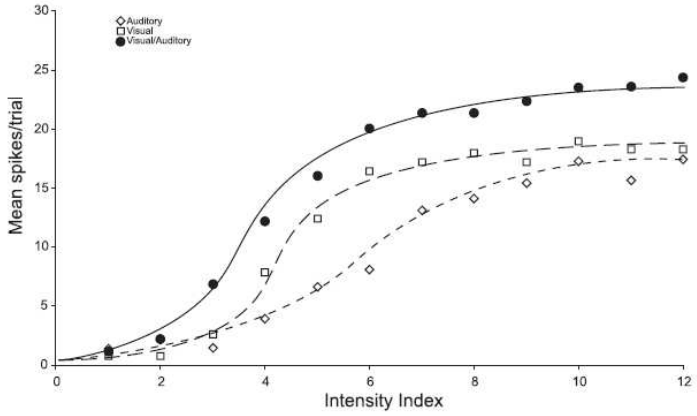


FIGURE 2.8: The dynamic ranges of a DSC neuron to visual (□), auditory (◊) and visual-auditory (•) stimuli. For visual and auditory stimuli, twelve intensity levels are presented, in which the first three intensity indexes determine the threshold and the last three intensity indexes determine the saturation. A bimodal stimulus is composed of visual and auditory unimodal stimuli at the same intensity level. Note that the dynamic range for bimodal response is larger than the two unimodal responses. From Perrault et al. (2005).

a significance level of 0.05 (when $z = \pm 1.96$) was adopted to decide whether a response was sub-additive ($z < -1.96$), additive ($-1.96 < z < 1.96$) or super-additive ($z > 1.96$). Fig. 2.9 shows the distribution of these z-scores, in which we can see that DSC multisensory neurons exhibit super-additive, additive and sub-additive responses. Moreover, Stanford et al. (2005) also showed that the probability of exhibiting super-additivity is the highest for near-threshold stimuli and it decreases when stimulus intensity increases. This corresponds well to the principle implicit in Perrault's results (Fig. 2.7). Among these three operations, additive responses are mostly observed, occupying 69.4% of the test cases. Since a majority of neurons yield multisensory responses consistent with a linear summation of the modality-specific responses, Stanford et al. (2005) suggested that an additive operator is a relatively good estimator for multisensory integration. However, the percentage of super-additivity and sub-additivity in the test cases is not negligible. A more natural way to interpret the results is to consider that a DSC neuron responds in a non-linear and saturating way. In this case, the reason that super-additivity and sub-additivity are less common than additivity is simply that these two operations occur on the tails of the response function, that is, the near-threshold region and the saturation region, respectively. The middle portion of the response function can be considered as approximately linear, or additive. Moreover, it should also be pointed out that the way Stanford et al. (2005) proposed to characterize additive responses is not statistically plausible. This will be discussed in more detail in Chapter 6.

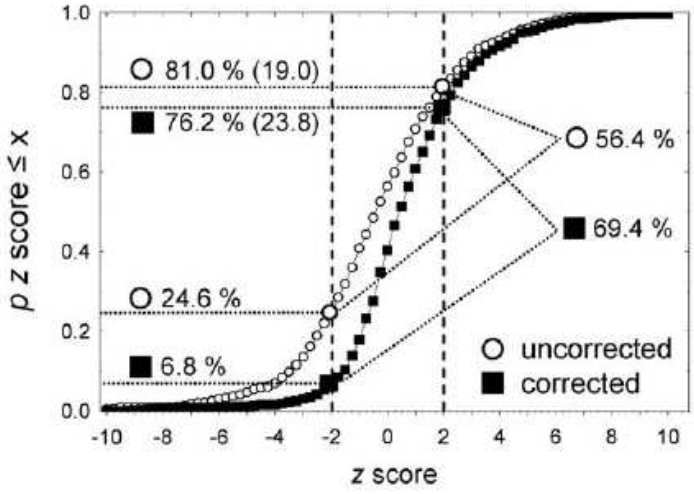


FIGURE 2.9: Cumulative density functions for all stimulus combinations with (filled squares) or without (open circles) spontaneous activity. The z score relates the actual multisensory responses to the distribution of a predicted sum of unimodal responses for 1482 stimulus combinations from 41 neurons. For the detailed procedure of evaluating this z score, please refer to Stanford et al. (2005). Vertical dotted lines indicate the transitions of operation modes from sub-additivity to additivity based on z score value of -1.96, and from additivity to additivity to super-additivity based on z score value of 1.96. The percentages of sub-additivity, additivity and super-additivity in the whole combinations are 24.6%, 56.4% and 19% respectively without correcting spontaneous activity. In the results for corrected spontaneous activity, the percentages of sub-additivity, additivity and super-additivity are 6.8%, 69.4% and 23.8%, respectively. From Stanford et al. (2005).

2.4.3 Cortical Influences

DSC neurons receive both ascending inputs from subcortical areas and descending inputs from cortical areas. This raises an interesting question of whether the descending and ascending projections play the same functional role in the response of DSC neurons. For this purpose, two cortical areas, the anterior ectosylvian sulcus (AES) and the rostral aspect of the lateral suprasylvian sulcus (rLS), have been particularly studied, since they provide the dominant descending projections to DSC (Wallace et al., 1993).

AES is perhaps the most studied multisensory cortical region. It is located in the fringe between the frontal, parietal and temporal lobes and is composed of three unimodal subregions, a visual region called AEV (Mucke et al., 1982), an auditory region called Field AES (or FAES) (Clarey and Irvine, 1986) and a somatosensory area SIV (Clemo and Stein, 1982). In contrast to SC, AEV and FAES neurons are not topographically organized, and only SIV neurons are well-organized to represent a map of the body (Clemo and Stein, 1982). Multisensory neurons are also found in AES, located between the borders of the unisensory regions (Clemo et al., 1991; Wallace et al., 1992). Similar to DSC multisensory neurons, these neurons are spatially registered for different sensory modalities. Furthermore, they also exhibit enhanced responses to bimodal stimuli com-

pared to unimodal stimuli (Wallace et al., 1992). Despite the existence of multisensory neurons in AES, projections from AES to SC are found to originate from its unisensory regions (Wallace et al., 1993). For example, visual afferents to DSC largely come from AEV. Although studies on rLS are less detailed than AES, it has been reported that the anatomy of rLS bears a resemblance to AES, containing a mixture of unisensory and multisensory neurons (Thompson et al., 1963; Toldi and Feher, 1984).

Electrophysiological experiments have studied the effect of AES/rLS deactivation on the responses of DSC multisensory neurons. Jiang et al. (2001) accomplished AES or/and rLS deactivation by inserted cooling coils placed between the sulcal walls and showed that DSC neuronal responses to bimodal stimuli become indistinguishable from unimodal responses after cortical deactivation (Jiang et al., 2001). Since multisensory enhancement is considered to characterize multisensory integration, DSC multisensory neurons that do not exhibit CME are considered as lacking the integrative capacity. Moreover, it has been shown that some DSC neurons lost their integrative capacity only after AES deactivation but not rLS deactivation, some only after rLS deactivation, and some are subject to both AES and rLS. In Fig. 2.10, enhancement disappears after the deactivation of AES, but is retained if rLS is instead deactivated; after reactivating AES by rewarming the area through the coils, multisensory response enhancement is recovered. Analogous disrupting effects in multisensory integration after cortical deactivation have also been demonstrated in behavioural studies. In the normal case, the orientation performance to a weak visual stimulus is enhanced by the presentation of a neutral auditory stimulus. After removing the influences of AES/rLS, orientation performance to visual-auditory stimulus reduces to be comparable to that with the visual stimulus alone (Jiang et al., 2002). It can therefore be concluded that AES and rLS are indispensable for multisensory integration.

Multisensory depression and MSS observed in DSC neurons are also considered to characterize multisensory integration (Kadunce et al., 1997). However, as shown by the experiments, multisensory depression and MSS exhibit dependencies on cortical influences different from CME, indicating that these three response properties are mediated by diverse circuits under different mechanisms. Jiang and Stein (2003) showed that deactivation of AES and rLS by cooling coils results in a decrease in multisensory depression for half the neurons. Fig. 2.11 shows that the depression of a DSC neuron is reduced when AES and rLS are deactivated. In this case, multisensory responses are disinhibited after cortical deactivation. For modality-specific suppression (MSS), although cortical deactivation reduces sensory responses of DSC neurons, the underlying computation remains relatively the same (Alvarado et al., 2007). As shown in Fig. 2.12, a sub-additive neuron remains sub-additive to the two visual stimuli after cortical deactivation. These are in contrast to multisensory enhancement which depends critically on cortical influences.

To demonstrate that AES and rLS are particularly involved in the integration of cross-

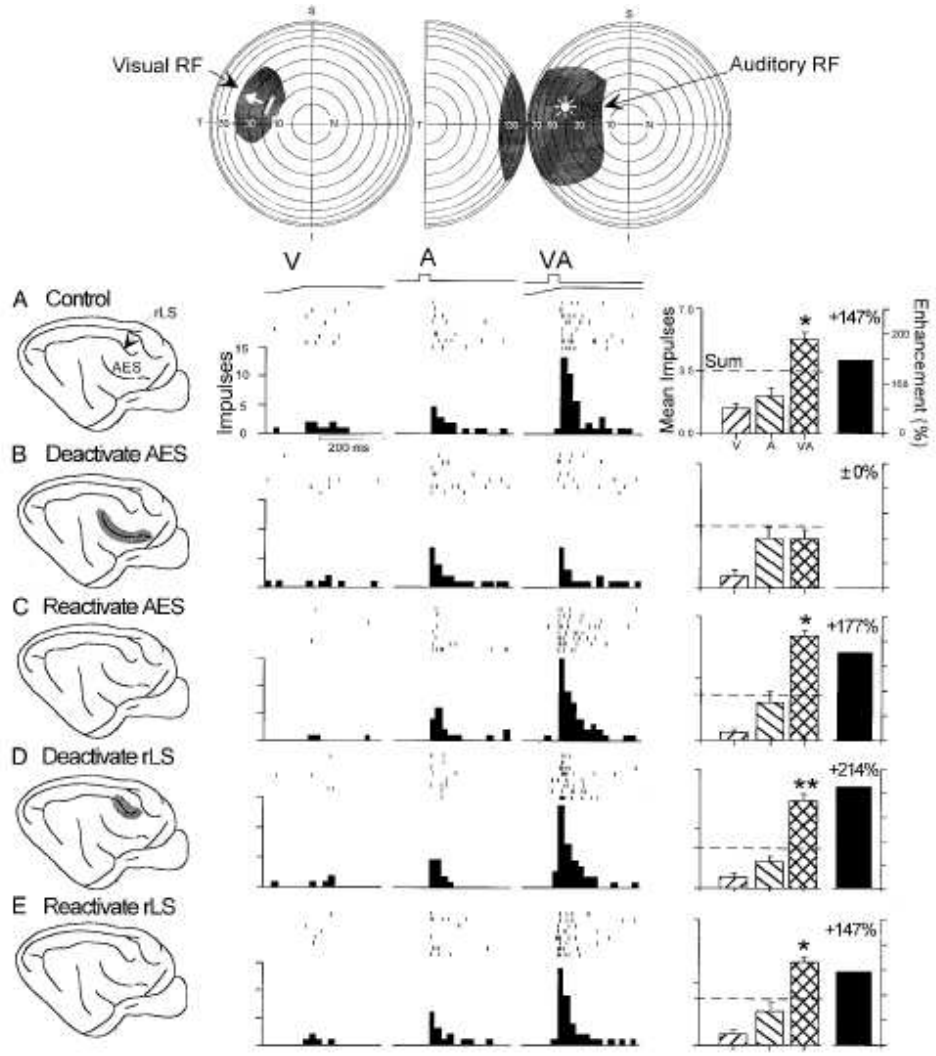


FIGURE 2.10: Influences from AES/rLS on multisensory enhancement. Deactivation of AES and/or rLS is indicated by shaded area in the schematic of the cortex. The conventions are the same as in Fig. 2.4. From Jiang et al. (2001).

modal inputs, cortical influences on DSC unimodal neurons have been studied. Alvarado et al. (2007) showed that AES and rLS deactivation has little effect on the responses of unimodal visual neurons both to a single visual stimulus and to two visual stimuli within their RFs. This demonstrated that AES and rLS do not affect unimodal neurons and play a role specifically in the responses of multisensory neurons.

Electrophysiological experiments have studied the morphology and postsynaptic distribution of both the brainstem and cortical somatosensory afferents, showing that trigeminal terminals from brainstem are relatively small and make synapses on distal dendrites of neurons in the intermediate layer of SC, and that corticotectal terminals are large and make multiple synaptic contacts with proximal dendrites (Harting et al., 1997). Based on this result, Harting et al. (1997) suggested that trigeminotectal projections endow SC neurons with the RF properties, and corticotectal inputs are involved in multisensory integration. Fig. 2.13 shows a proposed architecture that suggests a dialogue for AES

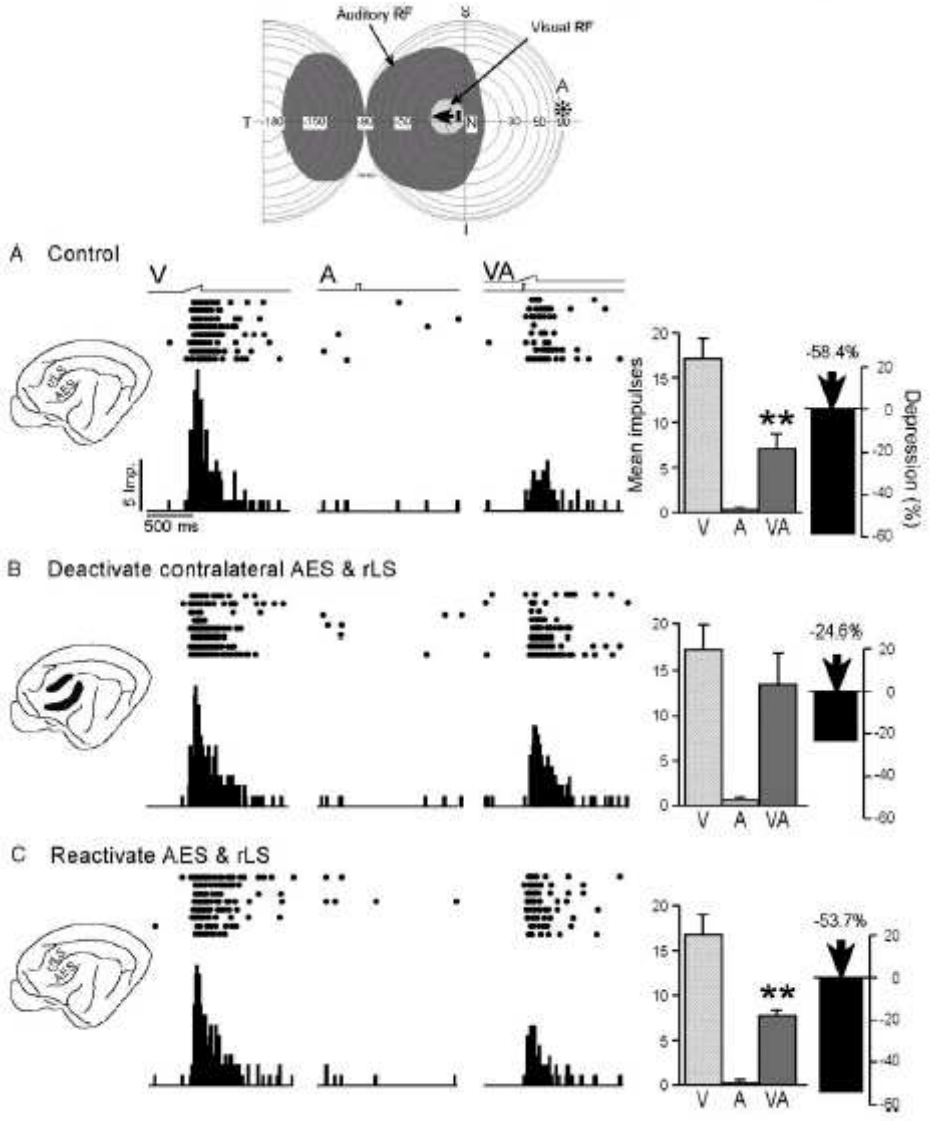


FIGURE 2.11: Influences from AES/rLS on multisensory depression. The conventions are the same as in Fig. 2.4. Taken from Jiang and Stein (2003).

and rLS to facilitate multisensory integration in SC neurons (Harting et al., 1997).

2.4.4 Development of Multisensory Integration

The development of multisensory processes in DSC is another important issue that gives hints of the mechanisms underlying multisensory integration. Neurophysiological experiments show that DSC neurons in cat are completely unisensory at birth (Wallace and Stein, 1997). The only stimuli DSC neurons respond to in the first several days after birth are tactile, and responses to sounds are found nearly at the end of the first postnatal week. As the prevalence of these two types of neurons gradually increases, auditory-somatosensory multisensory neurons appear. The first neurons responsive to

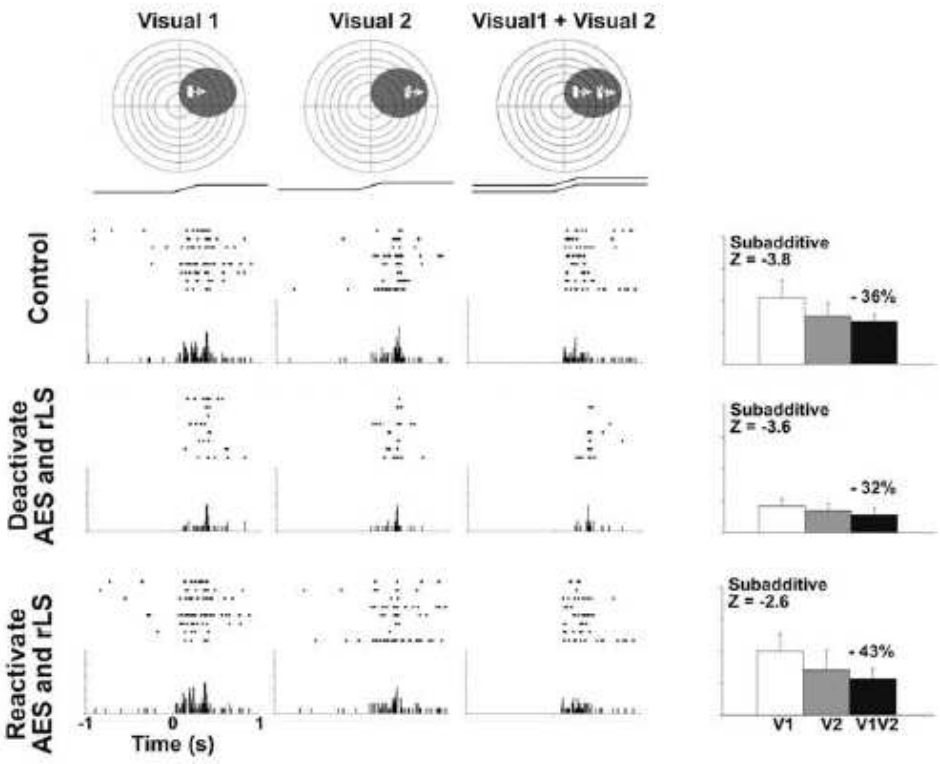


FIGURE 2.12: Influences from AES/rLS on modality-specific suppression. The conventions are the same as in Fig. 2.4. The calculation of z score is analogous to Fig. 2.9. From Alvarado et al. (2007).

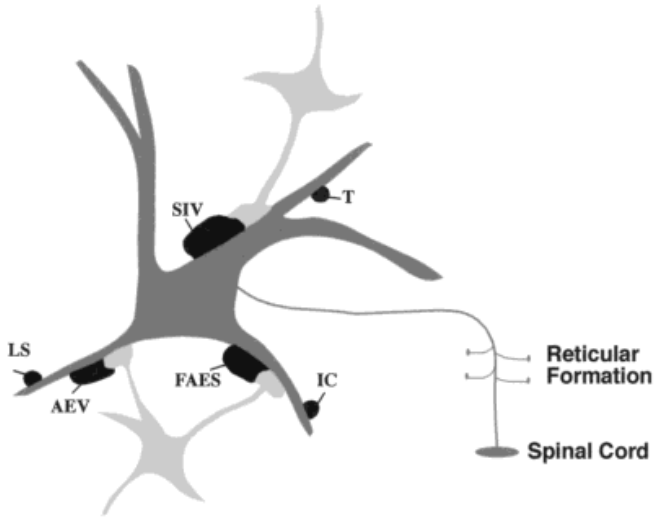


FIGURE 2.13: Schematic diagram of a proposed architecture of ascending and descending projections on a DSC neuron (medium gray). The corticotectal projections from AES include the somatosensory area (SIV), the auditory area (FAES) and the visual area (AEV). The terminals of these projections (black) are presynaptic to proximal dendrites of the neuron and are usually associated with pale vesicle-filled profiles (light gray). Terminals distributed distal to the AES cortical projections arise from trigeminal complex (T), the inferior colliculus (IC) and the lateral suprasylvian cortex (LS) and are suggested to bestow RF properties. From Harting et al. (1997).

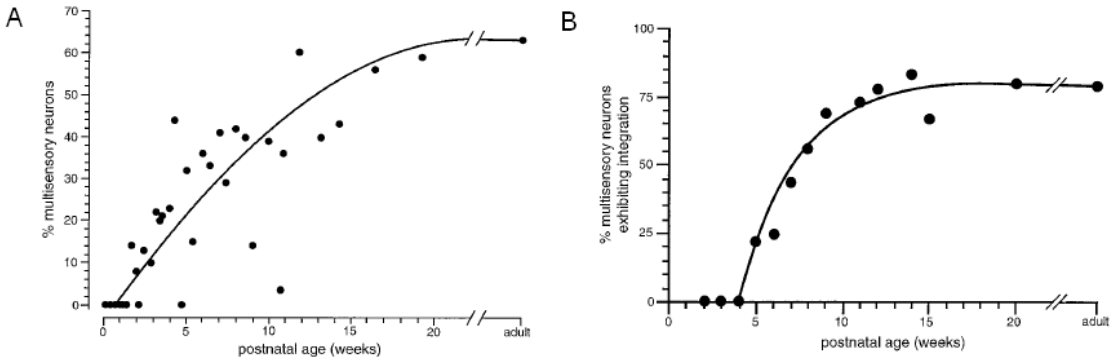


FIGURE 2.14: Development of multisensory neurons in DSC of cat. (A) The percentage of multisensory neurons in DSC. (B) The percentage of multisensory neurons exhibiting integrative capacity in multisensory population. From Wallace and Stein (1997).

visual stimuli appear in the third postnatal week, followed by the appearance of visually responsive multisensory neurons. Around 2 to 3 months, the percentage of multisensory neurons in DSC neurons reaches adult level, up to approximately 60%-70% of the whole population (Fig. 2.14A). Although the development of DSC neurons in monkey has not been studied as systematically as those in cats, some results indicate that DSC multisensory neurons in monkeys follow a developmental chronology different from cats. For example, in newborn monkeys, some DSC neurons are already responsive to visual, auditory and somatosensory stimuli, although with a lower percentage than adult level (Wallace and Stein, 2001).

Despite the existence of sensory-responsive neurons in newborn animals, their response characteristics are quite different from those in adult. Early unisensory neurons tend to have larger RF and longer response latency, and they do not exhibit the response selectivity observed in adult. As postnatal development progresses, these properties gradually develop to adult level (Wallace, 2004). For example, in cat, auditory response latencies are around 120 ms at 3 postnatal weeks, which then decrease to about 20 ms by 8-10 postnatal weeks. For visual processing, DSC neurons gradually acquire the selectivity to direction and speed. The size of DSC neurons' RFs show a steep decline and then consolidates to the adult level during development.

Intriguingly, in early developmental stages, the response properties of DSC multisensory neurons are different from those in adult. Wallace and Stein (1997, 2001) have shown that the responses of these early DSC multisensory neurons to cross-modal stimuli are indistinguishable from their unimodal responses, in dramatic contrast to multisensory enhancement exhibited by multisensory neurons in adult (Stein et al., 2004). Several weeks after the appearance of these early multisensory neurons, the capacity of multisensory integration is acquired within one or two days (Wallace and Stein, 1997). This abrupt transition from a non-integrative to an integrative state might occur as a consequence of the rapid opening of a biological gate (Wallace, 2004). Although having a much smaller temporal window, the response properties of these new integrative multisensory

neurons become quite similar to adult level. For example, they exhibit multisensory enhancement associated with inverse effectiveness, and multisensory depression with a magnitude comparable to the adult (Wallace and Stein, 1997). The percentage of integrative multisensory neurons in the population of multisensory neurons increases during development (Fig. 2.14B). Furthermore, Wallace and Stein (1997) showed that neurons with RFs similar in size to the adult level have a higher probability of being integrative. The consolidation of RFs thus appears to be correlated with the acquisition of integrative capacity.

As introduced previously, cortical areas AES and rLS play an important part in the integrative capacity of DSC multisensory neurons. Deactivation of these two areas by cooling coils largely eliminates the enhanced responses to bimodal stimuli compared to unimodal responses (Jiang et al., 2001). Since these non-enhanced responses accord with DSC multisensory neurons in newborn animals, it is interesting to see whether the gating effect in the acquisition of integrative capacity is related with AES and rLS influences. Wallace and Stein (2000) demonstrated that early multisensory neurons without the integrative capacity do not show obvious differences in their unimodal or bimodal responses before and after cortical deactivation. Fig. 2.15A shows that the responses and operation of a non-integrative multisensory neuron are not quite influenced by AES deactivation. As shown in Fig. 2.15B, however, for a neuron that just acquires the integrative capacity at 28 days postnatal, AES deactivation reduces its bimodal response, making it indistinguishable from unimodal responses. These newly developed integrative neurons, as in the adult, lose the integrative capacity. After reactivating the AES by rewarming, the same neuron recovers its capacity to integrate multisensory stimuli (Fig. 2.15B). On the other hand, Jiang et al. (2006) conducted neonatal ablation on animals at 3 weeks postnatal through subpial aspiration and studied the response properties of multisensory neurons when the animals were matured. The results showed that neonatal ablation of AES and rLS disrupted multisensory enhancement, whereas multisensory depression was preserved in some DSC neurons. These results strongly suggest that the lack of integrative capacity in DSC neurons in newborn animals might be the result of an absence of functional corticotectal afferents (Wallace, 2004).

In addition to the loss of integrative capacity, neonatal ablation of both AES and rLS also results in the misalignment of RFs for different sensory modalities (Jiang et al., 2006). This result further indicates that there is a link between the capability of multisensory integration, cortical influences and the refinement of RFs. For example, it may be possible that cortical influences are indispensable for the normal function of integration and the consolidation of RF. Interestingly, Jiang et al. (2006) showed that neonatal ablation of either AES or rLS does not result in large RF misalignment in the adult. It seems that, during early development, AES and rLS are able to compensate for the function of each other. This is also supported by the finding that early ablation of AES or rLS but not both has little effect on the development of multisensory processes (Stein,

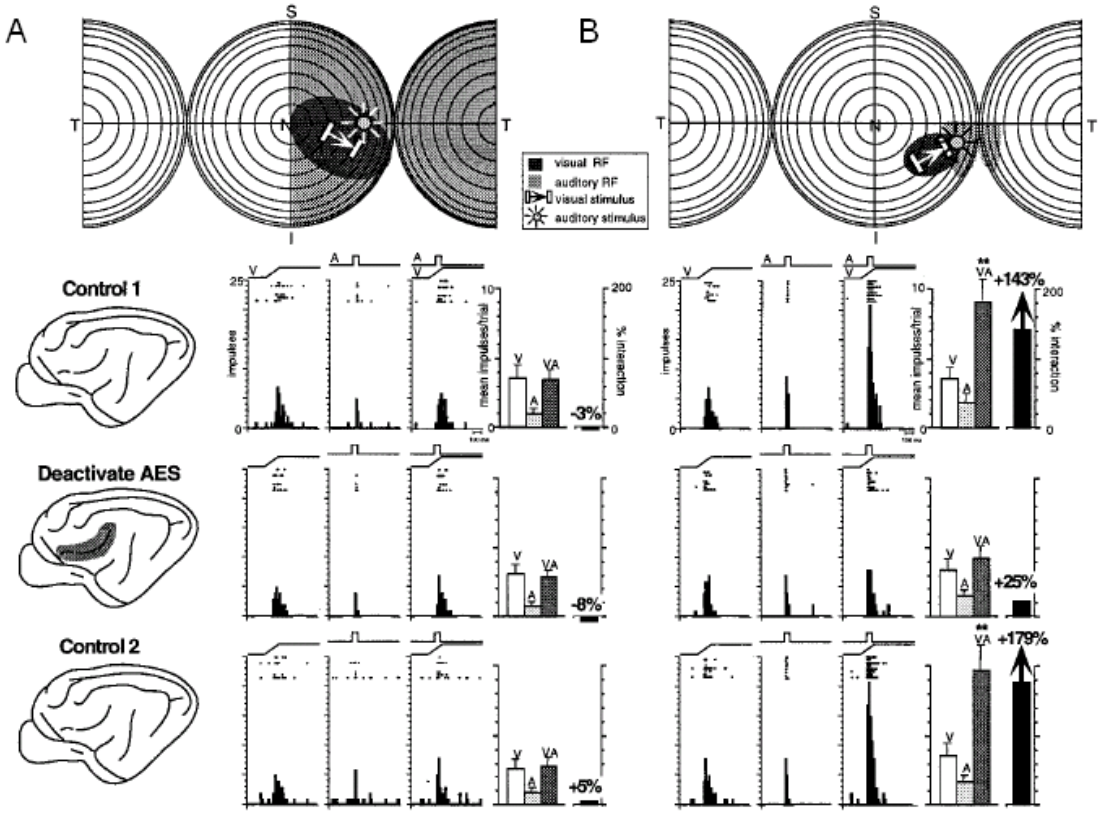


FIGURE 2.15: Cortical influences on multisensory non-integrative and integrative neurons. (A) multisensory non-integrative neuron. The unimodal responses and bimodal response of this neuron is not affected by the deactivation of AES, and CME remains relatively the same (see the "% Interaction"). (B) multisensory integrative neuron. The bimodal response becomes indistinguishable from the unimodal response after AES deactivation, and CME significantly reduces (see the "% Interaction"). The conventions are the same as in Fig. 2.4. From Wallace and Stein (2000).

2005).

Since AES and rLS are indispensable in the multisensory integration of DSC neurons, it is important to explore also whether the integrative capacity of DSC neurons is inherited from AES and rLS. Studies on the development of cortical multisensory integration provide strong evidence to answer this question. Recent experiments on cat suggest that multisensory neurons in AES follow the same developmental chronology as SC, developing from the unisensory to the multisensory state, and then acquiring the integrative capacity (Wallace et al., 2006). However, there is a delay between the chronology of SC and AES such that multisensory neurons in AES become integrative later than neurons in SC. Therefore, the integrative capacity in DSC neurons cannot be inherited from AES multisensory neurons. This also corresponds to the finding that projections from AES to DSC neurons originate from its unisensory regions (Wallace et al., 1993). The abrupt acquisition of integrative capacity might therefore be contributed by the unisensory neurons in AES. Moreover, since the maturation of corticotectal influences in DSC corresponds well with the peak period of cortical plasticity, Wallace and Stein (2000)

suggested that corticotectal projections for multisensory integration start to function only after the cortex is able to exert experience-dependent influences on SC. Wallace et al. (2004) showed that DSC neurons of dark-reared cats have abnormally large RFs and are not able to synthesize cross-modal information, indicating that visual experience is in fact important for multisensory integration. Similar influences on RFs have been observed in Syrian hamster (Carrasco et al., 2005). It might be possible that deprivation of visual experience affects the maturation of neurons in both SC and AES, which then exert a direct influence on multisensory integration.

2.5 Summary

Superior colliculus (SC) is a subcortical area in the midbrain that receives convergent visual, auditory and somatosensory inputs from both cortical and subcortical areas and mediates the orientation behaviour, such as saccade and head movements. It is the most studied multisensory area in the brain.

Since SC neurons are arranged in a way that neighbouring neurons have adjacent receptive fields, a population of SC neurons correspond to a sensory map of the space. With respect to different sensory modalities, SC neurons represent the corresponding visual, auditory and somatosensory spatial maps that align with each other to produce a coherent multisensory spatial map (Stein and Meredith, 1993). One consequence of the multisensory map in SC is that spatially coincident cross-modal sensory stimuli trigger the same SC neuron, producing responses different from those evoked by a single sensory stimulus. For example, when cross-modal stimuli are presented in time synchrony and space proximity, the response of a SC neuron is enhanced compared with its response to a unimodal stimulus, a property called cross-modal enhancement (CME) (Meredith and Stein, 1986). When the cross-modal stimuli are further apart in space, this enhanced response gradually changes to a suppressed one ("multisensory depression"). Compared with cross-modal stimuli, within-modality stimuli more often result in suppressed responses (Kadunce et al., 1997). To unravel the operation underlying DSC neuronal responses, systematic studies have been carried out with electrophysiological methods, showing that although most of the operations can be considered as an approximately linear summation of the unimodal effects, the occurrence of super-additive and sub-additive operations is not negligible (Stanford et al., 2005). It was also reported that DSC neurons might employ different operational modes to process cross-modal stimuli (Perrault et al., 2005). Furthermore, both neurophysiological and behavioural experiments have indicated that cortical influences from AES and rLS are indispensable for the normal function of multisensory integration such as cross-modal enhancement and multisensory depression (Jiang et al., 2001, 2002; Jiang and Stein, 2003). The development of multisensory processes in DSC has also been studied, revealing that different animals might follow a different developmental chronology (Wallace and Stein, 1997,

2001). Nevertheless, the CME response property is generally acquired at a later stage in the animals and would be disrupted by deactivating the AES and rLS (Jiang et al., 2006).

Despite the large number of studies, the mechanisms underlying multisensory integration in SC remain unclear. In this study, we propose an adaptive model that accounts for many reported operations underlying DSC neuronal responses, as will be discussed in later chapters.

Chapter 3

Neuronal Adaptation to Natural Statistics

Animals receive sensory signals from the surrounding and process these signals to form a perception of the environment. To understand the principles underlying this sensory processing, it is beneficial to study the statistics of natural signals as well as the response properties of sensory neurons. A large amount of research has explored the statistical properties of natural signals, studied the responses of neurons in unisensory systems to natural statistics and proposed information-theoretic principles to account for the underlying mechanisms. Although the results are from unisensory systems, it is not unnatural to assume that analogous properties and principles also exist in multisensory system. In the following, our discussion is broadly divided into three parts: natural signal statistics; adaptation of neuronal responses to input statistics; and potential information principles underlying these processes.

3.1 Natural Signal Statistics

The natural signals that an animal receives are diverse, originating from different sources in the environment and activating different sensory organs. Because of the complexity of the environment, natural signal statistics also appear to be very complicated, and it is a difficult task to understand their properties. Currently, there has been a large amount of research to investigate the statistics of signals in different sensory channels. Among various sensory stimuli, visual and auditory signals have received the most attention since they are experimentally accessible. In the following, we introduce natural image statistics and natural sound statistics.

3.1.1 Natural Image Statistics

Images are generally expressed as elements, for example, pixels, in a high-dimensional vector space. Within the space of all possible images, natural images occupy a rather small and restricted subset. It is possible that there are some general principles dominating this subspace. Research exploring the statistical properties of natural images has indicated some of these principles, which might help us to interpret the way sensory neurons process visual information.

Low-order Image Statistics

A traditional way to investigate the properties of natural images is to decompose them into different components and then focus on the low-order statistics of these components individually (Srivastava et al., 2003). For example, one popular method is principal component analysis (PCA), which considers only the second-order statistics of the signals from images and produces a set of orthogonal components corresponding to the eigenvectors of the covariance matrix (Hancock et al., 1992). If the image data are distributed according to multivariate Gaussian distribution, the resulting components are statistically independent.

Second-order statistics are very important properties in natural images, and have been suggested to be closely correlated with image categories and scene scales (Torralba and Oliva, 2003). It is also easy to see that the intensities at neighbouring spatial locations in a natural image are closely related. Generally, studies on natural images assume that their underlying processes are translation invariant, that is, the correlation of the intensities of two locations in an image is determined only by their relative position, but not the absolute positions. Consider an autocorrelation function, $C(\mathbf{x})$, where $\mathbf{x}$ is a two dimensional vector denoting the relative position between two locations of an image. According to Wiener-Khinchin theorem, the power spectrum $P(\mathbf{w})$ of the image is then given by $P(\mathbf{w}) = \int_{\mathbb{R}^2} C(\mathbf{x}) e^{-i\mathbf{w}\cdot\mathbf{x}} d\mathbf{x}$, where $\mathbf{w}$ represents the 2-dimensional spatial frequency. Many studies have shown that the power spectrum $P(\mathbf{w})$ decays according to a power law, that is, $1/|\mathbf{w}|^p$, where $|\mathbf{w}|$ is the magnitude of $\mathbf{w}$ and p is a constant (Field, 1987; Burton and Moorhead, 1987). Although the value of p changes for different types of images, it is typically near 2. As shown in Fig. 3.1, the power spectrum of a natural image averaged over all orientations is quite close to the curve for $1/f^2$, where f is the spatial frequency. The generality of the existence of power law spectra indicates that second-order statistics are very important properties in natural images.

Another property of image statistics is scale invariance, which refers to the invariance of image statistics when an image is scaled up or down (Burton et al., 1986; Field, 1987). Ruderman (1994) showed that the probability distribution of pixels from an image ensemble of woods environments at different scales has the same shape, and later explained it by a physical model (Ruderman, 1997). Studies on the histogram of wavelet

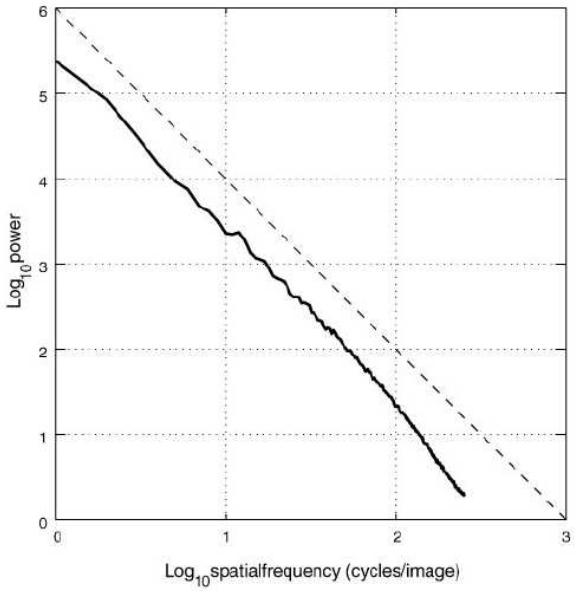


FIGURE 3.1: Power spectrum of a natural image (solid line) averaged over all orientations and the curve for $1/f^2$ (dashed line). From Simoncelli and Schwartz (1999).

decomposition of images also showed its invariance to image scaling (Zhu and Mumford, 1997). It should be pointed out that scale invariance is a property for a large ensemble of images, and the statistics of individual images do change for different scales. It is commonly believed that the power law spectrum is indeed the result of this scale invariance property. Spatially scaling the coordinate of an image by a factor of α leads to the scaling of its Fourier domain by a factor of $1/\alpha$. Under this scaling, only a power spectrum that conforms to a power law will retain the shape (Simoncelli and Olshausen, 2001).

Higher-order Image Statistics

Classical methods such as PCA work under the assumption that the natural images are second-order processes, and they fail for input distributions with significant higher-order statistics. Later research found that the statistics of natural images do not follow Gaussian distributions and contain higher-order statistics. Daugman (1989) showed that the response distributions of Gabor filters on natural images are quite non-Gaussian, with a sharp peak in the middle and long tails at two ends (kurtotic shape). Since a Gaussian-distributed image after the filters should produce Gaussian responses, the distribution of the original natural image is non-Gaussian. Mallat (1989) showed that the coefficients generated by decomposing natural images with multiscale, orthonormal wavelets can be described by a generalized Laplacian density, which is again non-Gaussian. Moreover, since whitening can decorrelate a Gaussian-distributed image to yield independent Gaussian responses, the whitened image should lose its structures, such as lines and edges. However, it turns out that a natural image after the whitening process retains some patterns (see Fig. 3.2), indicating that there is other information contained in the

original image, which thus can not be Gaussian-distributed. Furthermore, if the density of an image was Gaussian, its projections onto even- and odd- symmetric Gabor filters would be independent. However, Wegmann and Zetzsche (1990) found that the joint response distribution exhibits highly kurtotic margins. Buccigrossi and Simoncelli (1999) also showed that pairs of bandpass filters that eliminate the second-order correlations do not remove the dependencies in the response amplitudes for an image. These results indicate the existence of higher-order statistics in natural images.

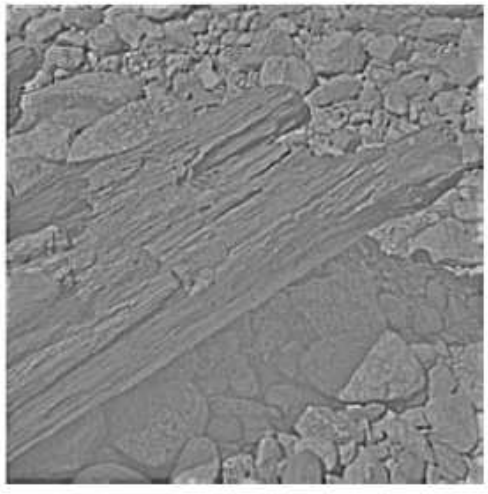


FIGURE 3.2: A whitened natural image. From Simoncelli and Olshausen (2001).

The non-Gaussian properties of natural images have motivated the application of some analyzing methods. For example, Huang (2000) showed that when an image is filtered by some random zero-mean filters, the response distribution exhibits kurtotic shape, suggesting that a maximization of the fourth-order statistics might facilitate the generation of independent components. This property inspired the implementation of independent component analysis (ICA), which, although not explicitly optimizing the independence, maximizes higher-order moments such as the kurtosis (Hyvärinen and Oja, 2000). Although ICA does not produce independent components in all cases, it achieves a good performance for natural images (Simoncelli and Olshausen, 2001). Field (1987) showed that the receptive fields of simple cells in the visual cortex are comparable to Gabor functions, leading to the application of Gabor filters, which have also been extensively used to decompose images both in space and frequency.

3.1.2 Natural Sound Statistics

The statistical properties of natural sound have also been explored. Interestingly, some of the identified properties bear a resemblance to those of natural images. As in the case of natural images, an investigation of natural sound could also contribute to our understanding of how the nervous system deals with sound signals since it is natural

to consider that the brain is configured according to the statistics of input signals to achieve efficient processing.

Low-order Sound Statistics

Temporal low-order statistics of natural sound have been studied as a first step to explore the statistical structures. In the study by Attias and Schreiner (1997), different sound ensembles containing 15 min sound, such as bird songs, pop music and speech, were passed through a bandpass filter. Both the amplitude probability distribution and the amplitude power spectrum of the filter responses in different frequency bands were then examined. Results showed that the histograms of the response amplitude remain approximately the same for different temporal resolutions from 0.75 ms to 150 ms, exhibiting analogous scale invariance property in natural images. Furthermore, the amplitude power spectrum $P_v(w)$ in a frequency band v exhibits a modified power law form, $P_v(w) \propto 1/(w_0^2 + w^2)^{\alpha/2}$ (Fig. 3.3), where $\alpha \in [1, 2.5]$ and $w_0 \in [10^{-4}, 1]$. This property is also analogous to the power law form of power spectrum in natural images, suggesting the presence of long-range temporal correlations in the amplitude. Similar power law spectra for natural sounds have also been shown by Voss and Clarke (1975). As indicated by these results, the amplitude statistics, or the low-order statistics, might be expressed as a function of simple parameters. This suggests that natural sounds occupy a small subspace within the space of all sounds and are highly redundant. It is therefore natural for auditory neurons in the nervous system to adapt to natural sound statistics in some way to process this redundant sound information more efficiently. Furthermore, since frequency bands correspond to different spatial locations on the basilar membrane (Pickles, 1988), identical distributions and power spectrum for different frequency bands also suggest the presence of translation invariance in auditory processing, analogous to that in natural images.

Higher-order Sound Statistics

Iordanov and Penev (1999) showed that, after applying PCA to sound ensembles, the resulting statistics exhibit non-Gaussian structure. For example, the marginal distributions of the PCA coefficients can be approximated by a general Laplacian distribution (Iordanov and Penev, 1999). Natural sounds thus also contain higher-order statistics. As in natural images, correlation-based methods again cannot deal with higher-order statistics of natural sounds. In fact, second-order statistics only provide information about the amplitude spectrum since the power spectrum is the square of the amplitude spectrum. It has been suggested that the power spectrum, or equivalently, the amplitude information is not the most informative part of natural signals (Bell and Sejnowski, 1996). This is illustrated by the example that if the amplitude information is removed, we can still recognize the signal although it is distorted. Instead, phase information appears to be more informative because the signal will appear as noise if its phase information is removed. In order to analyze phase information, algorithms sensitive to

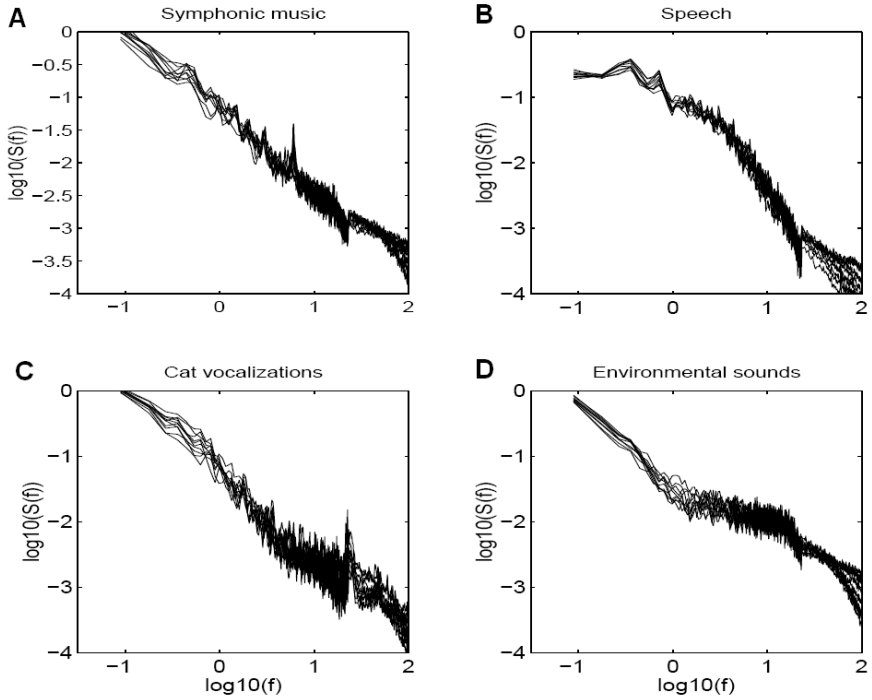


FIGURE 3.3: Amplitude power spectrum in different frequency bands for four sound ensembles, including (A) symphonic music, (B) speech, (C) cat vocalization and (D) environmental sounds. From Attias and Schreiner (1997).

higher-order statistics are required. To this end, Bell and Sejnowski (1996) proposed an algorithm based on ICA to yield independent basis functions. This algorithm produced results reflecting both the phase and frequency information, and it was suggested to be suitable to analyze the higher-order structure of natural sounds. However, since the algorithm is linear, it does not guarantee the production of independent components for the highly complex natural sound signals. To take a step further, Schwartz and Simoncelli (2000) showed that this dependence could be significantly reduced by a "divisive normalization" operation, in which the response of a filter is divided by a weighted sum of other rectified filter responses. Furthermore, their model with weights chosen to maximize the independence of filter responses accounts for non-linear response properties of auditory nerve.

3.2 Neuronal Adaptation to Input Statistics

Adaptation of sensory neurons to accommodate input statistics is crucial for sensory information processing in the brain. In natural environments, there are diverse signals that can activate different sensory organs. The magnitudes of some natural signals span a wide range. For example, the magnitude of ambient light varies over nine orders from day to night (Rushton, 1965). In contrast, the response range of a sensory neuron is limited and sometimes is far below the range of natural signals. The firing rates of optic

nerve fibres, for example, span less than two orders (Barlow, 1981). To cope with the wide range of natural signals with a limited operational range, sensory neurons must adapt their sensitivities and responses in order to have a good representation of the inputs. Currently, a large number of adaptation mechanisms have been discovered in sensory systems, under which sensory neurons adapt to different statistical properties or adapt to one statistical property in different manners and on different timescales. Since adaptation in the visual pathway is the most studied one in sensory systems, the following mainly focuses on visual adaptation.

3.2.1 Adaptation to the Mean

Light adaptation is one of the most studied forms of mean adaptation in the sensory systems. It refers to the adaptation of neuronal responses to the mean level of illumination. This adaptation typically centres the limited response range of a visual neuron, for example, a retinal ganglion cell, around the mean stimulus level (Shapley and Enroth-Cugell, 1984). If the response of a sensory neuron is modelled as a saturating transfer function, the effect of adaptation can to a first approximation be considered as the changes in semi-saturation constant of the transfer function according to the mean level of illumination. Graphically, changes in the semi-saturation constant correspond to the shifting of the input-output curve (Normann and Werblin, 1974). This "curve-shifting" effect has been reported in cat retinal ganglion cell (Sakmann and Creutzeeldt, 1969). Fig. 3.4 shows the response curves of a cat's retinal ganglion cell as a function of test spot luminance for different background luminances, in which, as the background luminance increases, the response curve shifts rightwards. In auditory processing, similar shifting effects have also been demonstrated in neurons in the inferior colliculus (IC) of guinea pigs for different mean sound levels (Dean et al., 2005).

Adaptation to the mean luminance can be contributed by the photoreceptor cells (Koutalos and Yau, 1996) and other mechanisms in the subsequent cells in the retina (Walraven et al., 1990). A number of important phenomena have been reported to associate with light adaptation. Battaglia et al. (2003) showed that, in photoreceptor cells, the presence of background illumination reduces the sensitivity of the neurons and speeds the time-to-peak in their responses. Matthews et al. (1988) demonstrated that when the internal Ca^{2+} concentration ($[Ca^{2+}]_i$) is maintained near its resting level, light-dependent changes in sensitivity of the responses of amphibian photoreceptor cells are abolished, indicating the important role of $[Ca^{2+}]_i$ in light adaptation. Around the same time, Nakatani and Yau (1988) also showed that amphibian rod and cone cells function over a very restricted range of intensities when the movement of Ca^{2+} is prevented. Despite the reports of these phenomena, the mechanisms underlying light adaptation have not yet been understood. In photoreceptor cells, at least eight distinct molecular mechanisms have been identified to be involved in light adaptation (Pugh et al., 1999). In the past,

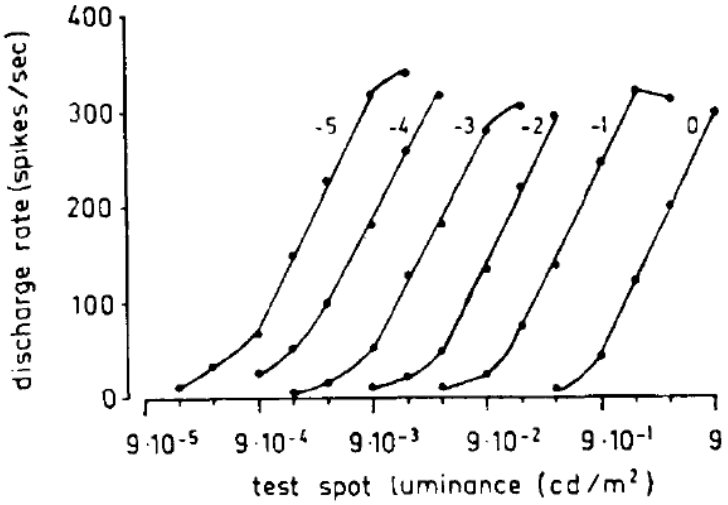


FIGURE 3.4: "Curve-shifting" effect in cat retinal ganglion cell to the changes of background luminances. The six curves correspond to the peak response of a cat retinal ganglion cell as a function of test spot luminance at six different background luminances. The number to the left of each curve is the background luminance in units of $\log cdm^{-2}$. Each point on the curve is the average of twenty responses. From Sakmann and Creutzeeldt (1969).

it was assumed that any one of these mechanisms might regulate all the aspects of light adaptation. However, as Pugh et al. (1999) suggested, only one of them exhibits all the properties expected for a light adaptation mechanism, such as predicted effects in the operating range and the sensitivity, and the dependence on $[Ca^{2+}]_i$. The contributions of each of these eight mechanisms are not entirely clear, not to mention other mechanisms that may yet to be identified.

3.2.2 Adaptation to the Variance

In addition to adapting to the mean, sensory neurons also adapt their responses to the variance, which describes the fluctuation of the signals around the mean. For example, adaptation to the variance has been observed in the auditory system, such as IC neurons (Dean et al., 2005). Among the reported cases of variance adaptation, a representative example is adaptation to the input contrast in visual neurons, under which the neuron could increase its sensitivity to improve the signal-to-noise ratio at low contrast and reduce its sensitivity to avoid response saturation at high contrast (Demb, 2002; Baccus and Meister, 2004).

Traditionally, experimental results showed that cells in visual system adapt to the changes in stimulus contrast on a timescale of tens to hundreds of milliseconds. This is generally referred to as "contrast gain control" (Shapley and Enroth-Cugell, 1984; Victor, 1987). Contrast gain control has been widely reported in retina ganglion cells of cats (Shapley and Victor, 1978, 1981) and salamanders (Baccus and Meister, 2002), and

also in cortical neurons of cat (Ohzawa et al., 1985) and monkey (Carandini et al., 1998). However, physiological studies found no contrast gain control in the lateral geniculate nucleus (LGN) of cat (Maffei et al., 1973; Movshon and Lennie, 1979; Ohzawa et al., 1985) or monkey (Derrington and Lennie, 1984). Later studies have identified a slower form of neuronal adaptation to contrast that occurs with a time course of tens of seconds. To differentiate this from the fast form, it is generally referred to as "contrast adaptation". Contrast adaptation has not been found in photoreceptor cells, and Rieke (2001) suggested that the first site of contrast adaptation is in the bipolar cells. Furthermore, contrast adaptation has been reported in the retinal ganglion cells of salamander (Smirnakis et al., 1997; Kim and Rieke, 2001), rabbit (Baccus and Meister, 2002) and macaque (Chander and Chichilnisky, 2001). In addition to the retina, contrast adaptation has also been reported in cortical neurons such as V1 neurons (Ohzawa et al., 1985; Carandini et al., 1998).

Despite the discovery of two forms of neuronal adaptation to contrast, the measure of contrast in visual system is not clear. One commonly used measure of contrast is the root-mean-square contrast, given by the standard deviation of light intensity divided by the mean (Shapley and Victor, 1981; Bonin et al., 2005),

$$C = \frac{\sqrt{\langle (L_S - L_{\text{mean}})^2 \rangle}}{L_{\text{mean}}}, \tag{3.1}$$

where L_S is the luminance of the stimulus, L_{mean} is the mean of luminance and $\langle\rangle$ represents the expected value. Another natural definition of contrast is the *Weber contrast* (Shapley and Enroth-Cugell, 1984), which is generally adopted in experiments with aperiodic stimuli such as uniform disks or bars on a background, given by

$$C_W = \frac{L_S - L_B}{L_B}, \tag{3.2}$$

where L_B is the luminance of the background. *Rayleigh contrast* has also been implicitly used by Rayleigh (1889) and is defined as

$$C_R = \frac{L_{\text{max}} - L_{\text{min}}}{L_{\text{max}} + L_{\text{min}}}, \tag{3.3}$$

where L_{max} is the maximum luminance and L_{min} is the minimum luminance. Rayleigh contrast is generally used in experiments using periodic spatial patterns such as sinusoidal gratings. In the above three measures, the Weber and the Rayleigh contrasts are defined specifically for spatial contrast, while the root-mean-square contrast is applicable to both spatial contrast and temporal contrast. In fact, in visual system, adaptation to both spatial and temporal contrasts has been widely reported.

Adaptation to spatial contrast has been revealed in many neurophysiological studies. For example, Shapley and Victor (1978) showed that cat retinal ganglion cells exhibit contrast gain control to standing sinusoidal gratings of adjustable spatial phase and spa-

tial frequency. Truchard et al. (2000) presented drifting sinusoidal gratings to binocular simple cells in V1 and fitted an adjustable simple cell model to the recorded neuronal responses, based on which they suggested that contrast gain control occurs primarily at a monocular site. In a recent study, Bonin et al. (2006) displayed white noise textures to anesthetized, paralyzed cats and found that the response gain of LGN neurons strongly depends on the standard deviation of the stimuli. This result also suggests that sensory neurons adopt the root-mean-square contrast as the measure for contrast gain control.

Adaptation to the temporal contrast has also been reported. For example, Rieke (2001) employed a light-emitting-diode as the stimulus and controlled the temporal contrast by adding a Gaussian fluctuation to the stimulus, and showed that salamander ON and OFF bipolar cells adapt to this stimulus. Baccus and Meister (2002) demonstrated both fast and slow form of adaptation to temporal contrast in bipolar, amacrine and the ganglion cells in salamander retina, using a rapidly flickering uniform filed with intensity conforming to Gaussian distribution. Smirnakis et al. (1997) showed that the response of salamander retinal ganglion cells to a step increase in the contrast of a spatially uniform flicker stimulation increases abruptly and then decays slowly to a lower value (Fig. 3.5). On the other hand, a step decrease in the stimulation results in an abrupt decrease in neuronal response which then gradually recovers (Fig. 3.5).

