Noise-induced reversals in bistable visual perception

Pedro Ernesto García Rodríguez

Centre de Recerca Matemàtica

Doctoral Thesis UPF / 2012

PhD supervision: Prof. Dr. Antoni Guillamon Grabolosa
Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya

Universitat Pompeu Fabra

Dipòsit Legal:
ISBN:

To my parents Georgina and Pedro,
my brother Aléxis
and my sisters Tania and Mileydis

Acknowledgments

Many people have contributed to this work, in many different ways. Here I want to express my gratitude to all of them, specially to:

- Johan Larsson: the best friend I found in Barcelona, with him I have enjoyed the most open conversations and many extra-work activities. But he also contributed many times with technical skills in *C* programming and use of the clusters of computers.

- Dr. Mario Pannunzi: great person, fantastic friend who is always willing to help, with informal but also technical conversations. Many times, he showed a remarkable talent to rapidly find mistakes in my reasoning :) I feel indebted to him for the interesting conversations on the weighted sum algorithm proposed in this thesis to optimize the fitting results. He became also a determining factor for closing this work, he knows very well, thanks a lot *Mariote*!

- José Raúl Naranjo, Albert-Ludwigs Univ., Germany: my dear colleague for more than 20 years, since we met in Cuba during our B.Sc. studies in Physics. Here, I wish to express my sincere thanks for the pleasant atmosphere during my research stay in Freiburg, when I had the opportunity to study and discuss the EEG studies in bistable perception recently reported by Dr. J. Kornmeier and co-workers. Naranjo transmits to me his enthusiasm for Cognitive Sciences every time that we meet. I want also to thank him for the personal and philosophical conversations that we had during my stay. He helped me a lot to deal with stress and frustration during my PhD and, more importantly, to cope being far from my family.

- Lydia Garcia: probably the best informed person of the Department staff (about PhD 'issues'), with an infinite patience to deal with unexpected but also expected problems in the Secretary of the Department. For any formality that you could need during your PhD, I recommend to contact her before any other person. She will offer you a solution, she is our dear *magician* for the bureaucracy stuff.

- Montse, Bea, Magda, and many others in the Secretary Dpt.: each of them is an expert in some field of the steps and formalities that one has to follow if you want a happy end when dealing with bureaucracy. Thanks a lot too!!

- Dr. Alexander Pastukhov, U. Magdeburg, Germany: great collaborator; I feel indebted to him for the code to optimally visualize most of the fitting results presented in this thesis, and also for the fastest algorithm to compute the cumulative history. Thanks also for the many references on psychophysics that you have sent anytime that I had no access to them.

- Joachim Haenicke, U. Berlin, Germany: he is the author of an efficient *C++* code that was firstly used in our joint paper, to simulate the adaptation-LC model in 3D parameter regions. I adapted his code to include the depression-LC and Moreno-Bote et al.'s models. The original version of a *MatLab* code, that was really useful to present the corresponding results in the form of clear figures, belongs to him too.

- Prof. Dr. Antoni Guillamon: my PhD advisor, who paid attention to many things from the beginning of this work, he read many papers I needed to define the next step or project in my work, and even helped with JC sessions!! Thanks for all the conversations on Dynamical System theory too!

- Dr. Andres Bulhmann: I discussed with him the possibility to study the models in subspaces larger than 2D parameter regions. Thanks also, OF COURSE, for helping me every time with Linux and cluster stuff, and for the friendship.

- Yota: amaizing greek girl, in every sense. She works like hell, and enjoys as well. You can talk to her about everything.

- Nuria and Cati: thanks a lot for the friendship, it is difficult to make friends outside the University...with this kind of work that demands your life, your brain , your everything. Nuria, thanks for helping to obtain the Spanish nationality as well.

- Andrea and Tim: for dealing efficiently, together with Mario, with the administration of the cluster now, even during weekends. And for many coffee-breaks too. Andrea, thank also for your sincere friendship!

- Marina: thanks for the relaxing atmosphere that you brought to the work, for many coffee-breaks, lunches and sincere conversations.

- Alveno Vitale: my dear colleague, also entering his PhD course after 35 years old! Are you crazy?? :) The best thing that you have is... your heart, man.

Besides, I want also to thank Dr. J. Braun and Dr. G. Deco for suggesting the starting point of this work, and for providing, since 2009, with the recent data on history effects reported now in Pastukhov & Braun (2011). Finally, I want to express my gratitude to Dr. Manuel Castellet and Dr. Joaquin Bruna, directors of Centre de Recerca Matemàtica, for the financial support and excellent conditions to work during more than three years that were provided by the center.

Abstract

In this thesis, a set of some prevailing rate-based models for bistable perception have been considered in order to find the implications of the novel results reported in Pastukhov & Braun (2011). These authors have quantified not only salient aspects of bistable perception (mean and dispersion of dominance distributions), but also some hidden hysteresis effects ignored up to now.

Extensive computational simulations of different prevailing models rigorously demonstrate that the history-dependence of the perceptual process shown by Pastukhov & Braun (2011), effectively constrains the region of the parameter space able to replicate the empirical data. Concretely, that just small regions residing inside a *bistable* or *two-attractor* region of the whole parameter space are actually adequate to reproduce the experimental results, both for BR and KDE displays. Remarkably, the results remain valid across all the different classes of models considered, regardless the details of the neuronal implementation.

The biological plausibility of the parameter region found for each of the models considered, is further tested with respect to the widely known Levelt's propositions. To that end, we make use of weighted sums across the parameter regions computed for each subject in the first part of this manuscript, an algorithm that constitutes an important improvement to the methodology proposed by Shpiro et al. (2007) to fit behavioral data by rate-based models.

It is shown how different neuronal mechanisms clearly differ in their suitability to replicate Levelt's propositions. For instance, models with a slow fatiguing process given by spike-frequency adaptation Wilson (2003); Shpiro et al. (2007), no matter if they are being described by linear Shpiro et al. (2007) or nonlinear Curtu et al. (2008)) functions of the activity, replicate quite well Levelt's second law. Oppositely, a notable discrepancy between model and empirical results is found when such negative feedback is described as a long-term depression affecting the synapses between the competing neurons representing the two alternative interpretations Laing & Chow (2002); Shpiro et al. (2007).

The present work finishes with a study about the capability of the mentioned models to reproduce the resonance effects happening when varying external *frequencies*, as shown by Kim et al. (2006). Importantly, a resonance respect to the noise dispersion (*i.e.*, a true *stochastic* resonance) is clearly demonstrated here for the first time. Previous estimations of noise dispersion ($20-30\%$ of the input) and its locus (adaptation variables) are questioned. It is demonstrated that increased sensitivity to even weak signals of the order of less than 10% can be obtained with the models considered, with the noise variable simply entering as part of the net input feeding the neuron.

Resumen

En este trabajo, son considerados una serie de modelos para frecuencia neuronal ampliamente aceptados en percepción bi-estable, con el objetivo de evaluar las implicaciones de los resultados recientemente reportados en Pastukhov & Braun (2011). Estos autores han cuantificado no solamente aspectos más conocidos sobre el fenómeno (media y dispersión de las distribuciones de dominancia), sino también efectos de historia que habían sido ignorados hasta el presente.

Por medio de simulaciones computacionales, se demuestra rigurosamente que la dependencia de la historia del proceso perceptual encontrada por Pastukhov & Braun (2011) efectivamente restringe la región válida de parámetros que es adecuada para reproducir los datos empíricos. Concretamente, que solamente pequeñas regiones del espacio de parámetros disponible, y que se encuentran dentro de una región dinámica *bi-estable* caracterizada por dos atractores, son realmente adecuadas, tanto para rivalidad binocular (BR) como para estímulos de estructura por movimiento (KDE). Resulta importante destacar que los resultados permancen válidos de un modelo a otro, independientemente de los detalles de implementación neuronal.

La plausibilidad biológica de la región de parámetros encontrada para cada modelo es entonces considerada, en el contexto de las ampliamente conocidas proposiciones de Levelt. Con tal objetivo, hacemos uso de un algoritmo de suma pesada para extraer valores medios de la regiónes de parámetros correspondientes a cada sujeto. Este algoritmo constituye una importante mejora a la metodología propuesta por Shpiro et al. (2007) para ajustar modelos de frecuencia neuronal a datos comportamentales de percepción bi-estable.

Es entonces mostrado como cada mecanismo neuronal considerado es claramente diferente en su capacidad para reproducir las proposiciones de Levelt. Por ejemplo, modelos conteniendo procesos lentos de retroalimentación negativa dados por adaptación de frecuencia de disparo Wilson (2003); Shpiro et al. (2007), sin importar si están descritos por funciones lineales Shpiro et al. (2007) o no lineales Curtu et al. (2008)) de la actividad, consiguen reproducir de modo razonable la segunda proposición de Levelt. Por el contrario, una notable discrepancia entre modelo y resultados empíricos es encontrada cuando tales procesos están dados por la presencia de depresión sináptica de larga duración.

El presente trabajo culmina con un estudio sobre la capacidad de los mencionados modelos para reproducir los efectos de resonancia que ocurren al variar la *frecuencia* externa de modulación Kim et al. (2006). Es de destacar que en nuestro caso, un efecto de resonancia es encontrado respecto a la dispersión del ruido, lo cual indica la presencia de una verdadera resonancia del tipo *estocástico*. Este efecto es claramente demostrado para estos modelos, por primera vez, en el presente trabajo. Previas estimaciones de la dispersión del ruido ($20-30\,\%$ de la señal de entrada) y su localización (variables de adaptación) son analizadas. Se demuestra que un incremento de la sensibilidad a incluso muy pequeñas señales de menos del $10\,\%$ puede ser encontrada en estos modelos, con sólo incluir la variable de ruido como parte de la corriente neta que alimenta la neurona.