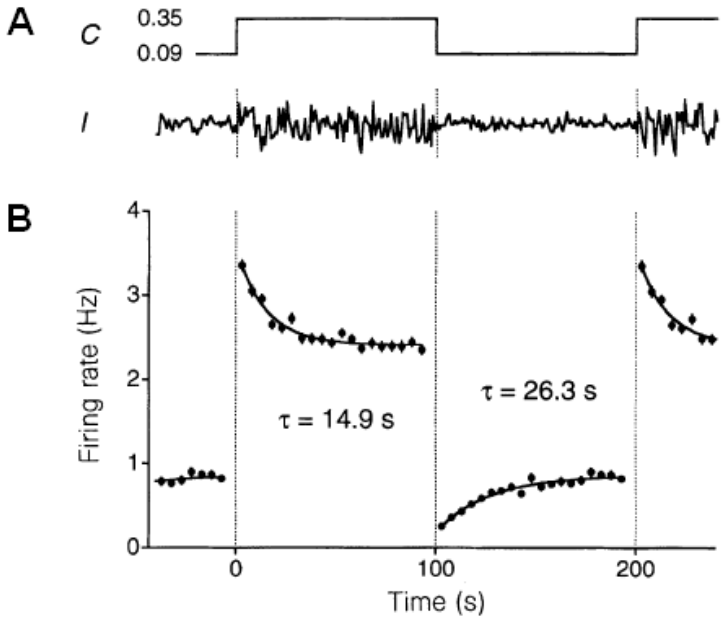


FIGURE 3.5: Responses of a salamander ganglion cell to temporally alternating contrast. Contrast is defined as the standard deviation of the stimulus divided by its mean. (A) The stimulus is a spatially uniform flickering light, alternating every 100 seconds between contrast values of 0.09 and 0.35. C represents the alternating contrast and I represents the intensity of the stimulus. (B) Neuronal responses are measured in the firing rate computed in 5 second time bin and averaged over 100 trials. Continuous lines are exponential fits of the data with decay time τ. The first segment and the last segment are periodic repetitions of the data. From Smirnakis et al. (1997).

Although adaptations to mean and contrast exist in sensory systems, Mante et al. (2005)

showed that in LGN neurons light adaptation exhibits the same effect for different contrasts, and contrast gain control exhibits the same effect for different mean luminances. Adaptations to the mean luminance and contrast therefore appear to be mediated by independent mechanisms. Neurophysiological experiments have studied the mechanisms underlying adaptation to the contrast, which, unlike mean adaptation, do not operate in the photoreceptors and have been reported in subsequent cells in the retina, thalamus and visual cortex (Smirnakis et al., 1997; Rieke, 2001). To study contrast gain control, Beaudoin et al. (2007) made extracellular and whole-cell recordings of guinea pig Y-type ganglion cells *in vitro* and suggested that the adaptation originates in the presynaptic bipolar cells. Shu et al. (2003) studied layer 5 pyramidal cells in prefrontal and visual cortical slices *in vitro*, and showed that background synaptic activity decreases the slope of the input-output curve. This suggested that gain control might be mediated by balanced barrages of excitatory and inhibitory synaptic activity. Furthermore, Sanchez-Vives et al. (2000a) showed that slow contrast adaptation occurs as a result of an intrinsic property of some cortical neurons, in which the afterhyperpolarization that recovers slowly over seconds suppresses the responses to stimuli. Since spiking can activate sodium influx that modulates potassium conductance to produce cortical afterhyperpolarization, spiking is sufficient to drive adaptation for a cortical cell (Sanchez-Vives et al., 2000a). However, slow adaptation in ganglion cells does not primarily result from an intrinsic property of the cell but from reduced glutamate release from presynaptic bipolar cells (Manookin and Demb, 2006). Currently, the mechanisms underlying adaptation to the contrast remain unclear.

3.2.3 Adaptation to Other Statistics

Sensory systems also adapt to input statistical properties other than the mean and the variance. Studies on LGN neurons showed that an increase in the correlation of the visual stimulus leads to decreases in both the gain and response selectivity (Lesica et al., 2007), indicating that LGN neurons also adapt to the correlations by gain control. Gain control based on correlation is thus comparable to that based on contrast in the sense that an increase in either correlation or contrast leads to decreased gain and selectivity. Although no evidence for sensory adaptation to correlation between cross-modal stimuli has been shown, it is not unnatural to assume that an analogous gain control based on cross-modal correlation could also be observed.

Bonin et al. (2006) showed that LGN neurons in cat only adapt to standard deviation, but not higher-order statistics, such as kurtosis or skewness. However, Sharpee et al. (2006) found that V1 neurons in cats adapt to stimulus statistics beyond the mean and variance with much longer timescales than that of contrast gain control, ranging from 40s to many minutes. Furthermore, Kvale and Schreiner (2004) demonstrated that neurons in IC of cat show adaptation not only to auditory stimuli with different variances of

their modulation depth distribution, but also to other higher moment statistics, such as kurtosis of the modulation envelope. Therefore, there are at least some sensory neurons that adapt to higher-order statistics.

3.3 Information Principles underlying Adaptation

3.3.1 Efficient Coding

Natural signals are inherently redundant. This redundancy partly comes from the autocorrelation of natural signals in time and space. The intensity at one location in a natural image is generally close to its surrounding intensities except in some special conditions such as when it is on an edge, and there is thus redundant information between neighbouring locations. Since neurons are physical devices with limited processing and computational capability, it is natural to consider that, in order to process enormously complicated natural signals more efficiently, they eliminate the redundancy in natural signals. This interpretation of sensory processing has led to a popular notion of a computational strategy known as "efficient coding" (Attneave, 1954; Barlow, 1961), which is proposed based on information theory.

Consider a message ensemble M with N messages m, each of which consists of l symbols $s_1, s_2, ..., s_l$ from an alphabet with N_s symbols. The quantity "redundancy" is defined based on information theory (Shannon and Weaver, 1949),

$$R = 1 - \frac{H(M)}{C} = \frac{1}{C}[C - \Sigma_{j=1}^{l} H(s_j)] + \frac{1}{C}[\Sigma_{j=1}^{l} H(s_j) - H(M)], \tag{3.4}$$

where C is the capacity of a channel given by $C = l \log_2 N_s$ and corresponds to the maximum information that l symbols from an alphabet with N_s symbols could possibly carry, $H(M)$ is the entropy of the message ensemble M and $H(s_j)$ is the entropy of a symbol s_j. R is called the Shannon redundancy. From this definition, we can see that the entropy $H(M)$ needs to be close to the capacity C in order to achieve low redundancy, or equivalently, high efficiency. Shannon redundancy R quantifies the performance of an efficient coding. A code that minimizes R is therefore called a minimum redundancy code. Moreover, to achieve $\Sigma_{j=1}^{l} H(s_j) = H(M)$, the symbols have to be statistically independent, in which case the joint probability $P(s_1, ..., s_j, ..., s_l)$ can be factorized to a product of $P(s_j)$ terms. Therefore, if a code merely minimizes the part $\Sigma_{j=1}^{l} H(s_j) - H(M)$, it is called a factorial code. For a more detailed introduction, refer to Atick (1992).

It has been suggested that sensory neurons adapt to eliminate the redundancy contained in the inputs (Barlow, 1961). Since factorial coding can be considered as a first step towards minimizing Shannon redundancy R, it has been extensively studied in neural systems, especially for neurons in early sensory processing stages, such as in the retina.

For example, Srinivasan et al. (1982) theoretically computed the inhibition required for neighbouring photoreceptors to cancel out the correlations in the natural images. This theoretical result matches well with the measurement from the interneurons in the compound eyes of a fly, indicating that there is a decorrelation process involved in these visual neurons. Moreover, Dan et al. (1996) have shown that the power spectrum of LGN neuronal responses are approximately white and therefore signals from LGN to visual cortex are decorrelated. In addition to neurons in the early visual pathway, efficient coding has also been used to account for direction-selective property of receptive field in cortical neurons such as V1 neurons (van Hateren and van der Schaaf, 1998). Simoncelli and Schwartz (1999) suggested a divisive normalization approach to account for the way cortical neurons deal with non-Gaussian statistics. In this divisive normalization, linear response of a basis function to image signals is rectified and divided by a weighted sum of these rectified signals in the neighbouring neurons. Maximizing the independence of the resulting responses produce results that matched well with the neurophysiological data about the suppression effect on the neuronal response by non-optimized stimuli. In the auditory system, the non-linearities in the response characteristics and several basic tuning properties of auditory nerve fibers can also be accounted for by efficient coding (Schwartz and Simoncelli, 2000; Lewicki, 2002). Efficient coding therefore appears to be widely implemented by neurons in early sensory pathway.

In addition to factorial coding, some neurons could also minimize the redundancy in the inputs by an equal use of their response levels. Under this condition, information carried by the response, or the response entropy, is maximized. Provided that the response is noiseless, this minimum redundancy coding can be achieved when the response function of a neuron corresponds to the cumulative density function (CDF) of the input (Laughlin, 1981). In other words, a neuron should implement gain control in a way that the sensitivity of the cell, defined as dr/di where r is the response and i is the input, is simply the probability density function $P(i)$ of the input to attain an optimal coding. Laughlin (1981) showed that large monopolar cells in insect compound eyes respond in a way that agrees well with the CDF of the input (Fig. 3.6), suggesting that those neurons maximize response entropy and achieve an optimal coding.

3.3.2 Other Principles

Although efficient coding successfully accounts for many physiological properties of sensory neurons, such as photoreceptor cells, LGN neurons and V1 neurons, this principle does not take into account the noisiness of the nervous system. In such a noisy system, redundancy will be required when the signal-to-noise ratio is low in order to remove the effect of noise and convey the meaningful information.

Physiological studies on visual system have suggested some information optimization principles for neuronal adaptation. For example, Wainwright (1999) proposed an opti-

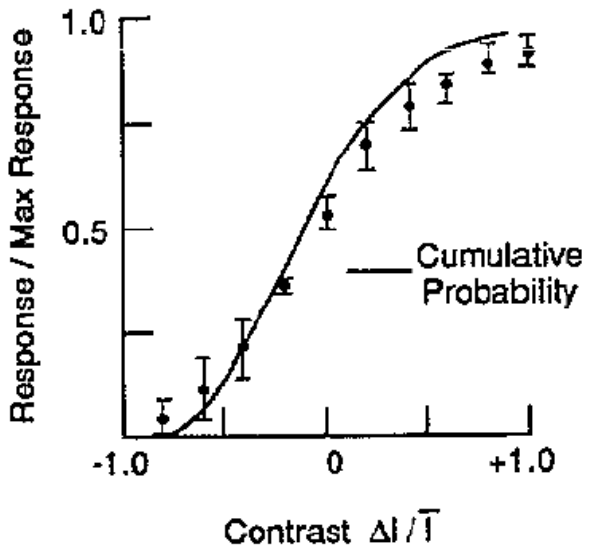


FIGURE 3.6: A comparison between the normalized contrast-response measured by Laughlin (1981) and the cumulative probability of input contrast for a large monopolar cell in the blowfly compound eye. From Atick (1992).

mal information transmission principle, in which the neurons adapt their responses to input statistics to maximize information transmission. This information-theoretic model is supported by its success in accounting for several psychophysical phenomena in visual neurons such as changes in contrast threshold, the tilt effect and changes in orientation discrimination thresholds. Brenner et al. (2000) tested the responses of H1 neurons of the blowfly, which are sensitive to the horizontal motion across the visual field, to an ensemble of horizontal velocity stimulus drawn from a distribution of zero mean and different standard deviations, and studied the principle of information transmission in these H1 neurons. As shown in Fig. 3.7A, the gain of neuronal responses decreases when the standard deviation is increased from $\sigma_1 = 2.3°/s$ to $\sigma_2 = 4.6°/s$. Interestingly, normalizing the response rate by its time-average value and stimulus velocity by the standard deviation aligns the two response curves (Fig 3.7B). These results not only indicate that the root-mean-square contrast is the measure for variation about the mean in neurons, but also suggest that standard deviation might exert a divisive effect in the responses through gain control. Brenner et al. (2000) also demonstrated that the input/output relation rescales to the changes of input statistics in such a way that the information transmission is maximized. As shown in Fig. 3.8, maximal information transmission is achieved when the stretch factor $\lambda = 1$, which corresponds to the measured data from the H1 neuron. On the other hand, Baddeley et al. (1997) showed that the response distributions of visual neurons in V1 and inferior temporal area are exponential. Given that the average firing rate is fixed, the response distribution of a neuron that maximizes the entropy is exponential. The responses of V1 neurons and those in the inferior temporal area might therefore reflect optimal information processing under some constraints.

In addition to optimal information principles, there are physiological studies suggesting

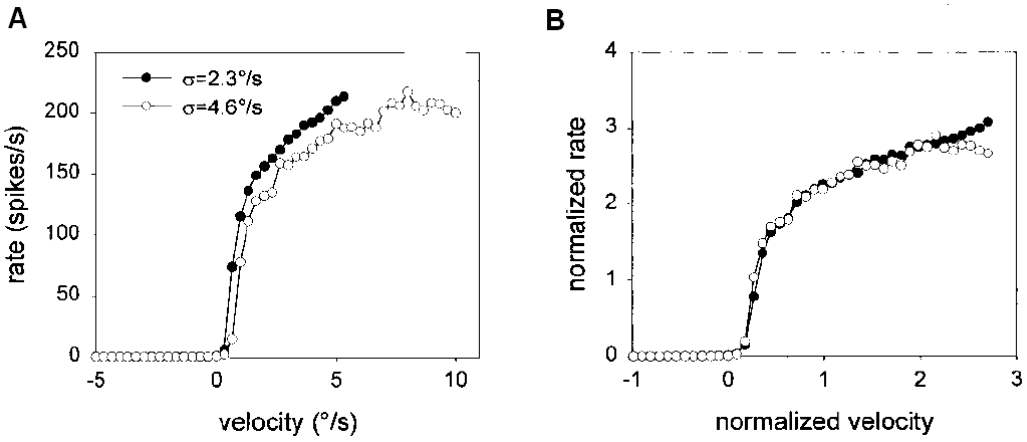


FIGURE 3.7: Responses of H1 neurons in the visual system of a blowfly to two stimulus ensembles, $\sigma_1 = 2.3°/s$ (closed circles) and $\sigma_2 = 4.6°/s$ (open circles). (A) Response as a function of stimulus velocity. (B) Normalized response as a function of a normalized stimulus velocity. The response firing rate is normalized by the time-averaged firing rate, and the stimulus velocity is normalized by the standard deviation. From Brenner et al. (2000).

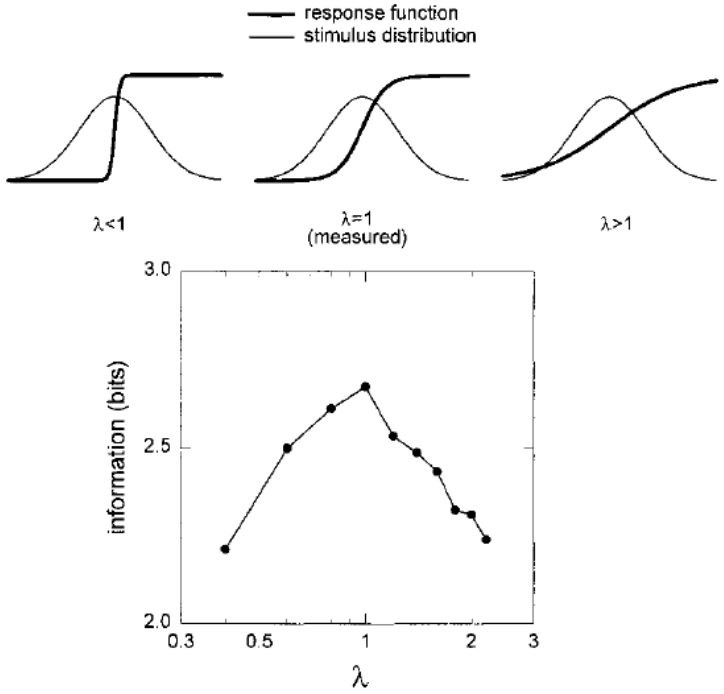


FIGURE 3.8: Information transmission as a function of the stretch factor λ for a H1 neuron in the blowfly. The three panels on top illustrate the effect of stretch factor λ on the input/output relation. When $\lambda = 1$, the response function corresponds to the measured data in the experiments. The bottom panel shows the calculated information transmission curve as a function of the stretch factor. From Brenner et al. (2000).

that in some areas in the brain, neurons actually function under a different principle. For example, Maravall et al. (2007) showed that neurons in the rat primary somatosensory "barrel" cortex adapt to repetitive whisker stimulation to maintain a constant information transmission. They suggested that the reason for information maintenance in barrel cortex might be that, in contrast to the many orders of magnitude in visual stimuli, whisker stimuli themselves have a constrained dynamic range, and somatosensory neurons thus do not need to adjust their responses in the same way as some visual neurons. Ringach and Malone (2007) studied V1 neurons and proposed that cortical cells translate and rescale the generator potential to adjust their operating point dynamically, which is defined by the mean and standard deviation of the generator signal, so that the neurons only respond when the generator potential is much larger than its mean value. This again appears to be distinct from the optimal information principle.

As introduced before, there are different information principles underlying neuronal processing along the sensory pathways. As information is transmitted from early sensory processing stage to a later stage, the functional requirements might vary, for example, from elementary pixel-to-pixel processing in the retina, to high-level feature-based processing in visual cortex. Furthermore, along different sensory channels, the range and content of the stimuli are also diverse. It is therefore natural to observe different principles underlying the processing of these stimuli. The proposed information principles, such as efficient coding and maximum information transmission, merely correspond to sensory processing at different stages in the brain, and none of them can act as a general principle. In Chapter 7, we will suggest another interpretation of the computational strategies of sensory neurons.

3.4 Summary

Many results show evidence that adaptation to input statistics is prevalent in unimodal neurons in the sensory systems. To explore this adaptation, it is beneficial to study the statistical properties of natural stimuli. Natural image statistics has been under intensive study, showing the existence of some general properties. For example, it has been found that the power spectra of natural images comply with a power law form (Field, 1987; Burton and Moorhead, 1987) and natural images exhibit the scale invariance property (Burton et al., 1986). Furthermore, studies on natural images showed that higher-order statistics exist (Field, 1987; Mallat, 1989). Analogous to natural images, natural sound statistics also exhibit a power law form of its amplitude power spectra (Voss and Clarke, 1975; Attias and Schreiner, 1997), and contain higher-order statistics (Bell and Sejnowski, 1996; Iordanov and Penev, 1999).

Receiving natural stimuli, sensory neurons adapt their responses according to the input statistics. In light adaptation, for example, visual neurons adapt to the mean level of

luminance, centring their response ranges around this mean level (Normann and Werblin, 1974; Shapley and Enroth-Cugell, 1984). Moreover, sensory neurons also adapt to the variance of the stimuli. For example, visual neurons exhibit both "contrast gain control" (Shapley and Enroth-Cugell, 1984) and "contrast adaptation" (Smirnakis et al., 1997; Kim and Rieke, 2001). In addition to neuronal adaptation to the mean and the variance, it has been found that sensory neurons also adapt to the correlation of the stimulus (Lesica et al., 2007) and to higher-order statistics (Kvale and Schreiner, 2004; Sharpee et al., 2006).

Many studies have proposed possible principles underlying these adaptations based on information theory. Efficient coding, for example, suggests that the redundancy in natural signals is removed in sensory neurons (Attneave, 1954; Barlow, 1961), which is supported by finding that interneurons in the compound eye of a fly appear to decorrelate the input stimuli (Srinivasan et al., 1982). Moreover, large monopolar cells in insect compound eyes respond in a way to maximize response entropy (Laughlin, 1981) and achieve an optimal coding. Other information principles have also been proposed, such as maximizing information transmission (Brenner et al., 2000) and maintaining a constant information transmission (Maravall et al., 2007).

Chapter 4

Neural Models for Multisensory Integration

Multisensory integration in the brain, or particularly in the superior colliculus (SC) or the optic tectum (OT), has been widely studied by neurophysiological and behavioural experiments that, as a consequence, produce many significant results. Based on these experimental results, mathematical and computational models have been proposed to reproduce the reported properties of multisensory neurons and to account for the underlying mechanisms. In this chapter, we will first discuss neural models of the response properties of neurons in the deep layers of superior colliculus (DSC), and then introduce models that focus on other aspects of multisensory integration.

4.1 Models on Neuronal Responses

4.1.1 Patton and Anastasio's Model (2003)

As discussed in Chapter 2, neurophysiological experiments have revealed that DSC neurons exhibit both cross-modal enhancement (CME) (Meredith and Stein, 1983) and modality-specific suppression (MSS) (Kadunce et al., 1997). Neuronal mechanisms underlying these two contradictory response properties remain unknown. Patton and Anastasio (2003) proposed that DSC neurons, which are postulated to receive stochastic inputs from multiple sensory channels, compute the posterior probability of a target using Bayes' rule. They implemented Bayes' rule in an augmented perceptron model and showed that, by proper adjustment of the parameters, the perceptron model is able to perform a Bayesian computation. Furthermore, simulation results showed that the model is able to exhibit both CME and MSS, indicating that the Bayes' rule model might account for the mechanisms underlying DSC neurons responses.

Denote a target by a binary random variable T. $T = 0$ represents that the target is absent and $T = 1$ represents that the target is present. The prior probability of T being absent or present is thus $P(T = 0)$ or $P(T = 1)$, respectively. Inputs from a sensory modality i are labeled as a random variable $M_i (i = 1, 2, ..., k)$, which has a non-negative value m_i representing the firing rate. As Patton and Anastasio (2003) proposed, a neuron computes the probability of $T = 1$ given a certain input vector $\mathbf{m} = (m_1, ..., m_k)^T$ where T for transpose, that is, the posterior probability $P(T = 1|\mathbf{m})$. According to Bayes' rule,

$$\begin{aligned} P(T = 1|\mathbf{m}) &= \frac{P(\mathbf{m}|T = 1)P(T = 1)}{P(\mathbf{m})} \\ &= \frac{P(\mathbf{m}|T = 1)P(T = 1)}{P(T = 1)P(\mathbf{m}|T = 1) + P(T = 0)P(\mathbf{m}|T = 0)}, \end{aligned} \tag{4.1}$$

where $P(\mathbf{m}|T = 0)$ is the probability of inputs with the absence of the target (the "spontaneous" inputs), and $P(\mathbf{m}|T = 1)$ is the probability of inputs in the presence of the target (the "driven" inputs). In addition, in an augmented perceptron model that has product units or pi nodes, the weighted sum of the input u is computed as

$$u = b + \sum_{i=1}^{k} w_i m_i + \sum_{i=1}^{k} \sum_{j=i}^{k} \rho_{ij}(m_i m_j), \tag{4.2}$$

where b is the bias, w_i is the synaptic weight for input m_i and ρ_{ij} is the weight for a pi node. The weighted sum u is then passed through a sigmoidal function,

$$f(u) = \frac{1}{1 + e^{-u}}, \tag{4.3}$$

to evoke neuronal responses. A neuron that performs the Bayesian computation will then have $f(u) = P(T = 1|\mathbf{m})$, and u is accordingly written as

$$u = \ln \left[\frac{P(\mathbf{m}|T = 1)}{P(\mathbf{m}|T = 0)} \right] + \ln \left[\frac{P(T = 1)}{P(T = 0)} \right], \tag{4.4}$$

based on which the weights and the bias in Eq. (4.2) can be expressed in terms of the prior probabilities and the spontaneous and driven input statistics.

According to the classical central limit theorem, the distribution of a combination of many independent and identically-distributed signals will be gaussian (Fristedt and Gray, 1997). Since each DSC neuron receives inputs from a large number of different neurons, Patton and Anastasio (2003) assumed that the input to a DSC neuron is approximately gaussian, given by

$$P(\mathbf{m}|T = t) = \frac{1}{(2\pi)^{k/2}|\mathbf{\Sigma}_t|^{1/2}} \exp \left[-\frac{1}{2}(\mathbf{m} - \boldsymbol{\mu}_t)^T \mathbf{\Sigma}_t^{-1} (\mathbf{m} - \boldsymbol{\mu}_t) \right], \tag{4.5}$$

where $\boldsymbol{\mu}_t$ is a vector of the mean of each input when $T = t$ $(t = 0, 1)$, and $\mathbf{\Sigma}_t$ is a k

by k covariance matrix of the input channels. In the case of bivariate Gaussian inputs, a two-channel augmented perceptron model was studied (Fig. 4.1). Fig. 4.2 shows the CME produced by the perceptron model as a function of the sensory inputs, in which maximum enhancement occurs at relatively weak input and the enhancement gradually decreases when input increases, corresponding well to CME with inverse effectiveness.

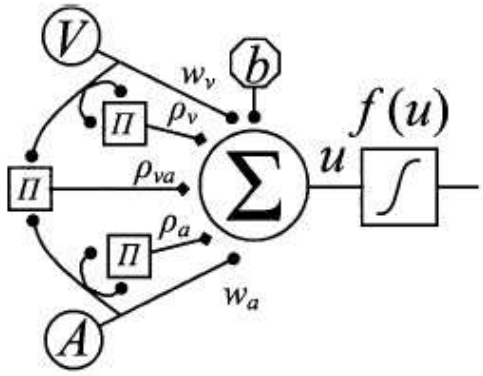


FIGURE 4.1: A two-channel augmented-perceptron model. From Patton and Anastasio (2003).

Furthermore, to illustrate the ability of the model to produce both CME and MSS, the two-channel augmented perceptron model was extended to three sensory channels. Fig. 4.3 shows this augmented perceptron, in which V and X are two visual channels and A represents the auditory channel. Patton and Anastasio (2003) assumed that the paired modality-specific stimuli to V and X channels activate separate but overlapping sets of visual neurons. This leads to receptive fields overlap in V and X channels, which then covary under driven conditions. Furthermore, as Patton and Anastasio (2003) suggested, the spontaneous activities of those overlapping visual neurons will cause V and X channels to covary under spontaneous conditions. Consequently, their spontaneous covariance σ^2_{vx0} and driven covariance σ^2_{vx1} are larger than those between V and A (σ^2_{va0} and σ^2_{va1}) or between X and A (σ^2_{xa0} and σ^2_{xa1}). In the model, therefore, within-modality covariances are set as $\sigma^2_{vx0} = 1.6$ and $\sigma^2_{vx1} = 3.6$, larger than the cross-

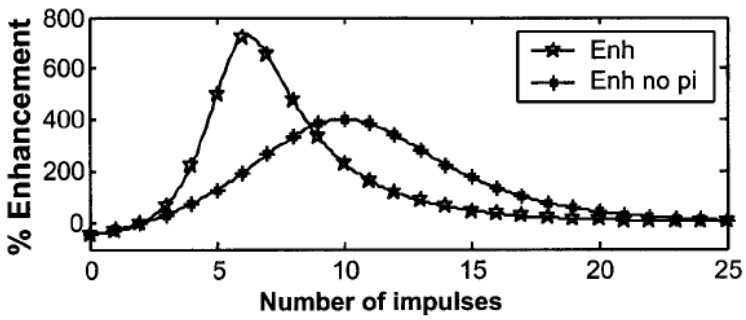


FIGURE 4.2: CME produced by an intact two-channel augmented-perceptron model (stars) or one that without the pi nodes (asterisks). Inputs V and A are set to be equal and varied from 0 to 25. CME is calculated under Eq. (2.1). From Patton and Anastasio (2003).

modality covariances set as $\sigma^2_{va0} = \sigma^2_{xa0} = 0.1$ and $\sigma^2_{va1} = \sigma^2_{xa1} = 2.8$. As shown in Fig. 4.4, V and A channels exhibit enhancement, whereas V and X channels exhibit response depression.

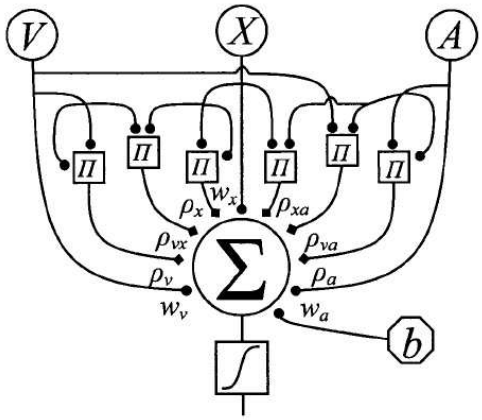


FIGURE 4.3: A three-channel augmented-perceptron model. From Patton and Anastasio (2003).

Patton and Anastasio (2003) have shown that suppression becomes evident when the spontaneous covariance σ^2_{vx0} approaches the spontaneous variance σ^2_{v0} or σ^2_{x0} ($\sigma^2_{v0} = \sigma^2_{x0}$ in their simulation), given that the driven variances are greater than the spontaneous variances. Based on these results, they suggested that suppression in DSC neurons occurs when these two conditions are satisfied, and the reason why DSC multisensory neurons exhibit MSS is that within-modality sensory channels are more likely to meet the two conditions. In addition, Patton and Anastasio (2003) suggested that the function of the pi nodes in the model might be implemented by N-methyl-D-aspartate (NMDA) receptors in DSC neurons. Binns and Salt (1996) applied NMDA receptor antagonist 2-amino-5-phosphonopentanoate (AP5) to multisensory DSC neurons in the cat and found that AP5 resulted in a greater percentage reduction in multisensory responses than unisensory responses, and thus a reduction in CME. This is comparable to the performance of the model after removing the pi nodes (Fig. 4.2).

This augmented perceptron model appears to unify the contradictory phenomena of CME and MSS under Bayes' rule. However, some studies have shown that multisensory responses and unisensory responses might involve different mechanisms. Kadunce et al. (1997) showed that MSS exhibit properties different from multisensory depression, indicating that MSS might be mediated at a different circuit from multisensory responses. Moreover, removal of cortical influences eliminates multisensory enhancement in DSC neurons, whereas the neurons still exhibit MSS (Jiang et al., 2001; Alvarado et al., 2007). It is therefore possible that CME and MSS are not the product of a unified mechanism as suggested by Patton and Anastasio (2003). In addition, as shown in the model, the spontaneous covariance σ^2_{vx0} needs to be large enough for the model to exhibit MSS. However, whether modality-specific sensory channels strongly covary under spontaneous conditions to result in large σ^2_{vx0} is unclear and requires further justification. To test

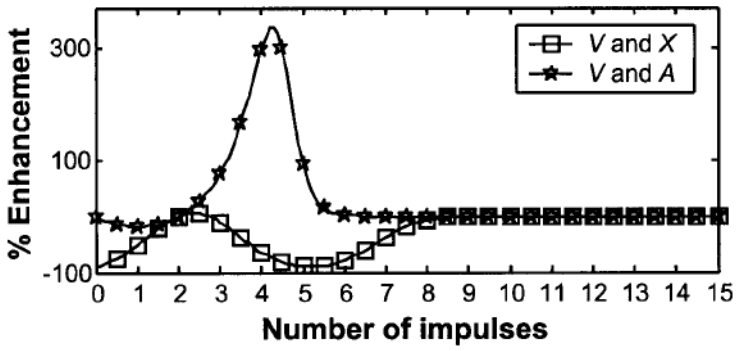


FIGURE 4.4: CME produced by the three-channel augmented-perceptron model. Both V and X are visual channels and A represents auditory channel. Within-modality covariances are set as $\sigma^2_{vx0} = 1.6$ (spontaneous case) and $\sigma^2_{vx1} = 3.6$ (driven case) and the cross-modality covariances are set as $\sigma^2_{va0} = \sigma^2_{xa0} = 0.1$ (spontaneous case) and $\sigma^2_{va1} = \sigma^2_{xa1} = 2.8$ (driven case). For cross-modal inputs (V and A), inputs in the X channel are fixed at their spontaneous rate 2, while V and A inputs are set to be equal and varied from 0 to 15 (abscissa). For within-modality inputs (V and X), inputs in the A channel are fixed at their spontaneous rate 2, while V and X are set to be equal and varied from 0 to 15. From Patton and Anastasio (2003).

whether their model plausibly accounts for DSC responses, we performed both numerical and analytical studies on the model, particularly focussing on the robustness of the model to different parameters and the role of σ^2_{vx0} in inducing MSS. In Chapter 5, we will present a detailed discussion of our analysis.

4.1.2 Anastasio and Patton's Model (2003)

Despite the fact that multisensory integration has been observed in DSC, not all DSC neurons are multisensory. For example, in cat, only about one-half of DSC neurons are multisensory; in monkey, only about a quarter are multisensory (Wallace and Stein, 1996). The reason why there is a mixture of unisensory and multisensory neurons in DSC remains unclear. On the other hand, descending projections from cortical areas AES/rLS to DSC appear to be an indispensable component for CME. Eliminating the effects of AES/rLS, CME disappears (Alvarado et al., 2007). The role of these cortical projections in DSC neuronal responses also remains unanswered. To provide possible answers to these two open questions, Anastasio and Patton (2003) proposed a corticotectal network model that develops according to a two-stage, unsupervised learning algorithm.

The network model consists of 100 DSC neurons arranged in a square 10×10 grid, each of which receives both primary ascending inputs X_j $(j = 1, 2, 3)$ from visual (V), auditory (A) and somatosensory (S) systems and modulatory descending inputs Y_k $(k = 1, 2, 3)$ from unimodal visual, auditory and somatosensory cortical areas (Fig. 4.5). A target T in the environment provides V, A, S sensory stimuli to DSC neurons. Consequently, the target has eight states, corresponding to a complete combination of the three types of sensory stimuli being absent or present. The target T being ab-

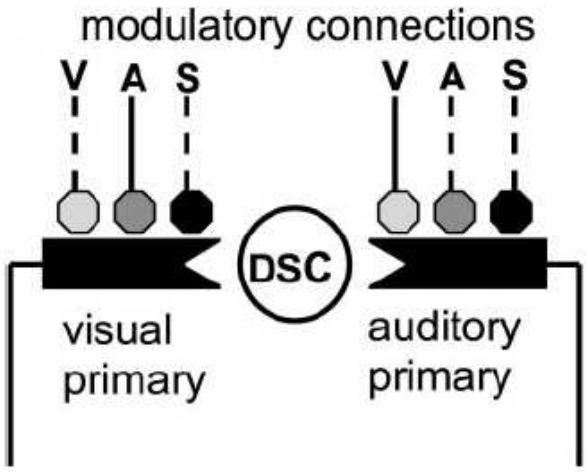


FIGURE 4.5: Schematic diagram for the corticotectal model of a DSC neuron. After stage-one training, a primary inputs might be modulated by V, A, S inputs from cortical areas (solid and dash lines). After stage-two training, the primary connections only receive modulatory inputs conforming to the "modality-matching" and "cross-modality" constraint (solid lines). From Anastasio and Patton (2003).

sent is represented by the state $(V = 0, A = 0, S = 0)$ with a probability arbitrarily set as $P(T = 0) = 1/2$. The remaining seven states correspond to $T = 1$ and are divided into two categories for clarity, one for unisensory target with probability p_s [(V=1,A=0,S=0),(V=0,A=1,S=0) and (V=0,A=0,S=1)], and the other for multisensory target with probability p_c [(V=1,A=1,S=0),(V=1,A=0,S=1),(V=0,A=1,S=1) and (V=1,A=1,S=1)], where we have $p_s + p_c = 1/2$. The instances of X_j and Y_k are denoted as x_j and y_k conforming to a binomial distribution $b(n, p)$, where n is for simplicity set as 20 and p is either the spontaneous probability or the driven probability depending on whether T is absent or present. In the driven case, p takes p_{x1} for primary inputs and takes p_{y1} for modulatory inputs; in the spontaneous case, p takes p_{x0} and p_{y0} for primary and modulatory inputs respectively. A weighted sum of the primary inputs is passed through a sigmoidal function to activate a DSC neuron with a response z_i,

$$z_i = \frac{1}{1 + \exp[-\gamma(\sum_j w_{ij} x_j - \theta)]}, \tag{4.6}$$

where θ is called the threshold and set as $\theta = 10$, γ is the gain adjusting the sensitivities of the response function and set as $\gamma = 1/5$, and w_{ij} represents the weight of connections between a primary input x_j and a DSC unit z_i. The weight w_{ij} is adjusted according to

$$w_{ij} = u_{ij} + \sum_k v_{ijk} y_k, \tag{4.7}$$

where u_{ij} represents the weight of the connection from the primary input j that is not modulated by cortical inputs, and v_{ijk} represents the weight of the projection from a modulatory input y_k.

Learning in the network model is composed of two stages, in which the primary weights and modulatory weights are adjusted respectively. In the first stage, the modulatory weights v_{ijk} are set to 0, so $w_{ij} = u_{ij}$, and the primary weights u_{ij} are initially set to random values uniformly distributed between 0 and 0.1 and are then adjusted with a self-organizing map algorithm (Kohonen, 1982). Stimulations to the model are generated according to the probability of the eight states of target T, after which the primary inputs are determined by the binomial distribution $b(n, p)$, where $p = p_{x1}$ when T is present ($T = 1$) and $p = p_{x0}$ when T is absent ($T = 0$). The activation of DSC units z_i is then determined by the primary inputs according to Eq. (4.6). The neuron with maximum z_i is identified as the winner and a neighbourhood h of a 5×5 square centred on the winner is determined, in which the winning DSC unit has activity 1, the nearest eight have activity 0.3 and the other 16 have activity 0.1. For DSC neurons in the subset h, the primary weights are adjusted according to their activities,

$$u_{hj} = u_{hj} + \alpha z_h x_j, \tag{4.8}$$

where α is the learning rate and decreases from 0.1 to 0.01 during training to produce stable results. The primary weights to a DSC unit are normalized so that $\sqrt{\sum_j (u_{hj})^2} = 1$. The training contains 5000 iterations, after which a pruning process occurs so that if $u_{ij} < \theta_u = 0.4$, then $u_{ij} = 0$. A re-normalization of the weights is applied afterwards.

In the second stage, the primary weights u_{ij} are fixed, and the modulatory weights v_{ijk} are initially set to zero and then trained based on the correlation and anti-correlation between modulatory inputs, primary inputs and the responses of DSC units. The learning algorithm for the second stage is inspired by two findings in neurophysiological results. One is that the sensory modalities of inputs from the AES/rLS to a DSC neuron match with the modalities the neuron receives from other sources (Wallace et al., 1993), indicating a property of "modality-matching" (Fig. 4.5); the other is the evidence that descending AES/rLS cortical inputs are from unimodal areas and have significant influences only on cross-modal responses (Jiang et al., 2001), based on which Anastasio and Patton (2003) suggested a "cross-modality" constraint in which cortical inputs are modulatory and only affect primary inputs of a different modality (Fig. 4.5). To determine whether the inputs or the DSC units are active, thresholds θ_x and θ_y are defined for primary and modulatory inputs respectively according to the intersection point of the corresponding spontaneous and driven binomial distribution, and θ_z is determined empirically for DSC units. The second-stage learning algorithm is accomplished by means of dummy variables d_{ijk}:

$$d_{ijk} = d_{ijk} + \beta, \quad \text{if } y_k > \theta_y \ \&\ z_i > \theta_z \ \&\ x_j \le \theta_x, \tag{4.9}$$

$$d_{ijk} = d_{ijk} - \beta, \quad \text{if } y_k > \theta_y \ \&\ z_i > \theta_z \ \&\ x_j > \theta_x, \tag{4.10}$$

$$d_{ijk} = d_{ijk} - 2\beta, \quad \text{if } y_k > \theta_y \ \&\ z_i \le \theta_z, \tag{4.11}$$

$$d_{ijk} = d_{ijk}, \quad \text{if } y_k \le \theta_y. \tag{4.12}$$

Eqs. (4.9) and (4.10) ensure the cross-modality constraint by increasing modulation for inactive primary inputs and decreasing modulation for active primary inputs when both the modulatory inputs and DSC units are active. According to Eq. (4.11), the modulation is decreased if a modulatory input is active but a DSC unit is not, ensuring the modality-matching constraint. The modulatory weight v_{ijk} is set to d_{ijk} only when $d_{ijk} > 0$ and with no change otherwise. During the training, the modulatory weights are constrained to be positive with an upper limit of one.

After stage-one training, a mixture of unimodal and multisensory DSC units emerges. The percentage of multisensory units is influenced by several factors. As shown in Fig. 4.6, an increase in θ_u, which is adopted in the pruning process for primary inputs, decreases the percentage of multisensory units. This is presumably because more weights are pruned and thus fewer neurons respond to stimuli from different modalities. Moreover, the proportion of modality-specific to cross-modal targets, which is determined by the probability of unisensory targets p_s, also affects the composition. Fig. 4.6 shows that a decrease in p_s increases the percentage of multisensory neurons, indicating that more neurons develop responses to multisensory stimuli in the presentation of more cross-modal targets during the training.

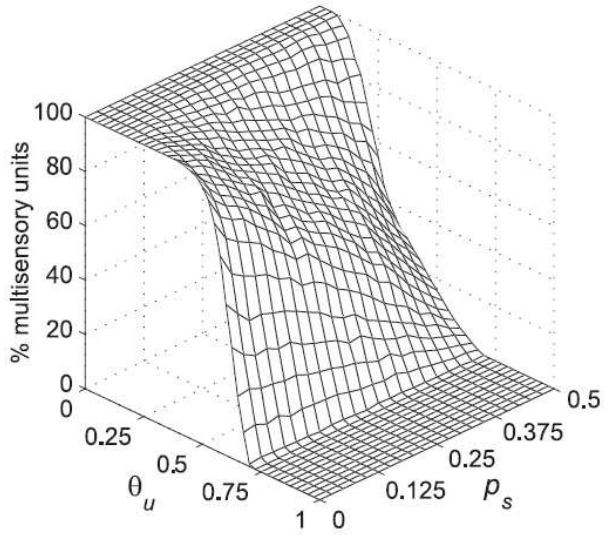


FIGURE 4.6: The percentage of multisensory neurons after stage-one training as a function of primary weight threshold θ_u and the probability of a unimodal target p_s. For each value of p_s, 10 networks are trained. Each of the 10 networks are thresholded at each value of the primary input threshold θ_u. The percentage of multisensory neurons is an average of the 10 networks. From Anastasio and Patton (2003).

Training in the second stage does not influence the percentage of multisensory neurons but results in significant multisensory enhancement. Figs. 4.7A-D shows that removing the modulatory connections in the model significantly reduced CME. Unimodal responses were also slightly reduced, because the removal of modulatory connections also eliminated the influence of spontaneous activities in modulatory inputs. These properties comply with the experimental results on the role of AES/rLS descending inputs in

DSC neuronal responses (Jiang et al., 2001). Furthermore, as shown in Figs. 4.7E-H, decreasing the spontaneous activity of primary inputs by reducing p_{x0} resulted in larger amount of CME, indicating that the spontaneous activities of sensory inputs also influence CME. Although experiments on anesthetized animals revealed large CME, Populin and Yin (2002) have shown that experiments on behaving cats failed to exhibit CME. Anastasio and Patton (2003) suggested the reason may be that spontaneous activities in DSC neurons were reduced by anesthesia which, as a consequence, induces larger CME.

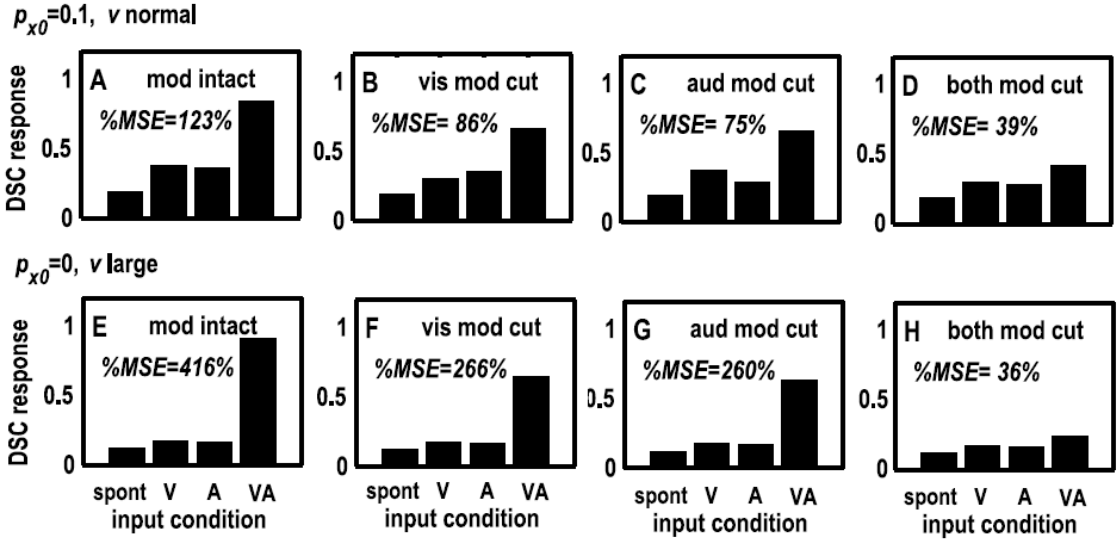


FIGURE 4.7: Multisensory enhancement (MSE) in the corticotectal model. MSE, or equally CME, is calculated under Eq. (2.1). The primary inputs that produced the maximum MSE is chosen. In the first row, the spontaneous activation probability for primary inputs is set as $p_{x0} = 0.1$. Responses and the CME for the intact model (A), the model without visual modulatory input (B), the model without auditory modulatory input (C) and the model without both visual and auditory modulatory inputs (D) are shown. In the second row, $p_{x0} = 0$ and (E)-(H) show the responses corresponding to (A)-(D). The modulatory weights have been increased by seven times (v large) compared with those in the first row (v normal). From Anastasio and Patton (2003).

The two-stage algorithm suggests possible answers to two important questions in the response properties of DSC neurons, that is, the existence of a mixture of unisensory and multisensory neurons in DSC and the role of unimodal cortical descending inputs in CME. In contrast to the model in Patton and Anastasio (2003) (Section 4.1.1), the learning algorithm does not endow DSC neurons with the capability to estimate target probabilities. To reconcile these two ideas, Anastasio and Patton (2003) suggested that some mechanisms like the two-stage algorithm might set up a coarse corticotectal circuitry, which is then tuned under some supervised learning algorithm to compute the target probabilities. However, the selections of some parameters in the model have not been justified, such as n in the input binomial distribution $b(n, p)$, the threshold θ and the gain γ in neuronal response function. Their influences on the performance of the model require further exploration. Furthermore, since enhancement in DSC neurons can also be naturally accounted for with a non-linear saturating response function, as will

be discussed in Chapter 6, the idea that cortical afferents act as modulatory inputs is not necessarily the only possibility. In fact, after removing the modulatory connections in the model, CME can still be observed since the model employs a sigmoidal response function (Fig. 4.7). It is therefore possible that cortical inputs affect DSC neurons in another way. We will discuss another interpretation of the impact of cortical influences on CME in Chapter 9.

4.1.3 Schauer and Gross's Model (2004)

Schauer and Gross (2004) proposed a computational model inspired by the response properties of DSC neurons to realize a robust integration of visual and auditory information. The general idea is to design a network model for visual-auditory integration and then determine the parameters in the model according to some biologically-inspired criteria. Stimuli are generated from a database containing recordings of sound and visual scenes that can be assembled to simulate reasonable cross-modal stimuli, which can be analyzed offline. With this model, they try to achieve low-level integration where visual neurons respond to changes in signal intensities and where auditory neurons process basic interaural time difference (ITD) information of auditory signals.

This computational model consists of both the visual and auditory sensory systems (Fig. 4.8). The visual system simply contains a model of the superficial layers of SC that codes the intensity differences for scene motion. The auditory system calculates ITD of the binaural signals for localization, and includes models of several principal neural structures, such as a cochlea model using an all-pole gammatone filter to simulate the mechanical properties of the basilar membrane (Slaney, 1988), a coincidence model for medial superior olive to cross-correlate the left and right signals (Jeffress, 1948), and a model of the external inferior colliculus (ICx) to produce a nontonotopic representation of the azimuthal locations based on ITD. Visual and auditory signals are then integrated in a model of DSC using a winner-take-all network with global inhibition. Specifically, a dynamic neural field of Amari type is adopted to integrate visual and auditory inputs (Amari, 1977),

$$\begin{aligned}\tau\frac{d}{dt}z(r,t) = &-z(r,t) + c_A x_A(r,t) + c_V x_V(r,t) \\ &-c_i \int y(z(r,t))dr + c_n \int w(r-r')y(z(r',t))dr',\end{aligned} \tag{4.13}$$

where $z(r,t)$ is the state of a neuron at a position r and at a time step t, $c_A x_A(r,t) + c_V x_V(r,t)$ is the weighted sum of auditory input x_A and visual input x_V, $c_i \int y(z(r,t))dr$ is a global inhibition from the whole network and $c_n \int w(r-r')y(z(r',t))dr'$ is the lateral feedback from neurons at neighbouring position r'. The output of a neuron is then

determined by a sigmoidal function,

$$y(z(r,t)) = \frac{1}{1+\exp[-\sigma \cdot z(r,t)]}. \tag{4.14}$$

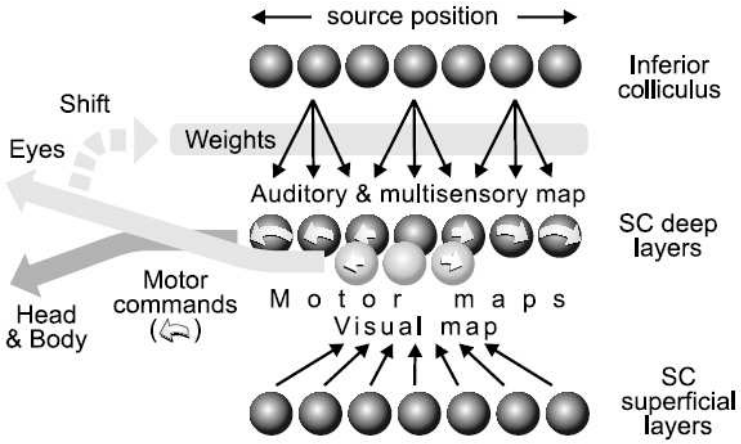


FIGURE 4.8: Schematic diagram of the auditory-visual integration model by Schauer and Gross (2004). From Schauer and Gross (2004).

This winner-take-all structure has a high-dimensional parameter space, which can be reduced under some constraints. For example, the time constant τ of the neural field was selected as 250 ms depending on the time-window in which interaction between cross-modal stimuli should occur in the experiments. Some other parameters, such as the weights of inputs, lateral feedbacks and global inhibition remain to be determined. To this end, Schauer and Gross (2004) proposed several evaluation criteria inspired by the response properties of DSC neurons, such as the maximum enhancement criterion that provides an upper limit for CME, the mean enhancement criterion that constraints the mean of all CME values and the single modality criterion that provides a lower bound for the sum of unisensory responses. Based on a search under these criteria, the suitable parameter sets could be obtained. For example, if the mean response enhancement criterion is taken as the final optimization criterion, the other criteria can be used to determine a subspace of the parameters before applying the mean criterion to get the final optimal parameter set. The neural model with parameters obtained through the optimization process exhibits many multimodal properties that correspond to neurophysiological results on DSC neurons (Fig. 4.9), such as response enhancement for spatially-coincident stimuli, inverse effectiveness, and response depression for spatially-disparate stimuli (Stein and Meredith, 1993). Schauer and Gross (2004) claimed that the methods can also be applied to other network types besides the Amari-type.

Although some parts of the model are comparable to the model by Rucci et al. (1997) (Section 4.2.1), the focus is not on the adaptation of multimodal representations but to design a neural model to perform a robust multimodal attention-mechanism for artificial systems. This is also in contrast to many mathematical and computational models

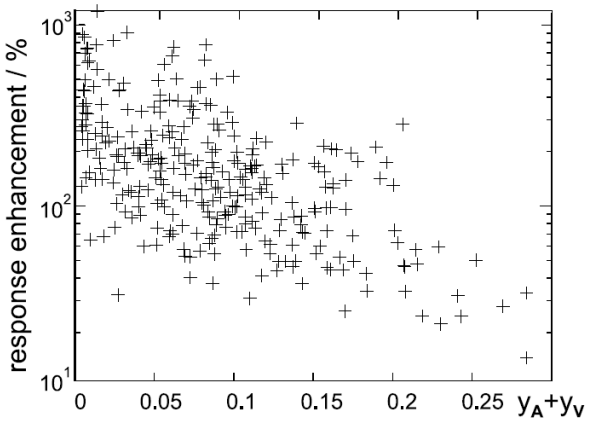


FIGURE 4.9: CME for 300 repetitive multimodal experiments with different temporal and spatial parameters. From Schauer and Gross (2004).

that aim to reveal the mechanisms underlying biological multisensory processes, such as Patton and Anastasio (2003), Anastasio and Patton (2003) and Cuppini et al. (2007). However, since the model exhibits comparable features to multimodal processing in DSC, some parts of the model might still reflect the mechanisms of multisensory integration in DSC neurons, such as the sigmoidal response function for multisensory enhancement and the lateral inhibition for multisensory depression.

4.1.4 Cuppini et al.'s Model (2007)

In the process of writing this thesis, we became aware of a neural network model that was proposed by Cuppini et al. (2007) to reproduce the response properties of DSC neurons, including CME, multisensory depression and MSS.

The neural network model consists of three layers, of which two represent visual and auditory unimodal areas and the other represents a multisensory area receiving inputs from the unimodal areas (Fig. 4.10).

In the unimodal areas, neurons have receptive fields determined by a Gaussian function and spanning approximately 10-15 deg and 20-25 deg in diameter for visual neurons and auditory neurons, respectively. These unimodal neurons also receive lateral connections, which are arranged according to a "Mexican hat" pattern (short-range excitatory and long-range inhibitory connections), from other neurons in the same area. Moreover, each neuron receives a one-to-one feedback from a DSC neuron according to the registration of cross-modal spatial maps in DSC neurons (Kadunce et al., 1997). The input u to a unimodal neuron indexed as ij is then written as

$$u_{ij}(t) = r_{ij}(t) + l_{ij}(t) + f_{ij}(t), \tag{4.15}$$

where r_{ij} represents the inputs aroused by a sensory stimulus, l_{ij} lateral inputs from

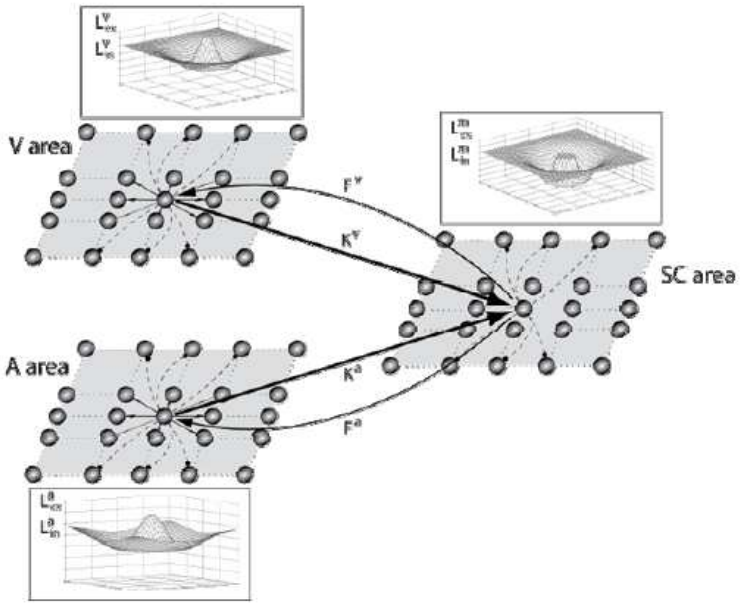


FIGURE 4.10: The general structure of the network model by Cuppini et al. (2007). V for visual area, A for auditory area and SC area for multisensory area consisting of DSC neurons. From Cuppini et al. (2007).

other neurons in the same area and f_{ij} the feedback from a DSC neuron. The activity x_{ij} of the neuron is then computed as

$$\tau \frac{d}{dt} x_{ij}(t) = -x_{ij}(t) + \varphi(u_{ij}(t)), \tag{4.16}$$

where τ determines the speed of responses to the stimulus, and φ represents a sigmoidal transfer function given by

$$\varphi(u(t)) = \frac{1}{1 + e^{-\gamma(u(t)-\theta)}} \tag{4.17}$$

with the threshold θ and the gain γ.

Each neuron in the multisensory area receives one-to-one inputs from each of the unimodal areas, according to the spatial registration of visual and auditory RFs in DSC neurons. Moreover, DSC neurons receive lateral connections from other DSC neurons. In contrast to neurons in the unimodal layer, a DSC neuron only receives inhibitory connections from distant DSC neurons in order to comply with experimental data suggesting multisensory depression for distant cross-modal stimuli and the absence of enhancement for proximal, within-modality stimuli. Accordingly, the overall input u to a DSC neuron is composed of two parts: inputs from the unimodal areas and lateral inhibitions from distant DSC neurons. The activity x of a DSC neuron is then computed in the same form as Eqs. (4.16) and (4.17).

Cuppini et al. (2007) showed that the model exhibits CME associated with IE, which changes to multisensory depression when the distance between the two stimuli increases. The property of MSS can also be produced in the model. Moreover, they demonstrated

that a DSC neuron in the model exhibits different dynamic ranges (the difference between neuronal activities at saturation and at threshold) for unimodal and bimodal stimuli, corresponding to the results shown by Perrault et al. (2005). The mechanisms underlying these properties have also been discussed based on the neural network model. For example, Cuppini et al. (2007) suggested that the non-linearity of neuronal response contributes to both the IE property of CME and the sub-additive/additive/super-additive responses in DSC neurons. They also indicated that the different dynamic ranges in DSC neurons for unimodal and bimodal stimuli are due to prior saturation in the unimodal neurons.

The neural network model explicitly or implicitly conveys several ideas on the operations underlying the responses of DSC neurons that are similar to our interpretation (see Chapter 6). For example, in the model, the mechanisms underlying CME, multisensory depression and MSS are different, where CME is mainly produced by the non-linearity of the response function, multisensory depression is contributed by lateral inhibitions in the multisensory layer and MSS is mediated in the unimodal layers. Moreover, the different dynamic ranges for unimodal and bimodal stimuli are generated by prior saturation in unimodal neurons, exactly the same as that which we will discuss in Chapter 6. However, although the structure of the model is sensible, the integration of cross-modal stimuli in the model is essentially a sum of the inputs from visual and auditory layers, which has probably oversimplified the operation of DSC neurons. Furthermore, Cuppini et al. (2007) only provided a brief discussion on the non-linearity of the response function to produce CME and additivity response properties. We will present in Chapter 6 a detailed and thorough discussion of the same issue. Cuppini et al. (2007) do not provide many parameters, such as those for lateral connections in the unimodal and multisensory layers and the threshold θ and the gain γ in the sigmoidal response function, of the neural network model, so a further exploration of the properties of the model is not possible.

4.2 Other Models

4.2.1 Rucci et al.'s Model (1997)

In the OT of barn owl, visual and auditory spatial representations closely align with each other. Experiments have revealed a high degree of plasticity in the external nucleus of the inferior colliculus (ICx) when this spatial map alignment is disrupted (Brainard and Knudsen, 1993; Knudsen, 1994). However, the mechanisms by which multisensory spatial map is constructed and maintained remain unclear. Rucci et al. (1997) proposed a model of the principal neuronal structures involved in the spatial localization and orientation behaviour of a barn owl to visual and auditory stimulation.

In the barn owl, visual signals are transmitted through direct retinotectal projections

to OT. The pathway for auditory azimuth localization involves several neural structures starting from the magnocellular cochlear nuclei, to nucleus laminaris (NL), the central nucleus of the inferior colliculus (ICc) and to the external nucleus of the inferior colliculus (ICx), and finally to OT (Konishi et al., 1988). Receiving inputs from both the left and right monaural structure magnocellular nuclei, neurons in NL are sensitive to specific frequency bands and respond maximally to specific values of ITD in binaural stimuli (Konishi, 1993). In the model, the outputs of NL neurons are simulated as the ITD corresponding to a given stimulus position. Driven by outputs from NL, ICc neurons have narrow sensitivities to both the frequency and ITD of the signals and are tonotopically arranged according to the frequency along one direction and according to ITD along another perpendicular direction (Wagner et al., 1987). Auditory signals across different frequency bands from ICc are then combined in ICx to form an auditory representation of the space (Knudsen et al., 1977), after which the auditory information is integrated with the visual information in the optic tectum. The architecture of the system is shown in Fig. 4.11.

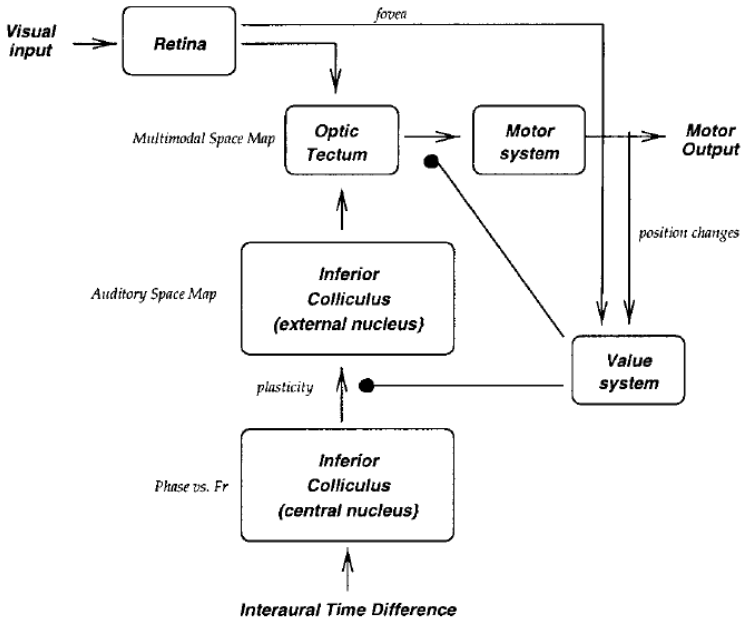


FIGURE 4.11: Schematic diagram of the components of the neural model by Rucci et al. (1997). From Rucci et al. (1997).

The main focus of the model is on the emergence and plasticity of the multisensory spatial representation in the OT. It has been suggested that the main site of plasticity in spatial map realignment is the ICx (Brainard and Knudsen, 1993). Although the mechanism remains unclear, Rucci et al. (1997) proposed that this plasticity is mediated by a learning strategy called "value-dependent learning", in which the synaptic changes are modulated by the saliency of sensorimotor event through modulatory systems such as the monoaminergic and cholinergic systems (Hasselmo, 1995; Mirenowicz and Schultz, 1996). Based on this notion, they implemented the modulatory system as a single value unit v that receives inputs from both the foveation area F_v and the two motor neurons

M_k $(k = 1, 2)$:

$$V(t) = \mathscr{F}\left(\sum_{k \in F_v} \rho_k R_k + \sum_{k=1,2} \chi_k M_k + \delta_V V(t-1) + n_v\right), \tag{4.18}$$

where $V(t)$ is the activation of the value unit, R_k is the input from the receptors of the retina, ρ_k is the weight of the projections from the retina, χ_k is the weight of afferents from the motor neurons, n_v represents the noise contribution and δ_V is a time constant. $\mathscr{F}(x)$ is a piecewise linear approximation of a sigmoidal function with an offset 0, slope 1 and saturation at $x = 1$ (Fig. 4.12).

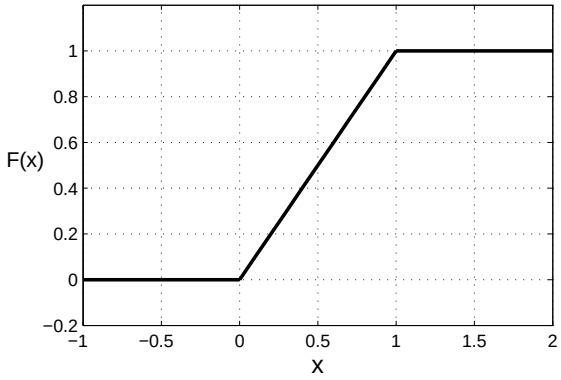


FIGURE 4.12: A piecewise linear approximation of a sigmoidal function with an offset 0, slope 1 and saturation 1.

From Eq. (4.18), the value unit tends to fire when a stimulus is in the foveation area and the motor neurons are activated, and then modulates the connections between ICc and ICx through diffuse projections λ_k. At each time step t, the weight w_{ijk} of the connection between a ICc neuron u_{ij} and a ICx neuron a_k is modified according to

$$w_{ijk}(t+1) = w_{ijk}(t) + \Phi_L(E_p). \tag{4.19}$$

In Eq. (4.19), $\Phi_L(x)$ is a piecewise linear approximation of the plasticity function proposed by Artola and Singer (1993) and is characterized by two thresholds θ_{LTD} and θ_{LTP} (Fig. 4.13),

$$\Phi_L(x) = \begin{cases} 0, & \text{if } x < \theta_{LTD}, \\ -k_1(x - \theta_{LTD}), & \text{if } x \in [\theta_{LTD}, \theta'), \\ +k_2(x - \theta_{LTP}), & \text{if } x \in [\theta', \theta_{LTP}), \\ +k_3(x - \theta_{LTP}), & \text{if } x \geq \theta_{LTP}, \end{cases} \tag{4.20}$$

where $\theta' = (k_2\theta_{LTP} - k_1\theta_{LTD})/(k_1 + k_2)$, and k_1, k_2 and k_3 are positive real numbers determining the slope of the piecewise lines. E_p is calculated based on two parts, one is a local factor depending on the activation of pre- and postsynaptic neurons, and the other is a global factor shared by all the synapses depending on the activation of the

value system:

$$E_p = \epsilon_1 \mathscr{H}(U_{ij}, A_k) + \epsilon_2 \lambda_k V, \tag{4.21}$$

where $\mathscr{H}$ can be a product of U_{ij} and A_k (the classical Hebbian term) or a sum of them, and ϵ_1 and ϵ_2 are constants that determine the relative contributions of the two factors. According to $\Phi_L(x)$, the weight w_{ijk} can remain unchanged, potentiate or depress depending on E_p. The thresholds θ_{LTD}, θ_{LTP} were selected in a way that potentiation occurs when the correlation between pre- and postsynaptic activities and the activation of the value system are high [$x > \theta_{LTP}$ in $\Phi_L(x)$], corresponding to the case of a successful orientation to bring the stimulus into the fovea. In contrast, w_{ijk} undergoes depression for highly correlated units when the orientation fails to bring the stimulus into the fovea [$\theta_{LTD} < x < \theta_{LTP}$ in $\Phi_L(x)$].

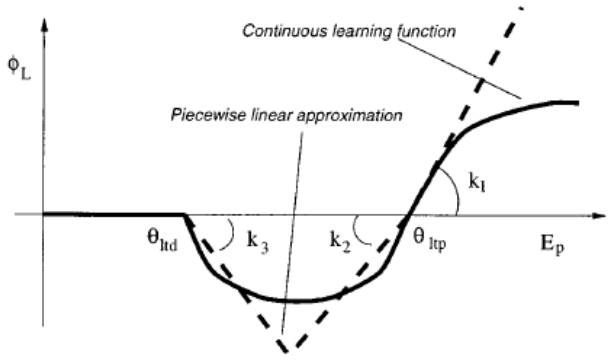


FIGURE 4.13: A piecewise linear approximation of the curve proposed by Artola and Singer (1993) for the modification of synaptic weights. From Rucci et al. (1997).

Based on this learning rule, Rucci et al. (1997) has simulated a neural system both under normal sensory experience and under altered visual experience. There is a bias in the initial connections between ICc and ICx so that the afferents received by ICx from a corresponding ICc area follows a probability that decreases with the distance from the area. This initial connection pattern forms a coarsely topographical organization. Trained with normal sensory experience, the development of a system composed of 320×50 units in ICc and 100 units in all of the other maps is observed. As shown in Fig. 4.14A, after the presentation of 15,000 audio-visual stimuli, the auditory and visual receptive fields are aligned precisely. The foveation error also decreases significantly both in the mean values and the standard deviation (Fig. 4.14B). If a constant shift is applied to the visual field by 20° at the initial stage, the same system also develops precisely registered spatial maps and exhibits decreased foveation error after the presentation of 15,000 stimuli. In addition, plasticity of the system towards altered visual experience has also been shown, in which normal visual stimulation is applied to a system well-adapted to a 20° shift visual field. As shown in Fig. 4.15A, immediately after the removal of visual shift, the auditory and visual maps in OT are misaligned with a displacement approximate to the visual shift. After an exposure to 20,000 visual-auditory stimuli, the two spatial maps return to good alignment (Fig. 4.15A). The foveation error after the

learning is significantly reduced compared with its value right after the removal of visual shift (Fig. 4.15B).

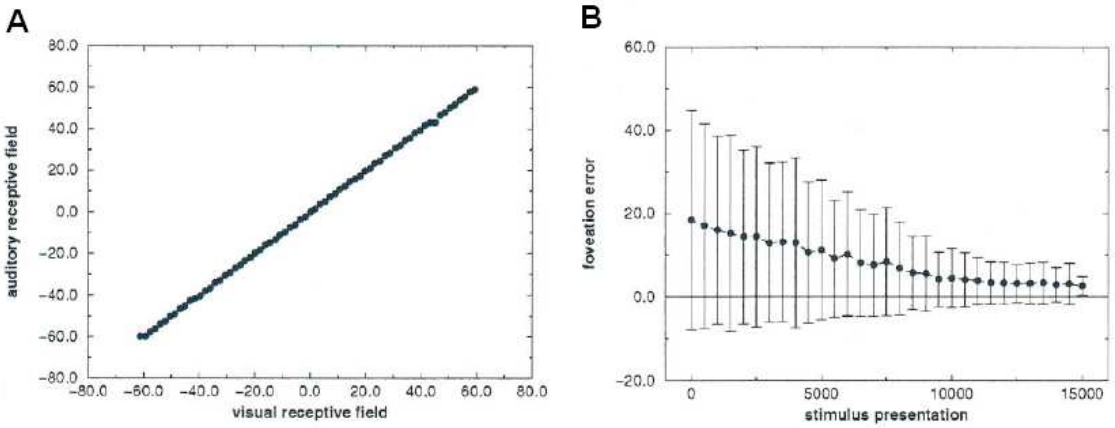


FIGURE 4.14: Performance of the system after the presentation of 15,000 audiovisual stimuli. (A) Relative alignment of the visual and auditory receptive fields. (B) Mean values of the foveation error for all the spatial locations of the stimuli. From Rucci et al. (1997).

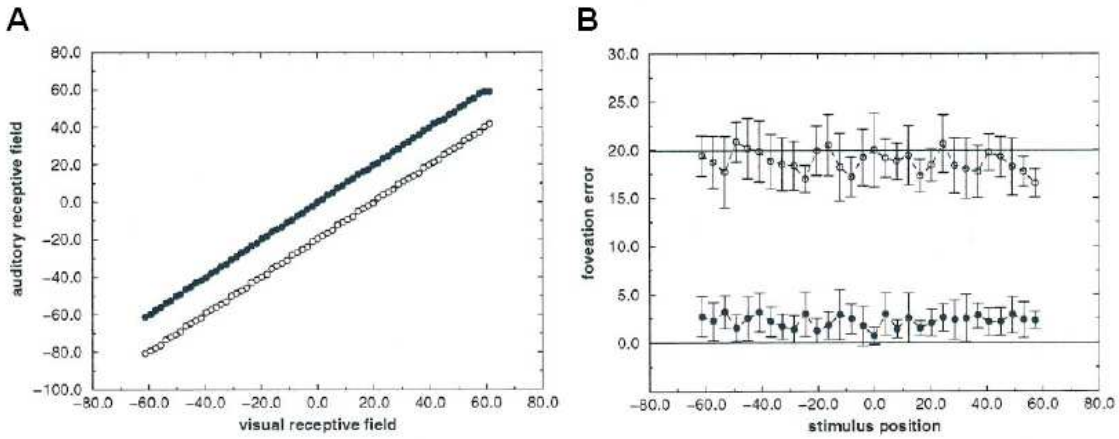


FIGURE 4.15: Performance recovery of the system after the removal of the 20° visual shift. (A) Relative alignment of the visual and auditory receptive fields right after the removal of the visual shift (open circles) and after an additional presentation of 20,000 stimuli (filled circles). (B) Mean values of the foveation error immediately after the removal of the visual shift (open circles) and after an additional presentation of 20,000 stimuli (filled circles). From Rucci et al. (1997).

The model nicely reproduces a number of experimental findings on the emergence and plasticity of multimodal spatial representation in OT both under normal sensory experience and under altered sensory experience. The results support the notion that auditory and visual spatial map alignment in OT is mediated through value-dependent learning. Furthermore, the model has been applied in a robotic system to control its orientating behaviour (Rucci et al., 1999). Under the model, the robotic system not only exhibits accurate orientation in normal sensory and motor conditions, but also recovers good performance with alteration in sensory inputs or motor outputs. However, the main

focuss of the study is adaptation in multisensory spatial maps in OT, and the response properties of multisensory neurons have not been explored. An obvious way to extend the work is therefore to include properties reported in DSC neurons, such as multisensory enhancement associated with inverse effectiveness, to achieve robust orientating behaviour for weak or noisy sensory stimuli.

4.2.2 Deneve and Pouget's Model (2004)

One problem concerning multisensory integration is to reconcile the different reference frames of the sensory modalities. For example, visual information is represented by neurons with receptive field on retina and is eye-centred, whereas auditory information is represented by neurons with receptive field around the head and is head-centred. Consequently, a movement of the eye induces a change in the reference frame of visual information, while the reference frame of auditory information remains unchanged. To combine cross-modal information consistently, the brain thus needs to accommodate the changes of the reference frames for different sensory modalities dynamically. Deneve and Pouget (2004) suggested that a Bayesian framework provides an optimal solution to this issue and proposed a basis function neural network with multidimensional attractors that implements cross-modal links and performs optimal Bayesian multisensory integration.

Consider an object at a position x that provides both visual and auditory stimulation. Since neural responses in the visual channel $\mathbf{r}_v$, a vector representing the firing rate of a large population of visual neurons, is noisy, a strategy to estimate the position x is to compute the posterior probability $P(x|\mathbf{r}_v)$. According to Bayes' rule, $P(x|\mathbf{r}_v)$ can be computed from the response distribution $P(\mathbf{r}_v|x)$, prior knowledge of the object position $P(x)$ and prior knowledge of neural response $P(\mathbf{r}_v)$,

$$P(x|\mathbf{r}_v) = \frac{P(\mathbf{r}_v|x)P(x)}{P(\mathbf{r}_v)}. \tag{4.22}$$

Since $P(\mathbf{r}_v)$ is independent of x which is the focus of the study, it is ignored. Furthermore, the prior probability $P(x)$ is assumed to be a constant for simplicity. Accordingly, this leads to

$$P(x|\mathbf{r}_v) \propto P(\mathbf{r}_v|x). \tag{4.23}$$

Similarly, in the auditory channel, we also have

$$P(x|\mathbf{r}_a) \propto P(\mathbf{r}_a|x), \tag{4.24}$$

where $\mathbf{r}_a$ is the vector representing the firing rate of a population of auditory neurons to the object at a position x. This proportional relation can also be extended to the bimodal case, so that $P(x|\mathbf{r}_v, \mathbf{r}_a) \propto P(\mathbf{r}_v, \mathbf{r}_a|x)$. Under the assumption that the noise in the visual channel is independent of that in the auditory channel, we have $P(\mathbf{r}_v, \mathbf{r}_a|x) =$

$P(\mathbf{r}_v|x)P(\mathbf{r}_a|x)$. This leads to

$$P(x|\mathbf{r}_v, \mathbf{r}_a) \propto P(\mathbf{r}_v|x)P(\mathbf{r}_a|x), \tag{4.25}$$

which, combined with Eqs. (4.23) and (4.24), gives

$$P(x|\mathbf{r}_v, \mathbf{r}_a) \propto P(x|\mathbf{r}_v)P(x|\mathbf{r}_a). \tag{4.26}$$

Therefore, the bimodal posterior distribution can be obtained from a product of the two unimodal posterior distributions. The estimation of object position given bimodal inputs $\hat{x_b}$ can be acquired with a maximum likelihood approach,

$$\hat{x_b} = \arg\max_x P(x|\mathbf{r}_v, \mathbf{r}_a). \tag{4.27}$$

If $P(x|\mathbf{r}_v)$ and $P(x|\mathbf{r}_a)$ are Gaussian distributions, the bimodal estimation $\hat{x_b}$ can be calculated as

$$\hat{x_b} = \frac{1/\sigma_v^2}{1/\sigma_v^2 + 1/\sigma_a^2}\hat{x_v} + \frac{1/\sigma_a^2}{1/\sigma_v^2 + 1/\sigma_a^2}\hat{x_a}, \tag{4.28}$$

where σ_v^2 and σ_a^2 are the variance of $P(x|\mathbf{r}_v)$ and $P(x|\mathbf{r}_a)$ respectively, and $\hat{x_v}$ and $\hat{x_a}$ are the maximum likelihood estimations in the unimodal case.

The basics function network model was introduced based on a linear computation between the eye-centred position of an object, x_r, and the head-centred position of the same object, x_a, through the eye position, x_e, where we have $x_a = x_r + x_e$. As shown in Fig. 4.16, the network for this additive computation consists of three one-dimensional input layers, corresponding to three populations of neurons to encode the information of x_a, x_r and x_e respectively, and an intermediate two-dimensional layer of basis function units. In the model, connections between the intermediate layer and the other three layers are reciprocal. Deneve et al. (2001) have demonstrated that this model is able to implement two types of tasks. One is "function approximation", in which given the noisy population codes for x_r and x_e, the model need to recover x_a; the other is "cue integration", in which the noisy neural responses of x_a, x_r and x_e are all provided and the model needs to combine these information to produce an optimal estimation $\hat{x_r}$, $\hat{x_e}$ and $\hat{x_a}$.

To test the performance of the model, Deneve et al. (2001) performed 100,000 trials, each of which contains three iterations to reach the stable states of the recurrent network, and compared the mean and the standard deviation of the results to those obtained from a maximum likelihood estimator based on Bayes' theorem. The results showed that the performance of the basis function network in function approximation and cue integration is comparable to the maximum likelihood estimator, and the model thus performs Bayesian multisensory integration. Moreover, the model was applied to estimate the results of Ernst et al.'s experiments, in which a human observer was required to discriminate the widths of visual bars, haptic bars or visual-haptic bars (Ernst et al.,

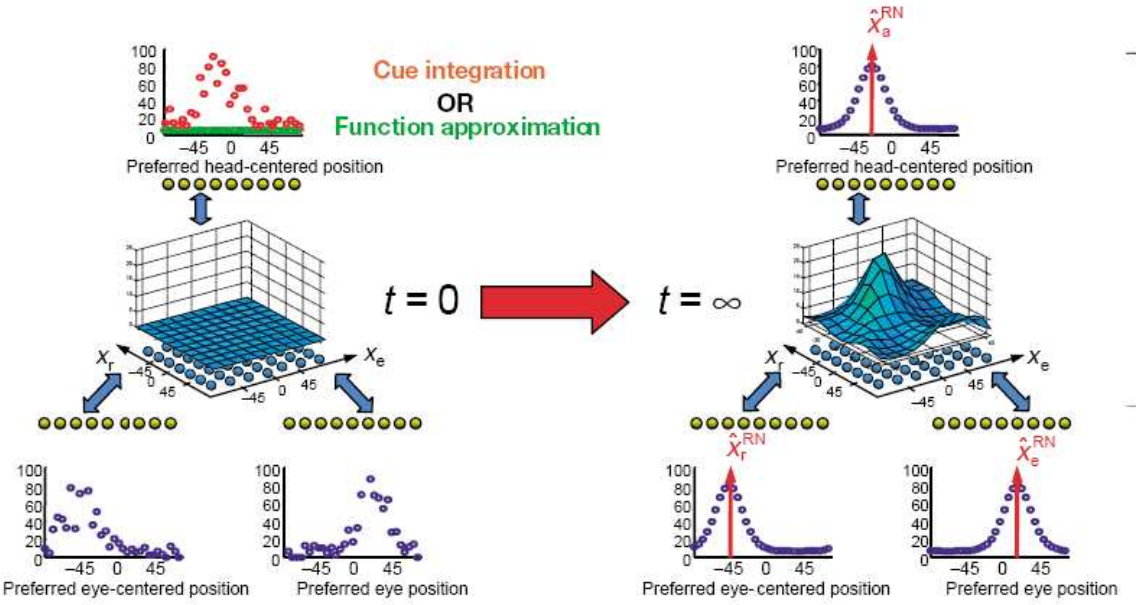


FIGURE 4.16: Architecture of the iterative basic function network. For function approximation task, there is no sensory input for the head-centred position x_a, while the task of cue integration provides noisy neural codes for x_a. For both the tasks, the network receiving noisy sensory inputs at $t = 0$ converges to a stable state at $t = \infty$. The position of the peak activity of the stable hills corresponds to the estimation of the network for $\hat{x_r}$, $\hat{x_e}$, $\hat{x_a}$. From Deneve et al. (2001).

2002). It turned out that the basis function network produces estimations that match well with both the experimental data and the predictions based on optimal Bayesian integration (Fig. 4.17).

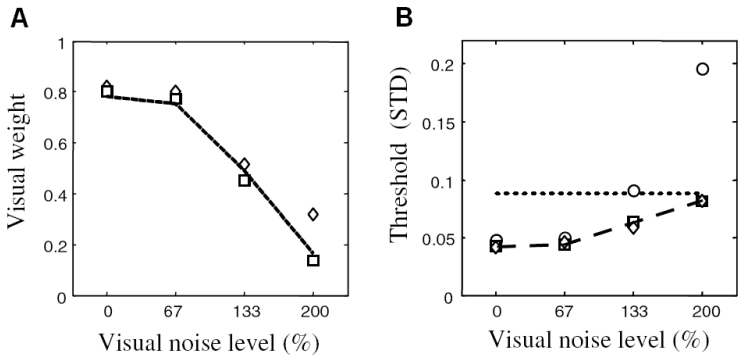


FIGURE 4.17: Application of the iterative basis function network to a visual-haptic integration task. (A) The weight of visual information in the width estimation task for a visual-haptic bar from experimental measurements by Ernst et al. (2002) (diamonds), predictions from the Bayesian model (dashed line) and from the iterative basis function network (squares). (B) Discrimination threshold for the width of visual bars (circles), haptic bars (dotted line), visual-haptic bars (diamond), predictions from the Bayesian model (dashed line) and from the iterative basis function network (squares). From Deneve and Pouget (2004).

This network model is designed to implement a spatial link between modalities and perform optimal estimation. Deneve and Pouget (2004) suggested that a suitable threshold

under which sensory input fails to drive the network is necessary to avoid the amplification of noise through reciprocal connections. This threshold, as suggested by Deneve and Pouget (2004), is sufficient to account for multisensory enhancement and the inverse effectiveness observed in DSC neurons. The reason is that bimodal inputs reach the threshold more easily than unimodal inputs, and enhancement thus occurs. However, this enhancement disappears when both unimodal and bimodal inputs exceed the threshold, corresponding to the inverse effectiveness property. Furthermore, Deneve and Pouget (2004) observed that the units in the intermediate layer have eye-centred receptive fields that exhibit partially shifting when the position of the eyes changes. For example, when the eyes move, the peak position in the response of an intermediate unit shifts in the same direction as a head-centred unit, but with only half the amplitude. These partially shifting receptive fields have also been reported in SC (Jay and Sparks, 1987; Peck et al., 1993). It might therefore be possible that the model to a certain extent accounts for some properties of DSC neurons.

4.3 Summary

We have reviewed several mathematical and computational models of multisensory integration in the brain, covering several important properties such as the alignment of multisensory maps (Rucci et al., 1997) and cross-modal spatial links (Deneve and Pouget, 2004). In particular, we focussed on those models dealing with neuronal responses in DSC neurons. For example, Patton and Anastasio (2003) proposed a perceptron model to simulate cross-modal enhancement based on a Bayesian approach; to account for the mixture of unisensory and multisensory neurons and the functional roles of cortical inputs in DSC neurons, Anastasio and Patton (2003) designed a two-stage unsupervised learning algorithm to construct a corticotectal model; Schauer and Gross (2004) built a computational model of a Amari-type competitive neural field to realize biologically-comparable multisensory enhancement through neuronally inspired criteria; Cuppini et al. (2007) proposed a three-layered neural network model to reproduce properties such as CME, multisensory depression and MSS. In spite of the ability of these models to reproduce some properties of DSC neurons, they either are not biologically plausible or still require further rigorous justification. It is thus possible that the principles underlying these models do not correspond to those underlying multisensory processes in DSC neurons. In the next chapter, we will discuss the model proposed by Patton and Anastasio (2003) in more detail since our main focus is the basic response properties of DSC neurons.

Chapter 5

The Patton and Anastasio Model of DSC Neuronal Responses

A representative model of the response properties of DSC neurons was proposed by Patton and Anastasio (2003) based on Bayes' rule. Although the model exhibits both CME and MSS, these response properties are sensitive to the parameters. In this chapter, we explore the parameters critical to the performance of the model using both numerical and analytical studies. The results show that the model does not plausibly account for the response properties of DSC neurons. Another interpretation of the mechanism underlying multisensory integration is thus required. The material presented in this chapter was briefly discussed in "On Natural Statistics in Multisensory Integration" (Elliott et al., 2008c).

5.1 Introduction

For completeness, we re-introduce the Patton and Anastasio model, so some of the materials coincide with that in section 4.1.1. The central assumption of the model is that a neuron computes the posterior probability of a target being present based on Bayes' rule. Patton and Anastasio (2003) also proposed a neuronal implementation of this rule using a perceptron model. Denote the target as $T = t$, where $t = 0$ corresponds to the absence of the target and $t = 1$ corresponds to the presence of the target, and the inputs as $\mathbf{m} = (m_1, m_2, \ldots, m_k)^T$, where m_i represents an input from a sensory modality labeled as i and k is the number of sensory channels. According to Bayes' rule, the posterior probability of a target being present given input $\mathbf{m}$ $P(T = 1|\mathbf{m})$ is computed as

$$P(T = 1|\mathbf{m}) = \frac{P(\mathbf{m}|T = 1)P(T = 1)}{P(\mathbf{m})}, \tag{5.1}$$

and we have the law of total probability

$$P(\mathbf{m}) = P(\mathbf{m}|T=1)P(T=1) + P(\mathbf{m}|T=0)P(T=0), \tag{5.2}$$

where $P(T=1)$ and $P(T=0)$ are prior probabilities, $P(\mathbf{m}|T=1)$ represents the input distribution when the target is present (the "driven" condition) and $P(\mathbf{m}|T=0)$ represents the input distribution when the target is absent (the "spontaneous" condition). Substituting Eq. (5.2) into Eq. (5.1), we obtain

$$P(T=1|\mathbf{m}) = \left[1 + \frac{P(T=0)P(\mathbf{m}|T=0)}{P(T=1)P(\mathbf{m}|T=1)}\right]^{-1}. \tag{5.3}$$

A sigmoidal function is adopted to describe the response of a neuron to a combined input u,

$$f(u) = \frac{1}{1+e^{-u}}. \tag{5.4}$$

Assuming the neuron computes the posterior probability $P(T=1|\mathbf{m})$, Patton and Anastasio (2003) suggested that $f(u) = P(T=1|\mathbf{m})$. Comparing Eq. (5.3) and Eq. (5.4), the expression of u is obtained,

$$u = \ln\left[\frac{P(\mathbf{m}|T=1)}{P(\mathbf{m}|T=0)}\right] + \ln\left[\frac{P(T=1)}{P(T=0)}\right]. \tag{5.5}$$

To implement a Bayes' computation, therefore, a neuron should be able to compute u as shown in Eq. (5.5).

When the input m_i are modelled as independent Poisson distributions, a perceptron is able to implement the computation in Eq. (5.5) if u is modeled as a weighted sum of the inputs,

$$u = b + \sum_{i=1}^{k} w_i m_j, \tag{5.6}$$

where b is the bias and w_i is the weight. This is because when we substitute the Poisson distributions into Eq. (5.5), the exponential terms in the distributions can be cancelled by the logarithm in Eq. (5.5), resulting in a linear combination of the inputs m_j. Patton and Anastasio (2003) have provided results showing that this perceptron model exhibits CME. However, as will be discussed in Chapter 6, a neuronal model with a non-linear, saturating response function is already capable of exhibiting CME. Since the perceptron model uses a sigmoidal response function [Eq. (5.4)], it is not surprising that it produces CME. Furthermore, independent Poisson distributed inputs in neurons are quite unlikely. Therefore, we do not consider this case.

According to the classical central limit theorem, the distribution of a combination of many identically-distributed and independent signals will be gaussian (Fristedt and Gray, 1997). Since inputs to a DSC neuron are derived by combining signals from a large number of neurons, Patton and Anastasio (2003) suggested that a more general

case is to consider the input as following a multivariate gaussian distribution, given by

$$P(\mathbf{m}|T=t) = \frac{1}{(2\pi)^{k/2}|\mathbf{\Sigma}_t|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{m}-\boldsymbol{\mu}_t)^T\mathbf{\Sigma}_t^{-1}(\mathbf{m}-\boldsymbol{\mu}_t)\right], \tag{5.7}$$

where $\mathbf{\Sigma}_t$ is the covariance matrix and $\boldsymbol{\mu}_t$ is a vector containing the means for different sensory channels ($t = 0$ for the spontaneous case and $t = 1$ for the driven case). Substituting Eq. (5.7) into Eq. (5.5), we have

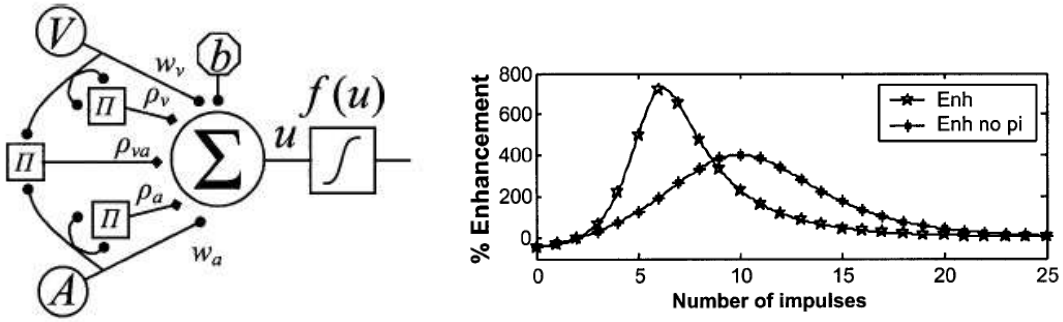
$$\begin{aligned} u = \frac{1}{2}[&\mathbf{m}^T(\mathbf{\Sigma}_0^{-1}-\mathbf{\Sigma}_1^{-1})\mathbf{m} + 2\mathbf{m}^T(\mathbf{\Sigma}_1^{-1}\boldsymbol{\mu}_1 - \mathbf{\Sigma}_0^{-1}\boldsymbol{\mu}_0) \\ &-\boldsymbol{\mu}_1^T\mathbf{\Sigma}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_0^T\mathbf{\Sigma}_0^{-1}\boldsymbol{\mu}_0] + \ln\left(\frac{|\mathbf{\Sigma}_0|^{1/2}P(T=1)}{|\mathbf{\Sigma}_1|^{1/2}P(T=0)}\right), \end{aligned} \tag{5.8}$$

in which the exponential term in Gaussian distribution, as in the Poisson case, is cancelled by the logarithm in Eq. (5.5), leaving the quadratic and multiplicative terms of the inputs. It is clear from Eq. (5.8) that a single perceptron [Eq. (5.6)] is insufficient to compute u for multivariate gaussian inputs. An augmented perceptron model with pi nodes is therefore required (see Figs. 5.1A for two-input case and 5.2A for three-input case), in which

$$u = b + \sum_{i=1}^{k} w_i m_i + \sum_{i=1}^{k}\sum_{j=i}^{k} \rho_{ij}(m_i m_j), \tag{5.9}$$

where ρ_{ij} is the weight for the pi nodes. Comparing this with Eq. (5.8), we can then determine the weights w_i, ρ_{ij} and the bias b, under which the neuron computes the posterior probability $P(T = 1|\mathbf{m})$. To demonstrate the performance of the model, a two-channel augmented perceptron that receives bimodal gaussian inputs from the visual (V) and auditory (A) channels (Fig. 5.1A) has been considered. The results of the model with parameters set as Table 5.1 are shown in Fig. 5.1B, in which we can observe that the maximum CME occurs for weak sensory stimuli and then decreases when the input increases, corresponding to the inverse effectiveness property of CME (Stein and Meredith, 1993; Stein et al., 2004). Furthermore, when the pi nodes are removed, CME still exists although with a smaller magnitude. Binns and Salt (1996) have reported that, after applying 2-amino-5-phosphonopentanoate (AP5) to block N-methyl-D-aspartate (NMDA) receptors of DSC neurons, the decrease in bimodal response is larger than that of unimodal responses, resulting in a decrease in the CME index. Patton and Anastasio (2003) proposed that this is comparable to the effect of removing pi nodes from the model (Fig. 5.1B), and the functional role of NMDA receptors in the responses of DSC neurons may be described by the pi nodes. However, blockade of NMDA receptors will reduce the overall activities of a neuron anyway, and it is not clear that NMDA receptors implement the same multiplicative operations as the pi nodes.

To demonstrate that the model is able to produce both CME and MSS, a three-channel augmented perceptron is also considered (see Fig. 5.2A), in which the V and X channels represent two visual channels and A represents the auditory channel. In this case, the



A B

FIGURE 5.1: Performance of a two-channel augmented-perceptron model. (A) An augmented-perceptron model for visual (V) and auditory (A) sensory channels. (B) The CME curves produced by an intact augmented-perceptron model (stars) and by the same model without the pi nodes (asterisks). CME is calculated as the percentage enhancement of bimodal response to the most effective unimodal response (see Eq. (2.1) in Chapter 2). The parameters are set as shown in Table 5.1. Inputs from V and A sensory channels are set to be equal and change from 0 to 25 in the number of spikes per 0.25s. From Patton and Anastasio (2003).

	μ_0	μ_1	σ_0^2	σ_1^2
V	2	6	5	6
A	2	6	5	6

σ_{VA0}^2	σ_{VA1}^2	P(T=1)	P(T=0)
0.1	2.8	0.1	0.9

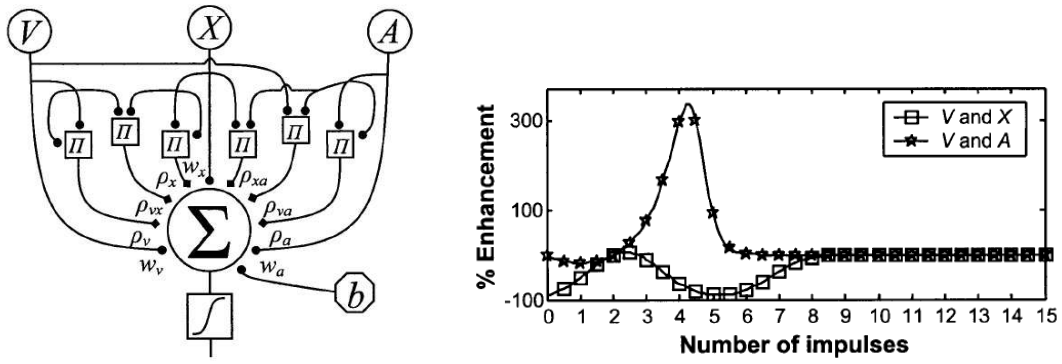
TABLE 5.1: Parameters for a two-input augmented perceptron as in Patton and Anastasio (2003). V stands for the visual channel and A stands for the auditory channel. μ_0 is the spontaneous mean, μ_1 is the driven mean, σ_0^2 is the spontaneous variance, σ_1^2 is the driven variance and σ_{va0}^2 and σ_{va1}^2 stand for spontaneous and driven covariance respectively.

covariance matrix $\mathbf{\Sigma}_t$ $(t = 0, 1)$ is a 3×3 matrix given by

$$\mathbf{\Sigma}_t = \begin{bmatrix} \sigma_{Vt}^2 & \sigma_{VXt}^2 & \sigma_{VAt}^2 \\ \sigma_{VXt}^2 & \sigma_{Xt}^2 & \sigma_{XAt}^2 \\ \sigma_{VAt}^2 & \sigma_{XAt}^2 & \sigma_{At}^2 \end{bmatrix}.$$

Patton and Anastasio (2003) assumed that the inputs from V and X channels to each DSC neuron derive from separate but overlapping visual neurons. Overlapping visual neurons result in an overlapping of the RFs of a DSC neuron for V and X channels. Under driven cases, therefore, inputs from V and X channels covary. Moreover, Patton and Anastasio (2003) argued that, under spontaneous conditions, these overlapping visual neurons, no matter whether they are silent or spontaneously active, lead to covariation in V and X. Based on this interpretation, Patton and Anastasio (2003) proposed that within-modality covariances are larger than cross-modality covariances, for example, $\sigma_{VX0}^2 > \sigma_{VA0}^2$ and $\sigma_{VX1}^2 > \sigma_{VA1}^2$. Fig. 5.2B shows the results produced with the parameters in Table 5.2, in which the spontaneous and driven means and variances of the three sensory channels are set the same. We can observe that when both V and A are active and X stays at its spontaneous mean μ_{X0}, CME occurs; on the other hand, when both

V and X channels are active and A is spontaneously active, MSS occurs. The model thus appears to simulate both CME and MSS observed in DSC neurons. Furthermore, Patton and Anastasio (2003) showed that, in order to exhibit MSS, the spontaneous covariance σ^2_{VX0} needs to approach the value of spontaneous variances $\sigma^2_{V0} = \sigma^2_{X0}$, and the driven variances need to be larger than the spontaneous variances. In Fig. 5.2B, for example, $\sigma^2_{VX0} = 1.6$ is close to $\sigma^2_{V0} = \sigma^2_{X0} = 2$. They also suggested that when the variances for different sensory modalities are different, similar relationships could be observed, although no evidence or results were given. In fact, the mean and variances for the three sensory channels are set the same in the sample they showed (see Tabel 5.2), which is a very special choice not likely to exist in the biological system. More important, the value of within-modality spontaneous covariance ($\sigma^2_{VX0} = 1.6$), which is critical to produce MSS, has not been clearly justified. Our analysis will focus on this aspect.



A B

FIGURE 5.2: Performance of a three-channel augmented-perceptron model. (A) An augmented-perceptron model for two visual channels (V and X) and one auditory (A) channel. (B) Cross-modality and within-modality CME curves produced by the three-channel augmented-perceptron model. The parameters are set as Table 5.2. For the cross-modality CME curve, inputs in the X channel are fixed at its spontaneous rate $\mu_{X0} = 2$, whereas the firing rates from V and A channels are set to be equal and change from 0 to 15. The CME curve is plotted as a function of the V and A impulses (stars). For the within-modality CME curve, inputs in the A channel are instead fixed at its spontaneous rate $\mu_{A0} = 2$, whereas the firing rates from the V and X visual channels are set to be equal and change from 0 to 15. The CME curve is plotted as a function of the V and X impulses (squares). From Patton and Anastasio (2003).

	μ_0	μ_1	σ^2_0	σ^2_1	σ^2_{VA0}, σ^2_{XA0}	σ^2_{VA1}, σ^2_{XA1}	σ^2_{VX0}	σ^2_{VX1}	P(T=1)	P(T=0)
V	2	6	2	6						
X	2	6	2	6	0.1	2.8	1.6	3.6	0.1	0.9
A	2	6	2	6						

TABLE 5.2: Parameters for three-input augmented perceptron as in Patton and Anastasio (2003). V, X stand for two visual channels and A stands for the auditory channel. The meaning of the other parameters follows the same convention as in Table 5.1. From Patton and Anastasio (2003).

5.2 Numerical Analysis

Patton and Anastasio (2003) have presented results for a two-channel Bayes' model for V and A inputs, and a three-channel model for V, X and A inputs. Comparing Table 5.1 and Table 5.2, we can find that, although the parameters for V and A channels between the two-channel and three-channel cases are generally the same, their spontaneous variances (σ^2_{V0} and σ^2_{A0}) are set to be different without providing a sound explanation. To investigate whether this difference influences the performance of the model, we set the spontaneous variances in the two-channel augmented perceptron to the values in the three-channel case, that is, $\sigma^2_{V0} = \sigma^2_{A0} = 2$, with all the other parameters unchanged. With these new parameters, we reproduced the responses and the CME indices of the intact model and of the model removing the pi nodes (Fig. 5.3). For the intact model, the bimodal responses are still larger than the unimodal responses (Fig. 5.3A), and the model exhibits CME with inverse effectiveness (Fig. 5.3B). On the other hand, for the model without the pi nodes, both the bimodal and unimodal responses decrease to nearly zero (Fig. 5.3A), indicating that the pi nodes dominate these two responses. In this case, some CME indices even reduce to negative values (Fig. 5.3B). The finding that removal of pi nodes from the model exhibits reduction but not an elimination of the CME is therefore merely induced by a specific choice of the parameters.

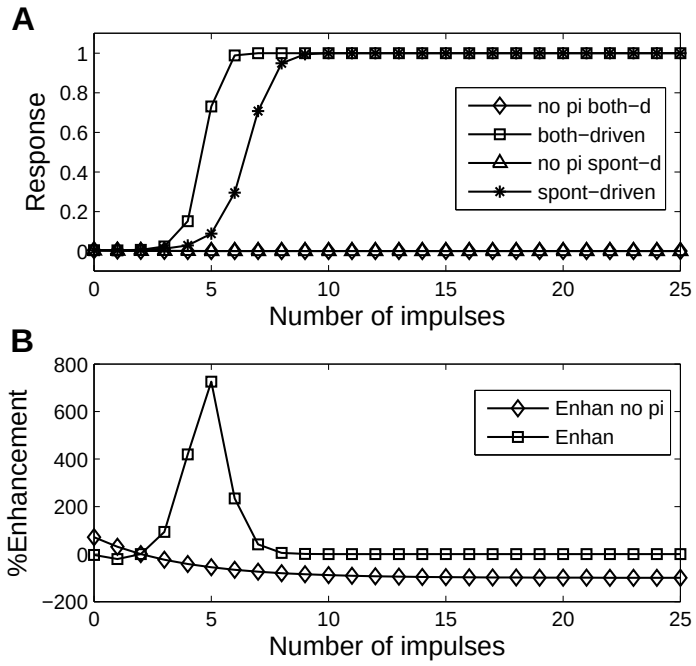


FIGURE 5.3: Response and CME curves for the two-channel augmented-perceptron with $\sigma^2_{V0} = \sigma^2_{X0} = 2$. Other parameters are set as Table 5.1. Conventions are the same as in Fig. 5.1(B). (A) The response curves as a function of the input for both-driven (squares), spontaneous (asterisks), both-driven without pi nodes (diamonds) and spontaneous without pi nodes (triangles) cases. (B) The CME curves as a function of the input for the intact model (squares) and the model without the pi nodes (diamonds).

As Patton and Anastasio (2003) have suggested, the value of spontaneous covariance σ^2_{VX0} is critical for the model to exhibit MSS. However, further study reveals that the spontaneous covariance $\sigma^2_{VX0} = 1.6$ in the model is actually a rather specific choice. Fig. 5.4A shows the unimodal responses where the X and A channels are at their spontaneous means and input from the V channel changes from 0 to 15 (V-alone case). When σ^2_{VX0} increases from 1.6 to 1.9, the V-alone response of the model to input $v = 0$ increases from nearly 0 to around 0.9. Obviously, this increase in neuronal response for zero input in the V sensory channel contradicts the neurophysiological data. Furthermore, we also calculate the CME indices when σ^2_{VX0} is decreased from 1.6 to 1.3 (Fig. 5.4B). For $\sigma^2_{VX0} = 1.6$, the model exhibits suppression and no evident enhancement can be observed. When σ^2_{VX0} is decreased to 1.3, however, we can observe response enhancement up to nearly 20% when the input v is around 3. This again is not consistent with the neurophysiological findings on MSS (Kadunce et al., 1997). Altogether, the model exhibits conflicting properties with what would be expected in a neuron when σ^2_{VX0} is either larger or smaller than 1.6. Therefore, $\sigma^2_{VX0} = 1.6$ is a specific selection to guarantee both no response for zero input in the V-alone case and no obvious enhancement to be observed in modality-specific interaction.

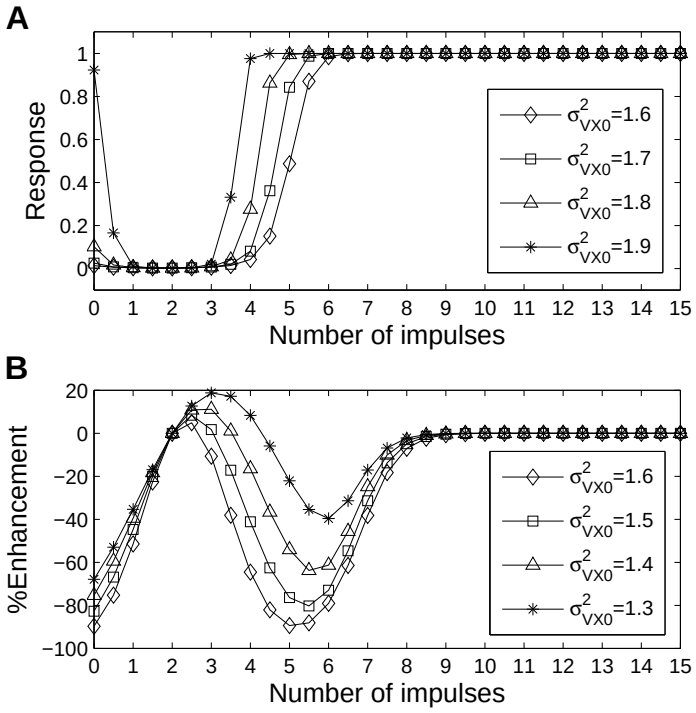


FIGURE 5.4: Response and CME curves for the three-channel augmented-perceptron with different values of σ^2_{VX0}. Other parameters are set as Table 5.2. Conventions are the same as in Fig. 5.1(B). (A) The unimodal response curves to the V-alone case for $\sigma^2_{VX0} = 1.6$ (diamonds), $\sigma^2_{VX0} = 1.7$ (squares), $\sigma^2_{VX0} = 1.8$ (triangles) and $\sigma^2_{VX0} = 1.9$ (asterisks). (B) The CME curves as a function of the input for $\sigma^2_{VX0} = 1.6$ (diamonds), $\sigma^2_{VX0} = 1.5$ (squares), $\sigma^2_{VX0} = 1.4$ (triangles) and $\sigma^2_{VX0} = 1.3$ (asterisks).

5.3 Mathematical Analysis

The Bayesian model has been shown to exhibit both CME and MSS (Patton and Anastasio, 2003). As we will discuss in Chapter 6, neurons with a sigmoidal response function are already capable of exhibiting CME. Since Patton and Anastasio (2003) also adopted a sigmoidal response function in the neuronal implementation of the Bayes' model, it is natural that CME can be observed. In the following mathematical analysis, therefore, we will not consider the CME property of the Bayes' model, but instead concentrate on how the model produces MSS, particularly on how the value of spontaneous covariance σ^2_{VX0}, or the spontaneous correlation coefficient ρ_{VX0}, contributes to MSS.

5.3.1 Spontaneous Covariance Matrix

As Patton and Anastasio (2003) suggested, the spontaneous covariance σ^2_{VX0} needs to approach the spontaneous variances $\sigma^2_{V0} = \sigma^2_{X0}$ to induce MSS in the model. Since the correlation coefficient is defined as

$$\rho_{VX0} = \frac{\sigma^2_{VX0}}{\sigma_{V0}\sigma_{X0}}, \tag{5.10}$$

when σ^2_{VX0} approaches $\sigma^2_{V0} = \sigma^2_{X0}$, the spontaneous correlation coefficient ρ_{VX0} approaches 1 and V and X channels become less and less distinguishable. In the extreme case that $\rho_{VX0} = 1$, V and X could be considered as the same channel under the spontaneous condition. In terms of the three-channel model, the spontaneous trivariate input distribution should therefore naturally reduce to a bivariate one when the spontaneous correlation coefficient $\rho_{VX0} \to 1$. However, when this happens, we have $\sigma^2_{V0} = \sigma^2_{X0}$, $\sigma^2_{VA0} = \sigma^2_{XA0}$ and the spontaneous correlation coefficient $\rho_{VX0} \to 1$, in which case the covariance matrix

$$\mathbf{\Sigma_0} = \begin{bmatrix} \sigma^2_{V0} & \sigma^2_{VX0} & \sigma^2_{VA0} \\ \sigma^2_{VX0} & \sigma^2_{X0} & \sigma^2_{XA0} \\ \sigma^2_{VA0} & \sigma^2_{XA0} & \sigma^2_{A0} \end{bmatrix}$$

becomes singular since its determinant $|\mathbf{\Sigma}_0| \to 0$. In this case, the inverse of the matrix $\mathbf{\Sigma}_0^{-1}$ does not exist. A trivariate Gaussian distribution for the inputs [Eq. (5.7)] is thus undefined and does not reduce to a bivariate one. As evident in Eq. (5.8), the combined input u also becomes undefined. This indicates that it is the near singularity of $\mathbf{\Sigma}_0$ that contributes to the response suppression in the model.

To investigate how this matrix singularity leads to MSS, we first analyze the performance of the model with bivariate Gaussian inputs when the spontaneous correlation coefficient ρ_{VX0} approaches 1, and then extend our analysis for inputs following bivariate Poisson distribution and Cowan exponential distribution (Cowan, 1987). The results suggest that, with the parameters chosen in the same way as Patton and Anastasio (2003),

the condition that $\rho_{VX0} \to 1$ always leads to MSS for two sensory inputs of equal magnitudes. Finally, we confirm this notion by providing a general proof.

5.3.2 Gaussian Distribution

In this section, we mathematically analyze the properties of the Bayesian model for Gaussian inputs [Eq. (5.7)] when the spontaneous correlation coefficient ρ_0 approaches 1. Since we are only considering MSS, a model dealing with only two sensory channels V and X is sufficient. The CME index is defined as

$$\%CME = \frac{f_{VX} - \max\{f_V, f_X\}}{\max\{f_V, f_X\}} \times 100\%, \tag{5.11}$$

where f_{VX} represents bimodal response, f_V and f_X represents unimodal responses. Therefore, the ratio $f_{VX}/\max\{f_V, f_X\} < 1$ corresponds to the occurrence of suppression. The general idea is to derive the formulae for f_{VX}, f_V and f_X, and study how the spontaneous correlation coefficient $\rho_0 \to 1$ results in MSS.

In the following, we will firstly derive a general expression for the response $f = P(T = 1|\mathbf{m})$ for any inputs $\mathbf{m}$, and then substitute $\mathbf{m} = (m_1, m_2)$ into f for bimodal response f_{VX}, and $\mathbf{m} = (m_1, \mu_0)^T$ and $\mathbf{m} = (\mu_0, m_2)^T$ for unimodal responses f_V and f_X respectively ($\mu_0 = \mu_{V0} = \mu_{X0}$ represents the spontaneous means of V and X channels). Since the variances are set to be equal (Patton and Anastasio, 2003), we denote $\sigma_{Vt}^2 = \sigma_{Xt}^2 = \sigma_t^2$ and the correlation coefficient $\rho_t = \sigma_{VXt}^2/\sigma_t^2$ ($t = 0$ represents the spontaneous case and $t = 1$ represents the driven case). We then rewrite the 2×2 covariance matrix as

$$\mathbf{\Sigma}_t = \begin{bmatrix} \sigma_{Vt}^2 & \sigma_{VXt}^2 \\ \sigma_{VXt}^2 & \sigma_{Xt}^2 \end{bmatrix} = \sigma_t^2 \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix}, \tag{5.12}$$

from which we can derive the matrix inverse

$$\mathbf{\Sigma}_t^{-1} = \frac{1}{\sigma_t^2(1-\rho_t^2)} \begin{bmatrix} 1 & -\rho_t \\ -\rho_t & 1 \end{bmatrix}, \tag{5.13}$$

and the determinant of $\mathbf{\Sigma}_t$

$$|\mathbf{\Sigma}_t| = \sigma_t^4(1-\rho_t^2). \tag{5.14}$$

Substituting Eqs. (5.13) and (5.14) into Eq. (5.7), we have

$$P(\mathbf{m}|T=0) = \frac{1}{2\pi\sigma_0^2\sqrt{1-\rho_0^2}} \exp\left[-\frac{(m_1-\mu_0)^2 - 2\rho_0(m_1-\mu_0)(m_2-\mu_0) + (m_2-\mu_0)^2}{2\sigma_0^2(1-\rho_0^2)}\right]. \tag{5.15}$$

Similarly, we can obtain the expression for the driven input probability $P(\mathbf{m}|T = 1)$. Substituting the expressions for the spontaneous input probability $P(\mathbf{m}|T = 0)$ and the

driven input probability $P(\mathbf{m}|T=1)$ into Eq. (5.3), we derive the formula for f,

$$f = P(T=1|\mathbf{m}) = \left[1 + \frac{P(T=0)}{P(T=1)}\frac{\sigma_1^2}{\sigma_0^2}\sqrt{\frac{1-\rho_1^2}{1-\rho_0^2}}\exp(C)\right]^{-1}, \tag{5.16}$$

where

$$\begin{aligned} C = &-\left[\frac{(m_1-\mu_0)^2 - 2\rho_0(m_1-\mu_0)(m_2-\mu_0) + (m_2-\mu_0)^2}{2\sigma_0^2(1-\rho_0^2)}\right] \\ &+\left[\frac{(m_1-\mu_1)^2 - 2\rho_1(m_1-\mu_1)(m_2-\mu_1) + (m_2-\mu_1)^2}{2\sigma_1^2(1-\rho_1^2)}\right]. \end{aligned} \tag{5.17}$$

We can then obtain the expressions for the bimodal response f_{VX} and unimodal responses f_V and f_X according to Eq. (5.16).

To calculate the neuronal response to bimodal VX inputs, we rewrite the input $\mathbf{m} = (m_1, m_2)^T$ as

$$\mathbf{m} = m_+\mathbf{n} + m_-\mathbf{e}, \tag{5.18}$$

where $m_+ = \frac{1}{2}(m_1+m_2)$, $m_- = \frac{1}{2}(m_1-m_2)$, $\mathbf{n} = (1,1)^T$ and $\mathbf{e} = (1,-1)^T$. In the implementation of the model, Patton and Anastasio (2003) adopted equal-valued inputs. By writing $\mathbf{m}$ as Eq. (5.18), we can clearly observe whether and how the difference between m_1 and m_2, represented by the m_-, influences the results. According to Eq. (5.18), we have $m_1 = m_+ + m_-$ and $m_2 = m_+ - m_-$. Substituting this into Eq. (5.16), we obtain the expression of the bimodal response f_{VX},

$$f_{VX} = \left[1 + \frac{P(T=0)}{P(T=1)}\frac{\sigma_1^2}{\sigma_0^2}\sqrt{\frac{1-\rho_1^2}{1-\rho_0^2}}\exp(C_{VX})\right]^{-1}, \tag{5.19}$$

where

$$C_{VX} = -\left[\frac{(m_+-\mu_0)^2}{\sigma_0^2(1+\rho_0)} + \frac{m_-^2}{\sigma_0^2(1-\rho_0)}\right] + \left[\frac{(m_+-\mu_1)^2}{\sigma_1^2(1+\rho_1)} + \frac{m_-^2}{\sigma_1^2(1-\rho_1)}\right]. \tag{5.20}$$

If $m_- \neq 0$, when the spontaneous correlation coefficient $\rho_0 \to 1$, both the terms $\sqrt{(1-\rho_1^2)/(1-\rho_0^2)}$ and $m_-^2/\sigma_0^2(1-\rho_0)$ go to infinity. However, since $m_-^2/\sigma_0^2(1-\rho_0)$ is in the exponent C_{VX}, it dominates the change in f_{VX}. When $m_-^2/\sigma_0^2(1-\rho_0) \to \infty$ (or C_{VX} goes to infinity), the exponential term in Eq. (5.19) goes to zero, driving the bimodal VX response f_{VX} to one. On the other hand, if $m_- = 0$, the term $m_-^2/\sigma_0^2(1-\rho_0)$ disappears. In this case, $\sqrt{(1-\rho_1^2)/(1-\rho_0^2)}$ dominates the change of f_{VX} when $\rho_0 \to 1$, driving the bimodal response f_{VX} to zero. That is,

$$f_{VX} \to \begin{cases} 0, & \text{if } m_- = 0, \\ 1, & \text{if } m_- \neq 0, \end{cases} \tag{5.21}$$

when the spontaneous correlation coefficient $\rho_0 \to 1$.