Contents

List of Figures

List of Tables

1 Introduction

Any person is constantly immersed in a flow of sensations, mental images and thoughts. The whole set of them, intermingled with each other through complex unknown relations, conforms the internal world that accompanies any individual every day. Such a subjective world, which is rich in events and evolves in time, is sometimes referred as a 'stream of consciousness' James (1981). For instance, a subject frequently interchanges a current idea with any other compatible thought or personal appreciation, when confronted with heterogeneous information.

Such unstoppable flow of consciousness, although private, is however accessible to experimental research, by confronting an observer with simpler and ambiguous (uncompleted) sensory information that allows various possible, although compatible, interpretations to compete for awareness Crick (1996); Cosmelli & Thompson (2007). A common situation that illustrates the complexity of such an internal world can be found in the continuous entrance and disappearance of multiple interpretations that a given visual stimulus scene can suggest to our brain.

Visual awareness constitutes a valuable candidate showing significant characteristics which probably reflects more general properties of the subjective experience in any individual Crick & Koch (1990, 1998). A classic psychophysical paradigm used to study spontaneous switching is *binocular rivalry*. This term refers to what happens when the visual system is dichoptically stimulated, *i.e.*, when each eye receives different images (*e.g.*, by means of a mirror stereoscope or color filters): after some time the perception is trapped in a continuous sequence of spontaneous switches back and forth between the rivaling stimuli that (in general) can not be fused in a single percept. For instance, a pair of sinusoidal gratings differing in orientation difference ($> 20°$) Brascamp et al. (2006), the images of a house and a face Tong et al. (1998) or of a sunburst-like pattern and a butterfly Sheinberg & Logothetis (1997) have been proven to induce visual rivalry.

With the exception of special situations (see below), the perception does not vacillate between any of the incongruent stimuli and the observer invariably reports a series of *flip-flop* changes in which a percept remains stable for a while and then finally fades, when the alternative representation comes into awareness Wheatstone (1838). Amazingly, the dynamics of this phenomenon seems to remain the same whether or not images are stabilized on the retina Logothetis et al. (1996), although it has been demonstrated that this holds only within a lim-

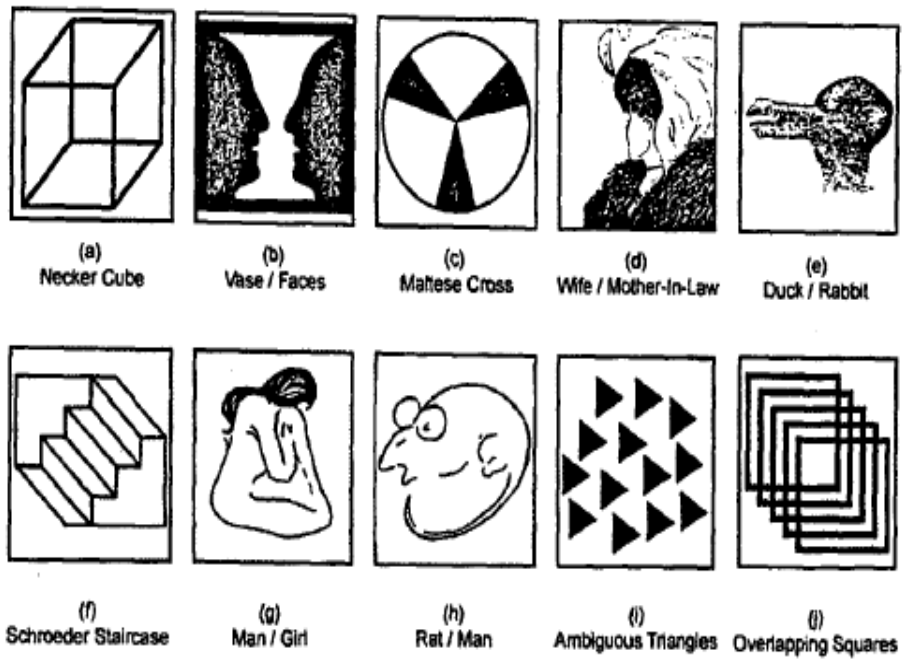

Figure 1.1: Common examples of ambiguous or reversible figures. Reprinted from Long & Toppino (2004).

ited range of spatial and temporal parameters of the stimuli and the experimental design Lee & Blake (1999).

Showing the subject a pair of dissimilar images to each eye is not the only one way to induce rivalry. Other kind of visual stimuli can also generate a perceptual conflict. Such is the case of the *reversible figures*, in which some degree of ambiguity can be introduced by uncompleted cues to avoid favoring any of the possible interpretations of the image shown. Significant examples of such kind of stimuli are Rubin's vase/face (that involves figure-ground reversals), the Necker cube or the Schroeder staircase (perspective inversions) and Boring's wife/mother-in-law (changes in meaning), among many others (see figure 1.1).

Monocular rivalry is also possible, by using two objects optically superimposed that are presented to a unique eye, although in this case the alternations become significantly slower Andrews & Purves (1997). Additionally, it is interesting to note that multi-stability has been observed in other sensory modalities, such as in the tactile and auditory systems (see for instance Box 1 of Sterzer et al. (2009) and references therein).

Rilvary's alternation cycles show a very rich dynamics. For example, the perceptual experience is characterized by a set of dominance durations that cannot be predicted at any time on the basis of the previous dominance evolution Borsellino

et al. (1972); Fox & Herrmann (1967); Walker (1975). According to the conventional point of view, such lack of correlation in the perceptual trace supports the idea of a pure stochastic-like behavior. No evidence of deterministic chaos has been found neither Lehky (1995), based not only on the usual autocorrelation techniques but on unsuccessful dimensionality reduction and unfruitful forecasting analysis of the percept durations series as well. However, previous reported measurements of such intrinsic correlations appear to be contaminated by duration impurities caused by reaction times, uncommon transitions and blinks, among other factors van Ee (2009).

The distribution of the dominance durations typically experienced by an observer, regardless of the the kind of ambiguous stimulation shown (binocular rivalry or ambiguous figures), has been found to be unimodal and right-skewed shaped with a long tail. It has traditionally been fitted by a Gamma function Fox & Herrmann (1967); Levelt (1967); Logothetis et al. (1996); Kovacs et al. (1996); Rubin & Hupé (2004)

$$f(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)},$$

although other mathematical formulae are possible, such as a log-normal distribution Lehky (1995)

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}},$$

or even a Weibull function Zhou et al. (2004)

$$f(x) = r\lambda^r x^{r-1} e^{-(\lambda x)^r}.$$

Although a bi-parametric function seems to be necessary to obtain a satisfactory fitting of empirical intervals distributions, a strong linear correlation between the two parameters describing the Gamma distribution of experimental data for the Necker cube and the Schroeder staircase has been reported by Borsellino et al. (1972). Additionally, is has been shown that the standard deviation $\sigma_{T_{dom}}$ and mean $< T_{dom} >$ of the dominance intervals are related by an average factor of 2 Levelt (1967); Walker (1975). Given the known mathematical relations between these statistical quantities and the parameters λ and r of the Gamma distribution Borsellino et al. (1972); Walker (1975):

$$\lambda = \frac{< T_{dom} >}{\sigma^2_{T_{dom}}} \qquad r = \frac{< T_{dom} >^2}{\sigma^2_{T_{dom}}},$$

one obtains $\lambda = r/ < T_{dom} >$ and an average value $r = 4$. Finally, if all the interval values are previously normalized by $< T_{dom} >$ as in Levelt (1967), the corresponding Gamma distribution obtained for the given observer should satisfy $\lambda = r$. A case of normalized dominance durations that has been fit by a Gamma distribution function can be seen in figure 1.2.

Interestingly, a Gamma distribution seems to describe the temporal dynamics of the bistable phenomenon induced by apparently distant stimulation conditions. Two remarkable examples are dichoptic masking (the two images are presented

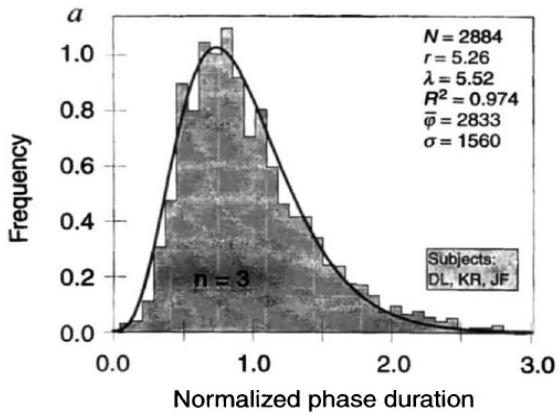

Figure 1.2: The successful fit of a Gamma distribution ($f(x) = \lambda^r/\Gamma(r)x^{r-1}\exp(-\lambda x)$) to the Logothetis et al. (1996) data. The ordinate axis shows relative phase durations, that is, phase durations expressed as a fraction of the mean dominance duration. The value $n = 3$ is the number of subjects whose data was used to create the histogram; N denotes the number of phases recorded from all observers; R^2 is the coefficient of determination (R^2 equals the ratio of the explained sum of squares of the frequency variable to the total sum of squares of this variable); and σ is the standard deviation of the relative durations. Reprinted from Logothetis et al. (1996).

in rapid succession, what leads to the perception of only one of them) van Boxtel et al. (2007) and motion-induced blindness (the continuous phenomenal appearance and disappearance of a small but salient object surrounded by a global moving pattern) Carter & Pettigrew (2003).

Other authors argue that a better characterization of the temporal dynamics of bistable perception can be obtained by fitting the perceptual alternation rates (*i.e.*, the inverse of each dominance duration value) instead the times spent in each percept Brascamp et al. (2005). In that work six subjects were presented with three kind of dichoptic stimuli and a classic perspective-reversible figure (Necker cube). In all the cases, the set of empirical instantaneous rate values showed less deviation to a Gamma function than the set of durations values, although a quotient of two Gamma functions (known as a Beta' distribution) was proven to be a reliable option. A combination of two Gamma functions (a Beta distribution) has been invoked previously by Borsellino et al. (1972) to mathematically model the set of predominance values (fraction of dominance time in each oscillation cycle) shown by any of the two possible interpretations of ambiguous figures.

On the other hand, although phenomenal changes in reversible figures are often experienced as sudden transitions from one compatible interpretation to the other, the case of dichoptic stimulation (*i.e.*, binocular rivalry) can generate a more complex perceptual behavior. For example, when the stimuli comprises a zone of about 1° or larger, each cycle of dominance can be followed by substantial

transition periods during which a compound of both (entire) original images is seen Bossink et al. (1993) or parts of the both images can emerge in different parts of the perceived pattern (*piecemeal* transitions). Besides, some of them can end up with dominance returning to the previously dominant eye (*return* or *failed* transitions) Mueller & Blake (1989) [1].

Finally, a set of empirical findings on binocular rivalry were early presented in a unified framework by Levelt (1966). Based on previous psychophysical studies and his own work, Levelt was able to enunciate various important propositions, concerning the perceptual behavior resulting from the manipulation of one or both external images. The central concept is the one of *stimulus strength*, a term that refers to the combined effect of stimulus parameters such as contrast, spatial frequency or amount of contour per area.

In his seminal work, Levelt presented four basic propositions. The first three laws concern the perceptual behavior when just one of the two stimuli is varying, while the other is held fixed. The fourth proposition describes the effect of a simultaneous (symmetrical) manipulation of both images:

(I) Increasing the strength of the image presented to one eye increases the predominance (fraction of total time) of that stimulus.

(II) Increasing the image strength presented to one eye does not affect the dominance duration of that stimulus.

(III) Increasing the image strength presented to one eye increases the rivalry alternation rate.

(IV) Simultaneous increasing of the stimulus strength decreases the average dominance duration of the percepts.

The first three Levelt's propositions can be summarized in a unique statement that can be derived from them; namely, that the decrease of the stimulus strength in only one eye just increases the dominance duration of the image presented to *other* eye. Such a statement is frequently refers in the literature as the Levelt's second proposition, and is the object of most of the analyses together with the fourth proposition. These psychophysical laws continue being the subject of currently active investigations to date, that in some cases have lead to important precisions of the basic Level's statements Brascamp et al. (2006) and in others to the extension of its validity to ambiguous figures Klink et al. (2008); Moreno-Bote et al. (2010).

The noise as an essential factor in bistable visual perception

The general theoretical framework usually used to explain some of the seemingly paradoxical results revised above consists of two (average) neuronal populations representing the possible perceptual outcomes, which compete for dominance in a

[1] In some contrast conditions, such uncompleted transitions to the other image may reach up about one half of all the transitions happened Brascamp et al. (2006).

process of mutual inhibition. Another essential ingredient is some kind of slow 'fatigue' process (*e.g.*, in the form of an adaptation current or a long-term synaptic depression) that acts as a negative feedback over the currently dominant population. Such a slow process shifts the balance from the currently dominant stimulus to the suppressed one every time some saturation level is reached. Examples of these models can be found elsewhere Mueller (1990); Laing & Chow (2002); Wilson (2003). Concerning the well known stochastic-like nature of bistable perception, this aspect is typically implemented by directly incorporating a source of noise in the equations for activities (firing-rates) of the simulated average neural units.

The adequate behavior of the models is frequently verified by successful fits of the dominance durations resulted from simulations to positively skewed frequency distributions Lehky (1988); Kalarickal & Marshall (2000); Laing & Chow (2002); Stollenwerk & Bode (2003). The resulting set of equations are then usually tested against the empirical behavior given by Levelt's propositions. That is, the non-monotonic curves that should describe the relation between the average dominance time and the current input strength (*e.g.*, contrast) shown in each eye or the common value impinging on both eyes Laing & Chow (2002); Shpiro et al. (2007); Curtu et al. (2008).

Nevertheless, neither the shape of the spontaneous dominance-duration distributions nor the profile of the mean phase duration as a function of the input strength provide sufficiently rigorous constraints to make a precise modeling. The unspecific character of many of the prevailing models can be seen in the fact that finding a convenient set of parameters representing adaptation (or depression), inhibitory interactions and noise, which reproduces appropriately positively skewed dominance-duration distributions with a given dispersion around a mean, is not a difficult task Shpiro et al. (2009). Fortunately, new empirical constraints have been obtained with innovative experimental procedures recently designed to collect more subtle aspects of the rivalry's dynamics Kim et al. (2006); Brascamp et al. (2006); Pastukhov & Braun (2011).

For instance, Brascamp et al. (2006) extensively studied the dominance and perceptual alternations characteristic of binocular rivalry by using a matrix of left-eye and right-eye contrast combinations, spanning the entire range from near the detection threshold to the theoretical maximum. Their results significantly extended the quantity and precision of data concerning dominance and transition durations and included new measures about the frequency of occurrence of the paradoxical return transitions (FRT). The transitions (as the dominance phases) showed to last for intervals of the order of seconds instead of being almost "instantaneous" events (as it is frequently treated in current models), and FRT, in some conditions, reached up about one half of all transitions starting from a given eye. Interesting tests with some current models for binocular rivalry concluded that all the present approaches exceed in the possible role that the slow adaptation variable might play during transitions, underestimating a potentially more fundamental role of fluctuations.

Additionally, Kim et al. (2006) also actively probed the dynamics of bi-stable perception by using a novel experimental paradigm in binocular rivalry, aimed

to uncover subtle relations between the deterministic (adaptation, inhibition) and stochastic forces involved. The experimental design was conceived to verify whether the timing of the perceptual switches could be explained by the presence of a special coupling between noise and the changing landscape created by the dynamics of the deterministic factors.

More specifically, these authors tried to prove the presence in bistable perception of a phenomenon already familiar in some physical and even biological systems: *s*tochastic resonance Gammaitoni et al. (1998). The phenomenon is typically considered as a purely noise-mediated one in which the sensitivity of a system to a relatively weak external periodic signal increases when the intrinsic average switching rate of the system is matched by the frequency of that signal.

Based on Levelt's proposition II (section 1.1.1), the luminance contrast was used as an image parameter able to alter the presumably existent landscape energy created by the deterministic factors and, consequently, the two perceptual states. By periodically changing in opposite phase the two image contrasts across a wide range of modulating frequency values, the authors successfully demonstrated resonance at the average (intrinsic) rate of the unperturbed perceptual switching. As additional signatures of resonance, Kim and co-workers also found higher-order resonance peaks (when modulation frequencies were appropriate) in the dominance-durations distributions at the odd-integer multiples of the modulating half-period Gammaitoni et al. (1998). More importantly, these authors made an interesting estimation of the magnitude of the relevant internal noise in terms of the contrast-modulation amplitude needed, a value between $20\% - 30\%$. Additionally, their work suggests that the magnitude of such a noise scales linearly with contrast, varying only with relative changes in the external stimuli, an observation that suggests the existence of some kind of gain control-type mechanism in bi-stable perception.

New intriguing effects have also been reported in Pastukhov & Braun (2011), who were able to measure subtle but significant correlation effects previously ignored in time series of bistable perception. In that work, a novel statistical measure - *cumulative history* - is proposed to account for the possible fatigue factors that modulate the temporal evolution of the perceptual process, by a sum of the past dominance and mixed phases that are weighted towards the most recent states. It is shown in binocular rivalry and ambiguous figures as well, that such a measure significantly correlates with the immediately next dominance period, in such a way that longer (shorter) dominance phases generally appear after shorter (longer) sequences of previous dominance periods. The correlations found are approximately twice larger than serial correlations between successive dominance phases, consistently revealing the presence of adaptation effects.

Nevertheless, Pastukhov & Braun (2011) also report that the time-scale of the relevant cumulative history correlates only loosely with the average dominance period shown by each observer. Additional evidence that fatiguing process do not determine solely the average perceptual behavior comes from the analysis of mixed ("patchy") phases during binocular rivalry. It is proven that both complete and failed (return) transitions can be dramatically affected by the eventual balance of the adapting states. That is, when no history bias exists between the

two competing percepts, patchy transitions last longer and the probability that return transitions occur notably increases. These findings concerning the transitional states, frequently ignored in the computation of sequential correlations (see van Ee (2009) as a rare exception), strongly suggest that perceptual reversals are fluctuation-driven in the absence of adaptive bias. Indeed, by computing the mutual information between past history and future dominance periods, Pastukhov & Braun (2011) have shown that adaptive forces could not explain more than 10% of the observed variability in reversal timing.