In the one-driven case, there are two possible inputs: $\mathbf{m}_V = (m_1, \mu_0)^T$ $(m_1 > \mu_0)$ or $\mathbf{m}_X = (\mu_0, m_2)^T$ $(m_2 > \mu_0)$, where μ_0 is the spontaneous mean of the input. To analyze CME, we need to calculate both f_V and f_X. Substituting $\mathbf{m}_V$ and $\mathbf{m}_X$ into Eq. (5.16), we obtain the expressions for f_V and f_X,

$$f_V = \left[1 + \frac{P(T=0)}{P(T=1)} \frac{\sigma_1^2}{\sigma_0^2} \sqrt{\frac{1-\rho_1^2}{1-\rho_0^2}} \exp(C_V)\right]^{-1}, \tag{5.22}$$

where

$$C_V = -\frac{(m_1-\mu_0)^2}{2\sigma_0^2(1-\rho_0^2)} + \frac{(m_1-\mu_1)^2 + 2\rho_1(\mu_1-\mu_0)(m_1-\mu_1) + (\mu_1-\mu_0)^2}{2\sigma_1^2(1-\rho_1^2)}, \tag{5.23}$$

and

$$f_X = \left[1 + \frac{P(T=0)}{P(T=1)} \frac{\sigma_1^2}{\sigma_0^2} \sqrt{\frac{1-\rho_1^2}{1-\rho_0^2}} \exp(C_X)\right]^{-1}, \tag{5.24}$$

where

$$C_X = -\frac{(m_2-\mu_0)^2}{2\sigma_0^2(1-\rho_0^2)} + \frac{(m_2-\mu_1)^2 + 2\rho_1(\mu_1-\mu_0)(m_2-\mu_1) + (\mu_1-\mu_0)^2}{2\sigma_1^2(1-\rho_1^2)}. \tag{5.25}$$

For f_V, the term $(m_1-\mu_0)^2/2\sigma_0^2(1-\rho_0^2)$ dominates the value of the unimodal response f_V when the spontaneous correlation coefficient $\rho_0 \to 1$, since it is in the exponent C_V. Therefore, when $\rho_0 \to 1$, the exponential term $\exp(C_V) \to 0$, driving the unimodal response to one. Similarly, we also have the unimodal response $f_X \to 1$ when the spontaneous correlation coefficient ρ_0 approaches 1. Combining this limit of unimodal responses with Eq. (5.21), we have

$$\frac{f_{VX}}{\max\{f_V, f_X\}} \to \begin{cases} 0, & \text{if } m_- = 0, \\ 1, & \text{if } m_- \neq 0, \end{cases} \tag{5.26}$$

when $\rho_0 \to 1$. According to Eq. (5.26), if $m_- = 0$, the ratio $f_{VX}/\max\{f_V, f_X\}$ goes to zero, indicating that the condition $\rho_0 \to 1$ leads to suppression. This corresponds to the way Patton and Anastasio (2003) chose the inputs $\mathbf{m} = (m_1, m_2)^T$ where $m_1 = m_2$. When $m_- \neq 0$, however, the limit of the ratio $f_{VX}/\max\{f_V, f_X\}$ goes to one, in which case neither enhancement nor suppression occurs.

The same results apply to the three-channel model described in Section 5.1, in which the input from auditory channel A is always set as its spontaneous mean μ_{A0} and the CME for V and X channels is considered. Similar to the bivariate case, the unimodal responses f_V and f_X have a limit of one when the spontaneous correlation coefficient approaches one. Moreover, for the VX driven bimodal response f_{VX}, the exponential term in f_{VX} also contains the term $m_-^2/\sigma_0^2(1-\rho_0)$ as in the bivariate case [Eq. (5.21)], which drives the response f_{VX} to zero when V and X inputs are different $(m_1 \neq m_2)$. On the other hand, if $m_1 = m_2$, the response f_{VX} goes to one. In this case, the ratio

$f_{VX}/\max\{f_V, f_X\}$ approaches 0 when $\rho_0 \to 1$, again leading to suppression, whereas neither enhancement nor suppression could be produced if $m_1 \neq m_2$. The conclusions are exactly the same as for the two-input model. As an example, we take the inputs to induce bimodal response f_{VX} as $\mathbf{m} = (m+5, m)^T$ instead of $\mathbf{m} = (m, m)^T$, and produce the CME curve based on this non-equal inputs in Fig. 5.5. We can observe that for input $a \geq 2$ the model indeed exhibits neither enhancement nor suppression. This confirms that the condition $\rho_0 \to 1$ only guarantees response suppression for equal-valued input stimuli. Therefore, the equal-valued inputs in Patton and Anastasio (2003) are merely special choices to induce suppression under the selected statistical parameters.

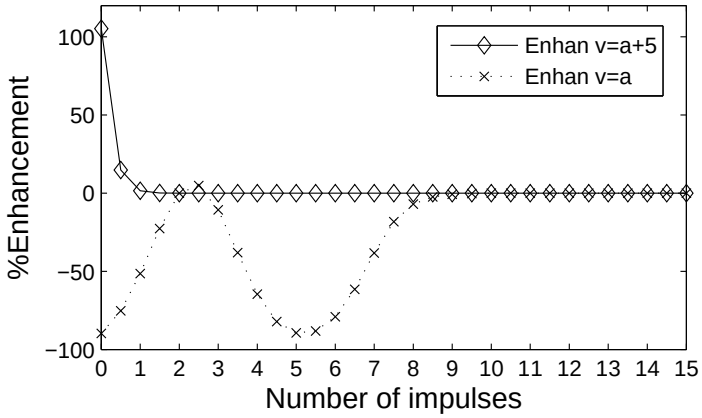


FIGURE 5.5: The CME curves for non-equal V and X inputs (diamonds) and for equal V and X inputs (crosses) in the three-channel augmented-perceptron model. In the case of equal V and X inputs, the parameters are set according to Table 5.2. Conventions are the same as in Fig. 5.1(B). In the case of non-equal V and X inputs, the parameters are also set according to Table 5.2, but the inputs v is set as $v = a + 5$. The abscissa in this case corresponds to input in the A channel.

A more general analysis of the model should consider different statistical parameters for the sensory channels V and X, for example, $\mu_{V0} \neq \mu_{X0}$ and $\sigma^2_{V0} \neq \sigma^2_{X0}$. It is obvious that these asymmetric parameters will influence the limit of the ratio $f_{VX}/\max\{f_V, f_X\}$ for equal-valued VX inputs when $\rho_0 \to 1$, because the term $1 - \rho_0$ is then unlikely to be removed from the exponential term for $m_1 = m_2$. For example, Fig. 5.6 shows the CME surfaces for V and X channels when they have equal means and equal variances (Fig. 5.6A), unequal means and equal variances (Fig. 5.6B), equal means and unequal variances (Fig. 5.6C) and unequal means and unequal variances (Fig. 5.6D), with the same spontaneous correlation coefficient $\rho_0 = 0.8$. As expected, we can observe from Fig. 5.6 that when the means or/and the variances of the V and X channels are unequal, equal-valued inputs do not guarantee the production of evident suppression in the model (see the diagonals in the figures). We will not further explore this case since it is now clear that MSS produced by the model does not naturally emerge from the Bayesian framework, but critically relies on the choices of statistical parameters. Furthermore, even provided equal means and variances for different sensory channels, the spontaneous covariance approaching variances only leads to MSS when the bimodal Gaussian inputs

are equal.

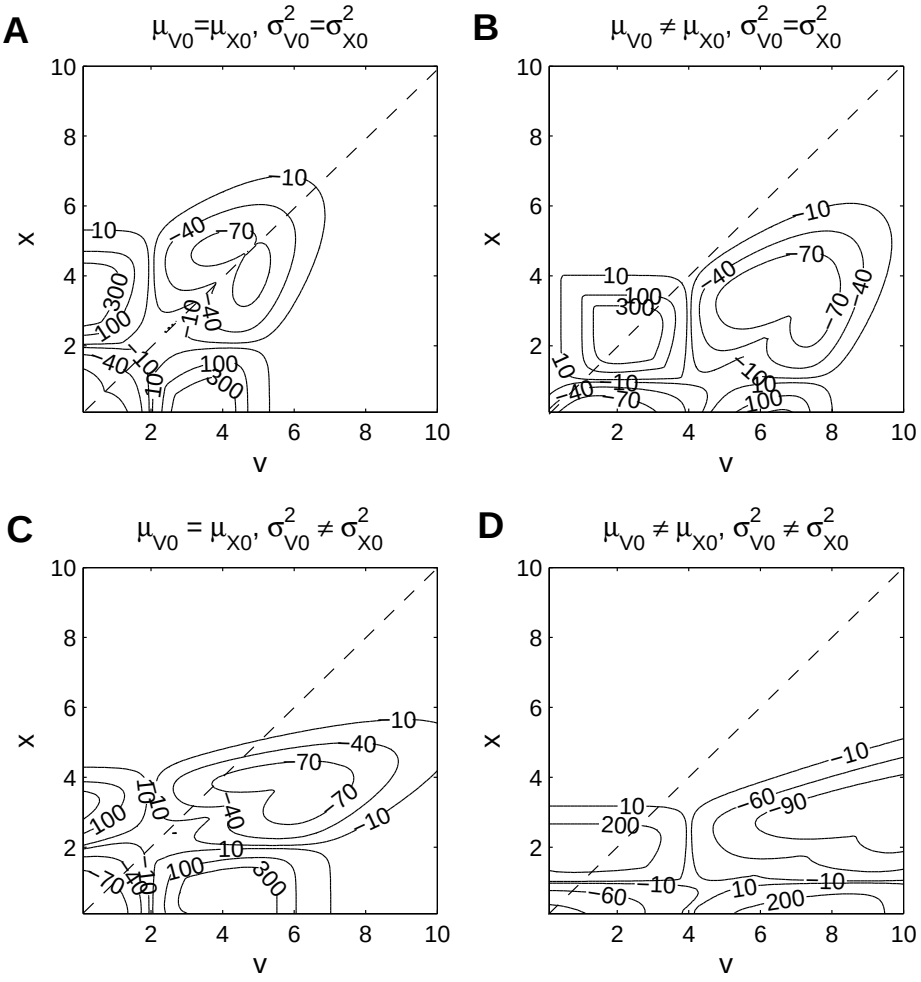


FIGURE 5.6: The CME surfaces on the v-x input plane for the two-channel perceptron model with different combinations of means and variances when $\rho_0 = 0.8$. The statistical parameters are (A) $\mu_{V0} = \mu_{X0} = 2$ and $\sigma^2_{V0} = \sigma^2_{X0} = 2$; (B) $\mu_{V0} = 1$, $\mu_{X0=4}$ and $\sigma^2_{V0} = \sigma^2_{X0} = 2$; (C) $\mu_{V0} = \mu_{X0} = 2$, $\sigma^2_{V0} = 1$ and $\sigma^2_{X0} = 4$; (D) $\mu_{V0} = 1$, $\mu_{X0=4}$, $\sigma^2_{V0} = 1$ and $\sigma^2_{X0} = 4$. The other parameters are set as those of V and X channels in Table 5.2. The dashed lines represent the diagonal in the v-x input plane.

To test whether the forms of input distribution affect this result, we further study the case when the inputs follow a Poisson distribution or a Cowan exponential distribution. We select the Poisson distribution and the Cowan exponential distribution for two reasons. Firstly, they have a relatively simple expression to analyze, and secondly, their correlation coefficients range from 0 to 1.

5.3.3 Poisson Distribution

We firstly construct a bivariate Poisson distribution $P(v, x)$, where v and x are instances of two Poisson processes V and X:

$$V = Z_V + Z, X = Z_X + Z, \tag{5.27}$$

where Z_V, Z_X and Z are independent Poisson processes with rates λ_V, λ_X and λ respectively. Accordingly, the marginal means μ_V for V and μ_X for X are

$$\begin{aligned} \mu_V &= \lambda_V + \lambda, \\ \mu_X &= \lambda_X + \lambda, \end{aligned} \tag{5.28}$$

the variances σ_V^2 and σ_X^2 are

$$\begin{aligned} \sigma_V^2 &= \lambda_V + \lambda, \\ \sigma_X^2 &= \lambda_X + \lambda, \end{aligned} \tag{5.29}$$

and the correlation coefficient ρ is given by

$$\rho = \frac{\lambda}{\sqrt{(\lambda_V + \lambda)(\lambda_X + \lambda)}}. \tag{5.30}$$

A bivariate Poisson distribution $P(v, x)$ is then defined as:

$$\begin{aligned} P[V = v \ \& \ X = x] &= P[Z_V + Z = v \ \& \ Z_X + Z = x] \\ &= \exp[-(\lambda_V + \lambda_X + \lambda)] \sum_{z=0}^{\min(v,x)} \frac{\lambda_v^{v-z}}{(v-z)!} \frac{\lambda_x^{x-z}}{(x-z)!} \frac{\lambda^z}{z!}. \end{aligned} \tag{5.31}$$

According to the Bayes' model (Patton and Anastasio, 2003), two input distributions are required, one for the spontaneous condition and the other for the driven condition. We denote the spontaneous input distribution as a bivariate Poisson distribution $P_0(v, x)$ with parameters λ_{V0}, λ_{X0}, λ_0 and ρ_0, and represent the driven input distribution as $P_1(v, x)$ with parameters λ_{V1}, λ_{X1}, λ_1 and ρ_1. Moreover, to comply with Patton and Anastasio (2003)'s selection of statistical parameters, we assume that $\mu_{Vt} = \mu_{Xt}$ and $\sigma_{Vt}^2 = \sigma_{Xt}^2$, where $t = 0$ corresponds to the spontaneous condition and $t = 1$ corresponds to the driven condition, and we thus have $\lambda_{Vt} = \lambda_{Xt}$. Consequently, Eq. (5.30) can be reduced to

$$\rho_t = \frac{\lambda_t}{\lambda_{Vt} + \lambda_t}, \tag{5.32}$$

which gives us

$$\lambda_{Vt} = \left(\frac{1-\rho_t}{\rho_t}\right)\lambda_t. \tag{5.33}$$

Substituting $\lambda_{Vt} = \lambda_{Xt}$ and Eq. (5.33) into Eq. (5.31), we can rewrite the bivariate Poisson distribution as

$$P_t(v,x) = \exp\left[\frac{2-\rho_t}{\rho_t}\lambda_t\right]\sum_{z=0}^{\min(v,x)}\left(\frac{1-\rho_t}{\rho_t}\right)^{v+x-2z}\frac{\lambda_t^{v+x-z}}{(v-z)!(x-z)!z!}. \tag{5.34}$$

According to Eq. (5.3), the neuronal response $f = P(T=1|\mathbf{m})$ where $\mathbf{m} = (v,x)^T$ is then

$$f = \left\{1 + \frac{P(T=0)}{P(T=1)}\exp\left[-\left(\frac{2-\rho_0}{\rho_0}\right)\lambda_0 + \left(\frac{2-\rho_1}{\rho_1}\right)\lambda_1\right]\frac{C_0}{C_1}\right\}^{-1}, \tag{5.35}$$

where

$$C_0 = \sum_{z=0}^{\min(v,x)}\left(\frac{1-\rho_0}{\rho_0}\right)^{v+x-2z}\frac{\lambda_0^{v+x-z}}{(v-z)!(x-z)!z!}, \tag{5.36}$$

and

$$C_1 = \sum_{z=0}^{\min(v,x)}\left(\frac{1-\rho_1}{\rho_1}\right)^{v+x-2z}\frac{\lambda_1^{v+x-z}}{(v-z)!(x-z)!z!}. \tag{5.37}$$

To analyze the CME index, we need to calculate bimodal response f_{VX} and unimodal responses f_V and f_X. In the both-driven case, we only consider the case when V and X inputs are equal, as in Patton and Anastasio (2003), and denote the inputs as $\mathbf{m} = (m,m)^T$. Substituting $\mathbf{m}$ into Eq. (5.35), we obtain

$$f_{VX} = \left\{1 + \frac{P(T=0)}{P(T=1)}\exp\left[-\left(\frac{2-\rho_0}{\rho_0}\right)\lambda_0 + \left(\frac{2-\rho_1}{\rho_1}\right)\lambda_1\right]\frac{C_0^{VX}}{C_1^{VX}}\right\}^{-1}, \tag{5.38}$$

where

$$C_0^{VX} = \sum_{z=0}^{m}\left(\frac{1-\rho_0}{\rho_0}\right)^{2(m-z)}\frac{\lambda_0^{2m-z}}{(m-z)!(m-z)!z!}, \tag{5.39}$$

and

$$C_1^{VX} = \sum_{z=0}^{m}\left(\frac{1-\rho_1}{\rho_1}\right)^{2(m-z)}\frac{\lambda_1^{2m-z}}{(m-z)!(m-z)!z!}. \tag{5.40}$$

We can rewrite the expression of C_0^{VX} as

$$C_0^{VX} = \frac{\lambda_0^m}{m!} + \sum_{z=0}^{m-1}\left(\frac{1-\rho_0}{\rho_0}\right)^{2(m-z)}\frac{\lambda_0^{2m-z}}{(m-z)!(m-z)!z!}, \tag{5.41}$$

from which it is clear that as the spontaneous correlation coefficient $\rho_0 \to 1$, $C_0^{VX} \to$

$\lambda_0^m/m!$. With this limit, we can acquire the limit of bimodal response f_{VX},

$$f_{VX} \to \left\{1 + \frac{P(T=0)}{P(T=1)} \exp\left[-\lambda_0 + \left(\frac{2-\rho_1}{\rho_1}\right)\lambda_1\right] \frac{\lambda_0^m}{C_1^{VX} m!}\right\}^{-1}. \tag{5.42}$$

Since the denominator of f_{VX} is larger than one, we know that the limit of f_{VX} is less than one.

In the one-driven case, we denote the unimodal V response as $\mathbf{m}_V = (m, \mu_0)^T$, where $m > \mu_0$ and $\mu_0 = \lambda_0/\rho_0$ represents the spontaneous mean for X channel. Substituting $\mathbf{m}_V$ into Eq. (5.35), we obtain unimodal response f_V,

$$f_V = \left\{1 + \frac{P(T=0)}{P(T=1)} \exp\left[\left(\frac{2-\rho_1}{\rho_1}\right)\lambda_1 - \left(\frac{2-\rho_0}{\rho_0}\right)\lambda_0\right] \frac{C_0^V}{C_1^V}\right\}^{-1}, \tag{5.43}$$

where

$$C_0^V = \sum_{z=0}^{\mu_0} \left(\frac{1-\rho_0}{\rho_0}\right)^{\mu_0+m-2z} \frac{\lambda_0^{\mu_0+m-z}}{(m-z)!(\mu_0-z)!z!}, \tag{5.44}$$

and

$$C_1^V = \sum_{z=0}^{\mu_0} \left(\frac{1-\rho_1}{\rho_1}\right)^{\mu_0+m-2z} \frac{\lambda_1^{\mu_0+m-z}}{(m-z)!(\mu_0-z)!z!}. \tag{5.45}$$

Since $m > \mu_0$ and $z \leq \mu_0$, we have $\mu_0 + m - 2z \geq m - \mu_0 > 0$. It is then obvious that when the spontaneous correlation coefficient $\rho_0 \to 1$, we have $C_0^V \to 0$. According to this limit, we can observe that f_V approaches one [Eq. (5.43)]. Moreover, since the means and variances for V and X channels are the same, this analysis also applies to the unimodal response f_X, that is, f_X also approaches one. Combining these unimodal limits with that of the bimodal response f_{VX} [Eq. (5.21)], we know that the limit of the ratio $f_{VX}/\max\{f_V, f_X\}$ is less than one when $\rho_0 \to 1$. Consequently, the condition that $\rho_0 \to 1$ again leads to response suppression under the Bayesian model for bivariate Poisson distributed inputs with equal values.

5.3.4 Cowan Exponential Distribution

A bivariate Cowan exponential distribution for V and X sensory channels is defined as (Cowan, 1987)

$$P(v,x) = \frac{\lambda(1-\eta)}{2s^3} \quad \left\{4\eta vx + \lambda s\left[v^2 + x^2 + s(v+x) + 2vx(1-\eta)\right]\right\} \exp\left[-\frac{1}{2}\lambda(v+x+s)\right], \tag{5.46}$$

where $s^2 = v^2 + x^2 - 2vx\cos a$ and $\eta = (1 + \cos a)/2$. The correlation coefficient ρ between v and x can be expressed as:

$$\rho = \begin{cases} 1 & \text{if } a = 0, \\ -1 + \dfrac{4}{1+\cos a}\left\{1 - \dfrac{1-\cos a}{1+\cos a}\log\left(\dfrac{2}{1-\cos a}\right)\right\} & \text{if } 0 < a < \pi, \\ 0 & \text{if } a = \pi. \end{cases} \tag{5.47}$$

When $a \to 0$, $1 - \cos a$ goes to zero. According to Eq. (5.47), we can see that the term $(1 - \cos a)/(1 + \cos a)$ dominates the logarithm term, so $\frac{1-\cos a}{1+\cos a}\log(\frac{2}{1-\cos a})$ goes to zero and $\rho \to 1$. The marginal means are $E(V) = E(X) = 1/\lambda$. Similarly to the previous analysis, two Cowan exponential distributions are considered, one for the spontaneous condition $P_0(v, x)$ with parameters a_0, λ_0, s_0 and ρ_0, and the other for the driven condition $P_1(v, x)$ with parameters a_1, λ_1, s_1 and ρ_1. Substituting $P_0(v, x)$ and $P_1(v, x)$ [Eq. (5.46)] into Eq. (5.3), we acquire the expression of neuronal response,

$$f = \left\{1 + \frac{P(T=0)}{P(T=1)}\frac{\lambda_0(1-\eta_0)s_1^3}{\lambda_1(1-\eta_1)s_0^3}\frac{C_0}{C_1}\exp(C)\right\}^{-1}, \tag{5.48}$$

where

$$C_0 = 4\eta_0 vx + \lambda_0 s_0[v^2 + x^2 + s_0(v + x) + 2vx(1 - \eta_0)], \tag{5.49}$$

$$C_1 = 4\eta_1 vx + \lambda_1 s_1[v^2 + x^2 + s_1(v + x) + 2vx(1 - \eta_1)], \tag{5.50}$$

and

$$C = -\frac{1}{2}[\lambda_0(v + x + s_0) + \lambda_1(v + x + s_1)]. \tag{5.51}$$

In the both-driven case, the inputs from the V and X channels are taken the same value, following Patton and Anastasio (2003), and are expressed as $\mathbf{m} = (m, m)^T$. Substituting $\mathbf{m}$ into Eq. (5.48) and, after some transformations, we have

$$f_{VX} = \left\{1 + P\sqrt{\frac{1-\cos a_1}{1-\cos a_0}}\frac{\lambda_0}{\lambda_1}\frac{C_0^{VX}}{C_1^{VX}}\exp(C^{VX})\right\}^{-1}, \tag{5.52}$$

where

$$C_0^{VX} = 2(1 + \cos a_0)m^2 + 2\sqrt{2(1 - \cos a_0)}\lambda_0 m^3 + 4m^3(1 - \cos a_0)\lambda_0, \tag{5.53}$$

$$C_1^{VX} = 2(1 + \cos a_1)m^2 + 2\sqrt{2(1 - \cos a_1)}\lambda_1 m^3 + 4m^3(1 - \cos a_1)\lambda_1, \tag{5.54}$$

and

$$C^{VX} = -\frac{1}{2}\lambda_0[2m + m\sqrt{2(1 - \cos a_0)}] + \frac{1}{2}\lambda_1[2m + m\sqrt{2(1 - \cos a_1)}]. \tag{5.55}$$

Eq. (5.52) contains the term $\sqrt{(1 - \cos a_1)/(1 - \cos a_0)}$ that goes to infinity as $a_0 \to 0$ (or $\rho_0 \to 1$), and therefore the bimodal VX response $f_{VX} \to 0$.

For f_V, substituting the unimodal input $\mathbf{m} = (m, 1/\lambda_0)^T$ into Eq. (5.48), we acquire

$$f_V = \left[1 + P\frac{(1-\cos a_0)\lambda_0 s_1^3}{(1-\cos a_1)\lambda_1 s_0^3}\frac{C_0^V}{C_1^V}\exp\{C^V\}\right]^{-1}, \tag{5.56}$$

where

$$C_0^V = \frac{2(1+\cos a_0)m}{\lambda_0} + \lambda_0 s_0\left[\frac{1}{\lambda_0^2} + m^2 + s_0(m+\frac{1}{\lambda_0}) + \frac{m(1-\cos a_0)}{\lambda_0}\right], \tag{5.57}$$

$$C_1^V = \frac{2(1+\cos a_1)m}{\lambda_0} + \lambda_1 s_1\left[\frac{1}{\lambda_0^2} + m^2 + s_1(m+\frac{1}{\lambda_0}) + \frac{m(1-\cos a_1)}{\lambda_0}\right], \tag{5.58}$$

and

$$C^V = -\frac{1}{2}\lambda_0(m+\frac{1}{\lambda_0}+s_0) + \frac{1}{2}\lambda_0(m+\frac{1}{\lambda_0}+s_1). \tag{5.59}$$

When $a_0 \to 0$, or equally $\rho_0 \to 1$, we have $s_0 \to |(m-1)/\lambda_0|$, in which case C_0^V/C_1^V and the exponential term in Eq. (5.56) are finite. It is then clear from Eq. (5.56) that the term $1-\cos a_0$ will drive the unimodal response f_V to one. Moreover, because of the symmetric means and variances in V and X, we also have the unimodal response $f_X \to 1$ when the spontaneous correlation coefficient $\rho_0 \to 1$. With both the limits for bimodal response and unimodal responses, we now know that the ratio $f_{VX}/\max\{f_V, f_X\}$ approaches 0 as $\rho_0 \to 1$, again leading to suppression.

5.3.5 Bivariate Distribution

At this stage, we have shown that, when the spontaneous correlation coefficient ρ_0 between the two channels is close to 1, the Bayesian model exhibits suppression for equal-valued bimodal inputs, no matter whether they are Gaussian, Poisson or Cowan exponential distributed. In the following, we will prove that the same result applies to any bivariate distribution when the means and the variances of the two sensory channels are the same.

Consider a bivariate probability distribution $P(v, x)$, where the means of v and x are labeled as μ_V and μ_X, and the variances as σ_V^2 and σ_X^2 respectively. We have

$$\begin{aligned} E[(V-X)^2] &= E[V^2 - 2VX + X^2] \\ &= E[(V-\mu_V)^2] + E[(X-\mu_X)^2] + E[2\mu_V V + 2\mu_X X - \mu_V^2 - \mu_X^2 - 2VX] \\ &= \sigma_V^2 + \sigma_X^2 + \mu_V^2 + \mu_X^2 - 2E[VX]. \end{aligned} \tag{5.60}$$

If the means of V and X are equal, that is, $\mu_V = \mu_X = \mu$, Eq. (5.60) can be transformed

as

$$\begin{aligned} E[(V-X)^2] &= \sigma_V^2 + \sigma_X^2 - 2(E[VX] - \mu^2) \\ &= \sigma_V^2 + \sigma_X^2 - 2\text{cov}(V,X) \\ &= \sigma_V^2 + \sigma_X^2 - 2\sigma_V\sigma_X\rho. \end{aligned} \tag{5.61}$$

Furthermore, if $\sigma_V = \sigma_X = \sigma$, Eq. (5.61) can be rewritten as

$$E[(V-X)^2] = 2\sigma^2(1-\rho). \tag{5.62}$$

According to Chebyshev's inequality, we have

$$P[|V-X| \geq \varepsilon] \leq \frac{E[(V-X)^2]}{\varepsilon^2}. \tag{5.63}$$

Substituting Eq. (5.62) into Eq. (5.63), we obtain

$$P[|V-X| \geq \varepsilon] \leq \frac{2\sigma^2}{\varepsilon^2}(1-\rho), \tag{5.64}$$

which shows that the probability of V and X being different approaches 0 when ρ approaches 1. Graphically, the probability $P(v,x)$ is a surface spanning the two-dimensional V-X plane. Eq. (5.64) indicates that the surface converges to the line $v = x$ as the correlation coefficient $\rho \to 1$. As an example, we produced the figures for bivariate Gaussian, Poisson and Cowan exponential distributions for $\rho_0 = 0.1$ and $\rho_0 = 0.95$ (Fig. 5.7). We can see that the surfaces of these three probability distributions indeed converge to the diagonal when ρ_0 is increased to 0.95. In the extreme case that the correlation coefficient $\rho = 1$, we have $E[(V-X)^2] = 0$ according to Eq. (5.62). Consequently, for the bivariate distribution $P(v,x)$, we have

$$P(v,x) = \begin{cases} 0, & \text{if } v \neq x, \\ > 0, & \text{if } v = x, \end{cases} \tag{5.65}$$

when $\rho = 1$.

Consider a bimodal response f_{VX} to equal-valued inputs $\mathbf{m} = (m,m)^T$. According to Eq. (5.3), we have

$$f_{VX} = \frac{1}{1 + \dfrac{P(T=0)}{P(T=1)} \dfrac{P((m,m)^T|T=0)}{P((m,m)^T|T=1)}}, \tag{5.66}$$

According to Eq. (5.65), we have $P((m,m)^T|T=0) > 0$ when the spontaneous correlation coefficient $\rho_0 \to 1$. Accordingly, the bimodal response $f_{VX} < 1$. On the other

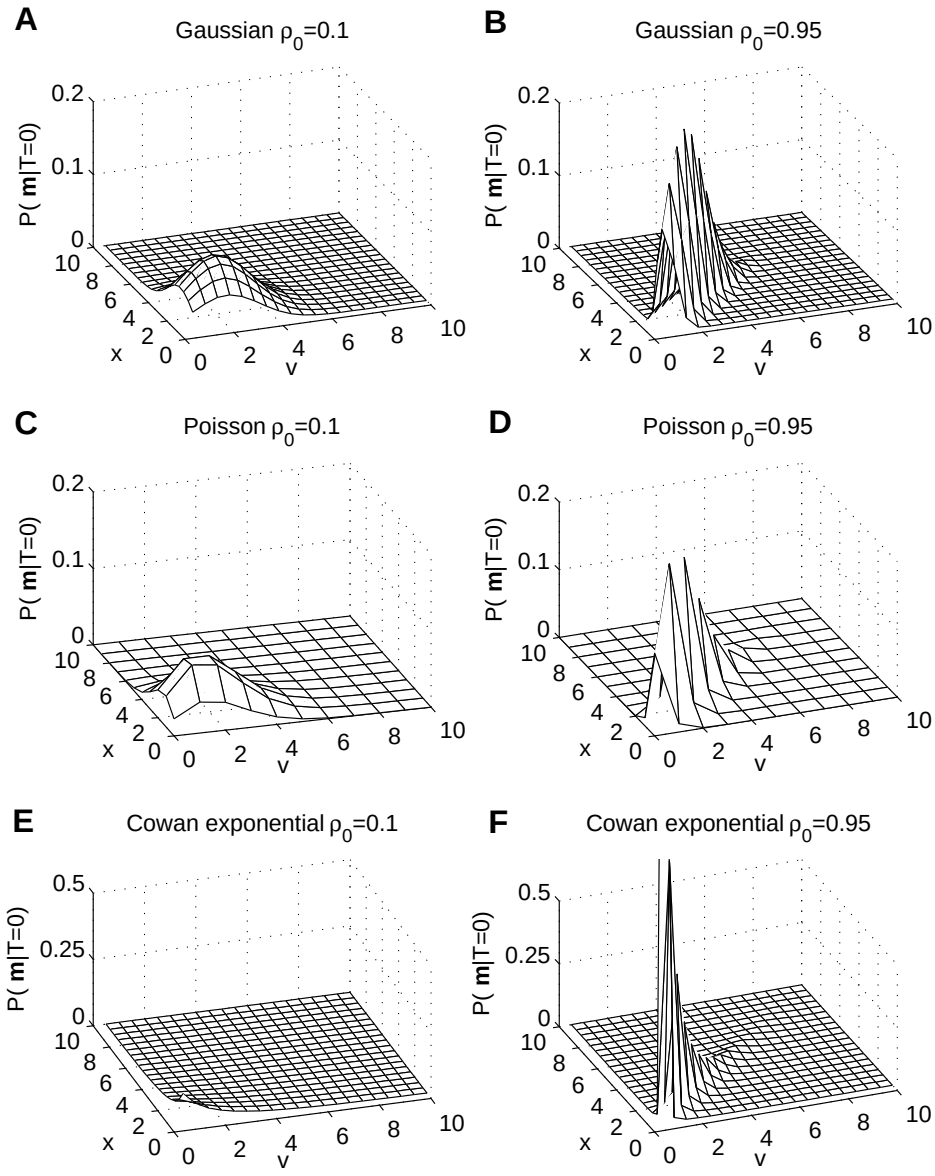


FIGURE 5.7: Probability distributions $P(\mathbf{m}|T = 0)$ when $\rho_0 = 0.1$ and $\rho_0 = 0.95$ for two sensory channels V and X. Two Gaussian distributions are shown in (A) for $\rho_0 = 0.1$ and (B) for $\rho_0 = 0.95$. In (A), $\sigma^2_{VX0} = 0.2$, and in (B) $\sigma^2_{VX0} = 1.9$. The other statistical parameters are set as those of V and X channels in Table 5.2. Two Poisson distributions are shown in (C) for $\rho_0 = 0.1$ and (D) for $\rho_0 = 0.95$. In (C), $\lambda_V = \lambda_X = 1.8$ and $\lambda = 0.2$; In (D), $\lambda_V = \lambda_X = 0.1$ and $\lambda = 1.9$. For the meaning of the parameters, refer to Section 5.3.3. Two Cowan exponential distributions are shown in (E) for $\rho_0 = 0.1$ and (F) for $\rho_0 = 0.95$. In (E), $a = 2.07$ and $\lambda = 0.5$; In (D), $a = 0.155$ and $\lambda = 0.5$. See Section 5.3.4 for the meaning of the parameters.

hand, the unimodal response f_V to the input represented by $\mathbf{m} = (m, \mu_0)^T$ is

$$f_V = \frac{1}{1 + \dfrac{P(T=0)}{P(T=1)} \dfrac{P((m,\mu_0)^T|T=0)}{P((m,\mu_0)^T|T=1)}}. \tag{5.67}$$

In this case, according to Eq. (5.65), we have $P((m, \mu_0)^T|T = 0) \to 0$ when the spontaneous correlation coefficient $\rho_0 \to 1$. The unimodal response f_V thus approaches one. This also applies to f_X since the mean and variance of X and V are the same. As a result, we know that the ratio $f_{VX}/\max\{f_V, f_X\}$ is less than 1, and thus the model exhibits suppression. Fig. 5.8 shows the surface of neuronal response to V and X inputs. When $\rho_0 = 0.1$, a bimodal response f_{VX} to $\mathbf{m} = (m, m)^T$ is larger than a unimodal response, for example, f_V to $\mathbf{m} = (m, \mu_0)^T$ (Fig. 5.8A). However, when $\rho_0 = 0.95$, we can see that a bimodal response f_{VX} to $\mathbf{m} = (m, m)^T$ is always smaller than a unimodal response such as f_V to $\mathbf{m} = (m, \mu_0)^T$ (Fig. 5.8B), in which case suppression is observed.

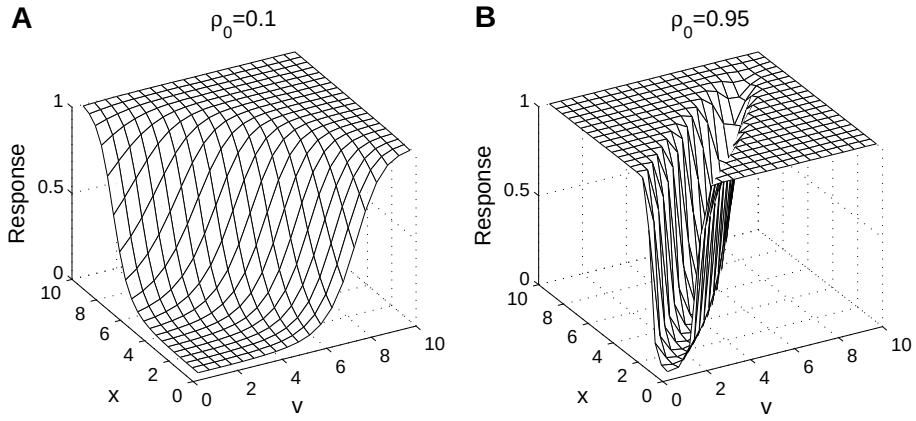


FIGURE 5.8: Neuronal response ($P(T = 1|\mathbf{m})$) for Gaussian inputs when $\rho_0 = 0.1$ and $\rho_0 = 0.95$. The other parameters are set as the V and X channels in Table 5.2.

As proven above, given that the means and the variances are the same for two sensory channels, the Bayesian model in fact produces MSS when the spontaneous correlation coefficient $\rho_0 \to 1$ independent of the specific form of input probability distribution. However, this only works for bimodal inputs with equal values. For an input $\mathbf{m} = (m_1, m_2)^T$ where $(m_1 \neq m_2)$, for example, both the bimodal response f_{VX} and the unimodal responses f_V and f_X approach 1 as the spontaneous correlation coefficient $\rho_0 \to 1$. The extreme of the ratio $f_{VX}/\max\{f_V, f_X\}$ in this case will be 1, corresponding to the situation that neither enhancement nor suppression is produced (also see Fig. 5.5). In fact, the Bayesian model proposed by Patton and Anastasio (2003) makes use of this very special properties of a bivariate distribution to produce the phenomenon of MSS, and it is thus not surprising to find that the model is very sensitive to the statistical parameters.

To sum up, the capability of Patton and Anastasio (2003)'s model to produce MSS critically depends on the choices of statistical parameters as well as the equality of

the inputs. It is therefore not a plausible model of the multisensory responses of DSC neurons. The hypothesis that each DSC neuron carries out a Bayesian computation is thus also questionable, and we need to interpret the mechanisms underlying DSC neuronal responses in another way.

5.4 Summary

The Patton model is a representative model to account for the response properties, such as CME and MSS, of a DSC neuron. We showed in this chapter, however, that the capability of the model to exhibit MSS critically depends on the parameters. Patton and Anastasio (2003) suggested that a large enough spontaneous correlation coefficient is crucial for the model to exhibit MSS. In numerical studies, we demonstrated that such a spontaneous correlation coefficient leads to some properties that conflict with those of a neuron. To further explore this, we mathematically analyzed the behaviours of the model when the inputs follow a Gaussian, Poisson or Cowan exponential distribution. Our results indicate that the MSS property of the model is produced by special choices of the statistical parameters as well as the inputs. This is further supported by a general proof for inputs following arbitrary distribution.

Our studies suggest that the Patton model is not an appropriate model for DSC neuronal responses. In the following chapters, we propose a new interpretation of the mechanisms underlying multisensory integration in DSC neurons based on a sigmoidal response function adapted according to the input statistics.

Chapter 6

Sigmoidal Response Function for Multisensory Integration

In the previous chapter, we have shown that Patton and Anastasio (2003)'s Bayesian model does not plausibly account for the response properties of DSC neurons. In this chapter, we instead focus on the role of a non-linear, saturating response function, such as the sigmoidal function, in reproducing the experimental results, such as the CME associated with the inverse effectiveness and the additive operations. We will discuss how the sigmoidal response function could account for these response properties, based on which we propose a simple interpretation for the mechanisms underlying the complicated experimental data.

6.1 Introduction

Neuronal responses are generally measured as the spiking rate of neurons. Subject to physical constraints, such as the refractory period lasting for around 2 ms, the spiking rate of a neuron has an upper limit of about 100 to 200 Hz. When neuronal responses reach this upper limit, an increase in the input will not induce a stronger response. Accordingly, neuronal responses are sometimes modeled by a saturating function. Popular choices of response function in the literature include the hyperbolic ratio function,

$$r(x) = s\frac{x^{4\gamma\theta/s}}{\theta^{4\gamma\theta/s} + x^{4\gamma\theta/s}}, \tag{6.1}$$

where s represents the bound of responses, θ determines the point of semi-saturation, $r(\theta) = s/2$, and γ the gain at semi-saturation. For example, the hyperbolic ratio function has been widely adopted to model neuronal response for contrast gain control in visual system (Naka and Rushton, 1966; Albrecht et al., 1984; Wilson and Humanski, 1993). Although not a saturating function, a linear model with rectification is a good model of

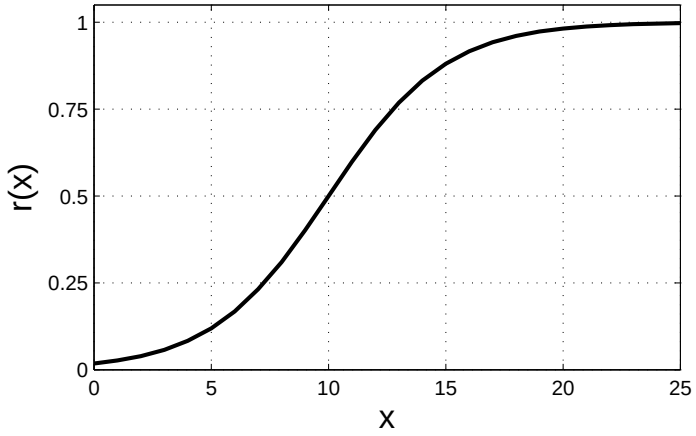


FIGURE 6.1: A sigmoidal function with $\gamma = 0.1$, $\theta = 10$ and $s = 1$.

V1 neuronal responses,

$$r(x) = \gamma[x - \theta]_+, \tag{6.2}$$

where $[]_+$ denotes the positive part, θ is here the threshold for response onset, and γ the gain. It is popular in the modelling of contrast adaptation in retinal circuitry (Chander and Chichilnisky, 2001; Kim and Rieke, 2001; Zaghloul et al., 2005). However, we will instead use the sigmoidal response function,

$$r(x) = \frac{s}{1 + \exp\left[-\dfrac{4\gamma}{s}(x - \theta)\right]}, \tag{6.3}$$

where θ and γ are, as for the hyperbolic ratio function, the semi-saturation constant and the gain, respectively. For γ large, θ determines the transition from no response to saturated response, so we refer to it as the threshold. The sigmoidal function can also be written in a hyperbolic form,

$$r(x) = \frac{s}{2}\left[1 + \tanh\frac{2\gamma}{s}(x - \theta)\right]. \tag{6.4}$$

Fig. 6.1 shows a sigmoidal function with $\gamma = 0.1$, $\theta = 10$ and $s = 1$. We employ the sigmoidal response function because it is defined for $x < 0$ (unlike the hyperbolic ratio function for most values of γ) and everywhere differentiable (unlike the linear rectified response).

Although the sigmoidal function in Eq. (6.3) or Eq. (6.4) is only for one sensory input x, it can be extended to deal with multimodal inputs. For example, in terms of a DSC neuron receiving bimodal inputs x_1 and x_2, we model the neuronal responses as

$$r(x_1, x_2) = \frac{s}{1 + \exp\left[-\dfrac{4}{s}(\gamma_1 x_1 + \gamma_2 x_2 - \theta)\right]}, \tag{6.5}$$

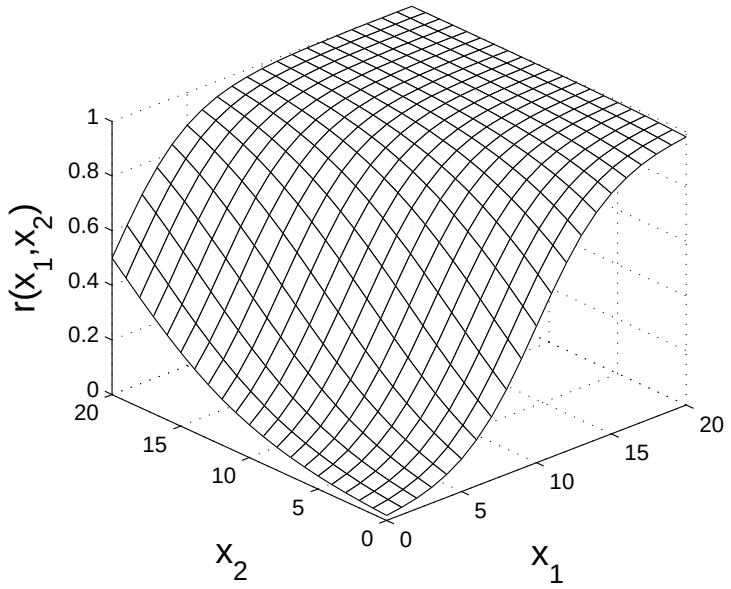


FIGURE 6.2: A bimodal sigmoidal function with $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$ and $s = 1$.

or equally, in the hyperbolic form,

$$r(x_1, x_2) = \frac{s}{2}\left[1 + \tanh\frac{2}{s}(\gamma_1 x_1 + \gamma_2 x_2 - \theta)\right], \tag{6.6}$$

where γ_1 and γ_2 represent the gains for x_1 and x_2 respectively, and θ represents the threshold for the neuron. With gains $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, the threshold $\theta = 1$ and $s = 1$, Fig. 6.2 shows the surface of a bimodal sigmoidal function. Stanford et al. (2005) indicated that a saturating response function is sufficient to account for the inverse effectiveness principle. We will extend this notion and show in the following sections that a neuron responding under the sigmoidal function is capable of exhibiting many observed response properties of DSC neurons, including the CME property and the additive responses. Moreover, we suggests a clear interpretation of the diverse operational modes reported by Perrault et al. (2005). For simplicity, we only consider the bimodal sigmoidal function in our discussion.

6.2 The CME Property

It has been widely reported that DSC neurons exhibit the property of CME quantified by Eq. (2.1) when bimodal visual and auditory stimuli are presented (Stein and Meredith, 1993; Stein et al., 2004). Moreover, the largest CME in DSC neurons is achieved with weak visual and auditory stimuli, a property called inverse effectiveness (IE). Presumably, this response property could improve the ability of an animal to detect weak signals.

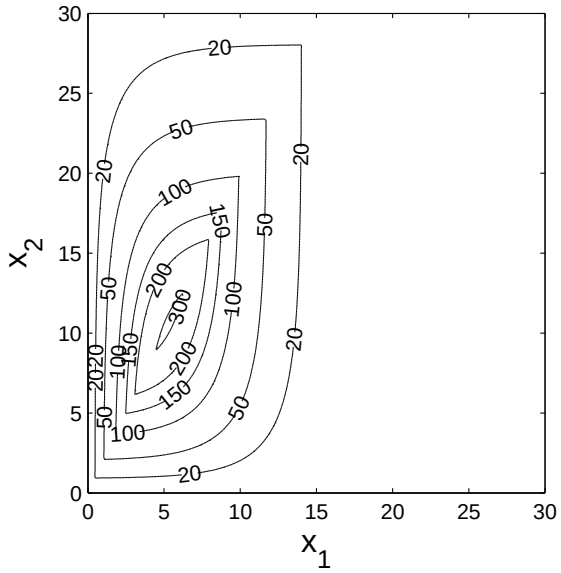


FIGURE 6.3: CME contours in the $x_1 - x_2$ input plane when $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$ and $s = 1$.

The occurrence of CME associated with IE in a neuron responding under a sigmoidal function is straightforward. The presence of a cross-modal stimulus x_2 in addition to a stimulus x_1 increases the combined input $\gamma_1 x_1 + \gamma_2 x_2$. According to Eq. (6.5), the resulting bimodal response $r(x_1, x_2)$ is thus larger than either $r(x_1, 0)$ or $r(0, x_2)$, leading to enhancement. Furthermore, since the sigmoidal function is a saturating function, a strong input stimulus x_1 could arouse a saturated response. For such a sensory stimulus x_1, the presence of a cross-modal sensory stimulus x_2 will not induce a stronger response. The enhancement is therefore approximately zero for strong sensory inputs. This indicates that the maximum CME occurs for weaker stimuli, corresponding to the IE property. To illustrate this, we produce the CME contours in the bimodal $x_1 - x_2$ input plane. In Fig. 6.3, the parameters for the sigmoidal function are arbitrarily set as $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$ and $s = 1$. We can observe that all the points in the $x_1 - x_2$ input plane exhibit CME. Moreover, the maximum CME occurs at around $x_1 = 5$, $x_2 = 10$. As x_1 and x_2 increase, the CME gradually decreases to zero (IE). The "kinks" in the contours are the result of a switching between the larger of the two unimodal responses in the definition of the CME index [see Eq. (2.1)].

As a characteristic property of CME, IE indicates that the maximum CME occurs for weak stimuli. To study how the sigmoidal function produce IE, we mathematically analyze the bimodal sigmoidal response function given in Eq. (6.5) for its maximum CME. In the calculation of the CME index, we consider the bimodal response $r(x_1, x_2)$ and the two unimodal responses $r(x_1, 0)$ and $r(0, x_2)$. Here we do not include spontaneous inputs in the unimodal responses, for example, using $r(x_1, x_2^0)$ where x_2^0 represents spontaneous x_2 input, since it does not bring any qualitative difference to the results. Without loss of generality, we further assume that $r(x_1, 0) \geq r(0, x_2)$, from which we

have $\gamma_1 x_1 \geq \gamma_2 x_2$ according to Eq. (6.5). The enhancement $E(x_1, x_2)$ given in Eq. (2.1) can then be written as

$$E(x_1, x_2) = \frac{r(x_1, x_2) - r(x_1, 0)}{r(x_1, 0)} = \frac{r(x_1, x_2)}{r(x_1, 0)} - 1. \tag{6.7}$$

To maximize $E(x_1, x_2)$ for a fixed x_1, we therefore need to maximize $r(x_1, x_2)$, which increases monotonically with the increase in x_2 [see Eq. (6.5)]. We have assumed that $\gamma_1 x_1 \geq \gamma_2 x_2$, so the maximum $r(x_1, x_2)$ is achieved when $x_2 = \gamma_1 x_1/\gamma_2$. We denote

$$E(x_1) = \frac{r(x_1, \gamma_1 x_1/\gamma_2)}{r(x_1, 0)} = \frac{1 + \exp\left[-\frac{4}{s}(\gamma_1 x_1 - \theta)\right]}{1 + \exp\left[-\frac{4}{s}(2\gamma_1 x_1 - \theta)\right]}. \tag{6.8}$$

It is obvious that maximizing $E(x_1, x_2)$ is equivalent to maximizing $E(x_1)$. To determine the input x_1 that produces the maximum CME, we then solve $dE(x_1)/dx_1 = 0$ for x_1. Substituting Eq. (6.8) into $dE(x_1)/dx_1 = 0$ and after some transformations, we obtain

$$e^{-\frac{4(2\gamma_1 x_1 - \theta)}{s}} + 2e^{-\frac{4\gamma_1 x_1}{s}} - 1 = 0. \tag{6.9}$$

Solving Eq. (6.9) for x_1, we have

$$x_1 = -\frac{s \ln\left[\frac{\sqrt{1+\exp(\frac{4\theta}{s})}-1}{\exp(\frac{4\theta}{s})}\right]}{4\gamma_1}, \tag{6.10}$$

for which input the CME is maximal. Although the expression for x_1 appears complicated, in the case that $\exp(\frac{4\theta}{s}) \gg 1$, Eq. (6.10) can be approximated as

$$x_1 \approx \theta/(2\gamma_1). \tag{6.11}$$

Since $\gamma_1 x_1 = \gamma_2 x_2$, we also have

$$x_2 \approx \theta/(2\gamma_2). \tag{6.12}$$

Accordingly, maximum CME can be observed at around $x_1 = \theta/(2\gamma_1)$, $x_2 = \theta/(2\gamma_2)$ in the $x_1 - x_2$ input plane, above which CME gradually decreases to zero. For the CME contours in Fig. 6.3, we have $\gamma_1 = 0.1$, $\gamma_2 = 0.05$ and $\theta = 1$. According to Eqs. (6.11) and (6.12), the bimodal input for the maximum CME are $x_1 \approx 5$ and $x_2 \approx 10$, consistent with Fig. 6.3. However, this approximate point for maximum CME works if $\exp(\frac{4\theta}{s}) \gg 1$ is satisfied, so, for example, when s is large, Eqs. (6.11) and (6.12) do not give us a good approximation.

6.3 The ADD Property

The ADD index has also been adopted in the literature to analyze the response property of DSC neurons, according to which the responses of neurons could be characterized as sub-additive (ADD<0), additive (ADD≈ 0) and super-additive (ADD>0) (Perrault et al., 2005; Avillac et al., 2007). Here we analyze how a neuron with the sigmoidal response function exhibits the additive responses defined by the ADD index in Eq. (2.2).

In a sigmoidal response function [Eq. (6.3)], when an input x is around the inflection point, or $x = \theta$, we can observe an approximately linear property since the second-order derivation of the sigmoidal function $f''(x)$ at $x = \theta$ is zero. According to Taylor's theorem, we have

$$f(\theta + \delta x) \approx f(\theta) + \ \delta x \ f'(\theta) + \frac{1}{2} \ \delta x^2 \ f''(\theta) = f(\theta) + \ \delta x \ f'(\theta), \tag{6.13}$$

Intuitively, this linear property of the sigmoidal function contributes to the additive responses of the neurons, so that sensory inputs in the region around θ induce additive responses, inputs on the left tail of the sigmoidal curve induce super-additive responses and those on the right tail sub-additive responses. We thus refer to the left tail of the sigmoidal function as the super-additive region, the area around the inflection point as the additive region and the right tail as the sub-additive region (see Fig. 6.4).

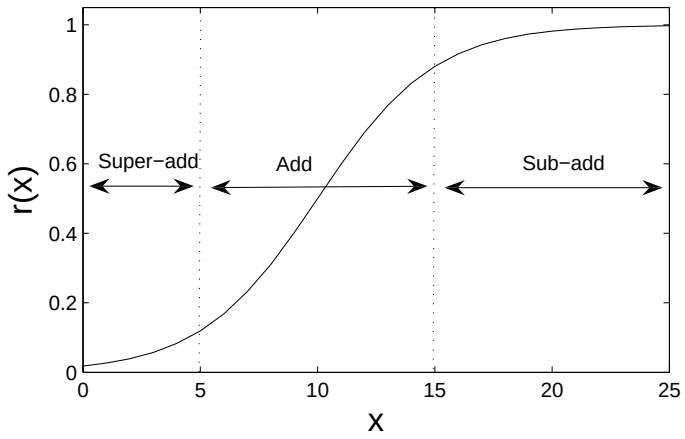


FIGURE 6.4: The super-additive, additive and sub-additive regions for a sigmoidal response function with $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$ and $s = 1$.

Consider the hyperbolic form of a bimodal sigmoidal response function in Eq. (6.6). According to the definition of the ADD index [Eq. (2.2)], the sign of

$$A(x_1, x_2) = r(x_1, x_2) + r(0, 0) - r(x_1, 0) - r(0, x_2) \tag{6.14}$$

determines whether a response is super-additive [$A(x_1, x_2) > 0$], additive [$A(x_1, x_2) \approx 0$] or sub-additive [$A(x_1, x_2) < 0$]. Substituting the expression for the sigmoidal response

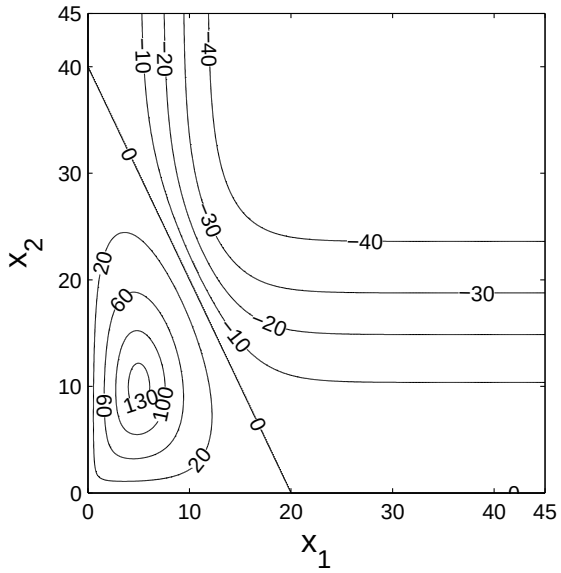


FIGURE 6.5: ADD contours in the $x_1 - x_2$ input plane when $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$ and $s = 1$.

function into Eq. (6.14), we obtain

$$A(x_1, x_2) = \frac{s}{2}[\tanh\frac{2}{s}(\gamma_1 x_1 + \gamma_2 x_2 - \theta) + \tanh\frac{2}{s}(-\theta) - \tanh\frac{2}{s}(\gamma_1 x_1 - \theta) - \tanh\frac{2}{s}(\gamma_2 x_2 - \theta)]. \quad (6.15)$$

For a hyperbolic tangent function, we have

$$\tanh x + \tanh y = \tanh(x + y)(1 + \tanh x \tanh y), \quad (6.16)$$

so Eq. (6.15) can be transformed to

$$A(x_1, x_2) = \frac{s}{2}\tanh\frac{2}{s}(\gamma_1 x_1 + \gamma_2 x_2 - 2\theta)[\tanh\frac{2}{s}(\gamma_1 x_1 + \gamma_2 x_2 - \theta)\tanh\frac{2}{s}(-\theta) - \tanh\frac{2}{s}(\gamma_1 x_1 - \theta)\tanh\frac{2}{s}(\gamma_2 x_2 - \theta)], \quad (6.17)$$

from which we see that $A(x_1, x_2) = 0$ is satisfied when

$$\tanh\frac{2}{s}(\gamma_1 x_1 + \gamma_2 x_2 - 2\theta) = 0, \quad (6.18)$$

which gives us

$$\gamma_1 x_1 + \gamma_2 x_2 - 2\theta = 0. \quad (6.19)$$

Therefore, in terms of the contours of the ADD index in the $x_1 - x_2$ input plane, we would expect to find the zero contour on the line $\gamma_1 x_1 + \gamma_2 x_2 - 2\theta = 0$. This is demonstrated in Fig. 6.5, in which the contours of the ADD index are shown for

$\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$ and $s = 1$. The zero contour is a line $2x_1 + x_2 - 40 = 0$ on the $x_1 - x_2$ input plane, complying with Eq. (6.19). Moreover, the area above the line $2x_1 + x_2 - 40 = 0$ corresponds to the sub-additive region where ADD< 0, and the area below $2x_1 + x_2 - 40 = 0$ corresponds to the super-additive region where ADD> 0.

6.4 Analysis of Stanford et al. (2005)'s Results

Stanford et al. (2005) have systematically explored the integration operations of DSC neurons. They designed an experimental protocol to analyze the data according to their statistical properties, based on which they showed that although DSC neurons exhibit sub-additive, additive and super-additive responses, the additive responses are most often observed. In the experiment, Stanford et al. (2005) recorded the responses of a DSC neuron to each fixed visual, auditory and bimodal visual-auditory stimulus 30 times. The general idea is to construct a distribution of expected additive responses based on the two recorded unimodal responses, and calculate the z-score of an observed bimodal responses based on the recorded bimodal responses for this distribution. The value of the z-score determines the additive properties of the responses. To construct the distribution of the expected additive responses, Stanford et al. (2005) firstly produced a set of 900 elements $E = \{x_1^E, \ldots, x_{900}^E\}$ by a complete combination of the unimodal visual responses and auditory responses, with each of the element x_i^E, where $i \in \{1, \ldots, 900\}$, being a sum of the two unimodal responses. They then randomly selected $n_t = 30$ trials X_j^E $(j = 1, \ldots, n_t)$ (with replacement) from the set E and averaged these trials to produce a sample X^L so that

$$X^L = \frac{1}{n_t} \sum_{j=1}^{n_t} X_j^E. \tag{6.20}$$

Under the same procedure, they generated $n_l = 10000$ samples to form a sample set $L = \{X_1^L, ..., X_{n_l}^L\}$, and consider it to represent a distribution of the expected additive responses for the DSC neuron. According to the central limit theorem, the 10,000 samples follow an approximately normal distribution since each sample X_k^L $(k = 1, \ldots, n_l)$ in the set L is an average of 30 trials. The sample mean

$$\bar{X}_L = \frac{1}{n_l} \sum_{k=1}^{n_l} X_k^L \tag{6.21}$$

and the unbiased estimator of the standard deviation

$$S_L = \sqrt{\frac{1}{n_l - 1} \sum_{k=1}^{n_l} (X_k^L - \bar{X}_L)^2} \tag{6.22}$$

of the set L are then calculated. In addition, the observed bimodal response x_{obs} is acquired by averaging the previously-recorded 30 bimodal responses. The z-score of x_{obs}

based on the sample set L is defined as

$$z = \frac{x_{obs} - \bar{X}_L}{S_L} \tag{6.23}$$

to characterize the response properties of a neuron. According to convention, a significance level 5% for a standard normal distribution, which corresponds to $z = \pm 1.96$, is adopted to decide whether the observed response x_{obs} is significantly different from the expected additive responses represented by the sample set L: when $z < -1.96$, the response is considered as sub-additive; when $-1.96 < z < 1.96$, the response is considered as additive; $z > 1.96$ characterizes a super-additive response.

However, we find that the z-score defined in Eq. (6.23) depends on the trial size $n_t = 30$ in their approach. Consider the trials $X_1^E, X_2^E, \ldots, X_{n_t}^E$ drawn from the set E with standard deviation denoted as σ_E. Since an expected response sample $X^L = \frac{1}{n_t}\sum_{j=1}^{n_t} X_j^E$, the standard deviation σ_L of the random sample X^L should in theory conform to

$$\sigma_L = \frac{\sigma_E}{\sqrt{n_t}}. \tag{6.24}$$

From Eq. (6.24), we can see that the larger the trial size n_t, the smaller the standard deviation σ_L, which results in a smaller sample standard deviation S_L for the set L. According to Eq. (6.23), a smaller sample standard deviation S_L leads to larger z-score. In the extreme case that $n_t = 900$, for example, we will have $S_L = 0$, and the z-score goes to infinity. This shows that the z-score proposed by Stanford et al. (2005) depends on the trial size n_t and it thus fails to characterize the intrinsic response properties of DSC neurons.

A better way to process the experimental data would be to use Welch's t test (Welch, 1947) to compare the mean of the set R containing the recorded bimodal responses and that of the set E containing the sum of the unimodal responses. Welch's t test is given by

$$t = \frac{\bar{X}_R - \bar{X}_E}{\sqrt{(S_R^2/n_r) + (S_E^2/n_e)}}, \tag{6.25}$$

where $\bar{X}_R$ and $\bar{X}_E$ are the sample means of the sets R and E respectively, S_R and S_E are the unbiased estimators of their standard deviations, and n_r and n_e are the sizes of the two sets. To use this test, it is required that the sample sizes n_r and n_e should be large enough to ensure that the distribution of the difference $\bar{X}_R - \bar{X}_E$ is approximately normal. In practice, n_r and n_e generally need to be no less than 30. The experimental data recorded by Stanford et al. (2005) satisfy this requirement since they have $n_r = 30$ and $n_e = 900$. The advantages of Welch's t test are obvious. On one hand, it is applicable regardless of the form of distribution in set R and E, and on the other hand, it takes into account the variances of both sets and avoids the dependence on n_t, in contrast to Stanford et al. (2005)'s approach. However, we have not re-analysed the

data since they are not available.

6.5 Correlation between the CME and ADD indices

We have shown that the sigmoidal response function can account for many response properties based on the CME index and the ADD index, both of which were proposed to characterize different aspects of multisensory integration (Meredith and Stein, 1983; King and Palmer, 1985; Avillac et al., 2007). The CME index differentiates responses between response depression (CME<0) and response enhancement (CME>0), whereas the ADD index categorises multisensory responses into sub-additive (ADD<0), additive (ADD≈ 0) and super-additive (ADD>0) responses. To combine the benefits of the two indices, some studies have employed both of them to analyze the experimental data on multisensory responses (Perrault et al., 2003; Avillac et al., 2007). However, we find that the two indices are inherently correlated, and a combination of them does not provide much more information about the characteristics of the response function, or the response properties of multisensory neurons, than using one of them.

Traditionally, the CME index and the ADD index are defined to have asymmetric negative and positive ranges from -100% to infinity [see Eqs. (2.1) and (2.2)]. Consequently, for example, a -50% and a 50% index under these definitions represent different degrees of change in the responses. To overcome this asymmetric range, Avillac et al. (2007) has adapted the traditional definitions of the CME and the ADD indices to normalized forms that ranges between -100% and 100% [see Eqs. (2.3) and (2.4)]. In their study, Avillac et al. (2007) recorded the responses of multisensory neurons in the ventral intraparietal (VIP) area in two behaving monkeys to both visual and tactile stimuli, and they calculated the normalized CME and ADD indices for each set of responses. Plotting the ADD index as a function of the CME index, they showed that these two indices are closely related (Fig. 6.6). For example, we can see that the data points in Fig. 6.6 are confined to a small region and most of the multisensory neurons exhibit both a depression and a sub-additive response. Furthermore, Fig. 6.6 shows that depression is more commonly observed than enhancement, and sub-additivity appears more common than super-additivity or additivity. These results are in contrast to experimental data suggesting that the enhancement and the additivity are most often observed in DSC neurons (Meredith and Stein, 1983; Perrault et al., 2003; Stanford et al., 2005). Avillac et al. (2007) suggested that these differences might be induced by the application of anesthesia in those experiments on DSC neurons.

In fact, the CME and ADD indices are inherently correlated. The correlations between the two indices observed by Avillac et al. (2007) are therefore not determined by the response properties of multisensory neurons in VIP. Consider the expressions for the CME and ADD indices in Eqs. (2.3) and (2.4). Assume that the unimodal response $|V|$

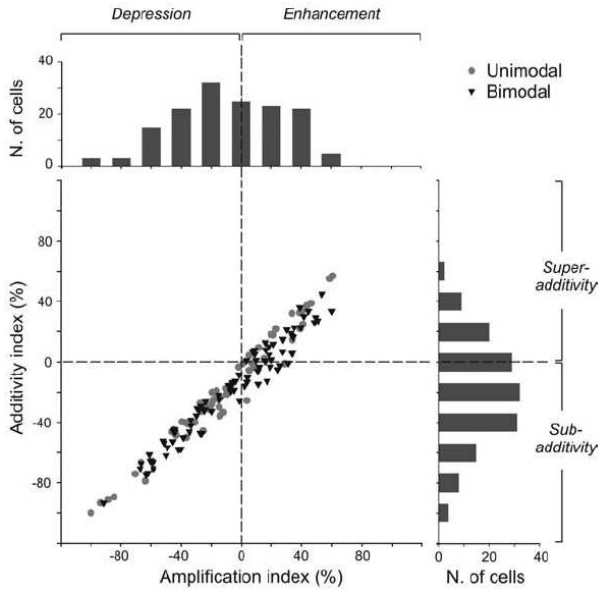


FIGURE 6.6: Distribution of the CME index and the ADD index calculated from the responses of 150 VIP neurons. The "Amplification factor" in the abscissa is equivalent to the CME index and the "Additivity factor" in the ordinate the ADD index. Black triangles represent bimodal neurons (n=87) and grey circles represent unimodal neurons (n=63). On the top and on the right of the plot show the distribution for each the index. From Avillac et al. (2007).

is larger than $|A|$ and let $|A| = a|V|$ where $0 \leq a \leq 1$. Relabeling the CME index as E, we can transform Eq. (2.3) to express the bimodal response $|Bi|$ by E and $|V|$,

$$|Bi| = |V|\frac{1 + (E/100)}{1 - (E/100)}. \tag{6.26}$$

Substituting Eq. (6.26) and $|A| = a|V|$ into Eq. (2.4), we rewrite the ADD index A as a function of E and a,

$$A = \left[\frac{2(E/100) + 2}{(1 - E/100)a + 2} - 1\right] \times 100. \tag{6.27}$$

It is obvious from Eq. (6.27) that A decreases monotonically when a increases. Therefore, for a given E, we have minimum A when $a = 1$ and maximum A when $a = 0$. Substituting $a = 1$ and $a = 0$ into Eq. (6.27), we then obtain

$$\begin{aligned} A_{min} &= \left[\frac{3(E/100) - 1}{3 - (E/100)}\right] \times 100, \\ A_{max} &= E. \end{aligned} \tag{6.28}$$

Accordingly, the additivity index as a function of the CME index is constrained by two curves, one defined by A_{min} as the lower bound and the other defined by A_{max} as the upper bound. In particular, we consider three conditions: $|Bi| < |A|$, $|A| < |Bi| < |V|$ and $|Bi| > |V|$. When $|Bi| < |A|$, according to Eq. (6.26) and $|A| = a|V|$, we directly have $\frac{1+(E/100)}{1-(E/100)} < a < 1$. Therefore, for the unimodal and bimodal responses satisfying

$|Bi| < |A|$, the ADD index as a function of the CME index is in a region I confined by

$$A^{\mathrm{I}}_{min} = \left[\frac{3(E/100)-1}{3-(E/100)}\right] \times 100,$$
$$A^{\mathrm{I}}_{max} = \left[\frac{(E/100)-1}{(E/100)+3}\right] \times 100. \tag{6.29}$$

When $|A| < |Bi| < |V|$, we obtain $a < \frac{1+(E/100)}{1-(E/100)}$ and $E < 0$, and the region II corresponding to this condition is given by

$$A^{\mathrm{II}}_{min} = \left[\frac{(E/100)-1}{(E/100)+3}\right] \times 100,$$
$$A^{\mathrm{II}}_{max} = E, \tag{6.30}$$

and $E < 0$. In the last case that $|Bi| > |V|$, we have $E > 0$, and the corresponding region III is defined by A^{III}_{min} and A^{III}_{max} which are the same as Eq. (6.28),

$$A^{\mathrm{III}}_{min} = \left[\frac{3(E/100)-1}{3-(E/100)}\right] \times 100,$$
$$A^{\mathrm{III}}_{max} = E, \tag{6.31}$$

for $E > 0$.

To illustrate this property, we produced 1000 sets of responses, each of which contains $|Bi|$, $|V|$ and $|A|$ that are randomly selected between 0 and 1, and calculated the CME and the ADD indices for each set. The resulting ADD index plotted as a function of the CME index is shown in Fig. 6.7, in which the data points are all confined in the sector determined by A_{min} and A_{max}, demonstrating the existence of an implicit correlation. Fig. 6.7 also shows the three regions I, II and III confined by Eqs. (6.29), (6.30) and (6.31) respectively. We can see that the correlation between the two indices for VIP neurons (Fig. 6.6) could be perfectly reproduced by using randomly-selected unimodal and bimodal responses (Fig. 6.7). For example, in Fig. 6.7, the depression and the sub-additivity are also more often observed, comparable to Fig. 6.6. Moreover, the sparse data points for CME>0 and ADD>0 in Fig. 6.7 merely reflect the statistical rule that it is less probable for one random variable $|Bi|$ to exceed the value of the larger or the sum of two other random variables $|V|$ and $|A|$. It is thus likely that those analogous properties observed in Fig. 6.6 for VIP neurons also reflect this same statistical rule, and are not caused by the absence of anesthesia.

To sum up, the CME and ADD indices are inherently correlated, and a combination of them to analyze multisensory responses will therefore not be much more informative than using only one of them.

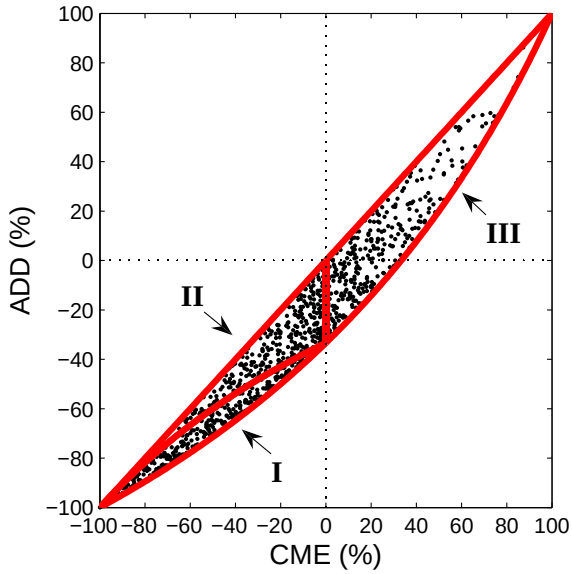


FIGURE 6.7: Distribution of the CME index and the ADD index calculated from 1000 sets of randomly-generated unimodal responses and bimodal response. The I, II and III regions correspond to the cases when $|Bi| < |A|$, $|A| < |Bi| < |V|$ and $|Bi| > |V|$ respectively.

6.6 Different Operational Modes

Perrault et al. (2005) have shown that there appears to be substantial differences in the operations adopted by DSC neurons to integrate cross-modal sensory inputs. In addition to the CME index, they applied the multisensory contrast, which is essentially similar to the ADD index, to study the response properties of DSC neurons, based on which four categories of neurons are identified according to their operational modes. The four categories include the super-additive neurons, the super-additive/sub-additive neurons, the sub-additive neurons and the additive/sub-additive neurons (Fig. 2.7). Although these four operational modes appear to be significantly different from each other, we will show in the following that they can be coherently accounted for by modelling DSC responses as sigmoidal functions.

To show how the sigmoidal response function explains the contrasting operational modes, we first demonstrate the role of a sigmoidal response function in generating the different response ranges between the unimodal and the bimodal cases. Perrault et al. (2005) defined the dynamic range of a neuronal response as the difference between the response threshold and saturation, and showed that the dynamic range of a DSC neuron to bimodal visual-auditory stimuli is larger than to either visual or auditory stimuli (Fig. 2.8). This finding, at first sight, appears to challenge the notion that the CME associated with IE can be explained by the ceiling effect, or equally the saturating responses, of DSC neurons (Perrault et al., 2005). These seemingly contradictory phenomena, however, can be naturally explained by saturated neuronal responses if we take into account response

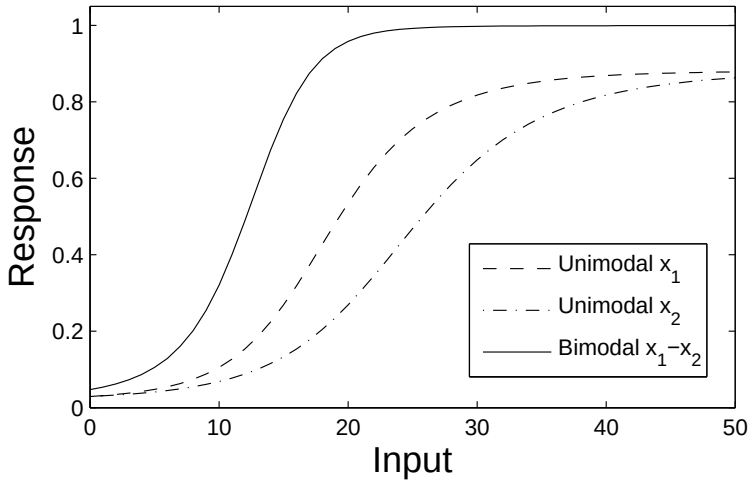
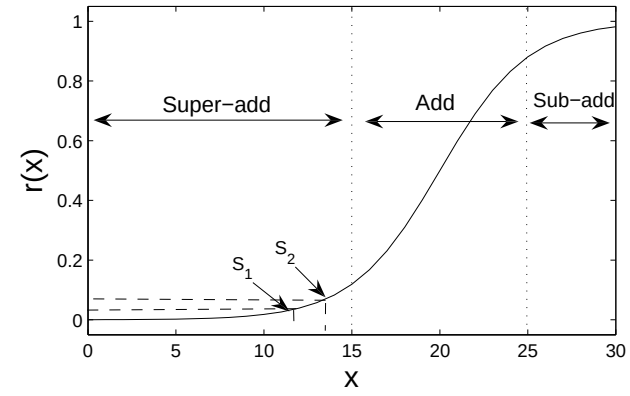


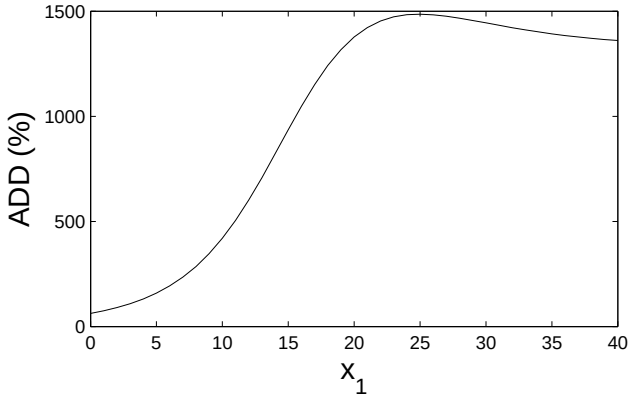
FIGURE 6.8: Different dynamic ranges for unimodal and bimodal responses in a DSC neuron by simulating the inputs to the DSC neuron as saturated outputs from earlier unimodal neurons. The "Input" in the abscissa represents the input x_1 or x_2 to unimodal neurons of which the outputs are the inputs of the DSC neuron, and the "Response" represents the response of the DSC neuron. In sensory channel 1 for x_1, the input to the DSC neuron is generated by a sigmoidal response function with $\gamma_{1i} = 0.6$, $\theta_{1i} = 15$ and $s_{1i} = 15$; In sensory channel 2 for x_2, the input to the DSC neuron is produced under a sigmoidal response function with $\gamma_{2i} = 0.9$, $\theta_{2i} = 20$ and $s_{2i} = 30$. The bimodal DSC neuron responds under a bimodal sigmoidal response function with $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$, and $s = 1$.

saturation in the unimodal neurons earlier along the sensory pathway. In this case, since the inputs to the DSC neurons are the outputs of earlier unimodal sensory neurons, it is possible that the inputs to DSC neurons have already been saturated, in which case the responses of DSC neurons *appear* to be at their saturation levels. However, the presence of a cross-modal input could augment the responses in DSC neurons, provided that the DSC neurons have not actually reached their saturation level. Consequently, we can observe different dynamic ranges for unimodal and bimodal responses. This can be shown in Fig. 6.8, in which input x_1 is passed through a saturated sigmoidal function with $\gamma_{1i} = 0.6$, $\theta_{1i} = 15$ and $s_{1i} = 15$ and input x_2 with $\gamma_{2i} = 0.9$, $\theta_{2i} = 20$ and $s_{2i} = 30$. The response of the bimodal DSC neuron is modeled under a bimodal sigmoidal function with $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 1$, and $s = 1$. Fig. 6.8 shows that the maximal bimodal response is larger than the maximal unimodal responses, comparable to the experimental results in Fig. 2.8.

The four operational modes reported by Perrault et al. (2003) for DSC neurons can also be explained by earlier saturations in the responses of unimodal neurons and response saturations in the DSC neurons. In the four types of neurons, super-additive neurons are those that only exhibit super-additive responses. In terms of the sigmoidal response function, this could be possible when all the inputs situate in the super-additive region of the response function. The constrained range in the inputs could be contributed by previous saturation in the unimdal neurons. As a result, cross-modal stimuli always



A



B

FIGURE 6.9: Illustration for super-additive neurons. (A) The inputs x_1 and x_2 are saturated at S_1 and S_2 respectively and are all in the super-additive region of the response function. (B) The ADD index as a function of x_1 for a super-additive neuron. Input x_1 is passed through a sigmoidal response function with $\gamma_{1i} = 0.6$, $\theta_{1i} = 15$ and $s_{1i} = 15$. Input x_2 is kept at $x_2 = 20$ and is passed through one with $\gamma_{2i} = 0.9$, $\theta_{2i} = 20$ and $s_{2i} = 30$. The super-additive neuron receives saturated x_1 and x_2, and responds based on a bimodal sigmoidal function with $\gamma_1 = 0.1$, $\gamma_2 = 0.05$, $\theta = 2.5$ and $s = 1$.

augment the response in a super-additive way. For demonstration purpose, we only consider the bimodal sigmoidal response function in Eq. (6.5) for $\gamma_1 = \gamma_2 = \gamma$. In this case, we have $\gamma_1 x_1 + \gamma_2 x_2 = \gamma(x_1 + x_2)$, and we can thus reduce the bimodal sigmoidal function to a unimodal one for a combined input $x = x_1 + x_2$. In Fig. 6.9A, inputs x_1 and x_2 saturate at S_1 and S_2 that are in the super-additive region of the sigmoidal response function. Consequently, the combined input $x = x_1 + x_2$ always exhibit super-additive response. Analogous analysis also applies to the bimodal case when $\gamma_1 \neq \gamma_2$. Moreover, an example of the super-additive neuron is given in Fig. 6.9B to plot the ADD index [Eq. (2.2)] as a function of x_1 when x_2 is set constant at 20, and we can see that all the responses are super-additive.

The super-additive/sub-additive neurons exhibit super-additive responses when the effectiveness of the stimulus is weak, and respond in a sub-additive way when the effec-

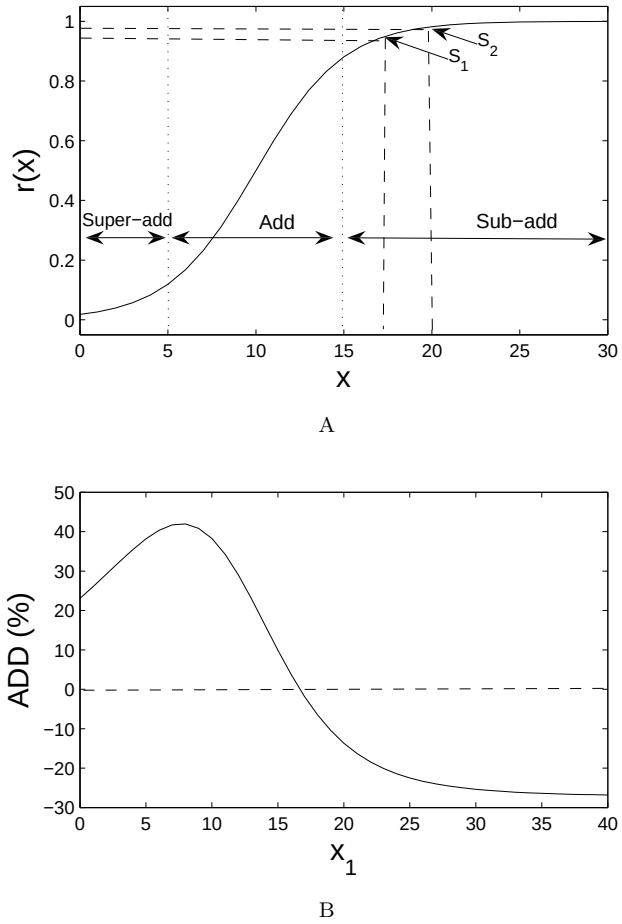


FIGURE 6.10: Illustration for a super-additive/sub-additive neuron. (A) The inputs x_1 and x_2 are saturated at S_1 and S_2 respectively, which span across the super-additive, additive and the sub-additive regions. (B) The ADD index as a function of x_1 for a super-additive/sub-additive neuron. The parameters in this simulation are the same as in Fig. 6.9(B) except that $\theta = 0.8$.

tiveness of the stimulus is strong (Fig. 2.7). In this case, it could be that the inputs can reach both the super-additive and sub-additive regions of the response function of a DSC neuron. Similar to the super-additive case, in Fig. 6.10A, we assume $\gamma_1 = \gamma_2$ to consider only the unimodal sigmoidal function, and we can see that the saturations S_1 and S_2 span across the super-additive, the additive and the sub-additive regions of the response function. Therefore, depending on the strength of the inputs, the DSC neuron could exhibit different types of additive responses. In Fig. 6.10B, we show an example of a simulated super-additive/sub-additive neuron that exhibits both super-additive and sub-additive responses under our model.

In the additive/sub-additive neurons, a transition from the additive to the sub-additive responses is seen when the input increases. To achieve this, their response function might be in a form that the inputs situate in the additive and sub-additive regions. In a unimodal sigmoidal response function, this could be achieved by shifting the response curve leftwards [decreasing θ in Eq. (6.3)]. For example, as shown in Fig. 6.11A, the

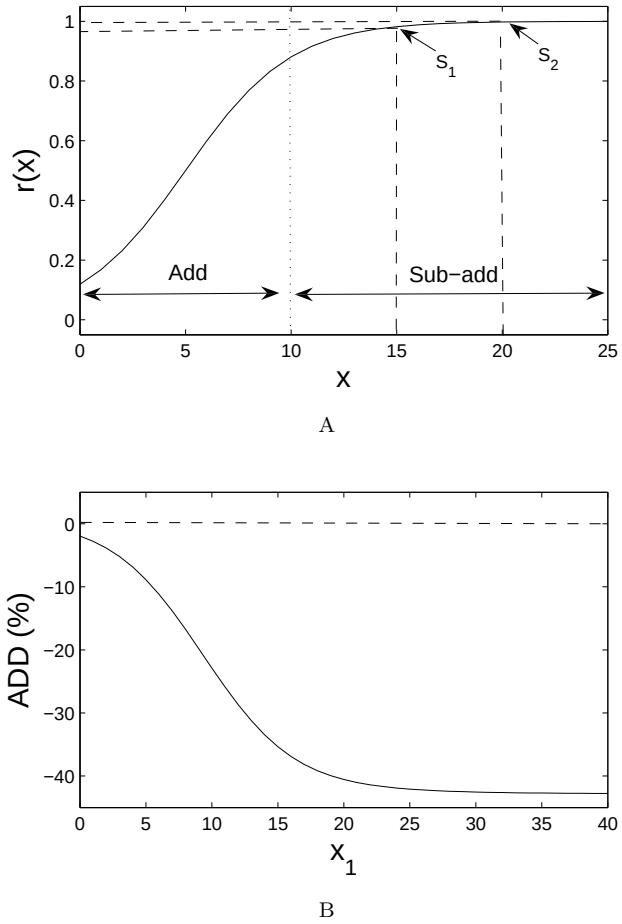


FIGURE 6.11: Illustration for a additive/subadditive neuron. (A) The inputs x_1 and x_2 are saturated at S_1 and S_2 respectively, which span across both the additive and the sub-additive regions. (B) The ADD index as a function of x_1 for a additive/sub-additive neuron. The parameters in this simulation are the same as in Fig. 6.9(B) except that $\theta = 0.4$.

sigmoidal function only leave the additive and sub-additive regions for positive inputs x_1 and x_2 that saturated at S_1 and S_2 respectively. Accordingly, the neuron exhibits either additive or sub-additive responses depending on the magnitudes of inputs. A additive/sub-additive neuron is simulated in Fig. 6.11B, in which the ADD index is plotted as a function of x_1, showing both the additive and sub-additive responses.

Similarly, in the sub-additive neurons, as we can see from Fig. 2.7, relatively weak sensory stimuli induce very strong responses and the neurons always exhibit sub-additive responses. This can be achieved when the sigmoidal response function of the DSC neuron is adjusted in a way that all the inputs situate in its saturation region. Compared with the additive/sub-additive neurons, the sigmoidal response curve need to be shifted further leftwards. Fig. 6.12A shows an illustration of the idea, and we can see that positive inputs x_1 and x_2 always situate in the sub-additive region of the response function. Therefore, the inputs x_1 and x_2 only arouse sub-additive responses. As an example, Fig. 6.12B shows a sub-additive neuron that only exhibit sub-additive responses

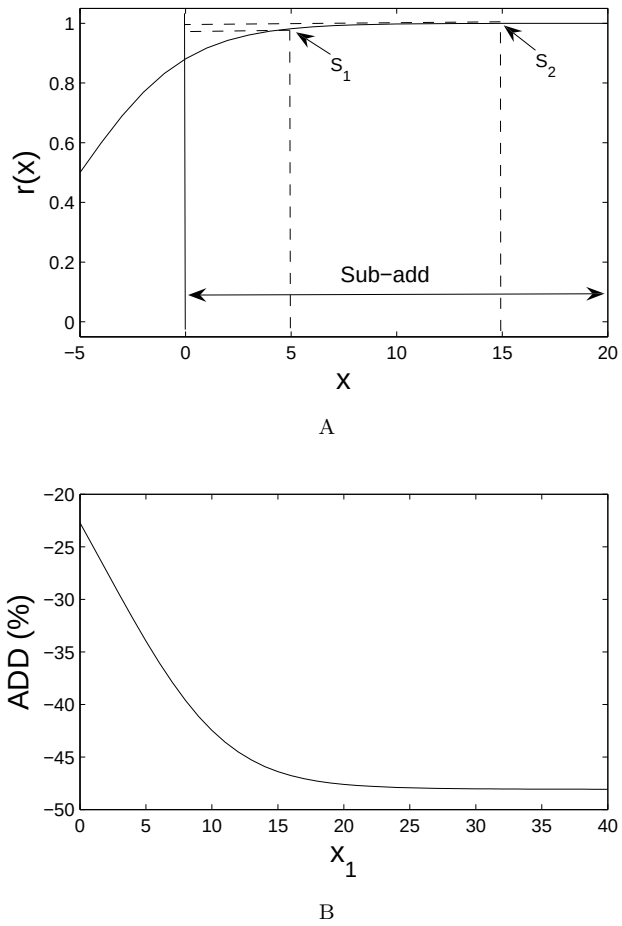


FIGURE 6.12: Illustration for a sub-additive neuron. (A) The inputs x_1 and x_2 are saturated at S_1 and S_2 respectively, which situate in the sub-additive region. (B) The ADD index as a function of x_1 for a sub-additive neuron. The parameters in this simulation are the same as in Fig. 6.9(B) except that $\theta = -0.1$.

for any input x_1.

As shown above, when we model the responses of DSC neurons as sigmoidal functions and take into account possible prior saturation in unimodal neurons earlier along the sensory pathway, the four apparently diverse operational modes naturally emerge. This simple interpretation captures all the dynamics in the four operational modes, suggesting that the complicated response phenomena observed in DSC neurons are direct consequences of the non-linear, saturating response function. We can also see that categorizing DSC neuronal responses to the four operational modes based on the ADD index is not appropriate since they might be produced by a coherent mechanism.

6.7 Summary

Experimental results on DSC neurons have indicated that the mechanisms underlying their responses are complicated. To characterize these mechanisms, some measures such as the CME index and the ADD index have been proposed. There has been considerable evidence showing that these two measures capture the response properties of most DSC neurons. Moreover, different DSC neurons are found to exhibit different operational modes characterized by the ADD index. We argued in this chapter, however, that all these diverse response properties can be accounted for by a non-linear, saturating response function, extending the notion of Stanford et al. (2005). We adopted a sigmoidal response function to show how such a response function can reproduce the reported response properties and unify the different operations observed in DSC neurons.

Chapter 7

Adaptation in Sensory Neurons for Maintaining the Operating Point

In the previous Chapter, we showed that a multisensory neuron under a sigmoidal response function is capable of exhibiting a wide range of response properties observed in DSC neurons. One remaining question is how the parameters in the sigmoidal function, such as the gain and the threshold, are determined. Inspired by experimental results on threshold adaptation and gain control in unimodal neurons, we propose an adaptation rule for unimodal neurons, and extend the rule for multimodal neurons receiving inputs from cross-modal sensory channels. This chapter is based mainly on our paper "An Invariance Principle for Maintaining the Operating Point of a Neuron" (Elliott et al., 2008b).

7.1 Introduction

We have described in Chapter 3 how neurons early in sensory pathways are believed to adapt their responses to the statistics of their inputs in order to maximise their coding efficiency, output entropy, or information rate (Attneave, 1954; Barlow, 1961; Laughlin, 1981; Atick, 1992; van Hatteren, 1992; DeWeese, 1996; Dan et al., 1996; Baddeley et al., 1997; Smirnakis et al., 1997; Wainwright, 1999; Brenner et al., 2000; Fairhall et al., 2001; Maravall et al., 2007). Other adaptive strategies have also been proposed for neurons later in sensory pathways (see, e.g., Carandini and Ferster (2000); Pena and Konishi (2002); Ringach and Malone (2007)). Despite the diversity of functional roles of neurons, and however their preferred operating points may be established over evolutionary or developmental timescales, it is possible that there exist adaptive principles, based on

input statistics, that enable neurons to maintain their preferred operating points without explicit reference to their functional roles.

We suggest one such principle. We define the operating point of a neuron by the cumulative distribution function (CDF) or probability density function (PDF) of its output spike rate. The operating point therefore embodies information about both the input-output transfer function of the neuron and the statistics of the environment from which the input is drawn. The operating point is not simply the transfer function, nor simply the environment, but both, united in the PDF of the neuron's output spike rate. We propose, in particular, the adaptive principle that a neuron adapts its transfer function in order to keep its output PDF invariant, or as invariant as possible, under changes in its input statistics. Such invariance would ensure that a neuron remains at its preferred operating point, regardless of how that point is set. One consequence of this view is that when a neuron's operating point is set by a principle such as maximum information rate, if a neuron can maintain an invariant output PDF, then it automatically remains at the point determined by that principle.

In this chapter, we develop a model of neuronal adaptation in which adaptation ensures that a neuron's output PDF remains invariant, or approximately invariant, under changes in its input statistics. We consider only sigmoidal transfer functions since it can reproduce most of the response properties in DSC neurons (see Chapter 6). We initially consider a unimodal neuron with input from a single sensory channel, and derive rules for threshold adaptation and gain control that seek to maintain an invariant output PDF. These rules are independent of how the preferred operating point of the neuron is set, but we also consider how the operating point may be set by a maximum entropy principle. The structure of our model permits generalisation to multimodal neurons receiving input from multiple, distinct sensory channels, and we extend our rules to this multimodal case. This extension necessitates the introduction of separate gains for each modality. Having developed the underlying approach to adaptation, we present several examples of a model neuron functioning at different preferred operating points and adapting to changing input statistics, both unimodal and bimodal.

7.2 An Invariance Principle for Adaptation

We first consider a neuron with an input from only a single sensory channel. Then we extend our approach to a neuron with inputs from multiple, distinct sensory channels.

7.2.1 Unimodal Inputs

Suppose that a neuron's unimodal input is determined by the (univariate) random variable X with PDF $f_X(x)$ and CDF $F_X(x)$, where X, without loss of generality, also may

absorb any input noise processes. Under the assumption of certain regularity conditions, X is uniquely determined by its moments, so that its moment generating function, $M_X(t)$, exists and its characteristic function, $\phi_X(t)$, is analytic on the real line (Feller, 1971).[1] We denote the moments of X by m_i, with mean $\mu = m_1$, and the central moments by μ_i, with variance $\sigma^2 = \mu_2$. We denote the transfer or output function of the neuron by $r = r(x)$ for input x drawn from the distribution X. The output is then a random variable, denoted by R, and its CDF and PDF are given by

$$F_R(r) = F_X(x(r)), \tag{7.1}$$

$$f_R(r) = \frac{dx(r)}{dr} f_X(x(r)), \tag{7.2}$$

respectively, where $x(r)$ is the inverse of $r(x)$, i.e. $x(r(x')) = x'$, and conforms to monotonicity. If the response is bounded, so that $r \in [0, s]$, and if $r(x) = sF_X(x)$, then $f_R(r) = 1/s$, so that R is uniform on $[0, s]$, and hence has maximum entropy, corresponding to Laughlin's result (Laughlin, 1981).

To model the neuronal responses, we use the sigmoidal response function,

$$r(x) = \frac{s}{2}\left[1 + \tanh\frac{2\gamma}{s}(x - \theta)\right], \tag{7.3}$$

as given in Chapter 6. We shall regard the saturation value s as fixed, since it is determined by the maximum firing rate of a neuron, which is limited by a neuron's refractory period. The sigmoidal response function thus endows a neuron with two parameters, the threshold and the gain, that it can change in an attempt to keep a neuron's output PDF invariant, or as invariant as possible. Clearly, then, for an arbitrary input distribution X, perfect adaptation is in general impossible, since a sigmoidal response $r(x)$ lacks the potentially infinite number of degrees of freedom required to adapt to the potentially infinite number of independent degrees of freedom in all the moments, m_i, of X. Furthermore, R cannot adopt the uniform, maximum entropy distribution, given the upper bound s on the response, except in the particular case that X obeys the logistic distribution, with CDF $F_X(x) = \frac{1}{2}\left[1 + \tanh\frac{2\gamma}{s}(x - \theta)\right]$, satisfying Laughlin's condition, $r(x) = sF_X(x)$. The choice of a two-parameter response function thus imposes limitations on the extent to which a model neuron can adapt to its input statistics. Although this may appear unsatisfactory, of course any finite-parameter response function will be so limited. Unless we hold to the view that a neuron's response function may be set arbitrarily, which seems unlikely, then such limitations are unavoidable. We could improve the adaptability of a model neuron by considering more complicated, multi-parameter response functions, but we have selected the two-parameter, sigmoidal function precisely because of its simplicity, so that the underlying approach, of adaptation to input

[1] Carleman's condition on the moments is sufficient to ensure uniqueness (Feller, 1971). The classic counter-examples to this condition are the Cauchy distribution, which has no moments, and the log-normal distribution, for which $M_X(t)$ does not exist despite all the moments existing.

statistics, is not obscured by unnecessary complexity.

Suppose that a neuron's initial response is defined by the parameter set $\{\theta, \gamma\}$, giving rise to the random output variable R with PDF $f_R(r)$ for the initial random input variable X. Suppose that the input statistics then change, giving a new input random variable X', inducing a new response parameter set $\{\theta', \gamma'\}$ and output random variable R' with PDF $f_{R'}(r)$. Since we wish to make the output PDF as invariant as possible under the change $X \to X'$, an obvious strategy would be to define a functional metric or distance $\mathcal{D}$ and explicitly find θ' and γ' such that $\mathcal{D}(f_R, f_{R'})$ is minimised. A popular candidate for $\mathcal{D}$ would be the relative entropy or Kullback-Leibler divergence between f_R and $f_{R'}$. Although explicit minimisation of $\mathcal{D}(f_R, f_{R'})$ is, mathematically speaking, a natural approach, it is potentially very expensive, computationally speaking. We prefer, therefore, to seek an alternative approach to determining θ' and γ', one that may be computationally simpler for a neuron to implement, but accept that it may not be statistically optimal under all circumstances.

Since our model neuron possesses only two degrees of freedom in θ and γ, it is reasonable to assume that these two parameters are modified in order to accommodate the most significant variations in X. Defining the Z-score as usual by $Z = (X - \mu)/\sigma$, the Chebyshev inequality,

$$P[|Z| \geq k] \leq \frac{1}{k^2}, \tag{7.4}$$

where k is any positive number, reveals a general property of the Z-score independent of X. That is, despite the distribution of X, its transformation into the Z-score based on the mean and standard deviation will nonetheless convert X to a category of distribution specified by Eq. (7.4). This suggests that, potentially, the lowest two moments are the most influential among all the moments of X. For an adaptive strategy that is applicable to an arbitrary distribution of X, they accordingly represent good targets for adaptive processes. We therefore propose that the parameters θ and γ should thus be adapted to accommodate changes in the mean, μ (Barlow and Mollon, 1982) and the standard deviation, σ (Meister and Berry, 1999; Smirnakis et al., 1997; Brenner et al., 2000; Fairhall et al., 2001), with the higher-order moments $m_3, \ldots$ playing second fiddle to m_1 and m_2 (Bonin et al. (2006); but see Kvale and Schreiner (2004)). More degrees of freedom in the response function would allow adaptation to moments higher than the second.

Consider an input PDF of the particular form

$$f_X(x) = \frac{1}{\sigma}\, g\left(\frac{x - \mu}{\sigma}\right), \tag{7.5}$$

where, by this, we mean that the dependence of f_X on μ and σ appears *only* through the combination $(x - \mu)/\sigma$ in f_X, together with the overall scale factor of $1/\sigma$, and that g does not depend on any other parameters. The normal distribution is an example of

such a PDF. Since

$$x(r) = \theta - \frac{s}{4\gamma} \log\left(\frac{s-r}{r}\right), \tag{7.6}$$

and

$$\frac{dx(r)}{dr} = \frac{1}{4\gamma} \frac{s^2}{r(s-r)}, \tag{7.7}$$

we then have

$$f_R(r) = \frac{1}{4\gamma\sigma} \frac{s^2}{r(s-r)} g\left(\frac{\theta-\mu}{\sigma} - \frac{s}{4\gamma\sigma} \log\left(\frac{s-r}{r}\right)\right). \tag{7.8}$$

We see immediately that $f_R(r) = f_{R'}(r)\ \forall r \in [0, s]$, and hence $\mathcal{D}(f_R, f_{R'}) \equiv 0$ for any metric $\mathcal{D}$, provided that

$$\frac{\theta-\mu}{\sigma} = \frac{\theta'-\mu'}{\sigma'}, \tag{7.9}$$

$$\gamma\sigma = \gamma'\sigma'. \tag{7.10}$$

Eq. (7.9) implies that a neuron adjusts the threshold θ to keep the z-score of θ invariant with respect to the input statistics. This ensures that a neuron's mean output is independent of its mean input, where this mean output is determined by a neuron's preferred operating point. Intuitively, Eq. (7.9) shifts the response function according to the average intensity of the inputs, so that a wide range of input signals can be effectively represented by the limited response range of a neuron. Eq. (7.10) then sets the output gain in inverse proportion to the input standard deviation, so that larger (smaller) standard deviations correspond to smaller (larger) gains. This ensures that the dynamic range of a neuron's input is mapped onto an invariant dynamic range of its output, where this output dynamic range is again set by a neuron's preferred operating point, keeping the output dynamic range constant. As the gain is inversely proportional to the standard deviation, the sensitivity of a neuron is increased to improve the contrast for inputs with small variations and decreased to avoid response saturation for inputs with large variations. This relationship between the gain of a neuron and the standard deviation of the input is a well-established experimental observation (see, e.g., Kvale and Schreiner (2004); Bonin et al. (2006); Maravall et al. (2007); and references therein), and is in effect analogous to the histogram equalization method for contrast adjustment.

Many standard distributions may be written in the form defined by Eq. (7.5), including the normal distribution, the exponential, Laplace, and doubly exponential distributions, and the logistic distribution, and in general an infinity of forms for the function g is available. For this class of input distributions, if a neuron sets θ and γ according to

$$\theta = \mu + \Theta\sigma, \tag{7.11}$$

$$\gamma = \Gamma\frac{s}{\sigma}, \tag{7.12}$$

where the constants Θ and Γ determine the preferred operating point of a neuron, then

a neuron remains at this operating point, with $f_R(r)$ exactly invariant in response to changes in the mean and variance of the input distribution X. We have made explicit a factor of s in Eq. (7.12) since scaling the response range s will scale the response gain γ by the same factor. With s made explicit, Γ will then be independent of s and scale-free.

For a completely general form of input distribution $f_X(x)$, Eqs. (7.11) and (7.12) will not suffice to maintain response PDF invariance. However, we propose that a neuron nonetheless sets θ and γ according to Eqs. (7.11) and (7.12). First, these rules are simple to implement. Second, they result in exact invariance for a large class of input distributions including the normal distribution. Third, they ensure adaptation to the lowest-order moments, μ and σ, of a distribution, albeit at the price of possible non-invariance due to the higher-order moments of completely general distributions. We might expect, in any event, that a neuron should not adapt perfectly to higher-order moments (but see Kvale and Schreiner (2004)), since such moments contain information of importance to learning algorithms such as independent component analysis (Hyvärinen et al., 2001).

The confounding influence of the higher-order moments $m_3, \ldots$ on the invariance of $f_R(r)$ in the presence of a general input distribution $f_X(x)$ can be confirmed by transforming to the Z variable,

$$\begin{aligned} f_R(r) &= \frac{dz(r)}{dr} f_Z(z(r)) \\ &= \frac{1}{4\gamma\sigma} \frac{s^2}{r(s-r)} f_Z\left(\frac{\theta-\mu}{\sigma} - \frac{s}{4\gamma\sigma}\log\left(\frac{s-r}{r}\right)\right). \end{aligned} \tag{7.13}$$

Of course, Eq. (7.13) is identical to Eq. (7.8), with the PDF of the Z-transformed input f_Z replacing g. However, for a general form of f_Z, Eq. (7.13) is *not* invariant under Eqs. (7.11) and (7.12) because the higher-order moments of Z are not invariant under changes in μ and σ. To see this, we write $f_Z(z)$ to exhibit explicitly all its moments,

$$f_Z(z) = f_Z(z; 0, 1, \tilde{m}_3, \tilde{m}_4, \ldots), \tag{7.14}$$

where 0 and 1 are the first and second moments of Z, by definition, and the ith moment of Z, $\tilde{m}_i$, is just the ith central moment of X divided by σ^i, $\tilde{m}_i = \mu_i/\sigma^i$. Making explicit the dependence of f_R on θ and γ, we then may write

$$f_R(r; \theta, \gamma) = \frac{1}{4\gamma\sigma} \frac{s^2}{r(s-r)} f_Z\left(\frac{\theta-\mu}{\sigma} - \frac{s}{4\gamma\sigma}\log\left(\frac{s-r}{r}\right); 0, 1, \tilde{m}_3, \ldots\right), \tag{7.15}$$

and under a change of statistics, $\mu \to \mu'$, $\sigma \to \sigma'$, $\tilde{m}_i \to \tilde{m}'_i$ and adaptive changes $\theta \to \theta'$, $\gamma \to \gamma'$, we have

$$f_{R'}(r; \theta', \gamma') = \frac{1}{4\gamma'\sigma'} \frac{s^2}{r(s-r)} f_Z\left(\frac{\theta'-\mu'}{\sigma'} - \frac{s}{4\gamma'\sigma'}\log\left(\frac{s-r}{r}\right); 0, 1, \tilde{m}'_3, \ldots\right). \tag{7.16}$$

Now, if a neuron implements Eqs. (7.11) and (7.12) in an attempt to maintain an invariant output PDF, we have

$$f_R(r;\theta,\gamma) = \frac{1}{4\Gamma}\frac{s}{r(s-r)} f_Z\left(\Theta - \frac{1}{4\Gamma}\log\left(\frac{s-r}{r}\right);0,1,\tilde{m}_3,\ldots\right), \quad (7.17)$$

$$f_{R'}(r;\theta',\gamma') = \frac{1}{4\Gamma}\frac{s}{r(s-r)} f_Z\left(\Theta - \frac{1}{4\Gamma}\log\left(\frac{s-r}{r}\right);0,1,\tilde{m}'_3,\ldots\right), \quad (7.18)$$

where Θ and Γ are, as above, constants, so that

$$f_R(r;\theta,\gamma) - f_{R'}(r;\theta',\gamma') = \frac{1}{4\Gamma}\frac{s}{r(s-r)}\sum_{i=3}^{\infty}\left(\tilde{m}_i - \tilde{m}'_i\right)\frac{\partial f_Z}{\partial \tilde{m}_i}\left(\Theta - \frac{1}{4\Gamma}\log\left(\frac{s-r}{r}\right);0,1,\xi_3,\xi_4,\ldots\right), \quad (7.19)$$

for some values of $\xi_i \in (\tilde{m}_i, \tilde{m}'_i)$. Hence, the attempt to maintain the invariance of f_R is contaminated by the possible changes in the higher-order moments of the Z-transformed input distribution.

This contamination may arise from a number of sources. For a general distribution f_Z, the higher-order moments $\tilde{m}_3, \ldots$ will in general change when μ and σ are changed. Although Eqs. (7.11) and (7.12) attempt to accommodate changes in μ and σ, these changes will leak into the higher-order moments. However, for the specific form of distribution $f_Z(z) = g(z)$ discussed above, the higher-order moments $\tilde{m}_3, \ldots$ are independent of μ and σ and hence are constants. Thus, for the form $f_Z = g$, the invariance of f_R is exact precisely because the higher-order moments are, by construction, also invariant. Another source of "contamination" could arise from distributions in which some of the higher-order moments are independent degrees of freedom that may be freely changed in an attempt to probe a neuron's ability to adapt to higher-order moments (Kvale and Schreiner, 2004; Bonin et al., 2006). In this case, of course, the non-invariance of f_R is explicit and direct, rather than implicit and indirect.

Despite the breakdown in exact invariance of $f_R(r)$ in the case of a completely general input distribution $f_X(x)$, we see from Eq. (7.19) that the magnitude of the change in the output PDF is controlled by the magnitude of the change in the higher-order moments. Unless an input distribution exhibits a high degree of sensitivity to its higher-order moments, we would thus not expect the non-invariance of f_R to be too severe. Of course, it is always possible to construct counter-examples to such arguments, but we might expect naturally-occurring input statistics to be reasonably well-behaved (Simoncelli and Olshausen, 2001).

So far, we have been concerned with how a neuron maintains its preferred operating point in the face of changing input statistics by adapting its threshold θ and gain γ so as to keep its output PDF f_R invariant, or approximately invariant. In our approach,

the operating point of a neuron is set by the two constants Θ and Γ in Eqs. (7.11) and (7.12). We have thus not been concerned with how Θ and Γ are set by a neuron. We now consider the possibility that Θ and Γ are set so as to maximise the entropy of the output distribution R, given the sigmoidal transfer function in Eq. (7.3).

The output entropy $S[R]$ is defined by

$$S[R] = -\int dr f_R(r) \log f_R(r), \tag{7.20}$$

from which we have

$$S[R] = -\int dx f_X(x) \log \left[\frac{dx}{dr(x)} f_X(x) \right]. \tag{7.21}$$

Writing

$$\gamma(x) = \frac{dr(x)}{dx}, \tag{7.22}$$

the instantaneous gain of the neuronal response $r(x)$ for input x, with $\gamma(\theta) \equiv \gamma$ at semi-saturation, we then have

$$S[R] = S[X] + \langle \log \gamma(x) \rangle_X , \tag{7.23}$$

where $S[X]$ is the entropy of the input distribution and $\langle\rangle_X$ means an average over the distribution X. If the response function is characterised by a set of adjustable parameters p_i, so that $r(x) = r(x; p_i)$, then we maximise the entropy $S[R]$ with respect to these parameters by evaluating the derivatives

$$\frac{\partial S[R]}{\partial p_i} = \left\langle \frac{1}{\gamma(x)} \frac{\partial \gamma}{\partial p_i}(x; p_j) \right\rangle_X \tag{7.24}$$

and setting them to zero. The derivative $\partial\gamma(x; p_j)/\partial p_i$ is the sensitivity of the instantaneous gain of the response function to the parameter p_i. For the sigmoidal function in Eq. (7.3), we have

$$\gamma(x) = \gamma \operatorname{sech}^2 \frac{2\gamma}{s}(x - \theta), \tag{7.25}$$

depending on the two parameters θ and γ, from which we obtain the two conditions

$$\frac{\partial S[R]}{\partial \theta} = \frac{4\gamma}{s} \int dx f_X(x) \tanh \frac{2\gamma}{s}(x - \theta) = 0, \tag{7.26}$$

$$\frac{\partial S[R]}{\partial \gamma} = \frac{1}{\gamma} - \frac{4}{s} \int dx\, x\, f_X(x) \tanh \frac{2\gamma}{s}(x - \theta) = 0, \tag{7.27}$$

and hence values of θ and γ that maximise $S[R]$. For the given values of μ and σ associated with input X, we can then fix Θ and Γ from Eqs. (7.11) and (7.12). Of course, in principle adaptation could take the form of determining θ and γ directly from Eqs. (7.26) and (7.27), and thus of always being optimal, in the sense of achieving maximum output entropy. However, if Θ and Γ are fixed once, perhaps over evolutionary or developmen-

tal timescales, Eqs. (7.11) and (7.12) then afford an arguably computationally easier method of remaining at, or near, maximum output entropy.

In the foregoing, we have allowed the possibility that X may contain noise from the environment, from signal transduction processes, and from neuronal transmission, but we have not considered the possibility that the output PDF could be contaminated by noise. In fact, under a simple additive model of output noise, in which the actual output distribution R is the sum of the noiseless output R_0 and a noise source N, $R = R_0 + N$, so that

$$\begin{aligned} f_R(r) &= \int dr_0 f_{R_0}(r_0) f_N(r - r_0) \\ &= \int dx f_X(x) f_N(r - r(x)), \end{aligned} \tag{7.28}$$

where f_N is the PDF of the noise process, it is a simple matter to show that the rules in Eqs. (7.11) and (7.12) follow directly after some transformations under the integral sign. For example, if the input PDF $f_X(x)$ has the particular form in Eq. (7.5), we will have $f_{R_0}(r_0)$ following Eq. (7.8). Substituting the expression of $f_{R_0}(r_0)$ into Eq. (7.28), we can write $f_R(r)$ as

$$f_R(r) = \int dr_0 \; \frac{1}{4\gamma\sigma} \frac{s^2}{r_0(s - r_0)} \, g\left(\frac{\theta - \mu}{\sigma} - \frac{s}{4\gamma\sigma} \log\left(\frac{s - r_0}{r_0} \right) \right) f_N(r - r_0), \tag{7.29}$$

from which it is obvious that the same rules in Eqs. (7.11) and (7.12) will maintain invariant $f_R(r)$. In the case of a general input distribution $f_X(x)$, we can write the PDF of noiseless output as $f_{R_0}(r_0; \theta, \gamma)$ and transform it to the Z variable [see Eq. (7.15)]. If a neuron implements the rules in Eqs. (7.11) and (7.12), we have $f_{R_0}(r_0; \theta, \gamma)$ following Eq. (7.17), and thus

$$f_R(r; \theta, \gamma) = \int dr_0 \; \frac{1}{4\Gamma} \frac{s}{r_0(s - r_0)} f_Z\left(\Theta - \frac{1}{4\Gamma} \log\left(\frac{s - r_0}{r_0} \right); 0, 1, \tilde{m}_3, \ldots \right) f_N(r - r_0), \tag{7.30}$$

where Θ and Γ are constants. Under a change in the statistics, $\mu \to \mu'$, $\sigma \to \sigma'$, $\tilde{m}_i \to \tilde{m}_i'$, the neuron adapts the threshold and gamma $\theta \to \theta'$, $\gamma \to \gamma'$ to maintain an (approximately) invariant output PDF,

$$f_{R'}(r; \theta', \gamma') = \int dr_0 \; \frac{1}{4\Gamma} \frac{s}{r_0(s - r_0)} f_Z\left(\Theta - \frac{1}{4\Gamma} \log\left(\frac{s - r_0}{r_0} \right); 0, 1, \tilde{m}_3', \ldots \right) f_N(r - r_0). \tag{7.31}$$

Accordingly, we have

$$\begin{aligned} & f_R(r;\theta,\gamma) - f_{R'}(r;\theta',\gamma') \\ &= \int dr_0 \, \frac{1}{4\Gamma} \frac{s}{r_0(s-r_0)} \\ &\quad \sum_{i=3}^{\infty} \left(\tilde{m}_i - \tilde{m}_i'\right) \frac{\partial f_Z}{\partial \tilde{m}_i} \left(\Theta - \frac{1}{4\Gamma} \log\left(\frac{s-r_0}{r_0}\right); 0, 1, \xi_3, \xi_4, \ldots\right) f_N(r-r_0), \end{aligned} \tag{7.32}$$

for some values of $\xi_i \in (\tilde{m}_i, \tilde{m}_i')$, analogous to the noiseless case. On condition that the input distribution are not highly sensitive to the higher-order moments, the neuron would maintain an approximately invariant operating point. The adaptation rules in Eqs. (7.11) and (7.12) are therefore also valid under the assumption of additive output noise.

This additive model is a very simplified form of noisy response process. In general, other forms are possible. For example, the noise term N might be a function of the mean of the noiseless output R_0. We have not considered these more general forms of noisy responses.

7.2.2 Multimodal Inputs

Many neurons, both subcortical and cortical, receive input not from just one sensory modality, but often from two or three distinct sensory modalities. For example, neurons in the deep layers of the superior colliculus can receive and integrate visual, auditory and somatosensory input. Each modality will be associated with its own intrinsic statistical parameters, and these may vary independently of those in other modalities. Moreover, different input sources may exhibit time-dependent correlations. A multimodal neuron may therefore be expected to exhibit adaptation to all its various input sources. It is natural, then, to consider extending the principles of adaptation in unimodal neurons to multimodal neurons.

We therefore now turn to the case in which a neuron receives input from at least two different sensory channels. For simplicity, we restrict to bimodal neurons, but our results generalise to the full, multimodal case. We therefore consider two input channel distributions, X_1 and X_2 with PDFs $f_{X_1}(x_1)$ and $f_{X_2}(x_2)$ and means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. The correlation coefficient between the two channels X_1 and X_2 is defined to be ρ. In general, of course, a joint PDF $f_{X_1X_2}(x_1, x_2)$ defines the joint input distribution, with $f_{X_1}(x_1)$ and $f_{X_2}(x_2)$ being the marginal distributions.

For a unimodal input, we defined the response function in Eq. (7.3) so that the argument of the tanh function is, up to factors, just $\gamma(x-\theta)$. We then derived rules for adapting θ and γ to the input statistics based on the invariance of the output PDF $f_R(r)$. We

have not discussed how a neuron instantiates its threshold θ and gain γ, and we have also not discussed the implementation mechanism leading to adaptive changes in θ and γ. Many mechanisms are implicated in adaptation to input statistics, both at the single neuron level and the circuit level (Sanchez-Vives et al., 2000b; Rieke, 2001; Baccus and Meister, 2002; Chance et al., 2002; Kim and Rieke, 2003; Shu et al., 2003; Dean et al., 2005; Ingham and McAlpine, 2005; Arganda et al., 2007). If adaptation occurs at the single neuron level, then information about the stimulus mean and standard deviation must be available locally. If adaptation to stimulus statistics occurs for a neuron with multimodal inputs, then locality demands that the statistics pertaining to a single input channel are available only at that channel's synapses onto the neuron, or at least at the local dendritic level, rather than the whole neuron level, at which presumably all the separate statistics for the individual input channels become merged and therefore lost.

Since we now wish to discuss the possibility of adaptation of a multimodal neuron to the separate statistics of its different input channels, we will therefore write the response function $r(x_1, x_2)$, for a bimodal neuron, in the form

$$r(x_1, x_2) = \frac{s}{2}\left\{1 + \tanh\frac{2}{s}\left[\gamma_1(x_1 - \theta_1) + \gamma_2(x_2 - \theta_2)\right]\right\}, \tag{7.33}$$

where the "thresholds" θ_1 and θ_2 and the "gains" γ_1 and γ_2 are now specific to each input distribution X_1 and X_2 and permit adaptation according to only locally-available information about each input. Of course, the θ_i and γ_i are no longer neuronal thresholds and gains in the conventional, unimodal sense, but we retain this nomenclature because of the clear analogy with the unimodal case.