Taken together, the results reported in Pastukhov & Braun (2011), Brascamp et al. (2006) and Kim et al. (2006) impose new stringent model constraints and allow to make new inferences on the mechanisms supporting bistable perception.

1.1 The neural substrate of multistable visual perception

The way that visual perceptual rivalry occurs remains a mystery, although significant advance has been made by means of diverse techniques. On the one hand, psychophysical studies permit to obtain simple but very easy to reproduced behaviors such the mathematical form of the dominance intervals distribution Lehky (1995); Borsellino et al. (1972); Fox & Herrmann (1967); Brascamp et al. (2005) and the Levelt's propositions, that are valid not only for binocular rivalry but in ambiguous patterns as well Klink et al. (2008); Moreno-Bote et al. (2010), proving to be of extreme generality.

At the same time, many hypotheses can be evaluated about the mechanisms that the neural substrate uses in multistable visual perception. As it was mentioned above, two of the most prominent proposals are (*i*) that exclusivity of dominance is mediated by mutual inhibition between the neurons responding to the resulting percepts, and (*ii*) the presence of a fatigue or adaptation process which is responsible of eventually curtailing the current dominance of a given percept, then shifting the balance to the suppressed one.

On the other hand, non-invasive neuroimaging techniques with high spatial precision (fMRI) or better temporal resolution (EEG, ERP, MEG) can be used to locate the neural site where the percepts are neurally represented and the winning (suppressed) interpretation emerges to (fades from) awareness. This issue is the field of an intense debate that initially leaned in the direction of pure sensory processes taking place at the early stages in the visual pathway (*e.g.*, monocular neurons in the lateral geniculate nucleus (LGN) of the thalamus and primary visual cortex (V1)) Blake (1989); Tong (2001). The alternative position that considers the ambiguity being resolved by the intervention of higher-level processes is followed in single-cell studies Logothetis (1998); Leopold & Logothetis (1999).

However, an intermediate view is currently dominating the scientific research, where no bottom-up or top-down approaches are valid solutions for the problem of perceptual rivalry. Indeed, the ambiguity initially introduced by the nonexistence of clear cues in the sensory input could be successfully interpreted with the help of many levels of processing stages, but between the sensory regions and the

cognitive ones Long & Toppino (2004); Tong et al. (2006). Below, we succinctly revise most general experimental evidences in favor of either point of view.

1.1.1 Psychophysical studies

Until the application of the modern neuroimaging techniques, most of the work in multistable perception was empirical in nature. DuTour was among the first in studying the characteristic alternations of binocular rivalry when the two stimuli differ only by their colors (DuTour (1760), English translation by O'Shea (1999)). This author thought that the brain is actually able to capture only one image at a given time, being the other one inevitable suppressed from awareness.

Such a hypothesis has come to be known as the 'suppression theory', and was posteriorly examined by Wheatstone, who discovered the binocular contour rivalry by using a mirror stereoscope Wheatstone (1838). Nevertheless, his work also demonstrated that, in some cases, a pair of images with shifted local features may be perceived as a unique (stable) pattern with a fused impression of stereoscopic depth. Hence, our mind might experience a combination of the information received on the two retinae by means of an organic (physiological) mixture of them. Notably, Wheatstone also found that some degree of voluntary attention, although incapable to influence the appearance of any of the images, did seem to increase the time during which the perception stabilizes.

Suppression theory later received a strong support by Levelt's work who considered binocular rivalry as the direct result of reciprocal inhibition between monocular channels Levelt (1965). Additionally, such a cross-inhibition adapts over time, allowing the eventual recovery of the currently suppressed eye. This notion was later implemented in several mathematical models for bistable visual perception Sugie (1982); Blake (1989); Lehky (1988); Mueller (1990), being able to reproduce a known Levelt's result: increasing the strength (*e.g.*, contrast) of one monocular image has no effect on its predominance in perception but, instead, decreases the predominance of the stimulus presented on the contralateral eye Levelt (1965, 1966). Modeling works also propose that phenomenal fusion of the two inputs Liu et al. (1992) would occur when inhibition is low Lehky (1988).

Subsequent psychophysical investigations further supported the idea of interocular competition. For instance, the sudden exchange of two orthogonal dichoptic gratings does not change the currently dominant eye, but instead exchanges the dominance between the two stimuli Blake et al. (1980). Eye-specific effects can be derived from other works where the sensitivity to detect non-related probes was studied. For instance, Wales & Fox (1970) demonstrated that suppression operates non-selectively during rivalry of two foveally located 1° disks, by using brief monocular flashes whose detectability was significantly reduced during suppression than during dominance of the probed eye. Similar elevation of detection thresholds had been previously found for moving contours Fox & Check (1968), and even the independence of such effect on the suppression duration was shown in Fox & Check (1972).

Psychophysical experiments with ambiguous or reversible figures (*e.g.*, Necker cube), although more susceptible to cognitive influences (see Long & Toppino

(2004) for a review of related work) than dichoptically presented images, also offer strong evidence in favor of passive sensory mechanisms. That is the case, for instance, of the work demonstrating the presence of neural adaptation or fatigue effects[2]. During the perception of such figures, the observer systematically reports a progressive shortening of the dominance periods for each percept over a presentation period of a few minutes, that invariably leads to an asymptotic rapid rate of alternations. These effects often show a localized character, strongly depending on the retinal region that is being stimulated, that further indicates a passive neural process Toppino & Long (1987). These authors demonstrated that changing the position in the visual field and/or the size of a Necker cube produced a return to baseline in the response pattern.

Another related fact is the possibility to induce the reversal in a desired direction, by just previously exposing the subject to an unambiguous version of the reversible figure. The basic assumption behind this procedure is that the prior stimulus adapts the neuronal channel that contributes to the perception of the same interpretation, allowing the unadapted neurons to then dominate. Consistently, the observer frequently reports to perceive the alternate version when the standard ambiguous figure is shown.

However, Brascamp and colleagues have recently found, in binocular rivalry with two orthogonally oriented gratings, that a prior unambiguous stimulus can sometimes generate a pronounced *facilitation* effect rather that satiation to the presented pattern Brascamp et al. (2007). The final result actually depends on the arithmetic product of the contrast and the duration of the prior stimulus, a quantity termed as its 'energy'. That is, higher probe-energy does cancel initial dominance, but can be facilitated with low contrast and/or short expositions to the prior pattern. Additionally, the authors showed that practically the same holds when the probe stimulus was eye-specific (*i.e*, unrelated with the two dichoptical gratings) and when one of the gratings was used as the test pattern for both eyes, making impossible to discard a pattern competition as a possible explanation.

It is worth mentioning that these mixed effects found by Brascamp and coworkers can only be observed, if a blank interval is used before presenting the standard stimulation. Shorter blanks weakens facilitation, and only suppression of the initial pattern is experienced by the observer, in agreement with the adapting effects found by previous researchers. Notably, that stabilization in the same percept increases with longer delays (in the order of seconds), has been also reported in discontinuous presentation of the *ambiguous* stimulus Brascamp et al. (2008), a finding that supports the existence of some memory trace in bistable perception Leopold et al. (2002); Maier et al. (2003); Pearson & Clifford (2004).

Nevertheless, the interocular competition hypothesis fails to explain some intriguing adapting effects that have been shown during binocular rivalry. Contrary to the intuition, visual after-effects Kohler & Wallach (1944) probably caused by cortical mechanisms (*i.e.*, binocular neurons in later stages of visual processing), can be created independently of whether the monocular adapting stimulus co-

[2]Sometimes also termed as *satiation* processes.

incides with dominance of suppression phases of rivalry. Rivalry is unable to influence (*e.g.*, detection threshold) spatial frequency shifts after-effects Blake & Fox (1974), tilt after-effects Wade & Wenderoth (1978) or motion after-effects Lehmkuhle & Fox (1975). Suppression is neither able to reduce inter-ocular transfer phenomena of motion after-effects O'Shea & Crassini (1981), suggesting again that rivalry is resolved after the first site of binocular convergence (V1) in the visual pathway.

Clearly, another related phenomenon that is difficult to explain in terms of the suppression theory is monocular rivalry. Two low contrast patterns optically superimposed and presented to a unique eye (or dioptically, *i.e.*, the two eyes receive identical stimulation), can indeed show a weaker form of perceptual rivalry (with less frequent and sometimes uncompleted transitions) Andrews & Purves (1997). Interestingly, the alternation rate of the mixed pattern behaved similarly as in conventional rivalry, when changing specific figures parameters (*e.g.*, spatial frequency and size of each stimulus and difference in hue and orientation between them). The fact that only one monocular channel is needed to induced rivalry may reflect the functioning of a *stimulus* or *pattern* competition mechanism, apart from the existence of competition between monocular neurons.

1.1.2 Neuroimaging techniques

The first studies that addressed the problem of finding direct evidence of the neural substrate underlying multistable perception used scalp electrodes at the occipital lobe to record (if any) the electrical activity related with the observer's reports of dominance and suppression phases Lansing (1964); Cobb et al. (1967). Although most of them found a significant reduction in the amplitude of the collected responses that could be associated to the suppressed stimulus, such electrical signals hid the origin of such modulations (*i.e.*, eye or stimulus) because they were obtained by time-averaging the signals pooled over both eyes.