Writing $U = \gamma_1 X_1 + \gamma_2 X_2$, $\gamma_U \equiv 1$ and $\theta_U = \gamma_1\theta_1 + \gamma_2\theta_2$, the bimodal response function in Eq. (7.33) can be reduced to an effective unimodal response function,

$$r(u) = \frac{s}{2}\left[1 + \tanh\frac{2\gamma_U}{s}(u - \theta_U)\right], \tag{7.34}$$

where the effective neuronal input u takes values from the effective input distribution U. The mean and variance of U are given by

$$\mu_U = \gamma_1\mu_1 + \gamma_2\mu_2, \tag{7.35}$$

$$\sigma_U^2 = \gamma_1^2\sigma_1^2 + \gamma_2^2\sigma_2^2 + 2\gamma_1\sigma_1\gamma_2\sigma_2\rho, \tag{7.36}$$

respectively, where the correlation coefficient ρ appears in the expression for σ_U^2. The output PDF $f_R(r)$ is therefore given by

$$f_R(r) = \frac{du(r)}{dr} f_U(u(r)), \tag{7.37}$$

where $u(r)$ is the inverse function of $r(u)$, given by an equation analogous to Eq. (7.6) under the replacements $\theta \to \theta_U$ and $\gamma \to \gamma_U$, and $f_U(u)$ is the PDF of the effective

unimodal input U, determined from the joint PDF $f_{X_1X_2}(x_1, x_2)$. Since $\gamma_U \equiv 1$, $f_R(r)$ can be written simply as

$$f_R(r) = \frac{s^2}{4r(s-r)} f_U\left(\theta_U - \frac{s}{4}\log\frac{s-r}{r}\right). \tag{7.38}$$

This reduction allows us to write down, by analogy, the rules according to which the separate channel thresholds and gains should be set to ensure the (perhaps approximate) invariance of $f_R(r)$.

Adapting to the effective, unimodal input distribution U, and hence to the actual, bimodal input distributions X_1 and X_2 is then achieved, according to our earlier rules in Eqs. (7.11) and (7.12), by setting

$$\theta_U = \mu_U + \Theta\sigma_U, \tag{7.39}$$

$$\gamma_U = \Gamma\frac{s}{\sigma_U}, \tag{7.40}$$

where the constants Θ and Γ determine, as usual, the preferred operating point of a neuron. We therefore have

$$\gamma_1(\theta_1 - \mu_1) + \gamma_2(\theta_2 - \mu_2) = \Theta\Gamma s, \tag{7.41}$$

$$\gamma_1^2\sigma_1^2 + \gamma_2^2\sigma_2^2 + 2\gamma_1\sigma_1\gamma_2\sigma_2\rho = \Gamma^2 s^2. \tag{7.42}$$

Eq. (7.42) defines, in general, an ellipse (or ellipsoid for the multimodal case) on which solutions γ_1 and γ_2 of this equation exist. The semi-major and semi-minor axes of this ellipse are set by σ_1 and σ_2, as well as by ρ. Thus, a general point on the ellipse will be influenced by both σ_1 and σ_2, and hence the solution for γ_1, say, will be influenced by σ_2. However, we have argued that information about the σ_i should only be available locally, at or near the site of the synapses associated with the inputs X_i. Hence, the gain associated with, say, input X_1, γ_1, should not be influenced, non-locally, by the statistics of X_2. We can achieve this by insisting that

$$\gamma_i\sigma_i = \Gamma_i s, \tag{7.43}$$

where the Γ_i are constants and, as usual, an overall scale is made explicit. In order to ensure that each modality is mapped onto the same output dynamic range, we set all these constants equal, so that $\Gamma_i = \Gamma'$, $\forall i$. Then, from Eq. (7.42), we have $\Gamma' = \Gamma/\sqrt{2(1+\rho)}$. Thus, the separate channel gains should be set according to

$$\gamma_i = \frac{\Gamma}{\sqrt{2(1+\rho)}}\frac{s}{\sigma_i}, \tag{7.44}$$

providing solutions of Eq. (7.42) satisfying the principle of locality. The channel-specific gains γ_i are therefore fixed locally, by the factor of $1/\sigma_i$, but the overall scaling also

depends on the correlation coefficient ρ. For the possible influence of correlations on gain control, although not in a multimodal context, see Sharpee et al. (2006) and Lesica et al. (2007). For the general, multimodal case, with n distinct channels, a sum of all $\frac{1}{2}n(n-1)$ separate correlation coefficients between all distinct pairs of channels appears in the denominator in Eq. (7.44). The overall scaling is a global factor affecting all gains equally, and this global modification of the local gains could be achieved at the whole neuron level, rather than at the local input level. Nonetheless, Eq. (7.44) does require knowledge of the correlation coefficients. We shall discuss this in Chapter 9.

Turning to Eq. (7.41), since $\theta_U \equiv \gamma_1\theta_1 + \gamma_2\theta_2$, we have

$$\theta_U = \gamma_1\mu_1 + \gamma_2\mu_2 + \Theta\Gamma s. \tag{7.45}$$

This equation defines the combined threshold θ_U and not θ_1 and θ_2 separately. However, we defined the θ_i separately only for convenience, for the analogy to the unimodal case. We see from Eq. (7.33) that only the combination θ_U is real, with θ_1 and θ_2 having no independent meaning. θ_U may be regarded as the actual threshold of the neuron, although in the multimodal case it does not specify the semi-saturation point, and Eq. (7.45) sets this threshold uniquely.

In summary, in the bimodal case, the response function of the neuron is given by

$$r(x_1, x_2) = \frac{s}{2}\left[1 + \tanh\frac{2}{s}\left(\gamma_1 x_1 + \gamma_2 x_2 - \theta_U\right)\right], \tag{7.46}$$

and if the separate channel gains γ_i are adapted locally according to Eq. (7.44) and the pseudo-threshold θ_U is adapted according to Eq. (7.45), then the output PDF $f_R(r)$ remains invariant, or approximately so. Thus, Eqs. (7.44) and (7.45) provide a means for a bimodal neuron to adapt to the changing input statistics of two distinct input channels, when these channels' statistics vary either separately or simultaneously. Furthermore, these results generalise directly to a multimodal neuron with more than two separate input channels. Consider multiple input channel distributions, $X_1, ..., X_n$ with means $\mu_1, ..., \mu_n$ and variances $\sigma_1^2, ..., \sigma_n^2$, respectively, where n is the number of input channels. The correlation coefficient between sensory channels X_j and X_k is denoted as ρ_{jk}. We can then write the response function $r(x_1, ..., x_n)$ for a multimodal neuron as

$$r(x_1, ..., x_n) = \frac{s}{2}\left[1 + \tanh\frac{2}{s}\left(\sum_{i=1}^{n}\gamma_i x_i - \theta_U\right)\right], \tag{7.47}$$

where

$$\gamma_i = \frac{\Gamma}{\sqrt{n + 2\sum_{j>k}\rho_{jk}}}\frac{s}{\sigma_i}\ , \tag{7.48}$$

$$\theta_U = \sum_{i=1}^{n}\gamma_i\mu_i + \Theta\Gamma s\ . \tag{7.49}$$

Accordingly, a multisensory neuron adapting its response function under Eqs. (7.48) and (7.49) maintains invariant or approximately invariant output PDF when the multimodal input statistics vary.

7.3 Examples of Adaptation in Model Neurons

First we consider the application of our adaptation rules to the case of a purely unimodal neuron and show examples of adaptation, both perfect and imperfect, to changes in the input mean and standard deviation. We then consider a bimodal neuron, and discuss in particular adaptation to changes in the correlation coefficient, which is not an available form of adaptation in the unimodal case.

7.3.1 Unimodal Inputs

We demonstrate the ability of the adaptation rules in Eqs. (7.11) and (7.12) to accommodate changes in the input statistics under noiseless condition. We consider a variety of different input distributions, one demonstrating only approximate invariance of $f_R(r)$.

7.3.1.1 Logistic Distribution

We first consider the logistic or sech-squared distribution, defined by the PDF

$$f_X(x) = \frac{\pi}{4\sigma\sqrt{3}} \operatorname{sech}^2 \frac{\pi}{2\sqrt{3}} \left(\frac{x-\mu}{\sigma} \right), \tag{7.50}$$

where μ and σ^2 are its mean and variance, respectively. Clearly $f_X(x)$ has a form for which $f_Z(z) = g(z)$, with $g(z) = \frac{\pi}{4\sqrt{3}} \operatorname{sech}^2 \frac{\pi}{2\sqrt{3}} z$, so perfect adaptation of θ and γ to changing μ and σ is possible. We consider the logistic distribution for two reasons. First, it is frequently employed as an alternative to the normal distribution, in order to simplify analysis (Johnson et al., 1995). Second, as mentioned above, it is the only input distribution for which it is possible, in the presence of the sigmoidal response function in Eq. (7.3), to generate an output distribution R that is uniform on $[0, s]$, and hence has the maximum entropy distribution, given the saturation constraint.

Since $f_R(r) = 1/s$ has maximum entropy on the bounded interval $[0, s]$, we deduce that Eqs. (7.26) and (7.27) are satisfied when

$$\theta = \mu, \tag{7.51}$$

$$\gamma = \frac{\pi}{4\sqrt{3}} \frac{s}{\sigma}, \tag{7.52}$$

from which we see that the constants Θ and Γ satisfy $\Theta = 0$ and $\Gamma = \frac{\pi}{4\sqrt{3}} \approx 0.4534$. The scaling of the gain γ with the range of the response s was anticipated earlier, giving rise to the scale-free definition of the constant Γ. This can be seen directly from Eqs. (7.26) and (7.27).

In Fig. 7.1A, we set $\Theta = 0$ and $\Gamma = \frac{\pi}{4\sqrt{3}}$, corresponding to the maximum entropy operating point. For $\mu = 10$ and $\sigma = 2$, Eqs. (7.11) and (7.12) then induce a threshold $\theta = 10$ and gain $\gamma = s\Gamma/2$, and we set $s = 10$ without loss of generality. For these values of the threshold and gain and fixed $\mu = 10$, we plot the output PDF $f_R(r)$ corresponding to three different values of σ: $\sigma = 2$, being the value to which θ and γ are adapted according to Eqs. (7.11) and (7.12); and $\sigma = 1$ and $\sigma = 3$, to which the selected values of θ and γ are not optimally adapted. For $\sigma = 2$, we see, as expected, that the output PDF is uniform, which is the maximum entropy distribution on a bounded interval. Changing σ away from $\sigma = 2$ without concomitant changes in θ and γ according to Eqs. (7.11) and (7.12) moves the neuron away from its preferred operating point, resulting in output PDFs that are non-uniform. We also set $\Theta = -0.5$ and $\Gamma = 0.4$ in order to compare adaptation and non-adaption for a preferred operating point away from the maximum entropy point. For the same input mean $\mu = 10$ and standard deviation $\sigma = 2$, the induced values of the threshold and gain at the preferred operating point are now given by $\theta = 9$ and $\gamma = 2$. With these values of θ and γ, in Fig. 7.1B we again plot the output PDF for the three values of σ, $\sigma = 1$, $\sigma = 2$ and $\sigma = 3$. The $\sigma = 2$ curve is, given Θ and Γ, again the preferred operating point of the neuron. Of course, this PDF is now no longer uniform. Without adapting θ and γ to accommodate the changes in σ, the $\sigma = 1$ and $\sigma = 3$ curves show how the PDF drifts away from the preferred operating point corresponding to the $\sigma = 2$ curve. Setting θ and γ appropriately would return the PDFs to the curve for $\sigma = 2$.

In Fig. 7.2, we therefore show how adaptation of θ and γ according to Eqs. (7.11) and (7.12) returns the output PDFs back to the two preferred operating points considered in Fig. 7.1. The two preferred operating points are shown, with thresholds and gains set appropriately for $\sigma = 2$. When σ changes to $\sigma = 1$, the PDFs change. To return the PDFs back to the preferred operating points, we must set θ and γ according to Eqs. (7.11) and (7.12). Incrementing θ and γ from their initial values at $\sigma = 2$ to their target values at $\sigma = 1$ shows how the output PDFs return to the preferred operating points. We see that adaptation is in this case perfect, and does not depend on the details of how the preferred operating point is set. In particular, when the preferred operating point is initially set according to a maximum entropy principle, as shown in Fig. 7.2A, adaptation restores the output PDF to the maximum entropy distribution in precisely the same manner that it returns the output PDF to a non-optimal preferred operating point, as shown in Fig. 7.2B. The use of Eqs. (7.11) and (7.12) thus avoids an explicit recomputation of the maximum entropy point following a change in the input statistics.

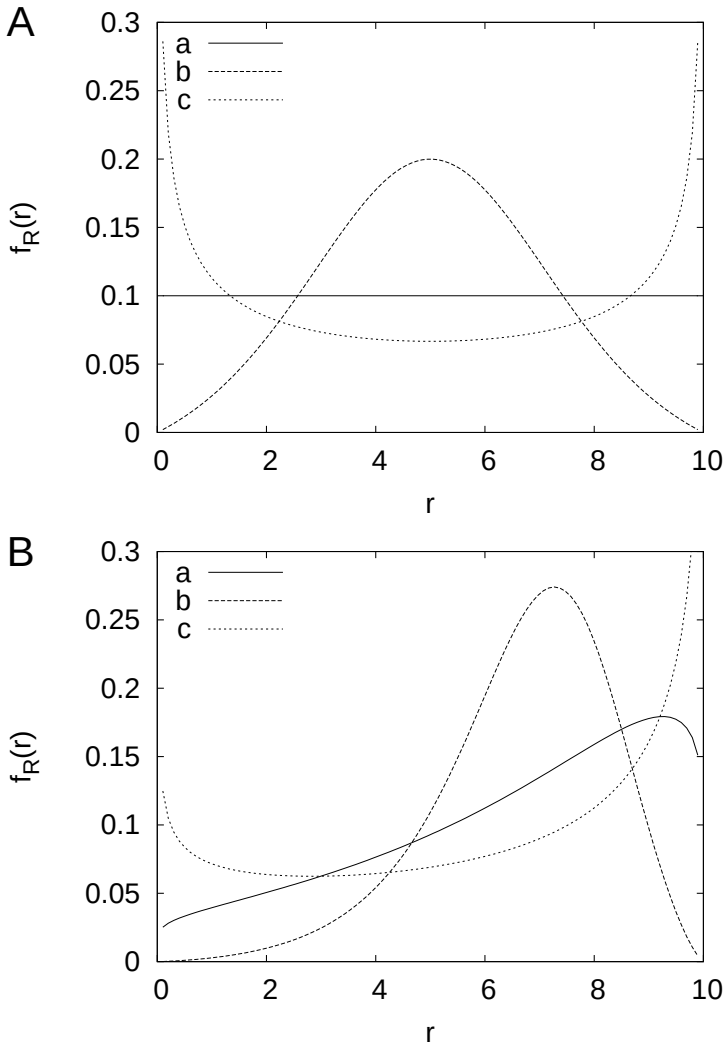


FIGURE 7.1: Output PDFs for a unimodal neuron in the presence of an input with a logistic distribution. (A) The preferred operating point of the neuron is set by $\Theta \equiv 0$ and $\Gamma = \frac{\pi}{4\sqrt{3}}$, corresponding to a maximum entropy distribution. For $\mu = 10$ and $\sigma = 2$, the induced threshold and gain at the preferred operating point are $\theta = 10$ and $\gamma \approx 2.27$. The output PDFs are shown with these values of θ and γ for $\sigma = 2$ (PDF a, optimal), $\sigma = 1$ (PDF b, non-optimal) and $\sigma = 3$ (PDF c, non-optimal). (B) The preferred operating point of the neuron is now set by $\Theta = -0.5$ and $\Gamma = 0.4$, corresponding to a preferred operating point away from maximum entropy. For $\mu = 10$ and $\sigma = 2$, the induced threshold and gain at the preferred operating point are $\theta = 9$ and $\gamma = 2$. PDFs a, b and c then show the resulting PDFs for the same values of σ in part A, with $\sigma = 2$ again being the desired adapted case.

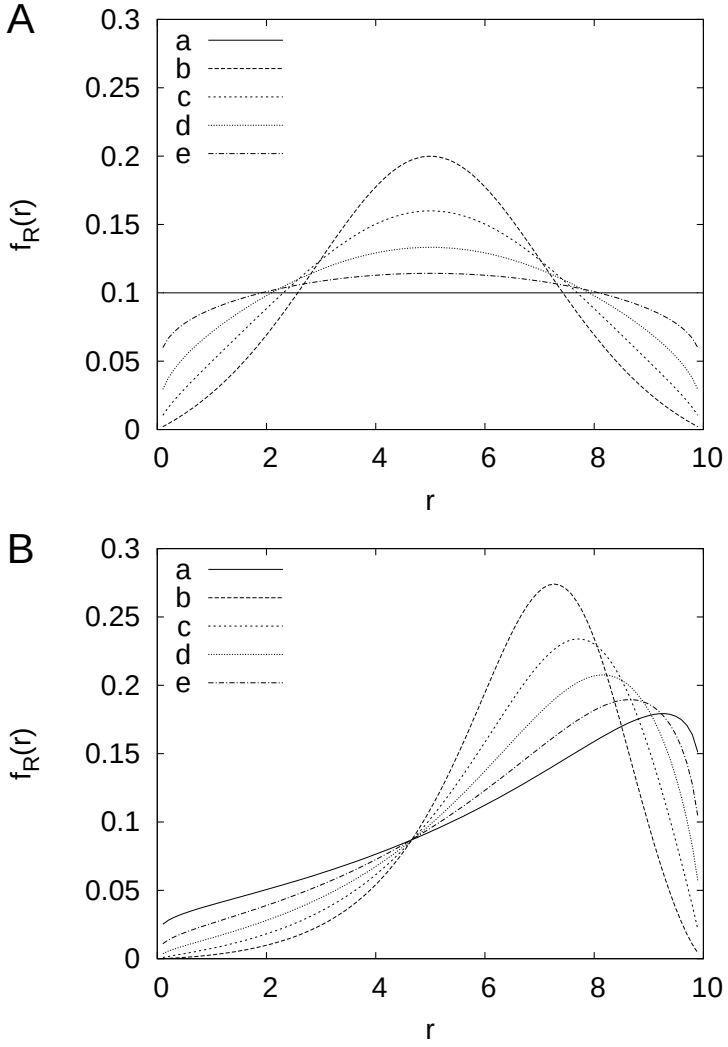


FIGURE 7.2: Adaptation of a unimodal neuron back to its preferred operated point following a change in its input variance, for an input with a logistic distribution. (A) For the maximum entropy operating point shown in Fig. 7.1A, the neuron adjust its gain to restore the PDF to the preferred distribution after σ changes from $\sigma = 2$ to $\sigma = 1$. Because $\Theta \equiv 0$ for this operating point, the threshold remains unchanged. PDF a shows the original, preferred PDF with $\sigma = 2$, $\theta = 10$ and $\gamma \approx 2.27$. PDFs b–e have $\sigma = 1$. The values of γ are: $\gamma = 2.27$ (PDF b); $\gamma = 2.83$ (PDF c); $\gamma = 3.40$ (PDF d); $\gamma = 3.97$ (PDF e). The PDF for $\gamma \approx 4.53$, the target value for $\sigma = 1$, is not shown, since it is identical to PDF a. (B) The operating point in now that shown in Fig. 7.1B, under the same change in σ as shown in part A. For this case, both the threshold and the gain change in order to return the neuron to its preferred operating point. PDF a shows the original, preferred PDF with $\sigma = 2$, $\theta = 9$ and $\gamma = 2$. PDFs b–e have $\sigma = 1$. The values of θ and γ are: $\theta = 9$, $\gamma = 2$ (PDF b); $\theta = 9.125$, $\gamma = 2.5$ (PDF c); $\theta = 9.25$, $\gamma = 3$ (PDF d); $\theta = 9.375$, $\gamma = 3.5$ (PDF e). The PDF for $\theta = 9.5$, $\gamma = 4$, the target values for $\sigma = 1$, is again not shown.

7.3.1.2 Normal Distribution

We now turn to the normal distribution, defined by the standard PDF

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \tag{7.53}$$

again with mean μ and variance σ^2, respectively. As with the logistic distribution, the normal distribution can clearly be written in the form $f_Z(z) = g(z)$, permitting perfect adaptation. Unlike the logistic distribution, the maximum entropy output distribution R, given the sigmoidal response function in Eq. (7.3), must be determined explicitly from the evaluation of the integrals in Eqs. (7.26) and (7.27). For the first integral, we have that

$$\frac{\partial S[R]}{\partial \theta} = \frac{1}{\sqrt{2\pi}} \frac{4\gamma}{s} \int_{-\infty}^{\infty} dz\, e^{-\frac{1}{2}z^2} \tanh \frac{2\gamma\sigma}{s}\left(z - \frac{\theta-\mu}{\sigma}\right). \tag{7.54}$$

Since $\exp(-z^2/2)$ is even around $z = 0$, the integral can be made to vanish when we choose the argument of the tanh function so that the tanh function is odd around $z = 0$. This is possible only if $\theta = \mu$. Indeed, it is clear that for *any* input distribution that is symmetric about its mean, we must set $\theta = \mu$ in order to satisfy Eq. (7.26), which implies that $\Theta \equiv 0$. With Eq. (7.26) satisfied, Eq. (7.27) reduces to

$$\Gamma \int_{-\infty}^{\infty} dz\, z\, e^{-\frac{1}{2}z^2} \tanh 2\Gamma z = \sqrt{\frac{\pi}{8}}, \tag{7.55}$$

from which we must determine the solution for Γ numerically, resulting in $\Gamma \approx 0.4372$ for the maximum entropy output PDF. Notice the similarity between this value of Γ for the normal distribution and the value $\Gamma = \frac{\pi}{4\sqrt{3}} \approx 0.4534$ for the logistic distribution, confirming the utility of replacing a normal distribution by a logistic distribution for the purposes of analytical tractability (Johnson et al., 1995).

Corresponding to the results in Fig. 7.1 for the logistic distribution, in Fig. 7.3 we exhibit the preferred operating point of a neuron under the two choices, $\Theta = 0$, $\Gamma = 0.4372$, corresponding to the maximum entropy point (Fig. 7.3A), and $\Theta = -0.5$, $\Gamma = 0.4$, corresponding to an operating point away from maximum entropy (Fig. 7.3B). We also show, as before, deviations from the preferred operating points as σ is changed without corresponding adaptations in θ and γ. Fig. 7.4 shows the return of the output PDFs to the two preferred operating points under adaptive changes in θ and γ. These results for the normal distribution are qualitatively and in fact quantitatively very similar to those for the logistic distribution, as expected from the similarity of these two distributions (Johnson et al., 1995). The closeness of the $\sigma = 2$ curve in Fig. 7.3A, corresponding to the maximum entropy distribution, to uniformity is remarkable.

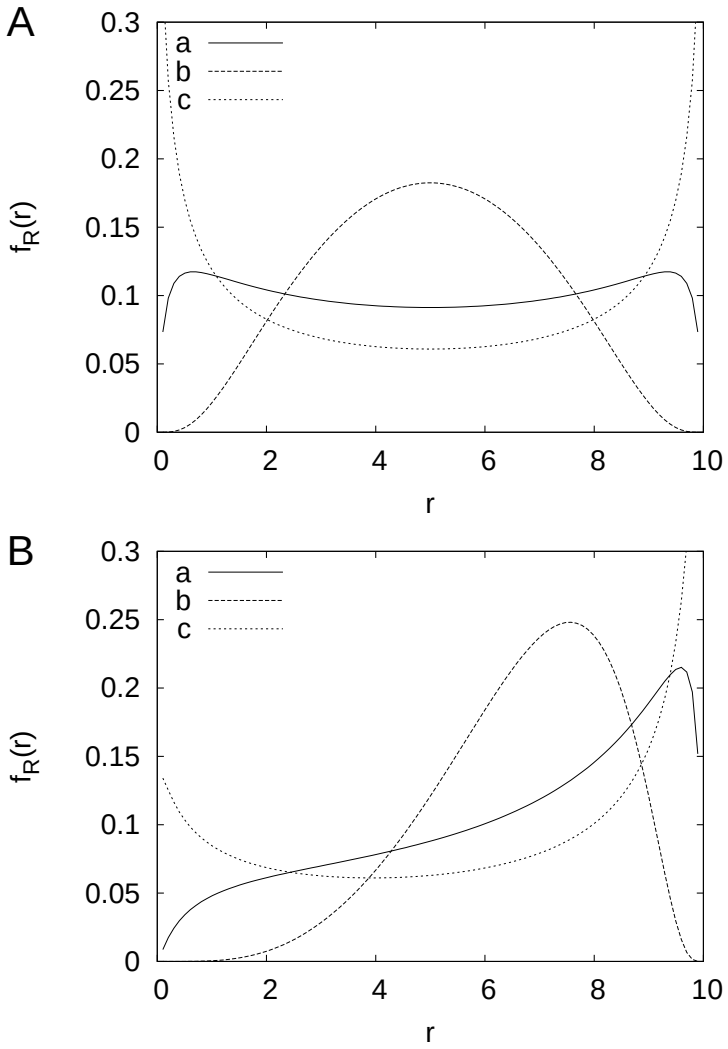


FIGURE 7.3: Output PDFs for a unimodal neuron in the presence of an input with a normal distribution. (A) The preferred operating point of the neuron is set by $\Theta \equiv 0$ and $\Gamma \approx 0.4372$, corresponding to a maximum entropy distribution. For $\mu = 10$ and $\sigma = 2$, the induced threshold and gain at the preferred operating point are $\theta = 10$ and $\gamma \approx 2.19$. The output PDFs shown are otherwise identical to those in Fig. 7.1A. The format and parameters in (B) are identical to those in Fig. 7.1B, except that the normal distribution replaces the logistic distribution.

7.3.1.3 Convolved Exponential Distribution

In considering a neuron with multimodal inputs below, it will be convenient to consider the distribution defined by the PDF

$$f_X(x) = \frac{\exp(-x/\lambda_+) - \exp(-x/\lambda_-)}{\lambda_+ - \lambda_-}, \quad x \geq 0, \tag{7.56}$$

which is the distribution of a variable X defined as the sum of two independent, exponentially-distributed variables $X_\pm$ with means $\lambda_\pm$, respectively, so that $X = X_+ + X_-$ and

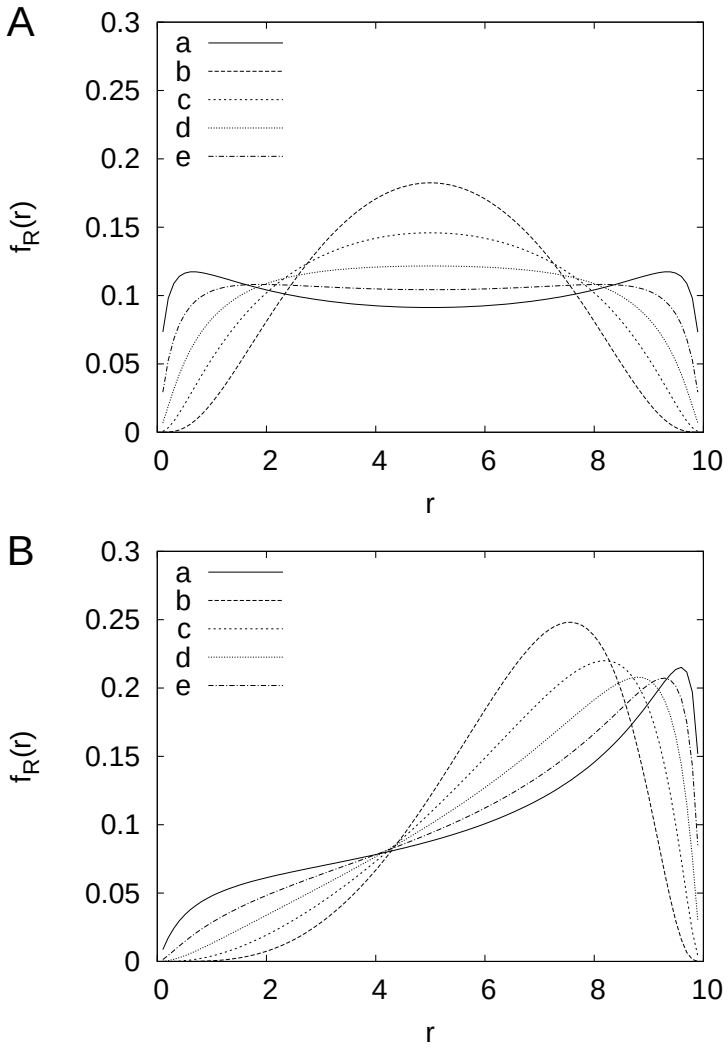


FIGURE 7.4: Adaptation of a unimodal neuron back to its preferred operated point following a change in its input variance, for an input with a normal distribution. (A) For the maximum entropy operating point shown in Fig. 7.3A, the neuron adjusts its gain to restore the PDF to the preferred distribution after σ changes from $\sigma = 2$ to $\sigma = 1$. PDF a shows the original, preferred PDF with $\sigma = 2$, $\theta = 10$ and $\gamma \approx 2.19$. PDFs b–e have $\sigma = 1$. The values of γ are: $\gamma = 2.19$ (PDF b); $\gamma = 2.73$ (PDF c); $\gamma = 3.28$ (PDF d); $\gamma = 3.83$ (PDF e). The format and parameters in (B) are identical to those in Fig. 7.1B, except that the normal distribution replaces the logistic distribution.

its PDF is obtained from the convolution of the two exponential PDFs,

$$f_X(x) = \int_0^x dy\, f_{X_+}(y) f_{X_-}(x-y). \tag{7.57}$$

The mean and variance of X are given by $\mu = \lambda_+ + \lambda_-$ and $\sigma^2 = \lambda_+^2 + \lambda_-^2$, respectively. Regarding μ and σ as the fundamental parameters, we can invert these relations to obtain

$$\lambda_\pm = \frac{1}{2}\left(\mu \pm \sqrt{2\sigma^2 - \mu^2}\right). \tag{7.58}$$

For this distribution to exist, we must have $\sigma < \mu < \sqrt{2}\sigma$, so that the mean can only take values defined in a range set by the standard deviation, and *vice versa*. Writing $q = \mu/\sigma$, so that $1 < q < \sqrt{2}$, and transforming to the Z variable, we have

$$f_Z(z) = \frac{\exp\left[\frac{-2(z+q)}{q+\sqrt{2-q^2}}\right] - \exp\left[\frac{-2(z+q)}{q-\sqrt{2-q^2}}\right]}{\sqrt{2-q^2}}. \tag{7.59}$$

We see that $f_Z(z)$ does not take the form $g(z)$ considered earlier, because the ratio q, depending on μ and σ, appears in $f_Z(z)$. Calculating directly the moments, we find, for example, that although $\tilde{m}_1 = 0$ and $\tilde{m}_2 = 1$ (by construction), we have

$$\tilde{m}_3 = q(3 - q^2), \tag{7.60}$$

which is not invariant under independent changes in μ and σ. Considering the bounds on q, $\tilde{m}_3$ exists in the range $[\sqrt{2}, 2]$. Given this non-invariance, adapting θ and γ according to Eqs. (7.11) and (7.12) will not, in general, result in an invariant output PDF $f_R(r)$. However, if μ and σ are varied in such a way that their ratio q is constant, then we do have invariance. Despite this general non-invariance, we will show, nonetheless, that adapting θ and γ according to Eqs. (7.11) and (7.12) results in only a mild breakdown in exact invariance.

We set $\mu = 10$ and take a value $\sigma = 8.5$ [approximately midway in the allowed range of σ of $(5\sqrt{2}, 10)$] and use Eqs. (7.26) and (7.27) to determine the values of θ and γ corresponding to the maximum entropy output PDF for the convolved exponential input PDF. For this input distribution, Eqs. (7.26) and (7.27) produce an assortment of hypergeometric functions, ${}_2F_1$ and ${}_3F_2$, so we perform direct numerical searches for the values of θ and γ that induce the maximum entropy output distribution for given input mean and standard deviation. For $\mu = 10$ and $\sigma = 8.5$, we find $\theta \approx 8.68$ and $\gamma \approx 0.60$, corresponding to values $\Theta \approx -0.16$ and $\Gamma \approx 0.51$.

In Fig. 7.5A, we thus set $\Theta = -0.16$ and $\Gamma = 0.51$, and for a fixed $\mu = 10$, plot the output PDF $f_R(r)$ for three different values of σ, $\sigma = 8.5$, giving the maximum entropy distribution and defining the preferred operating point of the neuron, and $\sigma = 7.5$ and $\sigma = 9.5$, moving the distribution away from the preferred operating point. Clearly, for the convolved exponential distribution, the output PDF is very far from uniform on $[0, s]$. Furthermore, we see that the full output range is not employed by the neuron. This is a consequence of the semi-bounded nature of the input distribution, for which we must have $x \geq 0$. Hence, the output response can never fall below $r(0)$, or

$$\frac{s}{2}\left[1 - \tanh\frac{2\gamma}{s}\theta\right]. \tag{7.61}$$

As the threshold and gain change, this lower limit changes, with this limit determined by the product $\gamma\theta$. In Fig. 7.5B, we define a preferred operating point for the neuron

away from the maximum entropy distribution by setting $\Theta = -0.5$ and $\Gamma = 0.4$ as usual. With threshold $\theta = 5.75$ and gain $\gamma = 8/17 \approx 0.47$, an input mean of $\mu = 10$ and standard deviation $\sigma = 8.5$ again ensures that the neuron functions at its preferred operating point. Shown is the output PDF corresponding to this preferred operating point, together with the output PDFs corresponding to $\sigma = 7.5$ and $\sigma = 9.5$, for the same, unadapted values of θ and γ. Because the product $\gamma\theta$ is larger in Fig. 7.5A than in Fig. 7.5B, we see that the output PDF is non-zero over a larger range of values in the former case compared to the latter.

In Fig. 7.6, we show how adaptation of θ and γ according to Eqs. (7.11) and (7.12) returns, or attempts to return the neuron to its preferred operating point when the input statistics change. For the convolved exponential distribution, we know that when the ratio $q = \mu/\sigma$ is held constant, perfect adaptation back to the preferred operating point is possible. This is demonstrated in Fig. 7.6A, where the preferred operating point is set by $\Theta = -0.5$, $\Gamma = 0.4$ with $\theta = 5.75$ and $\gamma = 0.47$ ensuring the preferred output PDF when $\mu = 10$ and $\sigma = 8.5$. Changing σ to $\sigma = 7.5$ with a concomitant change in μ to $\mu = 8.8$ to keep the ratio q constant then induces a threshold and gain set by $\theta = 5.05$ and $\gamma = 0.53$. Fig. 7.6A shows how the output PDF returns back to its preferred operating point as θ and γ are changed from their initial values at $\mu = 10$ and $\sigma = 8.5$ to their target values at $\mu = 8.8$ and $\sigma = 7.5$. As expected, exact restoration of the PDF is possible for fixed q. In Fig. 7.6B, however, we change σ to $\sigma = 7.5$ but hold μ fixed, so that the ratio q is not constant. Adaptation back to the preferred operating point should now be compromised by the non-invariance of the higher-order moments of the convolved exponential distribution. Under Eqs. (7.11) and Eq. (7.12), the target values for θ and γ with $\mu = 10$ and $\sigma = 7.5$ are $\theta = 6.25$ and $\gamma = 0.53$. We see, however, that adapting θ and γ to their target values defined by Eqs. (7.11) and (7.12) does not return the output PDF to the neuron's preferred operating point.

Despite this expected, general non-invariance of the output PDF $f_R(r)$ for convolved exponential input, it is natural to wonder whether adaptation according to Eqs. (7.11) and (7.12) is better than not adapting θ and γ at all. To this end, we determine the extent to which adaptation according to Eqs. (7.11) and (7.12) is able to track the maximum entropy distribution governed by the solutions of Eqs. (7.26) and (7.27), and compare this to the deviation induced in the absence of adaptation of θ and γ to changes in μ and σ. Thus, here, we regard the maximum entropy distribution as the preferred operating point of a neuron, and determine the deviations from this preferred operating point when either θ and γ are imperfectly adapted according to Eqs. (7.11) and (7.12) or θ and γ are not adapted at all. For $\mu = 10$ and $\sigma = 7.2$, the maximum entropy operating point, from the solution of Eqs. (7.26) and (7.27) is set by $\theta \approx 9.10$ and $\gamma \approx 0.66$, inducing values $\Theta \approx -0.13$ and $\Gamma \approx 0.48$. We increase σ up to its upper limit of $\mu = 10$ and for each value of σ, we compute three different output entropies. First, we determine the new, maximum output entropy values of θ and γ from Eqs. (7.26)

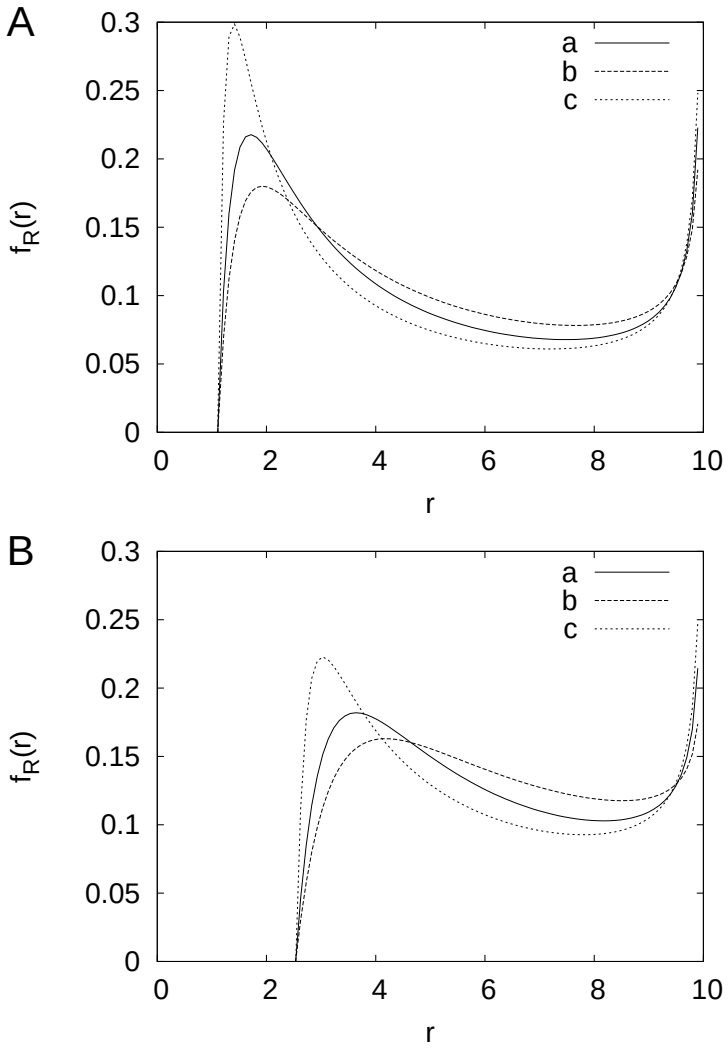


FIGURE 7.5: Output PDFs for a unimodal neuron in the presence of an input with a convolved exponential distribution. (A) The preferred operating point of the neuron is set by $\Theta = -0.16$ and $\Gamma = 0.51$, corresponding approximately to a maximum entropy distribution. For $\mu = 10$ and $\sigma = 8.5$, the induced threshold and gain at the preferred operating point are $\theta = 8.68$ and $\gamma = 0.60$. The output PDFs are shown with these values of θ and γ for $\sigma = 7.5$ (PDF a, non-optimal), $\sigma = 8.5$ (PDF b, optimal) and $\sigma = 9.5$ (PDF c, non-optimal). (B) The preferred operating point of the neuron is now set by $\Theta = -0.5$ and $\Gamma = 0.4$, corresponding to a preferred operating point away from maximum entropy. For $\mu = 10$ and $\sigma = 8.5$, the induced threshold and gain at the preferred operating point are $\theta = 5.75$ and $\gamma = 0.47$. PDFs a, b and c then show the resulting PDFs for the same values of σ in part A, with $\sigma = 8.5$ again being optimal.

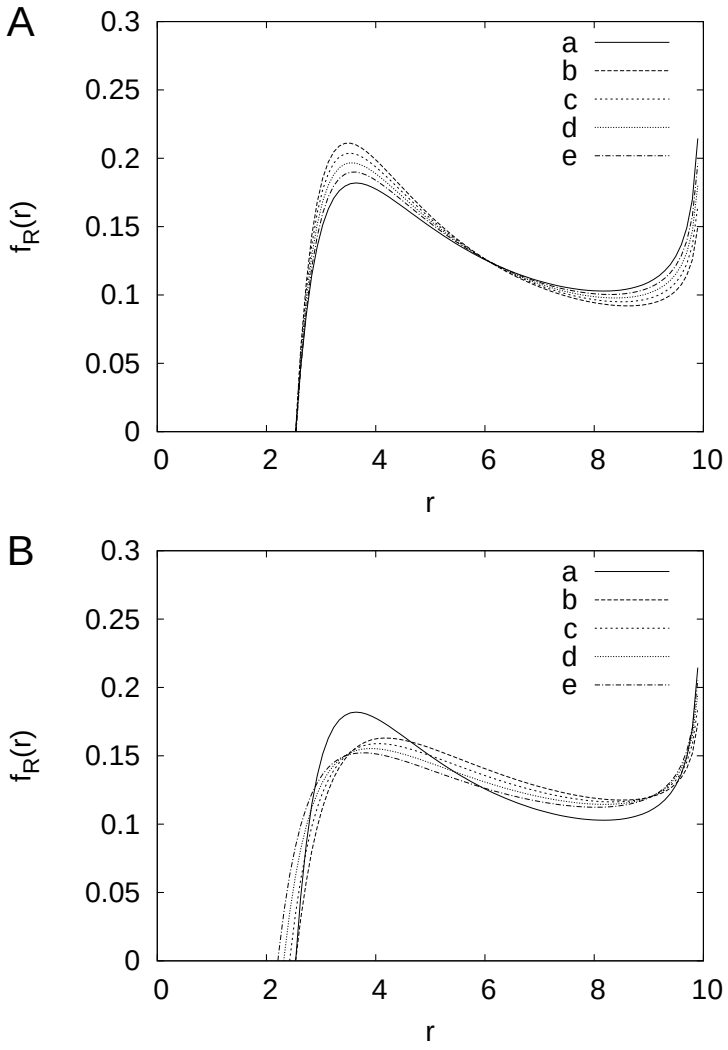


FIGURE 7.6: Examples of perfect and imperfect adaptation of a unimodal neuron in the presence of an input with a convolved exponential distribution. The operating point of the neuron is set as in Fig. 7.5B. (A) If the ratio μ/σ is held constant, then perfect adaptation is possible. For $\mu = 10$ and $\sigma = 8.5$, the induced values of the threshold and gain are $\theta = 5.75$ and $\gamma = 0.47$, corresponding to PDF a. With $\mu = 8.8$ and $\sigma = 7.5$, but θ and γ held constant, the PDF moves to b. If θ and γ are moved towards their target values for $\mu = 8.8$ and $\sigma = 7.5$, the PDF returns to that in a: $\theta = 5.575$, $\gamma = 0.485$ (PDF c); $\theta = 5.4$, $\gamma = 0.5$ (PDF d); $\theta = 5.225$, $\gamma = 0.515$ (PDF e). (B) If μ is instead held fixed at $\mu = 10$, while σ moves to $\sigma = 7.5$, then exact invariance of the output PDF is not possible. PDF a again shows the preferred operating point, with $\theta = 5.75$ and $\gamma = 0.47$. With θ and γ held constant, setting $\sigma = 7.5$ produces PDF b. Moving θ and γ to their target values for $\sigma = 7.5$ does not return the PDF to PDF a: $\theta = 5.875$, $\gamma = 0.485$ (PDF c); $\theta = 6.0$, $\gamma = 0.5$ (PDF d); $\theta = 6.125$, $\gamma = 0.515$ (PDF e).

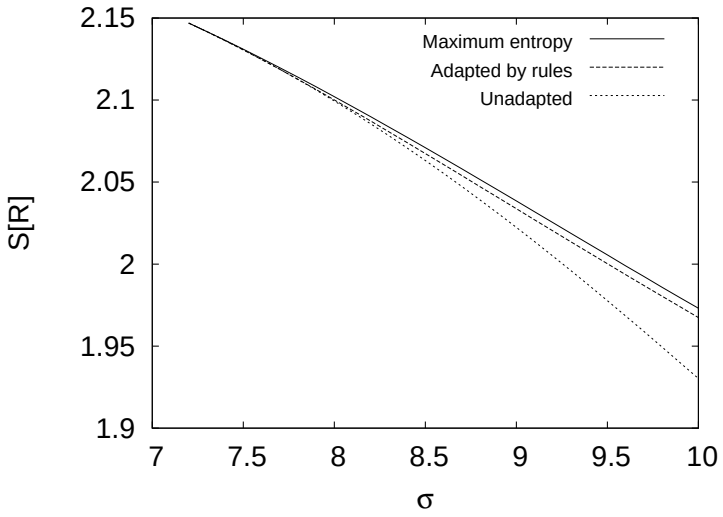


FIGURE 7.7: Tracking the maximum entropy output PDF by adaptation according to Eqs. (7.11) and (7.12), as σ varies, for a unimodal neuron with input drawn from a convolved exponential distribution. Here the preferred operating point of the neuron is defined functionally, as the maximum entropy output PDF. Adaptation according to Eqs. (7.11) and (7.12) in this case is imperfect, but tracks the maximum entropy distribution closely. In contrast, not adapting the output PDF quickly shifts the output PDF away from the preferred, maximum entropy operating point.

and (7.27), and the associated value of $S_{\max}[R]$. In this case, Θ and Γ are not fixed, but determined functionally. Second, we determine the entropy, $S_{\text{adapt}}[R]$, of the output PDF when θ and γ are instead adapted according to Eqs. (7.11) and (7.12), so that Θ and Γ are held constant. Finally, we calculate the output PDF entropy, $S_{\text{no}}[R]$ when θ and γ are held constant, so that there is no adaptation to the changing standard deviation. The results are shown in Fig. 7.7. We see that adapting θ and γ according to Eqs. (7.11) and (7.12), although not perfect, tracks the maximum entropy distribution very closely, while the non-adapted distribution deviates from the target distribution to a greater extent. It is worthwhile commenting that the maximum output entropy is here not constant precisely because the output distribution cannot be made exactly invariant. Were we to display a similar graph to this for the logistic distribution, however, $S_{\max}[R]$ would remain constant as σ is varied, and the perfect adaptation present in that case would ensure that $S_{\text{adapt}}[R] = S_{\max}[R]$ for all values of σ provided that equality is established for any one value of σ.

7.3.2 Multimodal Inputs

We have seen that the multimodal input case can be reduced, mathematically-speaking, to the unimodal input case by considering the effective input $U = \gamma_1 X_1 + \gamma_2 X_2$ and adapting the channel-specific gains γ_i and the pseudo-threshold θ_U according to Eqs. (7.44) and (7.45), respectively. Our presentation of the above results for the unimodal case therefore in general completely characterises the multimodal results too. We thus discuss

only briefly examples of the invariance of $f_R(r)$ in the presence of multimodal inputs for illustrative purposes. Novel to the multimodal case is the possibility of adaptation to the correlation coefficients between different modalities, so we focus on this case specifically.

7.3.2.1 Bivariate Normal Distribution

As perhaps the simplest, non-trivial example of a bivariate distribution, we suppose that the two channel inputs are drawn from a bivariate normal distribution with joint PDF $f_{X_1X_2}(x_1, x_2)$ given by

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi} \frac{1}{\sqrt{\det \Sigma}} \exp\left[-\frac{1}{2}(\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})\right], \tag{7.62}$$

where $\underline{x}^T = (x_1, x_2)$, the superscript T denoting the transpose, $\underline{\mu}^T = (\mu_1, \mu_2)$ and Σ is the covariance matrix,

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}, \tag{7.63}$$

where $\det \Sigma$ is its determinant. It is easy to see that the distribution of the effective, unimodal input U is normal, with mean and variance given by Eqs. (7.35) and (7.36), respectively. In order to determine the operating point corresponding to the maximum entropy output distribution in the presence of a bivariate normal input distribution, we must calculate the values of Θ and Γ determining this point for a univariate normal input distribution.

The examination of the unimodal, univariate normal input case in Section 7.3.1.2 suffices to understand the bimodal, bivariate normal input case. In the bimodal case, adaptation of γ_1 and γ_2 according to Eq. (7.44) and θ_U according to Eq. (7.45) leads to the exact invariance of $f_R(r)$ under changes in *all* the bivariate input statistics, including the correlation coefficient ρ. Since the correlation coefficient does not appear in the unimodal case studied earlier, we examine, in particular, adaptation to changes in ρ only, with the means μ_1 and μ_2 and the variances σ_1^2 and σ_2^2 held constant.

In Fig. 7.8, we as usual consider two different operating points, the first (Fig. 7.8A) corresponding to the maximum entropy distribution, with $\Theta \equiv 0$ and $\Gamma \approx 0.4372$ for a univariate normal distribution, and the second (Fig. 7.8B) defined by $\Theta = -0.5$ and $\Gamma = 0.4$. We fix $\mu_1 = 10$, $\sigma_1 = 2$, as for the univariate normal above, and then fix $\mu_2 = 8$ and $\sigma_2 = 4$ as representative values. We select an initial correlation coefficient of $\rho = -0.5$. The preferred operating points then determine γ_1, γ_2 and θ_U via Eqs. (7.44) and (7.45). We then decrease ρ to $\rho = -0.9$. Without adaptive changes in the γ_i and θ_U, the output PDFs moves away from the preferred operating points. As the γ_i and θ_U are restored to their induced values according to Eqs. (7.44) and (7.45), the output PDFs $f_R(r)$ return to the preferred operating points. Adaptation to changes in the correlation coefficient in this bivariate normal input case is perfect.

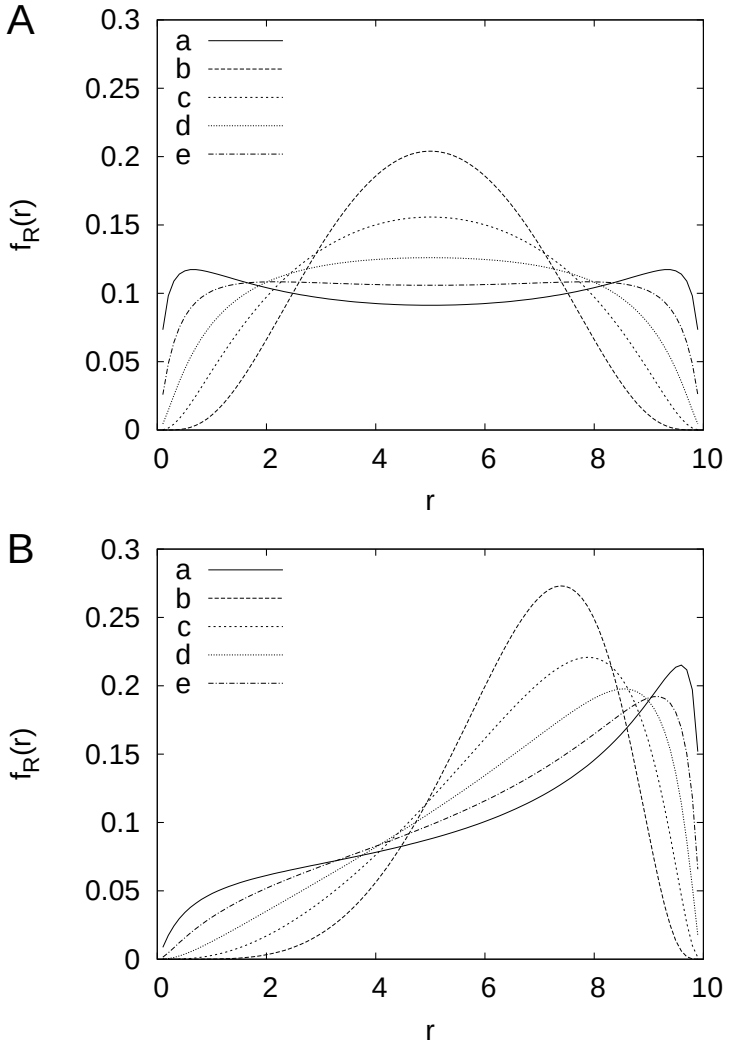


FIGURE 7.8: Adaptation of a bimodal neuron to changes in the correlation coefficient for inputs drawn from a bivariate normal distribution. (A) The preferred operating point is set by $\Theta \equiv 0$ and $\Gamma = 0.4372$, corresponding to a maximum entropy distribution. For an initial correlation coefficient $\rho = -0.5$, the induced pseudo-threshold and gains are set as $\theta_U = 30.60$, $\gamma_1 = 2.19$, $\gamma_2 = 1.09$ (PDF a), defining the preferred PDF. Changing the correlation coefficient to $\rho = -0.9$ while keeping the pseudo-threshold and gains constant shifts the neuron from its preferred PDF (PDF b). Moving the pseudo-threshold and gains to their target values for $\rho = -0.9$ restores the neuron's output PDF to its preferred point: $\theta_U = 40.07$, $\gamma_1 = 2.86$, $\gamma_2 = 1.43$ (PDF c); $\theta_U = 49.52$, $\gamma_1 = 3.54$, $\gamma_2 = 1.77$ (PDF d); $\theta_U = 58.98$, $\gamma_1 = 4.21$, $\gamma_2 = 2.11$ (PDF e). (B) The preferred operating point of the neuron is set by $\Theta = -0.5$ and $\Gamma = 0.4$, moving the neuron away from a maximum entropy distribution. For an initial correlation coefficient $\rho = -0.5$, the induced pseudo-threshold and gains are set as $\theta_U = 26$, $\gamma_1 = 2$, $\gamma_2 = 1$ (PDF a), defining the preferred PDF. Changing the correlation coefficient to $\rho = -0.9$ while keeping the pseudo-threshold and gains constant shifts the neuron from its preferred PDF (PDF b). Moving the pseudo-threshold and gains to their target values for $\rho = -0.9$ restores the neuron's output PDF to its preferred point: $\theta_U = 34.65$, $\gamma_1 = 2.62$, $\gamma_2 = 1.31$ (PDF c); $\theta_U = 43.31$, $\gamma_1 = 3.24$, $\gamma_2 = 1.62$ (PDF d); $\theta_U = 51.96$, $\gamma_1 = 3.85$, $\gamma_2 = 1.93$ (PDF e).

For the general multivariate normal input case with n inputs, there are n means, n variances, and $\frac{1}{2}n(n-1)$ correlation coefficients. Adaptation to all n first-order moments is accomplished through just one parameter, the pseudo-threshold θ_U. The n gains γ_i permit adaptation to the variances of the n inputs. Although it is an immediate consequence of the form of Eq. (7.44) in the general case, in which a sum of all the correlation coefficients appears in the denominator, it is nonetheless remarkable that adaptation to all $\frac{1}{2}n(n+1)$ second-order moments, consisting of n variances and $\frac{1}{2}n(n-1)$ correlation coefficients, is possible with just n gain parameters.

7.3.2.2 Bivariate Exponential Distribution

As a final example of a bivariate input distribution to a bimodal neuron, consider a bivariate exponential distribution, such as the Cowan distribution (Cowan, 1987). The CDF of U is given by

$$F_U(u) = \int_0^{u/\gamma_1} dx_1 \int_0^{(u-\gamma_1 x_1)/\gamma_2} dx_2 \, f_{X_1X_2}(x_1, x_2), \tag{7.64}$$

where $f_{X_1X_2}$ is now the Cowan bivariate exponential PDF, and the PDF of U is

$$f_U(u) = \frac{1}{\gamma_2} \int_0^{u/\gamma_1} dx_1 \, f_{X_1X_2}\left(x_1, \frac{u-\gamma_1 x_1}{\gamma_2}\right). \tag{7.65}$$

If the separate input channels X_1 and X_2 have means λ_1 and λ_2, then we have $\mu_U = \gamma_1\lambda_1 + \gamma_2\lambda_2$ and $\sigma_U^2 = \gamma_1^2\lambda_1^2 + \gamma_2^2\lambda_2^2 + 2\gamma_1\lambda_1\gamma_2\lambda_2\rho$. It is therefore clear that $\mu_U \neq \sigma_U$ unless $\rho \equiv 1$. Hence, with a bivariate exponential input, the effective unimodal input U is not itself exponentially-distributed, except for the very particular choice $\rho = 1$. This contrasts with the bivariate normal case considered above, for which U is always a univariate normal distribution.

In order to determine $f_U(u)$, for simplicity we set $\rho = 0$, so that $f_{X_1X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$, with X_1 and X_2 independent, exponentially-distributed inputs with parameters λ_1 and λ_2. From Eq. (7.65), we then obtain

$$f_U(u) = \frac{\exp[-u/(\gamma_1\lambda_1)] - \exp[-u/(\gamma_2\lambda_2)]}{\gamma_1\lambda_1 - \gamma_2\lambda_2}. \tag{7.66}$$

We recognise this as the PDF of the convolved exponential distribution, with parameters $\lambda_+ = \gamma_1\lambda_1$ and $\lambda_- = \gamma_2\lambda_2$, considered earlier. Hence, perfect adaptation is not possible for a bimodal neuron receiving uncorrelated, exponentially-distributed inputs, despite the fact that perfect adaptation is possible for a unimodal neuron receiving a single, exponentially-distributed input. In general, of course, the distribution of the effective input U will be very complicated, and f_U will not take the form g discussed above for perfect adaptation.

7.4 Summary

Sensory neurons adapt to changes in the natural statistics of the environment through processes such as gain control and firing threshold adjustment. It has been argued that neurons early in sensory pathways adapt according to information-theoretic criteria, perhaps maximising their coding efficiency or information rate. In this chapter, we drew a distinction between how a neuron's preferred operating point is determined and how its preferred operating point is maintained through adaptation. We proposed that a neuron's preferred operating point can be characterised by the probability density function (PDF) of its output spike rate, and that adaptation maintains an invariant output PDF, regardless of how this output PDF is initially set. Considering a sigmoidal transfer function, we derived simple adaptation rules for a neuron with one sensory input that permit adaptation to the lower-order statistics of the input, independent of how the preferred operating point of the neuron is set. Thus, if the preferred operating point is, in fact, set according to information-theoretic criteria, then these rules maintain a neuron at that point. Our approach generalises from the unimodal case to the multimodal case, for a neuron with inputs from distinct sensory channels. We also presented several examples of input distributions for a neuron functioning at different preferred operating points, both unimodal and bimodal.

Chapter 8

Impact of Adaptation on Multisensory Integration

We have proposed in Chapter 7 an adaptation rule for both unimodal and multimodal neurons. Based on this rule, we discuss in this chapter the possible influences of multimodal input statistics, including the mean, the variance and the correlation coefficient, on the responses of an adapted multisensory neuron. This chapter is based on our paper "Adaptation in Multisensory Neurons: Impact on Cross-Modal Enhancement" (Elliott et al., 2008a).

8.1 Introduction

We have described in Chapter 2 how superior colliculus integrates visual, auditory and somatosensory information. Most neurons in the deep layers of superior colliculus (DSC) are multisensory, robustly exhibiting cross-modal enhancement (CME). In CME, the response to a stimulus from one sensory modality is augmented by the presence of a spatially coincident stimulus from another modality (Meredith et al., 1987; Stein and Meredith, 1993; Meredith and Stein, 1996). Larger CME is generated by weaker stimuli, a property called inverse effectiveness (IE). It has been shown that deactivating the anterior ectosylvian sulcus (AES) and the rostral lateral/ suprasylvian (rLS) areas by cooling eliminates CME in DSC neurons (Jiang et al., 2001), indicating that these two areas are indispensable for the integrative responses of DSC neurons. Moreover, DSC neurons also exhibit modality-specific suppression (MSS), in which a response to a stimulus is suppressed by another stimulus from the same sensory channel (Kadunce et al., 1997).

The mechanisms underlying multisensory integration in DSC neurons remain unclear. In Chapter 6, we have shown that a thresholded, saturating response function is sufficient to

account for the generation of CME associated with IE and the additive response properties in DSC neurons. In a theoretical study, a model based on a Bayesian approach has been proposed to account for multisensory integration (Patton and Anastasio, 2003). However, the model explicitly employs a saturating, sigmoidal response function and this, by itself, is enough to account for CME. Furthermore, although the model is able to exhibit MSS, we have shown that the MSS property is quite sensitive to the statistical parameters and in fact derives from the invertibility of the covariance matrix (see Chapter 5). The Bayesian model therefore does not sufficiently and plausibly explain the multisensory responses of DSC neurons.

We proposed in Chapter 7 an adaptation rule for multimodal neurons to adjust their responses according to input statistics. Under this adaptation rule, a DSC neuron maintains an operating point precisely or approximately invariant in spite of the changes in input statistics. Based on the adaptation rule, we present in this chapter results showing the influences of prior adaptation to the input statistics, specifically, the mean, the standard deviation and the correlation coefficient, on multisensory enhancement. For the impact of adaptation to the mean, an increase of the mean is always associated with an increasing CME for any input stimulus. For the standard deviation, we show that the CME produced under a suitable selection of the input stimulus exhibits robust decreases when the standard deviation is increased. For the coefficient correlation, we demonstrate that the CME of a selected bimodal input decreases consistently when the coefficient correlation becomes larger. Due to the robustness of these results, the underlying adaptation notion can be naturally verified by neurophysiological experiments.

8.2 The Role of Adaptation in Multisensory Responses

As indicated by our adaptation rule, the response of a neuron is a function of its adaptive history, and therefore a neuron will respond differently to the same, given stimulus, for different statistical histories. As a result, CME is influenced by the prior input statistics to which a neuron has adapted. We now explore the impact of adaptation to prior input statistics on the CME exhibited by a bimodal neuron.

8.2.1 The CME Measure

In order to quantify the extent of enhancement in a bimodal neuron, we employ the standard CME index defined in Eq. (2.1),

$$\%CME = \frac{r(x_1, x_2) - \max\{r(x_1, 0), r(0, x_2)\}}{\max\{r(x_1, 0), r(0, x_2)\}} \times 100\%, \tag{8.1}$$

where $r(x_1, x_2)$ represents the bimodal response, and $r(x_1, 0)$ and $r(0, x_2)$ represent the two unimodal responses, respectively. Here the spontaneous firing rates of the two

sensory channels are not considered because they do not cause any qualitative difference to the properties of the measure. Although other measures have been proposed and employed (Populin and Yin, 2002; Laurienti et al., 2005; Stanford et al., 2005; Perrault et al., 2005), most are very closely related and none presents real advantages over the others (see Chapter 6).

Assuming that input X_1 evokes the same or a larger response than input X_2, i.e. $r(x_1, 0) \geq r(0, x_2)$, then we must have that $\gamma_1 x_1 \geq \gamma_2 x_2$, or, equally, $\sigma_2 x_1 \geq \sigma_1 x_2$ according to Eq. (7.46). Since we have $\max\{r(x_1, 0), r(0, x_2)\} \equiv r(x_1, 0)$, the CME index becomes

$$\%CME = \frac{r(x_1, x_2) - r(x_1, 0)}{r(x_1, 0)} \times 100\%. \tag{8.2}$$

In order to maximize the enhancement for a fixed value of x_1, we therefore need to make x_2 as large as possible, subject to the constraint $\sigma_2 x_1 \geq \sigma_1 x_2$. Maximum CME is therefore achieved on the line $x_1\sigma_2 = x_2\sigma_1$. The same conclusion is drawn if we instead assume that X_2 evokes a larger response. This observation is required to understand many of the results presented later.

To analyze how the CME index changes under the variation of the statistical parameters, such as the mean, the variance and the correlation coefficient, we differentiate the expression of the CME index with respect to these parameters. For this, we firstly rewrite Eq. (7.46) as

$$r(x_1, x_2) = \frac{s}{2}[1 + \tanh 2\Gamma(z - \Theta)], \tag{8.3}$$

where

$$z = \frac{1}{\sqrt{2(1+\rho)}} \left(\frac{x_1 - \mu_1}{\sigma_1} + \frac{x_2 - \mu_2}{\sigma_2} \right). \tag{8.4}$$

Denote $p \in \{\mu_i, \sigma_i, \rho\}$ where $i = 1, 2$, we have

$$\frac{\partial r}{\partial p} = s\Gamma \operatorname{sech}^2 2\Gamma(z - \Theta) \frac{\partial z}{\partial p}. \tag{8.5}$$

Since we have

$$\begin{aligned} \frac{s^2}{4}\operatorname{sech}^2\theta &= \left[\frac{s}{2}(1 + \tanh\theta)\right] \left[\frac{s}{2}(1 - \tanh\theta)\right] \\ &= \left[\frac{s}{2}(1 + \tanh\theta)\right] \left[s - \frac{s}{2}(1 + \tanh\theta)\right], \end{aligned} \tag{8.6}$$

Eq. (8.5) can be transformed into

$$\frac{\partial r}{\partial p} = \frac{4\Gamma}{s} r(s - r) \frac{\partial z}{\partial p}. \tag{8.7}$$

We also denote bimodal response $B = r(x_1, x_2)$, unimodal responses $U_1 = r(x_1, 0)$ and $U_2 = r(0, x_2)$ with

$$z_B = \frac{1}{\sqrt{2(1+\rho)}} \left(\frac{x_1 - \mu_1}{\sigma_1} + \frac{x_2 - \mu_2}{\sigma_2} \right), \tag{8.8}$$

$$z_1 = \frac{1}{\sqrt{2(1+\rho)}}(\frac{x_1-\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2}), \tag{8.9}$$

and

$$z_2 = \frac{1}{\sqrt{2(1+\rho)}}(-\frac{\mu_1}{\sigma_1} + \frac{x_2-\mu_2}{\sigma_2}) \tag{8.10}$$

respectively. According to Eq. (8.7), we then obtain

$$\frac{\partial B}{\partial p} = \frac{4\Gamma}{s}B(s-B)\frac{\partial z_B}{\partial p}, \tag{8.11}$$

$$\frac{\partial U_1}{\partial p} = \frac{4\Gamma}{s}U_1(s-U_1)\frac{\partial z_1}{\partial p} \tag{8.12}$$

and

$$\frac{\partial U_2}{\partial p} = \frac{4\Gamma}{s}U_2(s-U_2)\frac{\partial z_2}{\partial p}, \tag{8.13}$$

where $p \in \{\mu_i, \sigma_i, \rho\}$.

To study the two CME indices for $U_1 \geq U_2$ and $U_1 < U_2$, we consider $C_1 = \frac{B}{U_1}$ and $C_2 = \frac{B}{U_2}$ respectively. Since we are concerned with how the CME index changes with respect to the statistical parameters, we differentiate C_1 and C_2 with respect to p and obtain

$$\frac{\partial C_1}{\partial p} = \frac{U_1\frac{\partial B}{\partial p} - B\frac{\partial U_1}{\partial p}}{U_1^2} \tag{8.14}$$

and

$$\frac{\partial C_2}{\partial p} = \frac{U_2\frac{\partial B}{\partial p} - B\frac{\partial U_2}{\partial p}}{U_2^2}. \tag{8.15}$$

Accordingly, the signs of

$$N_1 = U_1\frac{\partial B}{\partial p} - B\frac{\partial U_1}{\partial p} \tag{8.16}$$

and

$$N_2 = U_2\frac{\partial B}{\partial p} - B\frac{\partial U_2}{\partial p} \tag{8.17}$$

determine whether the CME indices increase or decrease. In the following sections, we thus consider N_1 and N_2 when p refers to the mean μ_1, the standard deviation σ_1 and the correlation coefficient ρ.

8.2.2 General Properties of CME under the Adaptation Rule

We have shown in Chapter 6 that a sigmoidal function is capable of producing CME associated with the property of IE. We here demonstrate the same properties when the sigmoidal response function of a neuron is adjusted under our adaptation rule. According to Eqs. (7.46), (7.44) and (7.45), changes in the constant Θ shift the response threshold and changes in the constant Γ affect the slope. Since the choice of these two constants does not qualitatively influence the results, we set them as $\Theta = 0.5$ and $\Gamma = 0.4$ for

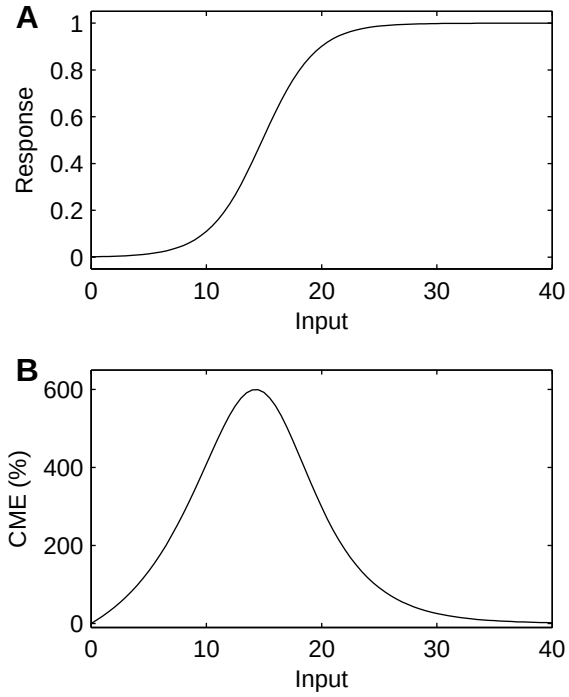


FIGURE 8.1: The neuronal response and CME as a function of the bimodal input stimulus for $x_1 = x_2$. The preferred operating point of the neuron is set to $\Theta = 0.5$ and $\Gamma = 0.4$. The input means are set by $\mu_1 = 15$ and $\mu_2 = 10$, the standard deviations by $\sigma_1 = 4$ and $\sigma_2 = 6$, and the correlation coefficient by $\rho = 0.2$. (A) The neuronal response curve. (B) The CME curve.

concreteness. In addition, because the results are also robust to the setting of the statistical parameters, we choose some typical values and set the means as $\mu_1 = 15$ and $\mu_2 = 10$, the standard deviations as $\sigma_1 = 4$ and $\sigma_2 = 6$, and the correlation coefficient as $\rho = 0.2$. The scale s is set without loss of generality as $s = 1$. According to Eqs. (7.44) and (7.45), we then have $\gamma_1 \approx 0.0645$, $\gamma_2 \approx 0.0430$, $\theta_U \approx 1.5986$. In Figs. 8.1A and 8.1B, we plot the curves of the response as well as the CME index with respect to equivalent inputs $x_1 = x_2$. As shown in Fig. 8.1B, the CME index for $x_1 = x_2$ reaches its maximum at around $x_1 = x_2 = 14$ and then gradually decreases as the magnitudes of x_1 and x_2 increase, in accordance with the property of IE observed in DSC neurons.

The changes in the CME index are complicated when several statistical parameters are varied at the same time. To demonstrate this, contours of the CME index in the $x_1 - x_2$ bimodal input plane are produced under the same parameters as Fig. 8.1. In Fig. 8.2A, we see a core of strongly enhancing input values, with the contours skewed by the relative difference in the standard deviations of the two inputs. It is evident from the contours that the point of maximum enhancement lies on the line $x_1\sigma_2 = x_2\sigma_1$. The "kinks" in the contours are due to the definition of the CME index switching between the larger of the two unimodal responses, leading to discontinuities in the first derivatives. Fig. 8.2B shows a new set of CME index contours produced with different statistical

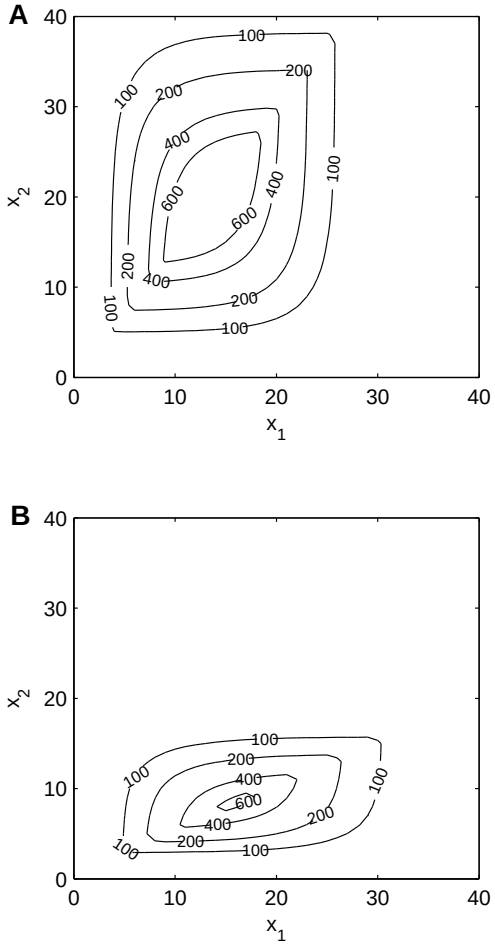


FIGURE 8.2: Two CME contour sets in the bimodal $x_1 - x_2$ input plane for differing statistical parameters. The preferred operating point of the neuron is set to $\Theta = 0.5$ and $\Gamma = 0.4$. (A) CME contours with the same statistical parameters as in Fig. 8.1. (B) CME contours with input means $\mu_1 = 15, \mu_2 = 5$, standard deviations $\sigma_1 = 8, \sigma_2 = 4$ and correlation coefficient $\rho = -0.4$.

parameters where the input means are set as $\mu_1 = 15$, $\mu_2 = 5$, the standard deviations are set as $\sigma_1 = 8$, $\sigma_2 = 4$ and the correlation coefficient is set as $\rho = -0.4$. Comparing Fig. 8.2A with Fig. 8.2B, we can see that, although the discontinuities are still along the maximum enhancement line $x_1\sigma_2 = x_2\sigma_1$, the position and appearance of the contours have changed dramatically and it is not easy to deduce the impact of input statistics on CME in this case. For clarity, therefore, we instead adopt the strategy of changing each of the input statistical parameters in turn (μ_1, σ_1 and ρ without loss of generality) and explore their individual influences on CME. Furthermore, in order to avoid clutter in figures, we specifically select a characteristic CME contour, for example, the 100% CME contour, and only examine the changes in this contour according to the variation of input statistics.

8.2.3 Impact of Adaptation to Mean on CME

Electrophysiological results have shown that retinal adaptation is regulated based on the mean level of illumination (Shapley and Enroth-Cugell, 1984). Similarly, in Eqs. (7.44) and (7.45), we can see that a neuron adapts its threshold according to the inputs' means. As a result, the two input means affect both the neuron's unimodal and bimodal responses, and consequently the CME measure as well. Here we present explicit results to demonstrate this effect.

Consider Eqs. (8.16) and (8.17) when $p = \mu_1$. To calculate N_1^μ for Eq. (8.16), we first obtain

$$\frac{\partial z_B}{\partial \mu_1} = \frac{\partial z_1}{\partial \mu_1} = \frac{\partial z_2}{\partial \mu_1} = -\frac{1}{\sigma_1\sqrt{2(1+\rho)}}. \tag{8.18}$$

Substituting Eqs. (8.11) and (8.12) for $p = \mu_1$ into Eq. (8.16), we get

$$\begin{aligned} N_1^\mu &= \frac{4\Gamma}{s}\left[U_1 B(s-B)\frac{\partial z_B}{\partial \mu_1} - U_1 B(s-U_1)\frac{\partial z_B}{\partial \mu_1}\right] \\ &= -\frac{4\Gamma}{s\sigma_1\sqrt{2(1+\rho)}} U_1 B(s-B-s+U_1) \\ &= \frac{4\Gamma}{s\sigma_1\sqrt{2(1+\rho)}} U_1 B(B-U_1). \end{aligned} \tag{8.19}$$

Similarly, we can obtain the formula for N_2^μ,

$$N_2^\mu = \frac{4\Gamma}{s\sigma_2\sqrt{2(1+\rho)}} U_2 B(B-U_2). \tag{8.20}$$

Since $B \geq U_1$ and $B \geq U_2$, we have $N_1^\mu \geq 0$ and $N_2^\mu \geq 0$. Therefore, CME always increases when μ_1 increases.

Since the choices of statistical parameters do not qualitatively influence the results below for adaptation to the mean μ_1, we set the other parameters as earlier in Fig. 8.1. In Fig. 8.3, we show three contours representing a CME index of 100% generated for three values of μ_1 corresponding to $\mu_1 = 10$, 15 and 20. From the figure, we can see an expanding contour as μ_1 increases. As shown in Fig. 8.2, the closer a point in the $x_1 - x_2$ input plane is to the centre of the core of the contours, the larger the CME index it induces. An expanding contour therefore indicates that every point in the $x_1 - x_2$ input plane has an increasing CME index as μ_1 increases. This corresponds to the above analysis. According to Eq. (7.46), in which the mean μ_1 affects only the threshold θ_U, as μ_1 increases, a fixed bimodal input stimulus is moved into the inverse effectiveness region of the sigmoidal function where larger enhancement is induced and thus the enhancement increases. Consequently, the contour representing the same value of the CME index expands.

We now show how adaptation to the mean μ_1 affects CME in a bimodal neuron for

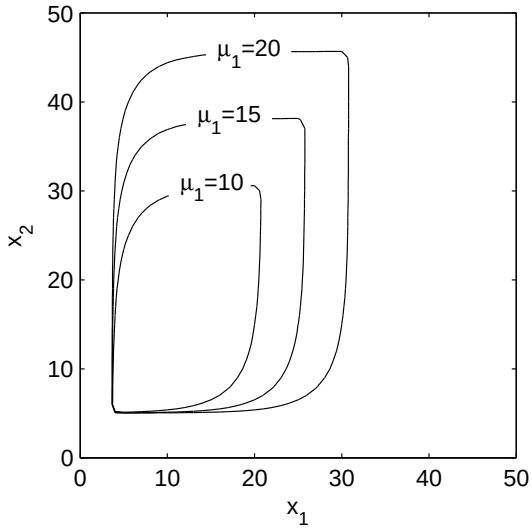


FIGURE 8.3: Contours representing a CME index of 100% in the bimodal input plane $x_1 - x_2$ for $\mu_1 = 10$, 15 and 20. The other parameters are set the same as in Fig. 8.1.

a given input stimulus. We first show how the CME index changes when a bimodal neuron has adapted to different μ_1 for an arbitrarily selected input stimulus. As shown in Fig. 8.4, we see that the CME index consistently increases as μ_1 increases from 0 to 25 for a bimodal stimulus $x_1 = 25, x_2 = 12$. To demonstrate that the same changes in the CME index are observed for other bimodal stimuli in the $x_1 - x_2$ input plane, we select four typical points corresponding to fixed values of the input stimuli x_1 and x_2, and produce the corresponding CME curves as μ_1 is increased. These four points are selected based on a comparison with the initial means to which the neuron has adapted (in this case, the initial means are $\mu_1 = 15$ and $\mu_2 = 10$). They represent four bimodal stimuli, one where x_1 is comparable to the initial mean $\mu_1 = 15$ and x_2 is comparable to the initial mean $\mu_2 = 10$ (a "mean-mean" bimodal stimulus); where x_1 is small compared with μ_1 and x_2 is large compared with μ_2 (a "small-large" bimodal stimulus); where both x_1 is large compared with μ_1 and x_2 is large compared with μ_2 (a "large-large" bimodal stimulus); where x_1 is large compared with μ_1 and x_2 is small compared with μ_2 (a "large-small" bimodal stimulus). We consider a mean-mean bimodal stimulus instead of a "small-small" bimodal stimulus because it is clear that when the magnitude of an input stimulus is much lower than the mean ($x_1 \ll \mu_1 = 15$ and $x_2 \ll \mu_2 = 10$), both the unimodal and bimodal responses to this stimulus vanish and are therefore experimentally undetectable. As shown in Fig. 8.5, all four bimodal input stimuli have an increasing CME when μ_1 is increased, demonstrating that this dependence of CME on adaptation is robust and independent of the input stimulus. Among the four typical examples of bimodal stimuli shown in Fig. 8.5, the mean-mean bimodal stimulus is preferable in an experiment because it produces evident CME for every μ_1 and the changes in CME are significant (Fig. 8.5A).

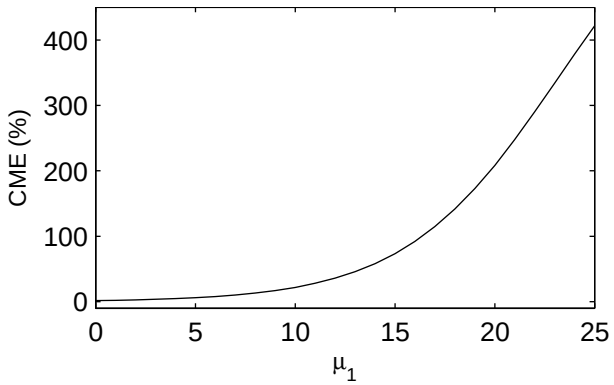


FIGURE 8.4: CME as a function of μ_1 for an arbitrarily selected bimodal stimulus given by $x_1 = 25, x_2 = 12$. The other parameters are set the same as in Fig. 8.1.

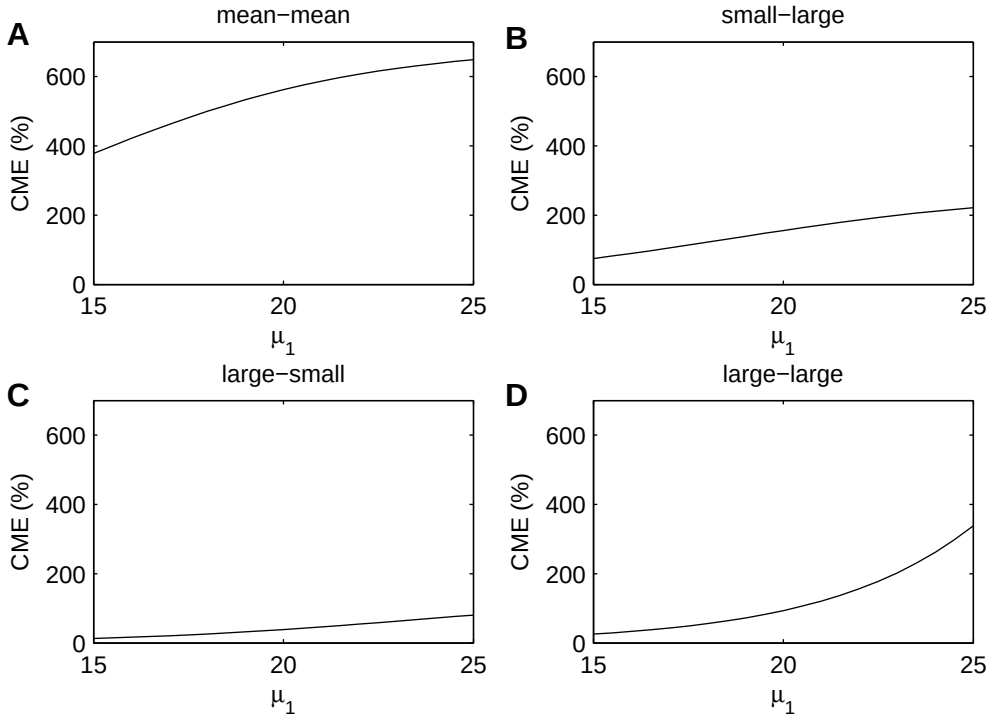


FIGURE 8.5: CME as a function of μ_1 for four typical bimodal stimuli. The other parameters are set as Fig. 8.1. The four bimodal stimuli are: (A) a mean-mean bimodal stimulus, $x_1 = 16, x_2 = 12$; (B) a small-large bimodal stimulus, $x_1 = 5, x_2 = 35$; (C) a large-small bimodal stimulus, $x_1 = 30, x_2 = 5$; (D) a large-large stimulus, $x_1 = 30, x_2 = 35$. Here "large" and "small" refer to a comparison with the initial means $\mu_1 = 15$ and $\mu_2 = 10$ for inputs x_1 and x_2, respectively. A mean-mean bimodal stimulus refers to one with magnitudes comparable to the initial means $\mu_1 = 15$ and $\mu_2 = 10$.

8.2.4 Impact of Adaptation to Standard Deviation on CME

Contrast gain control, demonstrating adaptation to the stimulus standard deviation, has been widely observed in the visual system (Truchard et al., 2000; Chander and Chichilnisky, 2001; Baccus and Meister, 2002; Solomon et al., 2004; Bonin et al., 2006). Here, we suggest that multisensory neurons adapt their modality-specific gains based on multimodal input statistics including the standard deviation of the sensory channels [Eqs. (7.44) and (7.45)]. We thus now explore the potential role that adaptation to the stimuli's standard deviations plays in multisensory integration of DSC neurons.

Consider Eqs. (8.16) and (8.17) when $p = \sigma_1$. Similar to the case of adaptation to the mean, we first derive

$$\frac{\partial z_B}{\partial \sigma_1} = \frac{\partial z_1}{\partial \sigma_1} = -\frac{x_1 - \mu_1}{\sigma_1^2\sqrt{2(1+\rho)}} \tag{8.21}$$

and

$$\frac{\partial z_2}{\partial \sigma_1} = \frac{\mu_1}{\sigma_1^2\sqrt{2(1+\rho)}}. \tag{8.22}$$

We then consider the sign of N_1^σ when $U_1 \geq U_2$, or equally, $x_2 \leq \frac{\sigma_2}{\sigma_1}x_1$. Substituting Eqs. (8.11) and (8.12) for $p = \sigma_1$ into Eq. (8.16), we have

$$\begin{aligned} N_1^\sigma &= \frac{4\Gamma}{s}U_1 B\left[(s-B)\frac{\partial z_B}{\partial \sigma_1} - (s-U_1)\frac{\partial z_1}{\partial \sigma_1}\right] \\ &= U_1 B\frac{4\Gamma}{s\sigma_1^2\sqrt{2(1+\rho)}}(B-U_1)(x_1-\mu_1). \end{aligned} \tag{8.23}$$

It is obvious that the sign of N_1^σ depends on $x_1 - \mu_1$. If $x_1 < \mu_1$, we have $N_1^\sigma < 0$ and CME decreases, and if $x_1 \geq \mu_1$, we have $N_1^\sigma > 0$ and CME increases, subject to $x_2 \leq \frac{\sigma_2}{\sigma_1}x_1$. On the other hand, if $U_1 < U_2$, or when x_2 exceeds $\frac{\sigma_2}{\sigma_1}x_1$, we need to consider N_2^σ. Substituting Eqs. (8.11) and (8.13) into Eq. (8.17), we obtain

$$\begin{aligned} N_2^\sigma &= \frac{4\Gamma}{s}U_2 B\left[(s-B)\frac{\partial z_B}{\partial \sigma_1} - (s-U_2)\frac{\partial z_2}{\partial \sigma_1}\right] \\ &= \frac{4\Gamma}{s\sigma_1^2\sqrt{2(1+\rho)}}U_2 B\left[-(s-B)(x_1-\mu_1) - (s-U_2)\mu_1\right] \\ &= \frac{4\Gamma}{s\sigma_1^2\sqrt{2(1+\rho)}}U_2 B\left[(B-s)x_1 - (B-U_2)\mu_1\right]. \end{aligned} \tag{8.24}$$

According to Eq. (8.24), the sign of N_2^σ depends on the linear function $f(x_1) = (B-s)x_1 - (B-U_2)\mu_1$. When $x_1 = 0$, we have $f(0) = -(B-U_2)\mu_1 \leq 0$. Moreover, since we have the slope $B - s \leq 0$, $f(x_1)$ decreases monotonically when x_1 increases. We therefore know that the linear function $f(x_1) \leq 0$ for $x_1 \geq 0$. Accordingly, we have $N_2^\sigma \leq 0$, indicating that the CME index decreases when $x_2 \geq \frac{\sigma_2}{\sigma_1}x_1$. As a result, we can characterize the dynamics of the changes in CME with respect to σ_1 into three regions. Fig. 8.6 shows these three regions I, II and III separated by the two lines $x_2 = \frac{\sigma_2}{\sigma_1}x_1$ and $x_1 = \mu_1$ with $x_2 \leq \frac{\sigma_2}{\sigma_1}x_1$.

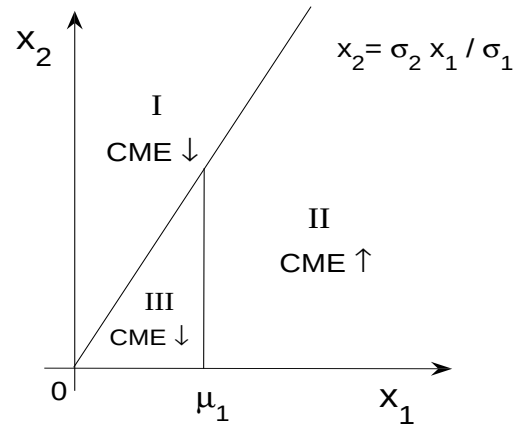


FIGURE 8.6: Three regions that characterize the dynamics of changes in CME with respect to σ_1.

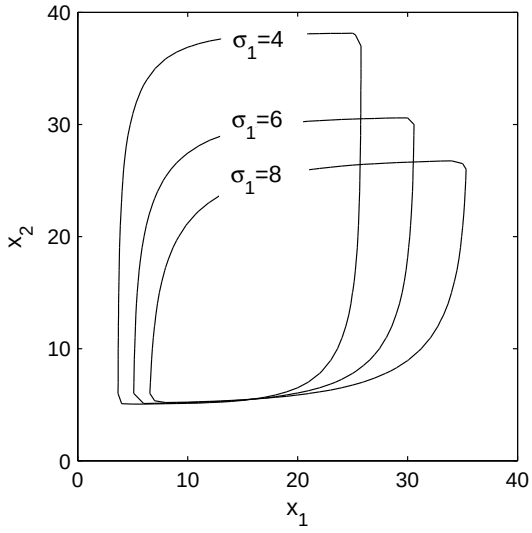


FIGURE 8.7: Contours representing a CME index of 100% in the bimodal input plane $x_1 - x_2$ for $\sigma_1 = 4$, 6 and 8. The other parameters are set the same as in Fig. 8.1.

We now show three contours representing a CME index of 100% for standard deviation $\sigma_1 = 4$, 6 and 8 in Fig. 8.7, with other parameters set as earlier in Fig. 8.1. The contour corresponding to 100% CME index rotates clockwise with the increase in σ_1. This property can be understood from the position of the line $x_1\sigma_2 = x_2\sigma_1$ on which maximum CME occurs. As σ_1 increases, the slope σ_2/σ_1 of the line decreases, and thus the maximum CME line rotates clockwise, corresponding to the rotation of the contour. Furthermore, we can see that there are crossings between the three contours, indicating that the changes in the CME index due to adaptation to the standard deviation are not consistent over the whole $x_1 - x_2$ input plane, in comparison to the contours for adaptation to the means in Fig. 8.3. These properties correspond to the three regions in Fig. 8.6.

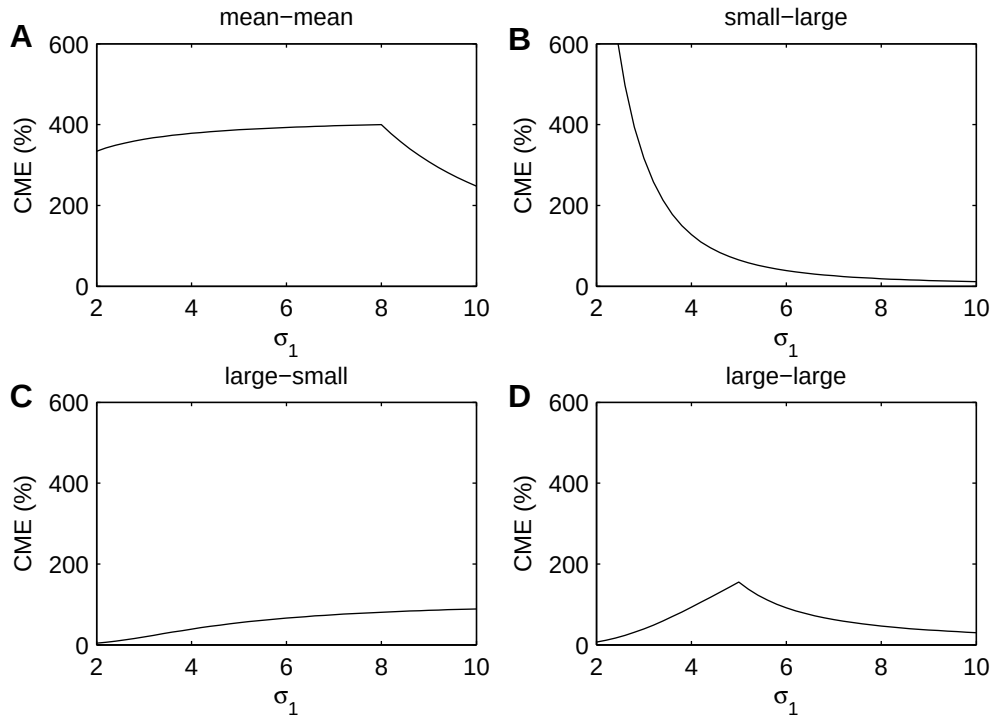


FIGURE 8.8: CME as a function of σ_1 for four typical bimodal stimuli. The other parameters are set as in Fig. 8.1. The four bimodal stimuli are: (A) a mean-mean bimodal stimulus, $x_1 = 16, x_2 = 12$; (B) a small-large bimodal stimulus, $x_1 = 5, x_2 = 30$; (C) a large-small bimodal stimulus, $x_1 = 25, x_2 = 5$; (D) a large-large bimodal stimulus, $x_1 = 25, x_2 = 30$. The meanings of these four bimodal stimuli are the same as indicated in the caption of Fig. 8.5.

To show the dynamics of the changes in CME induced by an increase in the standard deviation σ_1, we again produce the CME curves as a function of σ_1 for four typical points in the $x_1 - x_2$ input plane including a mean-mean bimodal stimulus ($x_1 = 16, x_2 = 12$); a small-large bimodal stimulus ($x_1 = 5, x_2 = 30$); a large-large bimodal stimulus ($x_1 = 25, x_2 = 30$); a large-small bimodal stimulus ($x_1 = 25, x_2 = 5$), where the means are $\mu_1 = 15, \mu_2 = 10$. As shown in Fig. 8.8, the dynamics of CME appear to be complicated. It is therefore necessary to decide which of these stimuli would be suitable as probes to study the impact of adaptation to standard deviation on CME in neurophysiological experiments. Here the suitability of a test stimulus is determined by its ability to produce a consistent increase or decrease in CME as σ_1 is varied, as well as the robustness of the results to changes in other parameters, leading to a set of reliable predictions relatively insensitive to parameter choices.

In Fig. 8.8A, we can see that the dependence of CME on σ_1 for the mean-mean bimodal stimulus is not monotonic. CME first increases, reaching its maximum at $\sigma_1 = 8$, and then decreases afterwards. For this mean-mean bimodal stimulus, we have $x_1 = 16 > \mu_1 = 15$. However, according to Fig. 8.6, we can see that the changes in the CME index with respect to σ_1 depends on where the bimodal stimulus is in the input plane.

When the bimodal stimulus (x_1, x_2) is in region I or III, increases in σ_1 always result in the decreases in CME. However, changes in CME might be non-monotonic if (x_1, x_2) is in region II. This is because an increase in σ_1 rotates the maximum line $x_1\sigma_2 = x_2\sigma_1$ clockwise, after which (x_1, x_2) might be in region I where the CME index decreases. In fact, Fig. 8.8A corresponds to this case. Furthermore, the CME curve for the large-large bimodal stimulus reaches its maximum point at $\sigma_1 = 5$ (Fig. 8.8D), also exhibiting non-monotonic changes. Similar to the mean-mean bimodal stimulus, this is because the bimodal stimulus $x_1 = 25, x_2 = 30$ is initially in region II. Of course, for a large-large bimodal stimulus that initially in region I, we would only observe decreases in the CME index when σ_1 increases (see Fig. 8.6). As a result, the non-monotonic dependence of CME on σ_1 renders the mean-mean bimodal stimulus and the large-large bimodal stimulus unsuitable as test stimuli, as no consistent trend in CME is observed in these cases.

In contrast, for the small-large bimodal stimulus, CME consistently decreases when σ_1 increases (Fig. 8.8B). As we have discussed earlier, the maximum CME line $x_1\sigma_2 = x_2\sigma_1$ rotates clockwise when we increase σ_1. However, it is obvious from Fig. 8.6 that a small-large bimodal stimulus will always be in region I, where the CME index always decreases, independent of the rotation of the maximum line. Therefore, for the points in the region corresponding to small-large stimuli, we would expect a consistent decrease in the CME index. Conversely, we see that, for a large-small bimodal stimulus, the CME curve exhibits a robust increase with increasing standard deviation σ_1 (Fig. 8.8C). This is also obvious from Fig. 8.6 since a large-small bimodal stimulus is in region II where the CME index increases. However, when the increase in σ_1 is large enough, a large-small bimodal stimulus can possibly be moved into region I, in which the CME index decreases, resulting in non-monotonic changes in CME. Moreover, the changes in the CME index for a large-small bimodal stimulus are relatively small (Fig. 8.8C). Therefore, the appropriate test stimulus to produce robust and evident experimental results is a small-large bimodal stimulus. Analogously, if the standard deviation σ_2 is increased instead of σ_1, then the suitable test stimulus would be a large-small bimodal stimulus. In the case that σ_1 or σ_2 is decreased rather than increased, we should still select a small-large test stimulus for σ_1 and a large-small test stimulus for σ_2 to observe, in this case, a robustly increasing rather than decreasing CME.

8.2.5 Impact of Adaptation to Correlation Coefficient on CME

Recent results indicate that correlations in natural visual scenes can play a role in the adaptation of neurons in the visual pathway (Sharpee et al., 2006; Lesica et al., 2007). We suggest that this phenomenon also applies to multisensory neurons, where gain control is regulated by the correlation coefficient between two sensory channels. Accordingly, multisensory integration in multisensory neurons is affected by the correlation coefficient,

and we now study its influence on CME in bimodal DSC neurons based on the adaptation rules given in Eqs. (7.44) and (7.45).

We now consider Eqs. (8.16) and (8.17) when $p = \rho$. We also obtain

$$\frac{\partial z_i}{\partial \rho} = -\frac{1}{2(1+\rho)} z_i, \tag{8.25}$$

where $i \in \{B, 1, 2\}$. Substituting Eqs. (8.11) and (8.12) for $p = \rho$ into Eq. (8.16), we obtain

$$\begin{aligned} N_1^\rho &= \frac{4\Gamma}{s} U_1 B \left[(s-B)\frac{\partial z_B}{\partial \rho} - (s-U_1)\frac{\partial z_1}{\partial \rho} \right] \\ &= -\frac{4\Gamma U_1 B}{2s(1+\rho)} \left[(s-B) z_B - (s-U_1) z_1 \right] \\ &= -\frac{4\Gamma U_1 B}{s[2(1+\rho)]^{3/2}} \left[(s-B)\left(\frac{x_1-\mu_1}{\sigma_1} + \frac{x_2-\mu_2}{\sigma_2} \right) - (s-U_1)\left(\frac{x_1-\mu_1}{\sigma_1} - \frac{\mu_2}{\sigma_2} \right) \right] \\ &= \frac{4\Gamma U_1 B}{s[2(1+\rho)]^{3/2}} \left[(B-U_1)\left(\frac{x_1-\mu_1}{\sigma_1} + \frac{x_2-\mu_2}{\sigma_2} \right) - (s-U_1)\frac{x_2}{\sigma_2} \right], \end{aligned} \tag{8.26}$$

subject to $U_1 \geq U_2$, or $x_2 \leq \frac{\sigma_2}{\sigma_1} x_1$. To find the contour for $\Delta\text{CME} = 0$, we set $N_1^\rho = 0$, which gives us

$$(B-U_1)\left(\frac{x_1-\mu_1}{\sigma_1} + \frac{x_2-\mu_2}{\sigma_2} \right) - (s-U_1)\frac{x_2}{\sigma_2} = 0. \tag{8.27}$$

The characterization of the dynamics for ΔCME with respect to ρ is not as clear as the cases of adaptation to the mean and variance. However, as an approximation, we consider when x_1 and x_2 are large enough so that $B \approx s$. Substituting this into Eq. (8.27), we can derive

$$x_1 \approx \mu_1 + \frac{\sigma_1}{\sigma_2}\mu_2 \tag{8.28}$$

for $x_2 \leq \frac{\sigma_2}{\sigma_1} x_1$. When $U_1 \leq U_2$, or $x_1 \leq \frac{\sigma_1}{\sigma_2} x_2$, we can derive N_2^ρ in an analogous way and get

$$x_2 \approx \mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1, \tag{8.29}$$

subject to $x_1 \leq \frac{\sigma_1}{\sigma_2} x_2$ for $\Delta\text{CME} = 0$. According to Eqs. (8.28) and (8.29), we can obtain a quadrilateral-shaped region in which CME decreases and outside of which CME increases after an increase in ρ (see Fig. 8.9).

To show the impact of adaptation to ρ on CME, we again set the other parameters as earlier in Fig. 8.1. Fig. 8.10 shows the three contours representing a CME index of 100% when ρ is set to -0.3, 0.2 and 0.7. We can see that there are crossings between the three 100% CME contours, in comparison to the case of adaptation to the mean. As in the investigation of adaptation to the mean and the standard deviation, we again study in the $x_1 - x_2$ input plane four typical probe stimuli representing a mean-mean bimodal stimulus, a small-large bimodal stimulus, a large-large bimodal stimulus and a

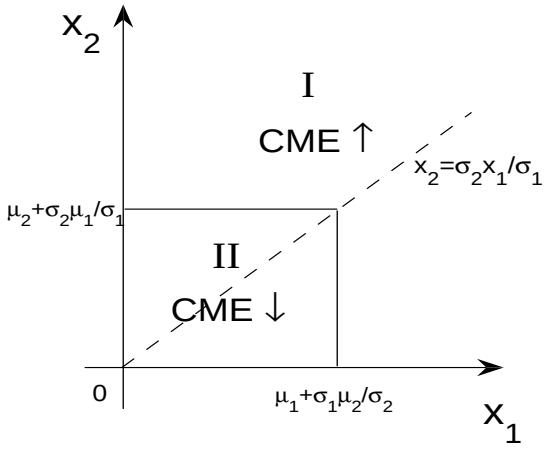


FIGURE 8.9: Two regions that approximately characterize the dynamics of changes in CME with respect to ρ.

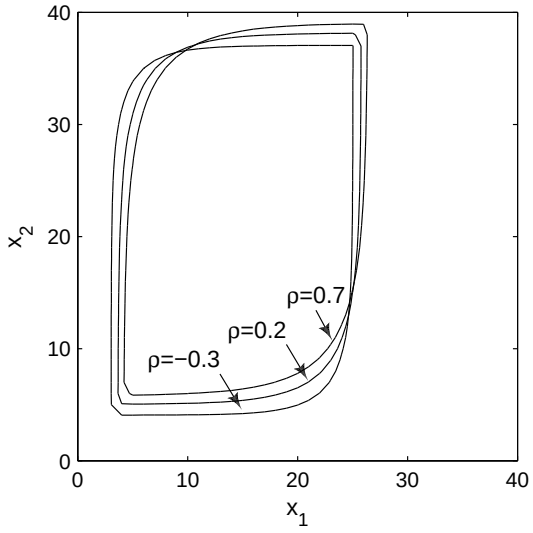


FIGURE 8.10: Contours representing a CME index of 100% in the bimodal input plane $x_1 - x_2$ for $\rho = -0.3$, 0.2 and 0.7. The other parameters are set the same as in Fig. 8.1.

large-small bimodal stimulus.

In the case of a small-large bimodal stimulus, the changes in CME are not consistent. We select here two small-large bimodal stimuli $x_1 = 5, x_2 = 25$ and $x_1 = 8, x_2 = 40$ and present the CME curves as a function of ρ (Fig. 8.11). For $x_1 = 5, x_2 = 25$, we observe a significant decrease in CME (Fig. 8.11A). However, for $x_1 = 8, x_2 = 40$, CME increases, although only slightly, when ρ becomes larger (Fig. 8.11B). As shown in Fig. 8.9, a small-large bimodal stimulus can be either in the region I or the region II, in which the changes in the CME index are different.

For a large-small bimodal stimulus, we would expect inconsistent changes in CME because it can also be in region I or region II (see Fig. 8.9). We select two large-small

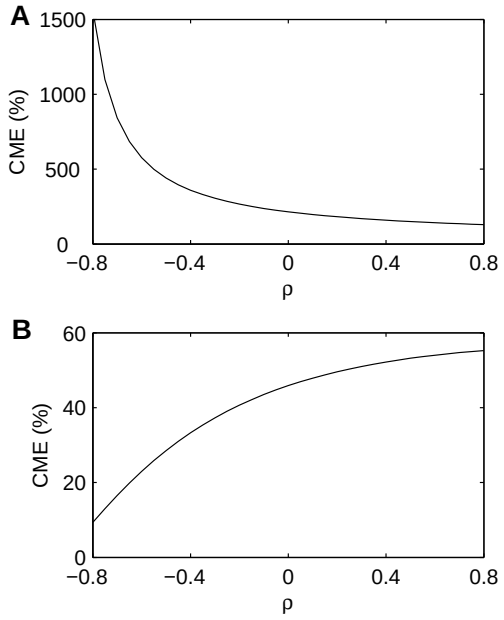


FIGURE 8.11: CME as a function of ρ for two small-large bimodal stimuli. The other parameters are set the same as in Fig. 8.1. Note the differences in CME scales. The two bimodal stimuli are given by: (A) $x_1 = 5, x_2 = 25$; (B) $x_1 = 8, x_2 = 40$. The meaning of the "small-large" bimodal stimuli is the same as indicated in the caption of Fig. 8.5.

bimodal stimuli $x_1 = 21, x_2 = 5$ and $x_1 = 28, x_2 = 6$ and produce the corresponding CME curves as a function of ρ in Fig. 8.12. For $x_1 = 21, x_2 = 5$, CME decreases when ρ is increased (Fig. 8.12A). For $x_1 = 28, x_2 = 6$, however, we observe an increasing CME (Fig. 8.12B). The trends in CME for these two stimuli are different, as expected.

Consider now the case for large-large bimodal stimuli. The CME curves as a function of ρ for two large-large bimodal stimuli $x_1 = 21, x_2 = 25$ and $x_1 = 25, x_2 = 30$ are shown in Fig. 8.13. In Fig. 8.13A, we see decreasing CME for the bimodal stimulus $x_1 = 21, x_2 = 25$. In Fig. 8.13B, however, we observe slightly increasing CME for $x_1 = 25, x_2 = 30$. This is again obvious from Fig. 8.9 since the bimodal stimulus $x_1 = 21, x_2 = 25$ is in region II and the bimodal stimulus $x_1 = 25, x_2 = 30$ is in region I.

The small-large, large-small and large-large bimodal stimuli are likely to produce inconsistent results and are thus not suitable as test stimuli. In contrast, mean-mean bimodal stimuli produce robust results. As an example, we choose two fixed, bimodal test stimuli, $x_1 = 14, x_2 = 10$ and $x_1 = 17, x_2 = 12$, and show their CME curves with respect to the changes of ρ. We can see that the CME index consistently decreases for these two stimuli (Fig. 8.14). This is because a mean-mean bimodal stimulus is guaranteed to be in region II, in which the CME index always decreases. In an experiment to study the influence of adaptation to the correlation coefficient on CME, we should therefore choose a mean-mean bimodal test stimulus.

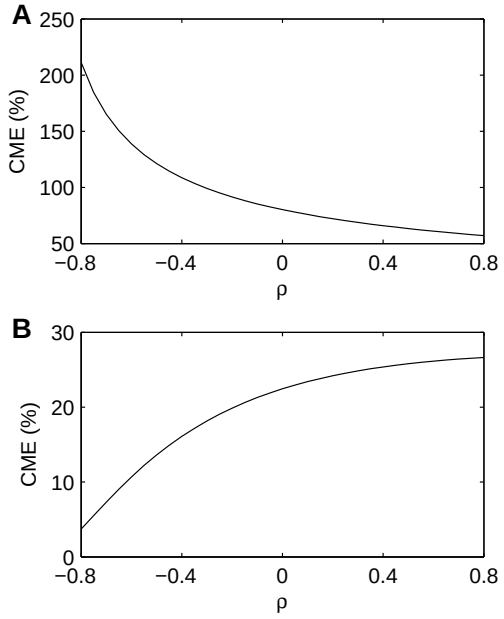


FIGURE 8.12: CME as a function of ρ for two large-small bimodal stimuli. The other parameters are set the same as in Fig. 8.1. Note the differences in CME scales. The two bimodal stimuli are given by: (A) $x_1 = 21, x_2 = 5$; (B) $x_1 = 28, x_2 = 6$. The meaning of the "large-small" bimodal stimuli is the same as indicated in the caption of Fig. 8.5.

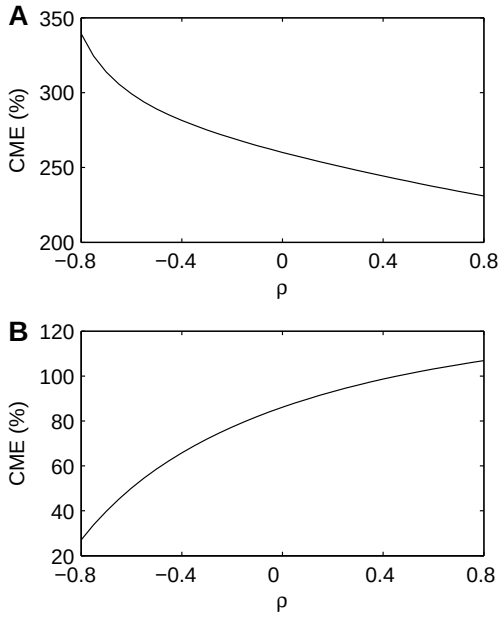


FIGURE 8.13: CME as a function of ρ for two large-large bimodal stimuli. The other parameters are set the same as in Fig. 8.1. Note the differences in CME scales. The two bimodal stimuli are given by: (A) $x_1 = 21, x_2 = 25$; (B) $x_1 = 25, x_2 = 30$. The meaning of the "large-large" bimodal stimuli is the same as indicated in the caption of Fig. 8.5.

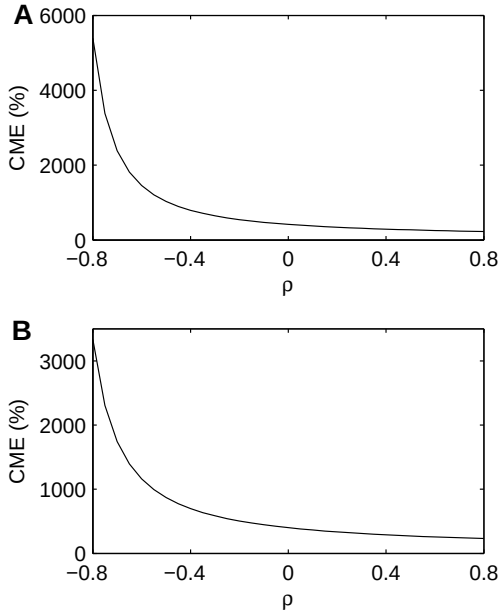


FIGURE 8.14: CME as a function of ρ for two mean-mean bimodal stimuli. The other parameters are set the same as in Fig. 8.1. Note the differences in CME scales. The two bimodal stimuli are given by: (A) $x_1 = 14, x_2 = 10$; (B) $x_1 = 17, x_2 = 12$. The meaning of the "mean-mean" bimodal stimuli is the same as indicated in the caption of Fig. 8.5.

8.3 Summary

Multisensory neurons in the deep layers of superior colliculus (DSC) integrate inputs from different sensory modalities. The response of a DSC multisensory neuron to a stimulus from a sensory channel is augmented in the presence of a stimulus from another sensory channel. We have suggested in Chapter 7 that DSC multisensory neurons adapt their responses according to the statistical properties of the multimodal inputs through processes such as gain control and firing threshold adjustment. Based on the simple adaptation rule proposed in Chapter 7 for the responses of DSC multisensory neurons, we showed in this chapter how multisensory enhancement is affected by changes in the input statistics to which a DSC neuron has adapted. We demonstrated, for example, that an increase in the correlation coefficient robustly leads to a decrease in multisensory enhancement for a suitable fixed, bimodal stimulus.

Chapter 9

Discussion

9.1 A Model of Multisensory Integration and its Extension

In this thesis, we have concentrated on the theoretical study of multisensory integration in DSC neurons. A large number of experiments have been conducted to explore the responses of DSC neurons, discovering properties such as cross-modal enhancement (CME), multisensory depression and modality-specific suppression (MSS) (Stein and Meredith, 1993; Stein et al., 2004). In order to account for these reported response properties, computational and mathematical models have also been proposed (Patton and Anastasio, 2003; Anastasio and Patton, 2003; Schauer and Gross, 2004; Cuppini et al., 2007). Despite extensive experimental and theoretical studies, the mechanisms underlying multisensory integration remain unclear.

We started by analyzing a representative model proposed by Patton and Anastasio (2003) to explain the CME and MSS properties of the responses of DSC neurons. The central assumption of their model is that a DSC neuron computes the posterior probability of a target being present based on Bayes' rule. A neural implementation of the model has also been proposed based on an augmented perceptron for cross-modal inputs following multimodal Gaussian distribution. In the augmented perceptron, the weights and bias are determined by the statistics of the inputs, and the perceptron is shown to exhibit CME. Patton and Anastasio (2003) argued that the covariances for within-modality inputs are larger than those for cross-modal inputs, both when a target is absent (the spontaneous case) and present (the driven case). Moreover, they suggested that this larger spontaneous covariance for within-modality inputs contributes to the occurrence of MSS, in which case no evident CME is exhibited. However, we found that the model is sensitive to the choice of parameters. For example, our numerical studies showed that the model evidently exhibits both CME and MSS for inputs with different magnitudes when the spontaneous covariance is slightly less than the selected value (see Chapter 5). In order to explore this result, we carried out a mathematical analysis and

found that the MSS property is produced by the near-non-invertibility of the spontaneous covariance matrix when the spontaneous covariance approaches the variances, or equally, the spontaneous correlation coefficient approaches one. Furthermore, we managed to show that spontaneous correlation coefficient approaching one induces MSS for bimodal inputs following any distribution, under equal spontaneous means, variances and equal-valued inputs for different sensory channels, consistent with the selections in Patton and Anastasio (2003). If the inputs or the statistical parameters are set to be different, however, we can not observe MSS. It is clear that the model is sensitive to the choice of parameters and inputs and is therefore not a biologically plausible model for DSC neurons.

A new interpretation of multisensory responses of DSC neurons is thus required. There have been many neurophysiological experiments indicating that the underlying operations of DSC neurons are complicated. For example, most of the neurons are found to exhibit the CME property associated with IE, and their bimodal responses could be sub-additive, additive or super-additive, compared to their unimodal responses (Stein et al., 2004). It has also been shown that the response range for a bimodal stimulus is larger than the corresponding unimodal stimuli for the same DSC neuron, and there are four operational modes in different DSC neurons (Perrault et al., 2005). Despite this diversity, we showed in Chapter 6 that a non-linear, saturating transfer function, such as a sigmoidal function, is capable of exhibiting all these response properties, extending the notion of Stanford et al. (2005). The CME property associated with IE, for example, could be produced by the saturating property of the sigmoidal function, and the additive properties are exhibited by the approximately linear response region around the inflection point of the function. Furthermore, the different dynamic ranges between unimodal and bimodal stimuli or the four operational modes observed in DSC neurons could be a direct consequence of prior saturation in the unimodal neurons earlier in the sensory pathway (see Chapter 6). Accordingly, we adopted a sigmoidal response function in our model for DSC neurons.

After confirming the importance of a sigmoidal response function in a DSC neuron to reproduce their response properties, our next question was how the parameters of the response function are determined. Inspired by the well-established experimental results that unimodal sensory neurons adapt to the statistics of the inputs (see Chapter 3), we proposed that the response function of a DSC neuron is also adjusted according to the input statistics. We have suggested an adaptation rule for both unimodal and multimodal neurons, under which a sensory neuron adapts its response to accommodate only the mean and standard deviation of the input statistics and neglects the higher-order moments. Defining the operating point as the output spike rate probability density function, we showed that our adaptation rule maintains an invariant operating point for a large class of input distributions, including the normal distribution, the exponential distribution and the logistic distribution. For other distributions, an approximately

invariant operating point can be achieved. One consequence of this idea is that, if a neuron is, for example, originally at an operating point with maximum entropy, then by adapting according to our adaptation rule, the neuron will stay precisely at or close to the maximum entropy operating point, even if the input statistics change. We presented several examples of input distributions for a neuron functioning at different operating points, in which case their operating points are exactly or approximately maintained under our adaptation rule.