The study by Brown & Norcia (1997) was probably the first that demonstrated a clear correlation between the fluctuations in visual evoked potentials (VEPs) at occipital areas and the usual cycles of dominance and suppression reported by observers exposed to binocular rivalry. Four active electrodes were positioned over the occipital pole of eight subjects to detect the electrical signals at the scalp. By continuously modulating the contrast of two orthogonally oriented gratings at slightly different rates (5.5 and 6.6 *Hz*), they were able to 'tag' the signal corresponding to each percept. The luminance of each rivaling stimuli was previously adjusted to produce dominance phases no shorter than the filter's time-constant of the recording device. Then, a significant 'physiological rivalry' was demonstrated by a pair of VEPs waveforms oscillating in counter-phase and notably phase-locked to the subject's reports of dominance and suppression (significant correlations larger than 0.37).

The fundamental drawback of using scalp electrodes resides in its poor capability to locate the exact place where such signals coupled with perceptual fluctuations might be originated. Both the dimensions usually accessed (in the order of mm) and the folded structure of the occipital cortex make the task

difficult. In the last decades, fMRI techniques, which provide a higher spatial resolution, have contributed to better clarifying the source of such signals.

For instance, Lumer et al. (1998) reported that blood oxygen level dependent (BOLD) signals were modulated at multiple cortical areas, in synchrony with perceptual alternations between a face and a grating stimuli dichoptically presented. Such event-related hemodynamic signals were detected at the extrastriate areas of the fusiform gyrus (a brain region implicated in face recognition), but they were absent in the striate cortex. Remarkably, Lumer et al. (1998) also reported rivalry-modulated hemodynamic activity outside the visual system, in frontoparietal areas of the right hemisphere that participate in spatial attention, a finding that demonstrated for the first time that areas traditionally associated with cognitive abilities may also be involved in multistable viewing. Unfortunately, Lumer et al. (1998) focused on brain activations correlating with the perceptual reversals and not in the sequence of successive rivalry states (dominance or suppression of the face and grating stimuli). Given the nature of the rivaling stimuli, it seems unjustified to discard a determinant role of early visual areas and to assume that such fluctuations might be instigated by the mentioned cognitive areas.

A subsequent study by Tong and colleagues investigated the involvement of a pair of functionally specialized extrastriate areas in bistable perception: again the fusiform face area (FFA), but in conjunction with the parahippocampal place area (PPA) Tong et al. (1998). The latter region selectively responds to indoor and outdoor scenes, such as houses. The brain of various observers where scanned with fMRI techniques, while viewing a house and a face dichoptically presented to induce binocular rivalry. Higher signals were then detected at the FFA (PPA) when the face (house) was perceived and the other stimulus faded from perception. More importantly, the strength of the BOLD signals appeared to be as strong as those measured when the house and the face were intermittently interchanged externally to mimic the conventional rivalry, which suggests that the neural representations of both the dominance and suppression events are already completed at these stages of processing. Consequently, the conflict in perception induced by the dichoptic stimulation might have been already resolved *before* the time the BOLD signals arose at FFA and PPA.

Subsequent work has given a strong support to the hypothesis that the neural events associated with dominance and suppression cycles of rivalry might indeed be happening within the boundaries of V1. That is the case of Polonsky et al. (2000), who were able to tag the activity corresponding to each image in a pair of orthogonally oriented gratings of different contrasts, by noting that BOLD signals increase monotonically with stimulus contrast. Then, the hemodynamic signal decreased when the subject perceived the image with lower contrast and increased with the entrance in perception of the higher one, leaving a clear trail of consecutive periods of dominance and suppression. The sensory input changed over time, by moving the gratings or reversing their contrasts, in order to avoid adaptation effects. During rivalry, BOLD signals in four subjects were found to be $45-83\%$ as large as those evoked by the physical alternation of the monocular images. Similar results were obtained by electrodes that monitored the activity at nearby visual areas (V2, V3 and V4), although it was not demonstrated by

complete data sets from all the observers. However, such findings by Polonsky et al. (2000) could suggest that neuronal events related with the subjective rivalry could indeed be initiated at V1 and then propagated to later areas, where the neuronal representations of the coherent percepts may be reinforced.

Soon after, another study firmly demonstrated the importance of monocular neurons in the brain's selection of the conscious visual inputs. Tong & Engel (2001) addressed this relevant problem by first isolating the V1 blind-spot representation, a relatively large region of around 10 $mm \times 5$ mm containing only monocular neurons that represent a retinal area of approximately $4 \times 6^{\circ}$ with no photoreceptors. Each V1 blind-spot representation exclusively receives direct input from the ipsilateral eye (what also excludes the blind-spot part of the contralateral eye), and is large enough to elicit measurable hemodynamic signals if the immediate blind-spot's surround is stimulated.

The experiment then consisted of showing the subject a red vertical grating and a green horizontal grating superimposed in the visual location of the right eye's blind spot, but permitting to stimulate cortical areas of the ipsilateral eye (through the contralateral retina). The two stimuli could be independently viewed by each eye through red and green filter glasses. V1 was scanned during both rivalry and pseudo-rivalry (*i.e.*, physical alternation of the gratings) conditions. Consistent with predictions of interocular suppression, Tong and Engel found increased activity in the V1 blind-spot region when a subject perceived a grating in the ipsilateral eye, and negligible hemodynamic activity when the competing stimulus was seen. All the observers showed statistically comparable BOLD signals fluctuations during rivalry and direct monocular stimulation, what suggests that the perceptual conflict could be entirely resolved among monocular neurons in V1.

Magnetoencephalography (MEG) studies have also been carried out to relate subjective rivalry and neural activity signals distributed across a large number of brain regions Tononi et al. (1998), presumably reflecting the synchronous spiking activity of large populations of cortical neurons involved in the process. As Brown & Norcia (1997), Tononi et al. (1998) used orthogonal gratings that flickered at slightly different frequencies (between $7 - 12$ Hz) with the aim to identify the neuromagnetic responses corresponding to each reported percept. A dense multichannels design (148 electrodes) allowed to detect significant activity related with both dominance and non-dominance of a particular percept that appeared at occipital but also at temporal and frontal areas. However, whenever any of the stimulus became suppressed from awareness, a reduction of $50 - 85\%$ in the amount of power at the associated tagging frequencies was found in the spectra of the signals recorded, relative to the signals measured during pseudo rivalry (*i.e.*, physical stimulus alternation) trials.

Another MEG study by the same group later demonstrated that dominance phases are reflected by marked increases of neuromagnetic responses arising in widely distributed regions such as the occipital, temporal, parietal and frontal lobes, the latter showing the larger levels of synchronization Srinivasan et al. (1999).

1.1.3 Single-cell recordings

Some of the first neurophysiological studies addressed to find the site of rivalry were focused in early visual processing stages: the lateral geniculate nuclei (LGN) of the thalamus and the primary visual cortex (V1). The laminar structure of LGN is particularly suitable to investigate the existence of interocular (inhibitory) competitions, since the retinal terminals from each eye project to different laminae in this subcortical center. Sengpiel et al. (1995); Sengpiel (1997) recorded the response of neurons in LGN and (simple and complex, monocular and binocular) cells in V1 of anesthetized cats, for drifting gratings of different orientation, spatial frequency and contrasts. However, only *binocular* cells in V1 showed reduced responses (compared to the one during monocular stimulation) to a grating presented to one eye when another (optimally) oriented grating was shown to the other eye, and over a wide range of spatial frequencies. Almost half of the LGN neurons studied exhibited modest inhibitory effects, but always irrespective of the orientation between the two gratings, what made them unlikely to serve as a neural basis for competition between differently oriented stimuli. However, a possible inter-inhibition in V1 between binocular neurons was not discarded, which might take place place at the level of their corresponding (adjacent) ocular dominance columns.

Subsequent single-cell recordings in the LGN of alert monkeys confirmed the Sengpiel *et al.*'s results, providing no empirical evidence for rivalry suppression at such subcortical level Lehky & Maunsell (1996). Those authors recorded the activity of (52) parvocellular and (27) magnocellular neurons and used the same kind of stimuli (drifting gratings at 2.0 Hz) under rivalry, congruent and monocular stimulation conditions. Both peri-stimulus time histograms (PSTHs) and the power spectra of the neuronal responses were analyzed. No significant differences were apparent among the distinct studied conditions.

More complete investigations were carried out by Logothetis and coworkers with behaving monkeys presented with bistable displays (see Logothetis (1998); Leopold & Logothetis (1999) for reviews). They investigated the neuronal activity elicited in single neurons located at several levels of the hierarchical organization of the brain, including V1, the extrastriate areas V2 and V4 Leopold & Logothetis (1996), as well as nervous cells in the superior temporal sulcus (STS) and the middle-temporal (MT) Logothetis & Schall (1989) and inferior temporal (IT) cortical regions Sheinberg & Logothetis (1997).

Their first attempt to uncover the neural substrate of bistable perception concerned the study of the middle temporal area (MT or V5), a region that shows pronounced motion direction selectivity Logothetis & Schall (1989). Three rhesus monkeys were trained to discriminate between two vertically drifting gratings dichoptically presented through a stereoscope. The mixed stimulus could be seen as moving in upward or downward directions. After replacing the rivaling stimuli by two spots on the left and on the right, the monkey reported the direction perceived by a saccadic movement to the left spot (for downward motion) or to the right one (for upward motion). Half of the trials consisted in congruent stimulations (*i.e.*, the two gratings drifting in the same direction) in order to

evaluate the directional selectivity of the cells.

In their joint work, Logothetis and Schalls found distinct unexpected neuronal behaviors that could mediate the phenomenon of binocular rivalry. Firstly, 32% of the recorded neurons that seemed to receive no inhibition connections, showed significant directionality in the vertical axis during non-rivalrous conditions, and continued responding *wherever* their preferred stimulus was present during rivalrous trials. Secondly, other 21% of the cells displayed (dynamic) vertical directionality during rivalry, although presenting a clear preference for *horizontal* or *oblique* directions of motion during standard stimulation. Finally, the activity of 22% of the recorded cells did appear to be modulated during rivalry. However, half of these neurons discharged only on those trials in which the monkey reported to perceive the motion in the cell's preferred direction, while the other half responded when its optimal stimulus was present in the *suppressed* eye Logothetis & Schall (1989).

Despite the complex scenario of interactions and information transmission suggested by Logothetis & Schall (1989), the presence of rivalry-modulated neurons could indicate that the perceptual conflict can be fully resolved at MT. However, the participation of lower/higher visual centers can not be discarded, given that such neuronal responses typically appear earlier than the monkeys reports. A subsequent study addressed the same problem, by recording neuronal activity from the areas V1/V2 and V4, while other two animals experience the binocular rivalry induced by two orthogonal gratings Leopold & Logothetis (1996). The (binocular) cells of these cortical areas differ in orientation preference and in their ocular dominance (*i.e.*, the extent to which the neuron is driven by each eye) as well.

As a typical behavior of cells of V1/V2 and V4, PSTHs aligned to the animal's reports showed that the firing rate of many neurons recorded increases immediately before the onset of dominance of the cell's preferred orientation, and decreases before the report of the orthogonal grating(for instance, see Fig. 3 of Leopold & Logothetis (1996)). Besides, other PSTHs demonstrates the existence of some neurons in V4 behaving in the reverse manner, discharging more with the onset of the suppression of its preferred orientation during congruent stimulation, a behavior resembling the one found in some cells of MT Logothetis & Schall (1989), and that may explain the presence of after-effects during rivalry (see subsection 1.1.1).

The proportion of neurons (all but one binocular cells) that showed rivalry-modulated activity was lower in V1/V2 ($\sim$ 18%) than in V4 ($\sim$ 38%), with around the half of them firing more in association with the suppression of the rivalrous orientation (only in V4). Such a high percent of modulated cells makes unlikely that competition could be explained by the interocular suppression theory (*i.e.*, competition of monocular channels at the striate cortex - V1). Importantly, some of them were non-selective neurons during congruent stimulation and just became selective during rivalry, as in Logothetis & Schall (1989).

A third study, Sheinberg & Logothetis (1997) has shed more light on the neural bases of visual awareness during rivalry, by analyzing the activity of single neurons at regions specialized in representation of visual objects: the superior

temporal sulcus (STS) and the inferior temporal cortex (IT). Binocular rivalry was instigated in the visual system of two macaques by presenting effective and ineffective stimuli to a different eye through a stereoscope, that were chosen from a set of hundred of visual images (for example a sunburst-like pattern, images of humans, monkeys, apes, butterflies, reptiles, among others)

Data collected in their work during a flash suppression paradigm, that ensures the onset of a not previewed pattern after being adapted to the competing stimulus, yielded an activity temporal profile very similar to the one recorded during rivalrous stimulation. The neurons discharged before and during the period that monkey reported seeing the effective stimulus, and ceased firing when the competing (ineffective) pattern become dominant.

In general, different activity patterns were found across the populations of neurons analyzed, such as a sustained discharging, a periodic burst or a clear transient response. Consequently, not only a conventional counting of spikes within a specified time window was used to characterize the cells responses, but complemented with dimensionality reduction techniques and statistical tests to reliably compare the cell's response two ineffective and effective stimulus. Nevertheless, 90% of the recorded cells in STS and IT were found to reliably predict the subjective state of the monkey in non-rivalrous, rivalrous and flash-suppression conditions.

1.2 Outline of the thesis

The present work is structured as follows. Next chapter provides a revision of some of the very recent experimental findings that conform the current view of bistable visual perception. Special attention receive those psychophysical works intended to uncover the neural mechanisms underlying the subjective experience of subjects presented with ambiguous sensory inputs. Specifically, we present in some detail the recent psychophysical results reported in Pastukhov & Braun (2011); Kim et al. (2006) and Brascamp et al. (2006), that provide new insights into the nature of bistable perception. All the studies refute current models which overestimate the role of the slow self-adaptation in deciding the occurrence of switches and of the cross-inhibition in abruptly precluding the dominance of the both images at the same time. Oppositely, these works point to a system in which the fluctuations could play an important role in determining the kind and the duration of transitions, as well as the periods of dominance for the perceived image. Chapter 2 ends with a succinct description of some of the prevailing models used in bistable perception, making emphasis in those previously tested against the recent experimental findings described in the chapter, and the ones that are considered in this thesis.

Chapter 3 contains the first block of results of this work. The methodology of Shpiro et al. (2009) is extended to fit various rate models to the experimental results of Pastukhov & Braun (2011), who have recently quantified the existence of hysteresis effects in unprecedented detail. The preference from among other modeling choices is based on the interest to evaluate dissimilar neuronal mecha-

nisms that could support the phenomenon of bistable perception. Consequently, different situations are considered, such as the presence of a slow fatigue process given by (linear and nonlinear) spike-frequency adaptation currents or a long-term synaptic depression, different architectures and activation functions.

The suitability of the parameter region found for each of the models considered in Chapter 3 is subsequently examined in Chapter 4. Then our results are further tested against known constraints given by behavioral responses found in experiments. That is, the biological plausibility is considered with respect to Levelt's propositions Levelt (1966) and the stochastic resonance effects reported by Kim et al. (2006).

The most important conclusions of this work, together with a brief digression about some possible directions of future research, are provided in Chapter 5.

2 State of the Art

In the following, the intense debate concerning the neural site where the perceptual conflict could be resolved is briefly revised, taking into account the most recent empirical evidences that favor either sensory or higher order processing stages. Then, we present in some detail the recent psychophysical results reported in Pastukhov & Braun (2011); Kim et al. (2006); Brascamp et al. (2006), that provide new insights into the nature of bistable perception. All the studies refute current models which overestimate the role of the slow self-adaptation in deciding the occurrence of switches and of the cross-inhibition in abruptly precluding the dominance of the both images at the same time. Oppositely, these works point to a system in which the fluctuations could play an important role in determining the kind and the duration of transitions, as well as the periods of dominance for the perceived image.

2.1 Where the perceptual ambiguity is resolved: interocular *versus* patterns competition

According to the results of the latest single-cell studies on the neural basis of multistable visual perception (subsection 1.1.3), the percent of neurons whose activity changes significantly correlate with the subject's phenomenal experience notably increases from lower to higher level of processing stages in the brain of behaving monkeys. In such scenario, only a limited role could be associated to the primary visual cortex, where just about 10% of the cells may participate in the characteristic successive cycles of suppression and dominance of multiple interpretations that are compatible with a unique physical stimulus Leopold & Logothetis (1996). In the same work, the authors have reported that later areas in the visual pathway (V2,V4) contain a larger percent ($\sim$ 20% and $\sim$ 40%, respectively) of neurons whose firing-rate is significantly modulated by the course of rivalry. Similarly, it was found in an early work that the activity of approximately 40% of the recorded cells in MT (or V5) appeared to be correlated with the perception of rivaling moving stimuli Logothetis & Schall (1989).

Subsequent primate single-unit recording studies by Logothetis and colleagues, have strongly suggested that temporal areas could indeed be the neural site where the phenomenal suppression of compatible interpretations could occur Sheinberg & Logothetis (1997). Then, when an alert behaving monkey was presented di-

choptically with an irrelevant and a preferred stimulus (a face), approximately 90% of the recorded cells from IT and the superior temporal sulcus (STS) were found to reliably predict the subjective state of the animal. These works further support a view of competition between incongruent interpretations of a given sensory input (*i.e.*, a pattern rivalry) rather than competition between monocular channels Leopold & Logothetis (1999).

Nevertheless, some neuroimaging studies provide important new findings supporting the conventional view of the suppression theory (section 1.1.2). For instance, the fMRI study by Tong & Engel (2001) clearly demonstrated the involvement a monocular neurons of V1 in the resolution of the perceptual conflict instigated by a pair of orthogonal gratings. Consistent with the view of interocular competition, Tong and Engel found increased activity in the V1 blind-spot region when a subject perceived a grating in the ipsilateral eye, and negligible hemodynamic activity when the competing stimulus was seen. All the observers showed statistically comparable BOLD signals fluctuations during rivalry and direct monocular stimulation, what suggests that the perceptual conflict could be entirely resolved among monocular neurons in V1.

Many factors have been indicated to explain the clear discrepancies between these neuromaging studies in humans and the single-unit recordings in behaving monkeys by Logothetis and co-workers. These include smaller neuronal samples in single-cell studies than in fMRI scanning, the unknown sequence of events that relates neuronal firing-rates and BOLD signals (that could also include subthreshold synaptic activity), anatomical differences between the two species; the reduced transients during possible blended states (transitions) that could be confounded with state-related signals recorded in single-cell studies with monkeys, but that were excluded in the neuroimaging studies; and ultimately the effects of inevitable eye movements in the signal recorded from neurons in V1/V2 with smaller receptive fields than the spatial window used in monkey's fixation Polonsky et al. (2000); Tong & Engel (2001).