With the adaptation rule, we were then able to investigate the impact of adaptation to input statistics on multisensory integration in DSC neurons. For simplicity, we only considered the bimodal case. We presented a mathematical analysis of the influences of adaptation to different statistical parameters, including the mean, the variance and the correlation coefficient, on the CME index, dividing the two dimensional input plane into regions in which the CME index undergoes robust increases or decreases when the corresponding statistical parameter increases. This indicates that, with a suitable choice of input stimuli, we would expect to observe consistent changes in the CME index when we manipulate the statistical properties of the stimuli. For example, we showed that an increase in the correlation coefficient leads to a decrease in the CME index for a suitable fixed bimodal stimulus. Because of the robustness of these results, they can be used to verify our adaptation notion in neurophysiological experiments.

An implicit assumption of our adaptation rule is that the neuron is capable of estimating the statistical parameters such as the means, the variances and the correlation coefficient of the input stimuli. We have not explored how this estimation can be performed in a sensory neuron. Analogous estimation problems have been considered in other fields such as signal processing for time series analysis. For example, Papoulis (1991) introduced a method to obtain an optimal estimation of the non-stationary mean by minimizing the mean-square error. Implementation of this estimation, however, requires the whole history of the signals and very complex computations at each time step. This is likely to exceed the computational capability of a single neuron. Moreover, Dean et al. (2008) have shown that auditory neurons in midbrain adapt faster to an increase in the mean sound level than to a decrease, indicating the existence of an asymmetric adaptation to the mean. It is thus desirable for a biologically plausible mechanism to reproduce this adaptation property. In terms of adaptation to the variance, Smirnakis et al. (1997) also showed that the adaptation time to an increased contrast is shorter than that to a decreased contrast. To account for this asymmetric dynamics in contrast adaptation, DeWeese and Zador (1998) proposed an optimal Bayesian estimator. In this estimator, the probability $P(\sigma_i|s_{j\leq i})$ of the variance σ_i given the previous signals $s_{j\leq i}$ is obtained based on Bayes' theorem, and the estimation of σ_i can then be computed under strategies such as computing the mean of σ_i or selecting the σ_i with the maximum probability according to $P(\sigma_i|s_{j\leq i})$. However, this estimator requires a neuron to adapt to several probability distributions, such as $P(s_i|s_{j<i}, \sigma_i)$ and $P(s_i|s_{j<i})$, in a short timescale in

order to calculate $P(\sigma_i|s_{j\leq i})$. This is again probably not realistic for a neuron with limited computational capability. On the other hand, "Brown's exponential smoothing" method only requires the instantaneous input rate and is applicable to the estimation of all the moments (Brown and Meyer, 1961). Despite this simplicity, there are no studies showing its ability to reproduce the adaptation properties in neurons, such as the asymmetric dynamics in contrast adaptation. To study the way neurons estimate the statistical parameters, more work is required.

To account for multisensory depression and MSS in DSC neurons, our adaptation rule for the response function of a single neuron is not sufficient. We can take the lateral interactions between DSC neurons into consideration and extend the model to a population level. By modelling the responses of each neuron in the population with a sigmoidal response function adjusted under our proposed adaptation rule, we can expect to observe both CME and multisensory depression in the model. Moreover, such a population model can be further extended to a multi-level one to include the unisensory input layers, in which the responses of unimodal neurons follow an adaptive sigmoidal function. In this multi-level model, MSS is mediated on the unisensory layers, and we suggest that the responses of unimodal neurons are also saturated so that DSC neurons exhibit different dynamic ranges for unimodal and bimodal stimuli, in agreement with the experimental results by Perrault et al. (2005). The structure of such a multi-level population model would be similar to the neural network model proposed by Cuppini et al. (2007).

Our adaptation model is also applicable to engineering implementations. One application is to assist the orientation behaviour of a robot towards visual and auditory stimuli. In contrast to the experiments by Rucci et al. (1999), in which the neural model maintains a consistent spatial mapping between different sensory modalities under altered sensory experience, the focus of our model would be the robust performance of the robot when the statistics of the inputs change. For example, when the noise in the visual signals increase, the robot could still integrate the visual and auditory signals in a sensible way so as to maintain a good orientation performance. This is also different from the model proposed by Schauer and Gross (2004), in which the statistical properties of the inputs are not considered. Another possible application of our model is image fusion, which refers to the process of combining several images to generate a new one that is more informative than any of the original sources. Since our model weights the inputs according to their statistical properties in the integration [see Eqs. (7.44) and (7.45)], it is probably a sensible choice to fuse source images with different dynamic ranges. These applications could reveal the advantages as well as the disadvantages of the model and provide feedbacks to biological multisensory integration.

9.2 Interpretation of Multisensory Responses

CME has been suggested to characterize multisensory integration in many experiments. To model a DSC neuron that shows CME responses, we discussed in Chapter 6 that nothing more complicated than a sigmoidal response function is required. A neuron under a sigmoidal response function is capable of exhibiting CME associated with IE, which naturally emerge from the intrinsic properties of the response function, comparable to the notion of Stanford et al. (2005). For example, a stronger input stimulus is closer to the saturating region of the response function than a weaker stimulus. The increase in the response induced by the presence of another stimulus for the stronger input would therefore be less than that for the weaker input.

Additivity properties in multisensory responses of DSC neurons have also been explored. Systematic studies by Stanford et al. (2005) showed that a majority of DSC neurons exhibit additive responses, that is, their responses to cross-modal stimuli are approximately the sum of their responses to individual unimodal stimulus. As discussed in Chapter 6, the method Stanford et al. (2005) adopted to characterize the additive properties is not statistically sound. Nevertheless, their results indicated that the responses of DSC neurons could be sub-additive, additive or super-additive. In addition to the CME properties, a neuron responding under a sigmoidal response function is also able to exhibit these three kinds of responses. Since the second derivative of a sigmoidal response function at the inflection point $x = \theta$ is zero [see Eq. (6.3)], the changes in the slope around $x = \theta$ is approximately zero. Inputs that are close to $x = \theta$ thus evoke an approximately linear response, that is, the response $f(\theta + \delta x) \approx f(\theta) + \gamma \delta x$, where $\gamma = f'(\theta)$ and is the gain at semi-saturation. This region could be considered as the additive response region. For x below the additive region, the slope of the response curve increases when x increases, and the responses evoked by stimuli in this region are therefore super-additive. On the other hand, for x above the additive region, the slope of the curve decreases to zero when x increases, and the responses are thus sub-additive. In fact, we have shown in Chapter 6 that the CME and the ADD measures are inherently correlated. It is therefore not surprising that the sigmoidal response function can naturally account for both the CME and additivity response properties.

Neurophysiological experiments by Perrault et al. (2005) showed that in DSC neurons the dynamic ranges of responses to multimodal stimuli are larger than those to unimodal stimuli. This result seems to contradict the notion that the responses of DSC neurons could be modelled by a consistent saturating response function since there appears to be dynamic modulations in the saturation level of a DSC neuron for unimodal and multimodal stimuli. In fact, if we take into account possible prior saturation in the unisensory neurons, whose outputs are the inputs to DSC neurons, the smaller response range for unimodal stimuli might be the result of the saturated inputs DSC neurons receive from those unimodal sensory neurons, in which case the DSC neurons themselves

have not reached the saturation level. This interpretation could also be applied to interpret the four seemingly diverse operational modes in DSC neurons, as proposed by Perrault et al. (2005), according to which neurons are categorized into super-additive neurons, super-additive/sub-additive neurons, sub-additive neurons and additive/sub-additive neurons. These diverse response properties could be induced by prior saturation in the unimodal sensory neurons that constrains the inputs to DSC neurons within certain operational regions of the response function (see Chapter 6). For example, if a unisensory neuron has a very small response range, the DSC neuron that receives inputs from it will only respond in the super-additive region, in which case neither additive nor sub-additive responses could be observed.

Although the sigmoidal response function itself is sufficient to account for most of the reported CME and additivity properties in the responses of DSC neurons, some phenomena such as multisensory depression and MSS cannot be explained only with a sigmoidal response function. As discussed before, multisensory depression occurs when cross-modal stimuli are spatially disparate, for example, when one stimulus is within the receptive field (RF) of a neuron and the other is outside the RF (Kadunce et al., 1997). The dependence on the spatial relation between cross-modal stimuli in the response properties can be explained by the excitatory and inhibitory regions of the RFs of DSC neurons. When both the stimuli are within the excitatory region of a DSC neuron, CME occurs. However, if one stimulus is within the excitatory region while the other is in the inhibitory region, multisensory depression occurs. Neurophysiologically, the excitatory and inhibitory regions are determined by the lateral connections between DSC neurons. For example, lateral connections between a neuron and its neighbouring neurons are excitatory, while those between this neuron and neurons further away are inhibitory. This short-range excitatory and long-range inhibitory form of lateral connections is typical in neural models of the cortex. In such a structure, we can see that the mechanisms underlying CME and multisensory depression are different: CME is mainly determined by the saturating response function of a neuron, and multisensory depression is mediated by lateral inhibitions within a population of neurons. This difference in the mechanisms is also supported by the experiments exploring cortical influences from anterior ectosylvian sulcus (AES) and the rostral aspect of the lateral suprasylvian sulcus (rLS) on DSC responses, in which CME and multisensory depression exhibit different dependences on AES and rLS: deactivation of AES/rLS results in the elimination of CME, but only reduces multisensory depression (Jiang et al., 2001; Jiang and Stein, 2003).

Some experiments have also indicated that modality-specific suppression (MSS) and multisensory depression are again mediated by different mechanisms. For example, Kadunce et al. (1997) found that MSS is more often observed than multisensory depression. They observed that DSC neurons that exhibit multisensory depression also exhibit MSS. However, DSC neurons that exhibit MSS do not necessarily show multisensory depression. Moreover, experiments on AES/rLS afferents to DSC neurons also indicate that the

mechanisms underlying multisensory depression and MSS are different. Although influenced by cortical deactivation, computation in DSC neurons for MSS remains relatively the same (Alvarado et al., 2007), in contrast to the reduction in multisensory depression (Jiang and Stein, 2003). As suggested by Kadunce et al. (1997), one possibility is that MSS is mediated at the unisensory level, from which DSC neurons inherit the MSS property, while multisensory depression is, as discussed previously, mediated at the multisensory level by lateral inhibitory connections between DSC neurons. Cuppini et al. (2007) has proposed a neural network model as a first attempt to implement an analogous idea.

In conclusion, a sigmoidal response function naturally accounts for many properties relating to CME. However, to consider other properties such as multisensory depression and MSS exhibited by DSC neurons, we need to extend the model from the single cell level to the population level.

9.3 Implications for Neuronal Computation

In our adaptation rule, we have considered adaption only to the input mean and variance because the sigmoidal response function employed here provides us with two degrees of response freedom, making changes in the input mean and variance suitable targets for adaptive changes in those two parameters (see Chapters 7). Evidence from neurons in the inferior colliculus suggests that neurons may be able to adapt to moments higher than the second (Kvale and Schreiner, 2004), although evidence from neurons in the lateral geniculate nucleus suggests that contrast gain control is sensitive only to the input mean and variance (Bonin et al., 2006). In order to admit adaptation to higher-order moments in our approach, it would be necessary to consider a more general response function characterised by more parameters. Of course, to allow perfect adaptation to the potentially infinite number of independent moments of a completely general input distribution, it would be necessary to consider an essentially arbitrary response function.

The extent to which a neuron can modify the functional relationship between its input spike rate and its output spike rate is, however, unclear. A neuron can modify its threshold and gain, but the input-output mapping is presumably not arbitrarily modifiable. It is therefore unrealistic to assume, in a moment-orientated approach to adaptation, that adaptation to more than a few moments is possible. The same issues, however, apply to information-theoretic approaches. Assuming, for example, that the output response is bounded, so that maximum output entropy is achieved with a uniform output probability distribution, optimal adaptation is achieved by setting a neuron's response function proportional to the CDF of the input distribution (Laughlin, 1981). Over evolutionary time, it is possible that such a mapping has been acquired by neurons in the presence of natural stimuli. But can such neurons adapt to a rapid, dramatic change in their

input distributions? An affirmative answer requires identical assumptions, in terms of the freedoms assumed to be available in the response function characterisation, as a moment-based view.

A view of adaptation based on moments of course requires that a neuron, or a circuit, can estimate the moments of its input distribution. We have not discussed how that could be achieved, nor the timescales on which changes in moments would affect a neuron's, or a circuit's, estimates of them. It is a simple matter to build a model based on running estimates of moments requiring access only to the instantaneous input rate, and we will pursue this in future work. The higher the moment, however, the more complicated the form for the running estimate, so it is probably unrealistic to assume that a neuron can estimate more than a handful of the lowest-order moments. Equally, however, in the above information-theoretic considerations, estimation of the input CDF will be tightly constrained by the computational resources available to a neuron, or a circuit.

We have therefore restricted to a consideration of adaptation to the lowest-order moments, and in particular to the mean and variance, of an input distribution, in order to maintain an invariant output distribution. Compared to the complexity of the full problem, requiring all the moments and an essentially arbitrary response function, restricting to the two lowest-order moments produces results, in terms of adapting the neuronal threshold and gain via Eqs. (7.11) and (7.12), that are easy to derive and simple to implement. Despite the fact that these rules ensure only approximate output PDF invariance for general input distributions, and although some functional properties of neurons may then escape our analysis (e.g. large deviation detection, for which the higher-order moments are clearly critical), we regard this simplicity as a virtue. Under the assumption that a neuron can indeed estimate the mean and variance of its input, Eqs. (7.11) and (7.12) provide an adaptive strategy that does not require a neuron to perform elaborate computations in order to adapt to changes in its input statistics. If we regard neurons not as perfect optimisers but rather as devices of extremely limited computational ability and resources, then adaptation according to Eqs. (7.11) and (7.12) may very well represent a solution for suboptimal, resource-constrained computation.

For multisensory neurons in general and DSC neurons in particular, we have extended the adaptation rule in Eqs. (7.11) and (7.12) to multimodal inputs and have written down a simple model of gain control and firing threshold adjustment [Eqs. (7.44) and (7.45)]. In the derivation of Eqs. (7.44) and (7.45), we require that each sensory channel has a separate gain. Each gain is influenced by both a local contribution, based on its channel's standard deviation, and a global contribution, based on the correlation coefficients between pairs of cross-modal sensory inputs. The local contributions to the gains arising from the modality-specific standard deviations likely cannot be accommodated by whole-neuron level processes, since such processes would presumably affect all gains equally. Such local contributions therefore probably require circuit-level processes involving the sensory afferents themselves. Whole-neuron level processes could in principle accommo-

date the global, common contribution to the gains due to the correlation coefficients. However, the estimation of correlation coefficients must be a non-local process, because they require a computation of pairwise activities between differing sensory channels. Thus, it is probable that both the local and global contributions to the gains require circuit level processes.

What might these circuits be? Data indicate that descending cortical projections from the AES and rLS are critically involved in multisensory integration in DSC neurons. In particular, when these two cortical areas are deactivated by cooling, CME is reversibly eliminated while unimodal responses, although perhaps somewhat attenuated, remain largely intact (Wallace and Stein, 2000; Jiang et al., 2001; Stein, 2005; Jiang et al., 2006; Alvarado et al., 2007). Furthermore, electron microscope studies suggest that there is a close juxtaposition of the descending, cortical projections with the ascending, sensory inputs on the dendrites of DSC neurons (Harting et al., 1997). Such a juxtaposition could underlie the computation of the correlation coefficients between different sensory channels' activities. Based on these data, we therefore speculate that the anatomical architecture of converging ascending and descending pathways on DSC neurons is organised so as to control the gains of the distinct, sensory pathways, and that cortical areas may participate in gain control in DSC neurons.

9.4 Model Predictions for Experimental Tests

We have examined in Chapter 8 the possible impact of adaptation to the input statistics on multisensory integration of DSC neurons under a simple adaptation rule. According to the adaptation rule, the response of a DSC neuron is affected by the input statistics to which the neuron has adapted. If we then consider probing a DSC neuron with a fixed and invariant stimulus to obtain a CME measure, then that measure can be systematically changed by prior adaptation of the neuron to different multimodal input statistics. For simplicity, we have considered in Chapter 8 separate adaptations to the stimulus means, the stimulus standard deviations, and the stimulus correlation coefficients.

Adaptation to the means is perhaps the simplest form of adaptation. Changes in the means affect the threshold and thus shift the response function. In an experiment to study adaptation to the mean in a DSC neuron, a fixed, probe stimulus can be moved between the neuron's different functional regimes, from super-additive to additive to sub-additive (Stanford et al., 2005; Avillac et al., 2007), depending on changes in the neuron's threshold induced by adaptation. Our results indicate that when an input's mean is increased, the corresponding CME always increases, independent of the other parameters and the test stimulus. Despite the robust changes in CME for arbitrary test stimuli, however, a suitable choice is a bimodal probe stimulus with strength comparable to the initial mean of the prior input statistics before being changed (a mean-mean

type bimodal stimulus). This type of bimodal stimulus guarantees the generation of detectable responses and evident CME.

Adaptation to the stimulus variances is also possible, although in this case the impact on CME is rather more complicated than that of the means. Despite these complex dynamics, we observe that bimodal stimuli given by input x_1 small compared to the mean μ_1, and input x_2 large compared to the mean μ_2 (a small-large type bimodal stimulus) exhibit consistent and clear decreases in the CME index when σ_1 is increased. According to the adaptation rule, the maximum CME index lies on the line $x_1\sigma_2 = x_2\sigma_1$. The $x_1 - x_2$ input plane can be divided by this maximum CME line and the line $x_1 = \mu_1$ with $x_2 < \frac{\sigma_2}{\sigma_1}x_1$ into three regions, in which the CME index changes consistently. A small-large bimodal stimulus is guaranteed to be in the region where CME decreases under an increase in σ_1, and it therefore results in a robust change in the CME index. Conversely, if σ_2 is instead increased, then the preferred test stimulus producing consistent and evident changes in the CME index is a large-small type bimodal stimulus.

It is possible that adaptation to the stimulus means and variances is not, in fact, observed in DSC neurons, if neurons earlier in the multiple sensory pathways have already adapted to these statistics. Perhaps, therefore, the most intriguing and robust form of adaptation is adaptation to changes in the correlation coefficient between a pair of sensory modalities, with all other statistics held constant. In this case, prior adaptation is unlikely, since the superior colliculus is one of the first sensory areas in which multimodal neurons exist. An experimental exploration of the possibility of adaptations to the various possible correlation coefficients therefore affords an opportunity for robust testing of our proposals without the possibility of contamination by prior adaptation in earlier sensory centres.

As with the stimulus variances, adaptation to the correlation coefficient has a complicated influence on CME. However, for a mean-mean type bimodal stimulus, increasing the correlation coefficient robustly decreases the CME exhibited by a multisensory neuron. This may initially seem counter-intuitive, but it is important not to confuse the (near-)simultaneity of presentation of two stimuli in different modalities, which evokes CME, with the statistical parameter that measures how often, on average, such bimodal stimuli are presented simultaneously. When the correlation coefficient is high (close to $\rho = 1$, say), two stimuli are nearly always presented simultaneously. In this case, one of the two channels is essentially redundant, from an information-theoretic perspective. Although CME still occurs in this case, we predict that the extent of CME is lower in well-correlated, and therefore more redundant sensory channels than in less well-correlated, and therefore less redundant sensory channels.

The predications obtained here for the possible role of adaptation to the input statistics on multisensory integration can be tested naturally in experiments. In such experiments, we would firstly let an animal adapt to a certain statistical context by providing appro-

priate stimulation. After that, fixed and invariant unimodal and bimodal test stimuli are presented in the receptive field of a DSC neuron, and unimodal and bimodal responses of this DSC neuron are recorded to produce the CME measure. To allow the animal to adapt to a new statistical context, another stimulation is provided. After sufficiently long period for the animal to finish adaptation, the same test stimuli are presented and the corresponding CME measure for the same DSC neuron is calculated. The same procedure can be iterated for several times, after which the changes of CME with respect to the changes of the statistics could be obtained. During the course of such an experiment on a single neuron, in which its multimodal receptive field is being probed with test stimuli, the neuron may adapt to these test stimuli, producing CME measures at the end of the experiment that differ from those at the beginning. Although we have not committed ourselves to a view about the timescale for adaptation in DSC neurons, it seems reasonable to assume that it is of the order of seconds to minutes, as in other sensory systems (Smirnakis et al., 1997). Such timescales are well within the periods for which single neurons may be held during electrode penetrations, and DSC neurons could thus easily adapt to different stimulation protocols several times during an experiment. Careful control of the protocols is therefore required in order to avoid any unwanted adaptation effects during experiments that seek to determine the mechanisms underlying multisensory integration. Equally, in order to observe adaptation in DSC neurons, carefully controlling especially the various possible correlation coefficients is critical.

9.5 Reflections on Experimental and Theoretical Studies

We have proposed in this study an adaptation rule for a sigmoidal transfer function to account for multisensory responses of DSC neurons. We have shown that, although simple, our model is capable of explaining many observed multisensory response properties in a natural way. This simple interpretation of the mechanisms underlying multisensory integration is in contrast to many experimental studies stating that complicated mechanisms and operations are involved.

It therefore appears to us that, although some experiments on the responses of DSC neurons have been set up and carried out in a sensible way to obtain valuable data on DSC neurons, these data are often not processed or interpreted in an appropriate way. For example, Stanford et al. (2005) have conducted a systematic recording of the responses of DSC neurons. However, their statistical method to process the data produces results that are critically dependent on one of the procedures (see Chapter 6). Consequently, their conclusion does not hold. This is due to insufficient understanding of the underlying statistical principles. In Perrault et al. (2005)'s studies, they observed different dynamic ranges for unimodal and bimodal responses in DSC neurons, based on which they directly suggested that the response ranges of DSC neurons might be under

a dynamic modulation for unimodal and bimodal stimuli. This seemingly straightforward interpretation indicates the existence of new and complicated operations in DSC neurons compared to other sensory neurons, which, although not impossible, might not be appropriate. In fact, with a careful consideration, we find that, to explain the same phenomenon, no elaborate mechanisms are required for DSC neurons. We only need to take into account possible prior saturation in the responses of unimodal neurons that project to the DSC neurons. In this case, the different dynamic ranges between the unimodal case and the bimodal case are a direct consequence of the saturated inputs that DSC neurons receive. This interpretation requires nothing more than traditional knowledge of the response properties of neurons, and thus appears to be more preferable than the idea of dynamic modulation in the response range. Moreover, Perrault et al. (2005) observed four different operational modes in DSC neurons and suggested that the DSC neurons can be categorized based on these operational modes into the super-additive neurons, the super-additive/sub-additive neurons, the sub-additive neurons and the additive/sub-additive neurons. Again, at first glance, these results indicate the existence of some intricate mechanisms in the responses of DSC neurons. Nevertheless, we find that these diverse operations can also be naturally explained by the typical and well-established response properties of neurons – prior saturation in the unimodal neurons that project to DSC neurons responding in a non-linear, saturating manner. Another example is the CME index and the ADD index, which have been proposed to characterize the response properties of DSC neurons. Certainly, these two indices capture some important properties in DSC neurons, as indicated by many experimental results (Stein et al., 2004). As we have discussed in Chapter 6 and earlier sections, however, nothing more complicated than a sigmoidal response function is required to explain all the related properties, such as the IE of CME, the super-additive, additive and sub-additive responses. Furthermore, we have shown that the two indices are inherently correlated, so that adopting a combination of the two in fact does not provide much more information than using only one. The reason that these somewhat redundant indices are proposed and well-accepted in the field is a lack of theoretical insight.

This is why theoretical studies are so important in this experimental field. Theoretical studies could provide novel aspects to interpret the experimental data, which, due to rigorous derivations and proofs, are potentially significant. Of course, sometimes theoretical studies may turn out to misinterpret the experimental data. For example, we have discussed in Chapter 5 that Patton and Anastasio (2003)'s model is an inappropriate explanation of the responses of DSC neurons. Although their model is based on a Bayesian approach, which is often adopted to explain neural processes, its ability to reproduce the response properties is merely a coincidence based on careful choices of parameters. A reliance on well-known theorems could lead to a negligent examination of the parameter regions and result in models that are misleading. One may therefore ask, what kinds of theoretical studies are valuable? In our opinion, a valuable theoretical model should firstly be constructed on sound theoretical bases and provide insightful ideas. Secondly,

it should reproduce solid experimental results while not too sensitive to the choices of parameters. Thirdly, a theoretical model should not place too much computational burden on a single cell. A neuron is a simple device with limited computational power. It is therefore unrealistic to expect it to carry out complicated computations. However, this aspect is commonly overlooked by many theoretical studies. For example, the maximum response entropy idea proposed by Laughlin (1981) requires a sensory neuron to adapt its response to the cumulative density function (CDF) of the inputs. Although theoretically interesting, this requirement probably exceeds the computational ability of a single neuron. Last but not least, it would be desirable if a theoretical model is able to produce some experimentally testable predictions so that the underlying notions can be verified.

In the course of our studies, we notice that a careful and clear interpretation of the existing experimental data in the field of multisensory integration in DSC neurons is still missing. To achieve this, both experimental and theoretical studies are required, since on one hand, experimental data offer the necessary material to inspire theoretical studies, and on the other hand, theoretical studies provide feedback to guide experimental studies. During this reciprocal process, a pivotal issue is to carefully consider the experimental data and interpret them in a natural and simple way. Having kept this in mind, we hope that the theoretical study presented in this thesis will act as valuable feedback to the experimental field of multisensory integration.

Bibliography

Albrecht, D. G., Farrar, S. B., and Hamilton, D. B. (1984). Spatial contrast adaptation characteristics of neurons recorded in the cat's visual cortex. *J. Physiol. (Lond.)*, 347:713–739.

Alvarado, J. C., Stanford, T. R., Vaughan, J. W., and Stein, B. E. (2007). Cortex mediates multisensory but not unisensory integration in superior colliculus. *J. Neurosci.*, 27(47):12775–12786.

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.*, 27:77–87.

Anastasio, T. J. and Patton, P. E. (2003). A two-stage unsupervised learning algorithm reproduces multisensory enhancement in a neural network model of the corticotectal system. *J. Neurosci.*, 23(17):6713–6727.

Arganda, S., Guantes, R., and de Polavieja, G. C. (2007). Sodium pumps adapt spike bursting to stimulus statistics. *Nature Neuroscience*, 10:1467–1473.

Artola, A. and Singer, W. (1993). Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends Neurosci.*, 16:480–487.

Atick, J. J. (1992). Could information theory provides an ecological theory of sensory processing? *Network: Comput. Neural Syst.*, 3:213–251.

Attias, H. and Schreiner, C. E. (1997). Temporal low-order statistics of natural sounds. In Mozer, M., Jordan, M., and Petsche, T., editors, *Advances in Neural Information Processing Systems*, volume 9. The MIT Press.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychol. Rev.*, 61:183–193.

Avillac, M., Hamed, S. B., and Duhamel, J.-R. (2007). Multisensory integration in the ventral intraparietal area of the macaque monkey. *J. Neurosci.*, 27(8):1922–1932.

Baccus, S. A. and Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36:909–919.

Baccus, S. A. and Meister, M. (2004). Retina versus cortex: contrast adaptation in parallel visual pathways. *Neuron*, 42:5–7.

Baddeley, R., Abbott, L. F., Booth, M. C., Sengpiel, F., Freeman, T., Wakeman, E. A., and Rolls, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. R. Soc. London Ser. B*, 264:1775–1783.

Baleydier, C., Kahungu, M., and Mauguiere, F. (1983). A crossed corticotectal projection from the lateral suprasylvian area in the cat. *J. Comp. Neurol.*, 214:344–351.

Barlow, H. and Mollon, J. D. (1982). *The senses*. Cambridge University Press, Cambridge, second edition.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In Rosenblith, W. A., editor, *Sensory Communication*, pages 217–234, Cambridge, MA. MIT Press.

Barlow, H. B. (1981). Critical limiting factors in the design of the eye and visual cortex. *Proc. R. Soc. London Ser. B*, 212:1–34.

Battaglia, P. W., Jacobs, R. A., and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A*, 20(7):1391–1397.

Beauchamp, M. S., Lee, K. E., Argall, B. D., and Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, 41:809–823.

Beaudoin, D. L., Borghuis, B. G., and Demb, J. B. (2007). Cellular basis for contrast gain control over the receptive field center of mammalian retinal ganglion cells. *J. Neurosci.*, 27(10):2636–2645.

Behan, M. and Appell, P. P. (1992). Intrinsic circuitry in the cat superior colliculus: Projections from the superficial layers. *J. Comp. Neurol.*, 315:230–243.

Bell, A. J. and Sejnowski, T. J. (1996). Learning the higher-order structure of a natural sound. *Network: Comput. Neural Syst.*, 7:261–266.

Bertelson, P. (1999). Ventriloquism: A case of cross-modal perceptual grouping. In Aschersleben, G., Bachmann, T., and Müsseler, J., editors, *Cognitive Contributions to the Perception of Spatial and Temporal Events*, pages 347–362. Elsevier, Amsterdam.

Berthoz, A., Grantyn, A., and Droulez, J. (1986). Some collicular efferent neurons code saccadic eye velocity. *Neuroscience*, 72:294–298.

Binns, K. E. and Salt, T. E. (1996). Importance of NMDA receptors for multimodal integration in the deep layers of the cat superior colliculus. *J. Neurophysiol.*, 75(2):920–930.

Bonin, V., Mante, V., and Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus. *J. Neurosci.*, 25:10844–10856.

Bonin, V., Mante, V., and Carandini, M. (2006). The statistical computation underlying contrast gain control. *J. Neurosci.*, 26(23):6346–6353.

Brainard, M. S. and Knudsen, E. I. (1993). Experience-dependent plasticity in the inferior colliculus: A site for visual calibration of the neural representation of auditory space in the barn owl. *J. Neurosci.*, 13(11):4589–4608.

Brenner, N., Bialek, W., and de Ruyter van Steveninck, R. R. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, 26:695–702.

Brown, R. G. and Meyer, R. F. (1961). The fundamental theorem of exponential smoothing. *Operations Research*, 9(5):673–687.

Buccigrossi, R. W. and Simoncelli, E. P. (1999). Image compression via joint statistical characterization in the wavelet domain. *IEEE Transactions on Image Processing*, 8(12):1688–1701.

Burton, G. J., Haig, N. D., and Moorhead, I. R. (1986). A self-similar stack model for human and machine vision. *Biol. Cybern.*, 53(6):397–403.

Burton, G. J. and Moorhead, I. R. (1987). Color and spatial structures in natural scenes. *Applied Optics*, 26(1):157–170.

Callan, D. E., Jones, J. A., Munhall, K. G., Callan, A. M., Kroos, C., and Vatikiotis-Bateson, E. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport*, 14:2213–2217.

Calvert, G. A. and Thesen, T. (2004). Multisensory integration: Methodological approaches and emerging principles in the human brain. *J. Physiol. (Paris)*, 98:191–205.

Carandini, M. and Ferster, D. (2000). Membrane potential and firing rate in cat primary visual cortex. *J. Neurosci.*, 20:470–484.

Carandini, M., Movshon, J. A., and Ferster, D. (1998). Pattern adaptation and cross-orientation interactions in the primary visual cortex. *Neuropharmacology*, 37:501–511.

Carrasco, M. M., Razak, K. A., and Pallas, S. L. (2005). Visual experience is necessary for maintenance but not development of receptive fields in superior colliculus. *J. Neurophysiol.*, 94:1962–1970.

Chance, F. S., Abbott, L. F., and Reyes, A. D. (2002). Gain modulation from background synaptic input. *Neuron*, 35:773–782.

Chander, D. and Chichilnisky, E. J. (2001). Adaptation to temporal contrast in primate and salamander retina. *J. Neurosci.*, 21:9904–9916.

Clarey, J. C. and Irvine, D. R. F. (1986). Auditory response properties of neurons in the anterior ectosylvian sulcus of the cat. *Brain Res.*, 386:12–19.

Clemo, H. R., Meredith, M. A., Wallace, M. T., and Stein, B. E. (1991). Is the cortex of cat anterior ectosylvian sulcus a polysensory area? *Society for Neuroscience Abstracts*, 17:1585.

Clemo, H. R. and Stein, B. E. (1982). Somatosensory cortex: A "new" somatotopic representation. *Brain Res.*, 235:162–168.

Cowan, R. (1987). A bivariate exponential distribution arising in random geometry. *Annals of the Institute of Statistical Mathematics*, 39(1):103–111.

Critchley, M. (1977). Ecstatic and synaesthetic experience during musical perception. In Critchley, M. and Henson, R., editors, *Music and the Brain: Studies in the Neurology of Music*, pages 187–205. Charles C Thomas, Springfield, Ill.

Cuppini, C., Magosso, E., Serino, A., Pellegrino, G. D., and Ursino, M. (2007). A neural network for the analysis of multisensory integration in the superior colliculus. *ICANN 2007, Part II*, pages 9–18.

Dalton, P., Doolittle, N., Nagata, H., and Breslin, P. A. S. (2000). The merging of the senses: Integration of subthreshold taste and smell. *Nature Neuroscience*, 3:431–432.

Dan, Y., Atick, J. J., and Read, R. C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational model. *J. Neurosci.*, 16:3351–3362.

Daugman, J. G. (1989). Entropy reduction and decorrelation in visual coding by oriented neural receptive fields. *IEEE Trans. Biomed. Eng.*, 36(1):107–114.

Dean, I., Harper, N. S., and McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, 8:1684–1689.

Dean, I., Robinson, B. L., Harper, N. S., and McAlpine, D. (2008). Rapid neural adaptation to sound level statistics. *J. Neurosci.*, 28(25):6430–6438.

Demb, J. B. (2002). Multiple mechanisms for contrast adaptation in the retina. *Neuron*, 36:781–783.

Deneve, S., Latham, P. E., and Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *Nature Neuroscience*, 4(8):826–831.

Deneve, S. and Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *J. Physiol. (Paris)*, 98:249–258.

Derrington, A. M. and Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque. *J. Physiol.*, 357:219–240.

DeWeese, M. (1996). Optimization principles for the neural code. *Network: Comput. Neural Syst.*, 7:325–331.

DeWeese, M. and Zador, A. (1998). Asymmetric dynamics in optimal variance adaptation. *Neural Comput.*, 10:1179–1202.

Edwards, S. B., Ginsburgh, C. L., Henkel, C. K., and Stein, B. E. (1979). Sources of subcortical projections to the superior colliculus in the cat. *J. Comp. Neurol.*, 72:282–287.

Eimer, M. (2004). Multisensory integration: How visual experience shapes spatial perception. *Curr. Biol.*, 14:115–117.

Elliott, E., Kuang, X., Shadbolt, N. R., and Zauner, K. P. (2008a). Adaptation in multisensory neurons: Impact on cross-modal enhancement. *Submitted*.

Elliott, E., Kuang, X., Shadbolt, N. R., and Zauner, K. P. (2008b). An invariance principle for maintaining the operating point of a neuron. *Network: Comput. Neural Syst.*, 19(3):213–235.

Elliott, E., Kuang, X., Shadbolt, N. R., and Zauner, K. P. (2008c). On natural statistics in multisensory integration. In *Computational and Systems Neuroscience 2008*, Salt Lake City, Utah.

Ernst, M. O., Banks, M. S., and Bulthoff, H. H. (2002). Human integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415:429–433.

Fairhall, A. L., Lewen, G. D., Bialek, W., and de Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412:787–792.

Falchier, A., Clavagner, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.*, 22:5749–5759.

Feldman, D. E. and Knudsen, E. I. (1997). An anatomical basis for visual calibration of the auditory space map in the barn owl's midbrain. *J. Neurosci.*, 17(17):6820–6837.

Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1:1–47.

Feller, W. (1971). *An introduction to probability theory and its applications*. Wiley.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12):2379–2394.

Frassinetti, F., Pavani, F., and Ladavas, E. (2002). Acoustical vision of neglected stimuli: Interaction among spatially converging audio-visual inputs in neglect patients. *J. Cog. Neurosci.*, 14(1):62–69.

Fristedt, B. and Gray, L. (1997). *A modern approach to probability theory*. Birkhäuser, Boston.

Gronenberg, W. and López-Riquelme, G. O. (2004). Multisensory convergence in the mushroom bodies of ants and bees. *Acta Biologica Hungarica*, 55:31–37.

Hancock, P. J. B., Baddeley, R. J., and Smith, L. S. (1992). The principal components of natural images. *Network: Comput. Neural Syst.*, 3:61–70.

Harting, J. K., Feig, S., and Van Lieshout, D. P. (1997). Cortical somatosensory and trigeminal inputs to the cat superior colliculus: Light and electron microscopic analyses. *J. Comp. Neurol.*, 388:313–326.

Hasselmo, M. E. (1995). Neuromodulation and cortical function: modeling the physiological basis of behavior. *Behav. Brain Res.*, 67:1–27.

Huang, J. (2000). *Statistics of natural images and models*. PhD thesis, Division of Applied Mathematics, Brown University.

Hyvärinen, A., Karhunen, J., and Oja, E. (2001). *Independent component analysis*. Wiley.

Hyvärinen, A. and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, 13:411–430.

Ingham, N. J. and McAlpine, D. (2005). Gabaergic inhibition controls neural gain in inferior colliculus neurons sensitive to interaural time differences. *J. Neurosci.*, 25:6187–6198.

Iordanov, L. G. and Penev, P. S. (1999). The principal component structure of natural sound. In Kearns, M., Solla, S., and Cohn, D., editors, *Advances in Neural Information Processing Systems*, volume 11. The MIT Press.

Jay, M. F. and Sparks, D. L. (1987). Sensorimotor integration in the primate superior colliculus. *J. Neurophysiol.*, 57:35–55.

Jeffress, L. A. (1948). A place theory of sound localization. *Journal of Comparative Physiological Psychology*, 41:35–39.

Jiang, W., Jiang, H., and Stein, B. E. (2002). Two corticotectal areas facilitate multisensory orientation behavior. *J. Cog. Neurosci.*, 14(8):1240–1255.

Jiang, W., Jiang, H., and Stein, B. E. (2006). Neonatal cortical ablation disrupts multisensory development in superior colliculus. *J. Neurophysiol.*, 95:1380–1396.

Jiang, W. and Stein, B. E. (2003). Cortex controls multisensory depression in superior colliculus. *J. Neurophysiol.*, 90:2123–2135.

Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W., and Stein, B. E. (2001). Two cortical areas mediate multisensory integration in superior colliculus neurons. *J. Neurophysiol.*, 85:506–522.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*. NY: Wiley.

Jones, E. G. and Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*, 93:793–820.

Kadunce, D. C., Vaughan, J. W., Wallace, M. T., Benedek, G., and Stein, B. E. (1997). Mechanisms of within- and cross-modality suppression in the superior colliculus. *J. Neurophysiol*, 78:2834–2847.

Kanaseki, T. and Sprague, J. M. (1974). Anatomical organization of pretectal nuclei and tectal laminae in the cat. *J. Comp. Neurol.*, 158:319–338.

Kawamura, K. and Konno, T. (1979). Various types of corticotectal neurons of cats as demonstrated by means of retrograde axonal transport of horseradish peroxidase. *Exp. Brain Res.*, 35:161–175.

Kim, K. J. and Rieke, F. (2001). Temporal contrast adaptation in the input and output signals of the salamander retinal ganglion cells. *J. Neurosci.*, 21:287–299.

Kim, K. J. and Rieke, F. (2003). Slow Na+ inactivation and variance adaptation in salamander retinal ganglion cells. *J. Neurosci.*, 23:1506–1516.

King, A. J. and Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurons in the guinea-pig superior colliculus. *Exp. Brain Res.*, 60:492–500.

King, A. J. and Parsons, C. (1999). Improved auditory spatial acuity in visually deprived ferrets. *Eur. J. Neurosci.*, 11:3945–3956.

Knudsen, E. I. (1982). Auditory and visual maps of space in the optic tectum of the owl. *J. Neurosci.*, 2:1177–1194.

Knudsen, E. I. (1994). Supervised learning in the brain. *J. Neurosci.*, 14(7):3985–3997.

Knudsen, E. I. and Brainard, M. S. (1995). Creating a unified representation of visual and auditory space in the brain. *Annu. Rev. Neurosci.*, 18:19–43.

Knudsen, E. I., Esterly, S. D., and du Lac, S. (1991). Stretched and upside-down maps of auditory space in the optic tectum of blind-reared owls: Acoustic basis and behavioral correlates. *J. Neurosci.*, 11:1727–1747.

Knudsen, E. I. and Konishi, M. (1978). Space and frequency are represented separately in auditory midbrain of the owl. *J. Neurophysiol*, 41:870–884.

Knudsen, E. I., Konishi, M., and Pettigrew, J. D. (1977). Receptive fields of auditory neurons in the owl. *Science*, 198:1278–1280.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybern.*, 44:59–69.

Konishi, M. (1993). Listening with two ears. *Sci. Am.*, 268:34–41.

Konishi, M., Takahashi, T. T., Wagner, H., Sullivan, W. E., and Carr, C. E. (1988). Neurophysiological and anatomical substrates of sound localization in the owl. In Edelman, G., Gall, W., and Cowan, W., editors, *Auditory Function: Neurobiological Bases of Hearing*, pages 721–744, New York. Wiley.

Koutalos, Y. and Yau, K. W. (1996). Regulation of sensitivity in vertebrate rod photoreceptors by calcium. *Trends Neurosci.*, 19:73–81.

Kvale, M. N. and Schreiner, C. E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. *J. Neurophysiol.*, 91:604–612.

Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch. C*, 36:910–912.

Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., and Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Exp. Brain Res.*, 166:289–297.

Leichnetz, G. R., Spencer, R. F., Hardy, S. G. P., and Astruc, J. (1981). The prefrontal corticotectal projection in the monkey: An anterograde and retrograde horseradish peroxidase study. *Neuroscience*, 6:1023–1041.

Lesica, N. A., Jin, J., Weng, C., Yeh, C., Butts, D. A., Stanley, G. B., and Alonso, J. (2007). Adaptation to stimulus contrast and correlations during natural visual stimulation. *Neuron*, 55:479–491.

Lessard, N., Pare, M., Leporé, F., and Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, 395:278–280.

Lewicki, M. S. (2002). Efficient coding of natural sounds. *Nature Neuroscience*, 5(4):356–363.

Lewkowicz, D. J. (1994). Development of intersensory perception in human infants. In Lewkowicz, D. and Lickliter, R., editors, *The Development of Intersensory Perception: Comparative Perspectives*, pages 165–203. Lawrence Erlbaum.

Macaluso, E., Frith, C., and Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289:1206–1208.

Maffei, L., Fiorentini, A., and Bisti, S. (1973). Neural correlate of perceptual adaptation to gratings. *Science*, 182:1036–1038.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693.

Manookin, M. B. and Demb, J. B. (2006). Presynaptic mechanism for slow contrast adaptation in mammalian retinal ganglion cells. *Neuron*, 50:453–464.

Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., and Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nature Neuroscience*, 8(12):1690–1697.

Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E., and Diamond, M. E. (2007). Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biol.*, 5:0323–0334.

Matthews, H. R., Murphy, R. L., Fain, G. L., and Lamb, T. D. (1988). Photoreceptor light adpatation is mediated by cytoplasmic calcium concentration. *Nature*, 334:67–69.

McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 256(5588):746–748.

McHaffie, J. G., Kruger, L., Clemo, H. R., and Stein, B. E. (1988). Corticothalamic and corticotectal somatosensory projections from the anterior ectosylvian sulcus (siv cortex) in neonatal cats: An anatomical demonstration with HRP and 3h-leucine. *J. Comp. Neurol.*, 274:115–126.

McIlwain, J. T. (1990). Topography of eye-position sensitivity of saccades evoked electrically from the cat's superior colliculus. *Vis. Neurosci.*, 4:289–298.

Meister, M. and Berry, M. J. (1999). The neural code of the retina. *Neuron*, 22:435–450.

Meredith, M. A. and Clemo, H. R. (1989). Auditory cortical projection from the anterior ectosylvian sulcus (Field AES) to the superior colliculus in the cat: An anatomical and electrophysiological study. *J. Comp. Neurol.*, 289:687–707.

Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons I. Temporal factors. *J. Neurosci.*, 7:3215–3229.

Meredith, M. A. and Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 221:389–391.

Meredith, M. A. and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J. Neurophysiol.*, 56(3):640–662.

Meredith, M. A. and Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *J. Neurophysiol.*, 75(5):1843–1857.

Meredith, M. A., Wallace, M. T., and Stein, B. E. (1992). Visual, auditory and somatosensory convergence in output neurons of the cat superior colliculus: Multisensory properties of the tecto-reticulo-spinal projection. *Exp. Brain Res.*, 88:181–186.

Mirenowicz, J. and Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*, 379:449–451.

Moschovakis, A. K. and Karabelas, A. B. (1985). Observations on the somatodendritic morphology and axonal trajectory of intracellularly hrp-labeled efferent neurons located in the deep layers of the superior colliculus of the cat. *J. Comp. Neurol.*, 239:276–308.

Movshon, J. A. and Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurons. *Nature*, 278:850–852.

Muchnik, C., Efrati, M., Nemeth, E., Malin, M., and Hildesheimer, M. (1991). Central auditory skills in blind and sighted subjects. *Scand. Audiol.*, 20:19–23.

Mucke, L., Norita, M., Benedek, G., and Creutzfeldt, O. (1982). Physiologic and anatomic investigation of a visual cortical area situated in the ventral bank of the anterior ectosylvian sulcus of the cat. *Exp. Brain Res.*, 46:1–11.

Naka, K. I. and Rushton, W. A. H. (1966). S-potentials from colour units in the retina of fish (cyprinidae). *J. Physiol. (Lond.)*, 185:536–555.

Nakatani, K. and Yau, K.-W. (1988). Calcium and light adaptation in retinal rods and cones. *Nature*, 334:69–71.

Newman, E. A. and Hartline, P. H. (1981). Integration of visual and infrared information in bimodal neurons of the rattlesnake optic tectum. *Science*, 213:789–791.

Normann, R. A. and Werblin, F. S. (1974). Control of retinal sensitivity. I. Light and dark adaptation of vertebrate rods and cones. *J. Gen. Physiol.*, 63:37–61.

Ohzawa, I., Sclar, G., and Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *J. Neurophysiol.*, 54:651–667.

Ojanen, V. (2005). *Neurocognitive mechanisms of audiovisual speech perception.* PhD thesis, Helsinki University of Technology.

Pandya, D. N. and Kuypers, H. G. (1969). Cortico-cortical connections in the rhesus monkey. *Brain Res.*, 13:13–36.

Papoulis, A. (1991). *Probability, random variables and stochastic processes.* McGraw-Hill, New York, third edition.

Patton, P. E. and Anastasio, T. J. (2003). Modeling cross-modal enhancement and modality-specific suppression in multisensory neurons. *Neural Comput.*, 15(4):783–810.

Peck, C. K., Baro, J. A., and Warder, S. M. (1993). Sensory integration in the deep layers of superior colliculus. *Prog. Brain Res.*, 95:91–102.

Pena, J. L. and Konishi, M. (2001). Auditory spatial receptive fields created by multiplication. *Science*, 292:249–252.

Pena, J. L. and Konishi, M. (2002). From postsynaptic potentials to spikes in the genesis of auditory spatial receptive fields. *J. Neurosci.*, 22:5652–5658.

Perrault, T. J., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2003). Neuron-sepcific response characteristics predict the magnitude of multisensory integration. *J. Neurophysiol.*, 90:4022–4026.

Perrault, T. J., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2005). Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J. Neurophysiol.*, 93(5):2575–2586.

Pickles, J. O. (1988). *An introduction to the physiology of hearing.* Academic Press, San Diego, CA, second edition.

Populin, L. C. and Yin, T. C. T. (2002). Bimodal interactions in the superior colliculus of the behaving cat. *J. Neurosci.*, 22(7):2826–2834.

Pugh, E. N., Nikonov, S., and Lamb, T. D. (1999). Molecular mechanisms of vertebrate photoreceptor light adaptation. *Curr. Opin. Neurobiol.*, 9:410–418.

Purves, D., Augustine, G. J., Fitzpatrick, D., and Hall, W. C. e. (2004). *Neuronscience, third edition.* Sinauer Associates Inc.

Rauschecker, J. P. (2004). Cross-modal consequences of visual deprivation in animals. In Calvert, G., Spence, C., and Stein, B. E., editors, *The Handbook of Multisensory Integration*, chapter 43, pages 695–702. The MIT Press.

Rauschecker, J. P. and Kniepert, U. (1994). Enhanced precision of auditory localization behavior in visually deprived cats. *European Journal of Neuroscience*, 6:149–160.

Rayleigh, J. W. S. (1889). On the limit to interference when light is radiated from moving molecules. *Phil. Mag.*, 27:298–304.

Rieke, F. (2001). Temporal contrast adaptation in salamander bipolar cells. *J. Neurosci.*, 21:9445–9454.

Ringach, D. L. and Malone, B. J. (2007). The operating point of the cortex: Neurons as large deviation detectors. *J. Neurosci.*, 27(29):7673–7683.

Rockland, K. S. and Ojima, H. (2001). Calcarine area V1 as a multimodal convergence area. *Society for Neuroscience Abstracts*, 27:511–520.

Roder, B. and Neville, H. J. (2003). Early vision impairs tactile perception in the blind. *Curr. Biol.*, 14:121–124.

Rucci, M., Edelman, G. M., and Wray, J. (1999). Adaptation of orienting behavior: From the barn owl to a robotic system. *IEEE Transactions on Robotics and Automation*, 15:96–110.

Rucci, M., Tononi, G., and Edelman, G. M. (1997). Registration of neural maps through value-dependent learning: Modeling the alignment of auditory and visual maps in the barn owl's optic tectum. *J. Neurosci.*, 17(1):334–352.

Ruderman, D. L. (1994). The statistics of natural images. *Network: Comput. Neural Syst.*, 5:517–548.

Ruderman, D. L. (1997). Origins of scaling in natural images. *Vis. Res.*, 37(23):3385–3398.

Rushton, W. A. (1965). Visual adaptation. *Proc. R. Soc. London Ser. B*, 162:20–46.

Sakmann, B. and Creutzeeldt, O. D. (1969). Scotopic and mesopic light adaptation in the cat's retina. *Pflügers Arch.*, 313:168–185.

Sanchez-Vives, M. V., Nowak, L. G., and McCormick, D. A. (2000a). Cellular mechanisms of long-lasting adaptation in visual cortical neurons *in vitro*. *J. Neurosci.*, 20:4286–4299.

Sanchez-Vives, M. V., Nowak, L. G., and McCormick, D. A. (2000b). Membrane mechanisms underlying contrast adaptation in cat area 17 *in vivo*. *J. Neurosci.*, 20:4267–4285.

Schauer, C. and Gross, H. (2004). Design and optimization of Amari neural fields for early auditory-visual integration. In *Proc. 2004 IEEE Int. Joint Conf. on Neural Networks*, volume 4, pages 2523–2528, Budapest.

Schwartz, O. and Simoncelli, E. P. (2000). Natural sound statistics and divisive normalization in the auditory system. In Leen, T., Dietterich, T., and Tresp, V., editors, *Advances in Neural Information Processing Systems*, volume 13. The MIT Press.

Sekuler, R., B., S. A., and R., L. (1997). Sound alters visual motion perception. *Nature*, 385:308.

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Shapley, R. and Enroth-Cugell, C. (1984). Visual adaptation and retinal gain control. *Prog. Retinal Res.*, 3:263–346.

Shapley, R. M. and Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *J. Physiol.*, 285:275–298.

Shapley, R. M. and Victor, J. D. (1981). How the contrast gain modifies the frequency responses of cat retinal ganglion cells. *J. Physiol. (Lond.)*, 318:161–179.

Sharpee, T. O., Sugihara, H., Kurgansky, A. V., Rebrik, S. P., Stryker, M. P., and Miller, K. D. (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature*, 439:936–942.

Shu, Y., Hasenstaub, A., Badoual, M., Bal, T., and McCormick, D. A. (2003). Barrages of synaptic activity control the gain and sensitivity of cortical neurons. *J. Neurosci.*, 23:10388–10401.

Simoncelli, E. P. and Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annu. Rev. Neurosci.*, 24:1193–1216.

Simoncelli, E. P. and Schwartz, O. (1999). Image statistics and cortical normalization models. In Kearns, M., Solla, S., and Cohn, D., editors, *Advances in Neural Information Processing Systems*, volume 11, pages 153–159.

Slaney, M. (1988). Lyon's cochlear model. Technical report, Advanced Technology Group, Apple Computer, Inc.

Smirnakis, S. M., Berry, M. J., Warland, D. K., Bialek, W., and Meister, M. (1997). Adaptation of retinal processing to image contrast and spatial scale. *Nature*, 386:69–73.

Solomon, S. G., Peirce, J. W., Dhruv, N. T., and Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. *Neuron*, 42:155–162.

Srinivasan, M. V., Laughlin, S. B., and Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. London Ser. B*, 216:427–459.

Srivastava, A., Lee, A. B., Simoncelli, E. P., and Zhu, S.-C. (2003). On advances in statistical modeling of natural images. *Journal of Mathematical Imaging and Vision*, 18:17–33.

Stanford, T. R., Quessy, S., and Stein, B. E. (2005). Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J. Neurosci.*, 25(28):6499–6508.

Stein, B. E. (2005). The development of a dialogue between cortex and midbrain to integrate multisensory information. *Exp. Brain Res.*, 166:305–315.

Stein, B. E. and Clamann, H. P. (1981). Control of pinna movements and sensorimotor register in cat superior colliculus. *Brain Behav. Evol.*, 19:180–192.

Stein, B. E. and Gaither, N. (1983). Receptive field properties in reptilian optic tectum: Some comparisons with mammals. *J. Neurophysiol.*, 50:102–124.

Stein, B. E., Jiang, W., and Stanford, T. R. (2004). Multisensory integration in single neurons of the midbrain. In Calvert, G., Spence, C., and Stein, B. E., editors, *The Handbook of Multisensory Integration*, chapter 15, pages 243–264. The MIT Press.

Stein, B. E. and Meredith, M. A. (1993). *Merging of the senses*. The MIT Press.

Stein, B. E., Meredith, M. A., Huneycutt, W. S., and McDade, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *J. Cog. Neurosci.*, 1:12–24.

Thompson, R. F., Johnson, R. H., and Hoopes, J. J. (1963). Organization of auditory, somatic sensory and visual projections to association cortex in the cat. *J. Neurophysiol.*, 26:243–364.

Toldi, J. and Feher, O. (1984). Acoustic sensitivity and bimodal properties of cells in the anterior suprasylvian gyrus of the cat. *Exp. Brain Res.*, 55:180–183.

Torralba, A. and Oliva, A. (2003). Statistics of natural image categories. *Network: Comput. Neural Syst.*, 14:391–412.

Truchard, A. M., Ohzawa, I., and Freeman, R. D. (2000). Contrast gain control in the visual cortex: monocular versus binocular mechanisms. *J. Neurosci.*, 20(8):3017–3032.

van Atteveldt, N., Formisano, E., Goebel, R., and Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, 43:271–282.

van Hateren, J. H. and van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. London Ser. B*, 265:3359–366.

van Hatteren, J. H. (1992). A theory maximizing sensory information. *Biol. Cybern.*, 68:23–29.

Victor, J. D. (1987). The dynamics of the cat retina X cell centre. *J. Physiol. (Lond.)*, 386:219–246.

Voss, R. F. and Clarke, J. (1975). 1/f noise in music and speech. *Nature*, 258:317–318.

Vroomen, J. and de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26:1583–1590.

Wagner, H., Takahashi, T. T., and Konishi, M. (1987). Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *J. Neurosci.*, 74:3105–3116.

Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Res.*, 39:3960–3974.

Wallace, M. T. (2004). The development of multisensory integration. In Calvert, G., Spence, C., and Stein, B. E., editors, *The Handbook of Multisensory Integration*, chapter 39, pages 625–642. The MIT Press.

Wallace, M. T., Carriere, B. N., Perrault Jr, T. J., Vaughan, J. W., and Stein, B. E. (2006). The development of cortical multisensory integration. *J. Neurosci.*, 26(46):11844–11849.

Wallace, M. T., Meredith, M. A., and Stein, B. E. (1992). The integration of multiple sensory inputs in cat cortex. *Exp. Brain Res.*, 91:484–488.

Wallace, M. T., Meredith, M. A., and Stein, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *J. Neurophysiol.*, 69:1797–1809.

Wallace, M. T., Perrault, J. T. J., Hairston, W. D., and Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *J. Neurosci.*, 24(43):9580–9584.

Wallace, M. T. and Stein, B. E. (1996). Sensory organization of the superior colliculus in cat and monkey. *Prog. Brain Res.*, 112:301–311.

Wallace, M. T. and Stein, B. E. (1997). Development of multisensory neurons and multisensory integration in cat superior colliculus. *J. Neurosci.*, 17(7):2429–2444.

Wallace, M. T. and Stein, B. E. (2000). Onset of cross-modal synthesis in the neonatal superior colliculus is gated by the development of cortical influences. *J. Neurophysiol.*, 83:3578–3582.

Wallace, M. T. and Stein, B. E. (2001). Sensory and multisensory responses in the newborn monkey superior colliculus. *J. Neurosci.*, 21(22):8886–8894.

Wallace, M. T. and Stein, B. E. (2007). Early experience determines how the senses will interact. *J. Neurophysiol.*, 97(1):921–926.

Wallace, M. T., Wilkinson, L. K., and Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *J. Neurophysiol.*, 76(2):1246–1266.

Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A., and Schnapf, J. L. (1990). The control of visual sensitivity. In Spillmann, L. and Werner, S., editors, *Visual Perception: the Neurophysiological Foundations*, pages 53–101. San Diego: Academic.

Wegmann, B. and Zetzsche, C. (1990). Statistical dependence between orientation filter outputs used in a human vision based image code. In *Proceedings of Visual Communication and Image Processing*, volume 1360, pages 909–922. Society of Photo-Optical Instrumentation Engineers.

Welch, B. L. (1947). The generalization of "student's" problem when several different population variances are involved. *Biometrika*, 34:28–35.

Welch, R. B. and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 3:638–667.

Wessinger, C. M., Fendrich, R., and Gazzaniga, M. S. (1999). A positron emission tomographic study of auditory localization in the congenitally blind. *J. Neurosci.*, 20:2664–2672.

Wilson, H. R. and Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vis. Res.*, 33:1133–1149.

Yin, T. C. and Chan, J. C. (1990). Interaural time sensitivity in medial superior olive of cat. *J. Neurophysiol.*, 64:465–488.

Zaghloul, K. A., Boahen, K., and Demb, J. B. (2005). Contrast adaptation in subthreshold and spiking responses of mammalian y-type retinal ganglion cells. *J. Neurosci.*, 24:860–868.

Zhu, S. C. and Mumford, D. (1997). Prior learning and Gibbs reaction-diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(11):1236–1250.