An interesting proposal have been made by Polonsky and colleagues in order to reconcile the evident differences between the two kind of studies Polonsky et al. (2000). The idea consists of a reanalysis of the data obtained by electrophysiological studies in behaving monkeys, based on the fact that fMRI signal depends not only on the number of neurons but also on their firing rates. Specifically, a more reliable estimation of the strength of rivalry modulations can be obtained by comparing the neuronal activity during (physical) stimulus alternation with the absolute value of the responses during rivalry. An estimation of each response magnitude is defined in single-unit studies by a "firing-rate modulation index", which is computed as $(R_p - R_n)/(R_p + R_n)$, where R_p and R_n stand for the response to the preferred and the non-preferred stimulus, respectively. Accordingly, a percentage of rivalry modulation may be given by the ratio of the two response indexes during rivalry and pseudo-rivalry conditions averaged across the given visual area Polonsky et al. (2000), a measure that is also termed as "rivalry modulation index" in Tong (2001).

Using these relative values of neuronal responses, Polonsky and collaborators have concluded that previous single-neuron measurements reported by Lo-

gothetis's group might have underestimate the average activity at early visual areas. In contrast to these studies, the average rate modulation in V1 was found to be 33% as large as that evoked by physical alternation of the two rivaling images. Moreover, comparable values of modulation indexes of 23% and 27% were obtained for V2 and V4, respectively, in agreement with roughly similar fMRI (relative) measures of 56%, 42% and 51% across the same areas. It is worth to say that, however, both hemodynamic characterization and the reanalysis of the electrophysiological data reported in Polonsky et al. (2000) did not distinguish among the subpopulations of neurons with different responses patterns found during the primate single-cell studies Logothetis & Schall (1989); Leopold & Logothetis (1996); Sheinberg & Logothetis (1997).

2.2 Psychophysical findings supporting a fundamental role of noise

Noise in activity levels associated with both percepts causes a random distribution of dominance intervals Levelt (1967); Logothetis et al. (1996). The histogram of dominance durations shows a unimodal and skewed distribution, with a long tail at long durations. Fits to the data of a Gamma function are commonly found Logothetis et al. (1996); Kovacs et al. (1996), but other kind of functions fit the experimental data well Laing & Chow (2002). Another important aspect concerning the stochastic nature of binocular rivalry is the lack of any strong correlation beyond zero lag Logothetis et al. (1996).

Stochastic resonance behavior - Kim et al. (2006)

Perceptual bistability can be thought of as a sequence of multiple spontaneous, apparently random, switchings between two marginally stable perceptual states. Such a dynamical framework can be more clearly conceptualized as a motion across an energy landscape J. Hertz & Palmer (1991). Under this theory, the evolution reported by a subject to a new perceptual state is considered to be analogous to the dynamical development of a physical system towards its minima of energy. Once they are reached, the time of residence in such stable fixed points (attractors) can be corresponded to a given dominance time. Additionally, some quantity of noise could eventually allow the system to overcome the energy barrier limiting the basin of attraction of the currently occupied state, making possible the occurrence of stochastic incursions to the empty well related with the next perceptual outcome.

In the framework explained above, slow neuronal fatigue processes (e.g. adaptation, synaptic depression), stimulus strength and cross-inhibition would play a secondary role as direct causes for the switches between the possible interpretations available for the stimuli presented. However, these deterministic factors do actively participate in the conformation of the energy landscape of the system, forming or even continuously modifying aspects that define its final structure, such as the position of the minima in the state space, the relative location be-

tween them and the energy depth of their attraction basins Moreno-Bote et al. (2007).

In order to test such an intriguing hypothesis, Kim et al. (2006) intensely probe the dynamics of bi-stable perception by using a novel experimental paradigm in binocular rivalry, aimed to uncover subtle relations between the deterministic and stochastic forces involved. The experimental design was conceived to verify whether the dynamics of the perceptual switches could be explained by the presence of a special coupling between noise and the changing landscape created by the dynamics of the deterministic factors. Concretely, these authors tried to demonstrate the presence in bistable perception of a phenomenon already familiar in some physical and even biological systems: *stochastic resonance* Gammaitoni et al. (1998).

Based on the Levelt's proposition II (section 1.1.1), the luminance contrast was used as an image parameter able to alter the presumably existent energy landscape. Due to the inverse relation between the increase of the contrast of one image and the dominance duration of the competing percept, a counter-phase periodic modulation of the contrasts of both images could be assumed to change accordingly the depth of the corresponding wells in opposite phase. This is, increasing (decreasing) the contrast of one image should make shallower (deeper) the well representing the other stimulus.

A pair of incongruent stimuli (Fig. 2.1) were then presented to three human subjects, with an additional weak periodical modulation of its contrasts, taking as reference a higher fixed contrast (baseline), and spanning a wide range of frequencies ($0.28 - 2.48$ Hz) within the range of spontaneous alternation rates. The procedure was expected to alter the underlying structure of the energy surface and, consequently, to affect the perceptual dynamics dramatically. If succeeded, a visible change in the probability distribution of dominance durations should be obtained when the frequency of the periodic perturbation matches the natural alternation rate, with an increased concentration of the probability around the modulation half-period (HP), a well known signal of stochastic resonance Gammaitoni et al. (1998).

Concerning the modulation of the contrast of the stimuli, a square-wave modulation was preferred to keep the impacts of the rising and falling components of the external signal constant across the range of frequencies tried [1]. The higher contrast is taken as the baseline reference for the modulation, being the lower contrast a variable parameter to obtain different contrast modulations. To calculate the contrast C of a given stimulus, the authors used the definition $C = \frac{L_{Stimulus} - L_{Background}}{L_{Stimulus} + L_{Background}}$, where L stands for luminance. A percent contrast modulation is then defined as $M = \frac{C_{Baseline} - C_{Lower}}{C_{Baseline}} \times 100\%$. The interval of frequencies examined were considered either in the ascending or descending order across sessions, keeping fixed the baseline contrast and the relative percent of modulation.

[1] For more details about this choice, see the explanation given in the footnote at page 396 of Kim et al. (2006)

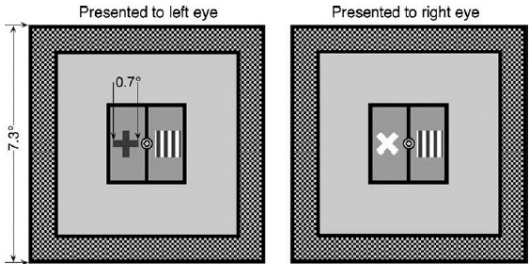

Figure 2.1: The stimuli used to detect stochastic resonance in binocular rivalry were presented dichoptically using a four-mirror stereoscope. A high-contrast texture frame was added to help in stable binocular alignment, and the rivaling patterns (a dark "+" shape and a light "x" shape) were presented along with a non-rivaling grating on the right side to balance the overall stimulus configuration and help also stabilize fixation. The rivaling shapes were chosen to be small ($< 1\,^\circ$ visual angle) to avoid perception of mixed patterns. The luminance contrasts of the rivaling stimuli were additionally modulated in opposite phase at several frequencies intending to find evidences of stochastic resonance. Adapted from Kim et al. (2006).

One of the authors and other two observers that were naïve to the purpose of the experiment, were presented dichoptically with a pair of small dissimilar stimuli, that facilitate exclusive dominance of only one pattern each time (Fig. 2.1). In general, no blinking was allowed to all observers during continuous stimulus presentation, but one of them was also tested permitting normal blinking in order to establish the independence of the results of such factor. The trials in this case lasted for 60 sec for that subject, and 16 secs with a short break in between for all the subjects in the no-blink condition. Finally, $C_{Baseline} = 0.50$ was always used in the blinking allowed sessions, with either $M = 40\%$ or $M = 20\%$. $C_{Baseline} = 0.50$ and $C_{Baseline} = 0.25$ were tested in the no-blink conditions (all observers), with both $M = 30\%$ or $M = 20\%$ in each case, in order to study the dependence of the results on the values chosen for those control parameters.

Additionally, a control condition was considered, *i.e.*, the dominance-duration distribution in the absence of an effective contrast modulation and at a frequency value slower than the maximum spontaneous switching rate (HP = 6 sec in the blink-condition, and HP = 8 sec otherwise). The static image contrast matched the baseline used in the given experimental condition under study. Each session lasted 1-2 h with at least 2-min break between trials, to allow the visual system to safely recover from adaptation.

Figure 2.2 (lower half) shows the resulting dominance duration distributions, when the contrast of rivaling shapes were subjected to a periodic modulation at a given frequency. Notably, a (primary) peak at the modulation HP grows in size as the frequency of modulation is increased, reaching a maximum at about 0.50 Hz, around the value of resonance. Higher-order peaks with diminishing gains turn

also to be visible at odd-integer multiples of HP when the frequency is increased beyond the resonance value, a quantitative result that is erroneously identified in this work as a characteristic signature of stochastic resonance. Actually, this fact must be considered as a less specific kind of input-output synchronization Gammaitoni et al. (1998). Such peaks would also reflect the tendency of the perception to oscillate in phase with the driving periodic signal incorporated in the images and consequently with the potential wells. But such peaks at odd multiples of HP (so longer than HP) can be explained as the intend of the perceptual state to switch at the next moments in time when the landscape energy offers the same favorable situation as at t = HP, once it failed to make the transition the first time.

Because the results were very similar across individual cases and distinct experimental conditions, although noisier due to smaller quantity of data points, all of them were averaged across the three subjects, the values 0.50 and 0.25 used for the baseline contrast, and the blink and no-blink conditions. The percent contrast modulation was $M = 30 - 40\%$, since a lower value (20%) was ineffective to induce stochastic resonance (see explanations below, related with Figure 2.3).

To estimate accurately the resonance frequency, Kim and co-workers used the coefficient of variation (CV), another index typically used to demonstrate the presence of stochastic resonance Pikovsky & Kurths (1997). Also known as the noise-to-signal ratio, since it is computed as the quotient between the dispersion and the mean of the dominance-duration distribution, it can easily indicate a high grade of periodicity by lower values of CV when plotted as a function of the driving frequency. Consequently, the presence of resonance is usually shown by a sharp dip at the natural frequency of the system.

Figure 2.3, showing the data corresponding to different observers and in various experimental conditions, clearly demonstrates that the spontaneous alternation rate (located at the vertical gray bands) is being successfully matched with the external frequency that produces the greatest periodicity. As before, the results remain valid for all observers and no matter the baseline contrast that is used in the modulation signal. However, a sharp dip in the CV profile is evident when the percent modulation was $M = 30\%$ (primary graphs), but visibly reduced or absent for $M = 20\%$ (insets in the first and second rows). The strength of the resonance, quantified by the the size of the dips, seems to be independent on the baseline contrast, given that the percent modulation is the same. The presence of such a kind of divisive normalization of the system to the baseline contrast was verified for a wide range of values $M \sim 0 - 100\%$.

Further quantitative analyses by using the P_1 index of the dominances distribution [2] was made in Kim et al. (2006), that definitely confirmed that 30% and 40% contrast modulation were effective to induce stochastic resonance, whereas the value of $M = 20\%$ was not.

The last finding lead the authors to make an intriguing estimation of the relevant noise underlying the perceptual switches: a value between $20\% - 30\%$.

[2] P_1 is defined as the proportion of probability between HP $\pm$ 1/2 HP. For more details, the reader is referred to Gammaitoni et al. (1998).

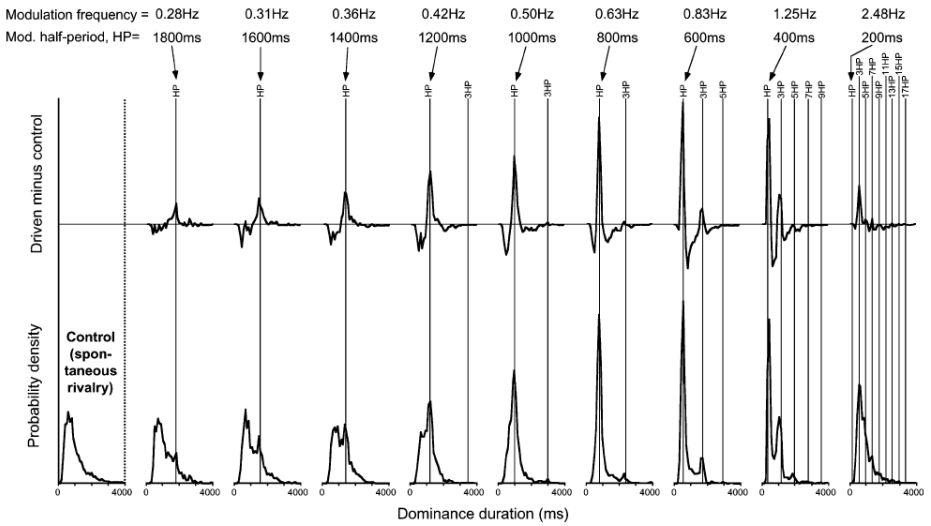

Figure 2.2: Distributions of perceptual dominance durations when the contrasts of the competing stimuli were modulated in opposite phase, sweeping the frequency in a range of 0.28-2.48 Hz (the corresponding driving half-periods (HP) are also shown). The bottom graphs show the presence of a primary peak at HP (indicated by the vertical lines) growing in size as the frequency approaches the spontaneous alternation rate, a signature of stochastic resonance. Higher-order peaks at peaks in the probability distributions at the odd-integer multiples of HP appears beyond resonance, as part of the synchronization between the system underlying binocular rivalry and the periodic signal. In the top half, the gain due to the periodic perturbation is isolated, by subtracting a control distribution from the lower plots (see text for more details). The distributions have been averaged across observers and experimental conditions to obtain smother curves. Adapted from Kim et al. (2006).

That speculation would be based in the following assumptions. The amplitude of the external modulation must be finely tuned to the amount of noise involved in the perceptual switches. Neither lower nor higher modulation signals could be adequate to elicit resonance, being the first case insufficient to influence the system dynamics and the second too high thus completely defining the rivalry dynamics. Then, finding the appropriate level of external signal eliciting stochastic resonance, would be equivalent to find an estimation of the magnitude of the internal noise participating in the perceptual switches.

Additional insights about the nature of the neural system underlying binocular rivalry were obtained in Kim et al. (2006), by attempting to reproduce theirs results with three of the current models mainly designed to reproduce statistical properties of bi-stable perception (Lehky (1988); Wilson (2003); Mueller (1990)). Intriguingly, only the first one, characterized by linear interactions among stim-

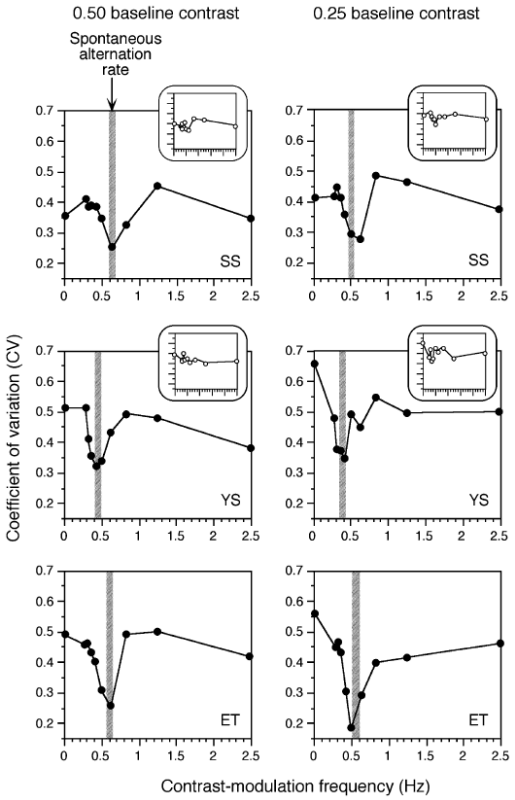

Figure 2.3: Coefficient of variation (CV) of the dominance-durations data as a function of the driving modulation frequency. The data shown correspond to baseline contrasts of 0.50 (left panels) and 0.25 (right panels) for three subjects (SS, YS and ET). The primary graphs show the results for $M = 30\%$ contrast modulation and the insets (observers SS and YS) refer to the results with $M = 20\%$ of percent modulation. The resonance (indicated by the sharp dips in the profiles) clearly happens at the spontaneous switching rate (located at the gray bands), but only when the modulation contrast is high enough, being $M = 20\%$ ineffective. Adapted from Kim et al. (2006).

ulus strength, adaptation, cross-inhibition and noise was able to fit the experimental data as in Figure 2.2. Importantly, this model also clearly differs from the others two, in a threshold mechanism to implement exclusive dominance of one of the competing images, and failed to reproduce the data if it is substituted to a winner-take-all algorithm [3] as is the case of Wilson (2003); Mueller (1990) (see section 2.3 for a brief presentation of some of these models). Curiously, the Lehky (1988) model is based on a Schmitt trigger, a model system that is

[3] See the footnote at page 404 of Kim et al. (2006).

known to exhibit stochastic resonance Melnikov (1993).

In summary, Kim and collaborators designed a novel experimental paradigm that proved for the first time the possibility that the machinery responsible of binocular rivalry can also produce stochastic resonance. Besides, they made a quantitative estimation of the quantity of noise that seems to be involved in the perceptual reversals, and demonstrated that the magnitude of such a noise scales linearly with contrast, varying only with relative changes in the external stimuli which points to the presence of some kind of gain control-type mechanism in bi-stable perception.

Next we continue with the exposition of another important work where the perceptual system is subjected to new stringent conditions, this time by presenting the subjects with images whose contrast spanned the whole range from detection threshold to maximum. This work allows to analyze the findings of Kim et al. (2006) from a new perspective, since it offers the first opportunity in more than thirty years to precise Levelt's work, which is the theoretical basis of the experimental design made by those authors to elicit stochastic resonance in binocular rivalry.

The time course of binocular rivalry - Brascamp et al. (2006)

In line with the notion of a nonlinear flip-flop is the finding that interactions between rivalrous percepts are strongly asymmetrical. Increasing the contrast in one eye may have a large effect on the duration of the percept associated with the other eye, whereas it does not affect durations for the ipsilateral percept (Levelt's second proposition) Levelt (1965, 1966). More recent work demonstrated small changes in the eye with the variable contrast as well Bossink et al. (1993).

Brascamp et al. (2006) contributed to a more complete view of the dynamics of binocular rivalry. In their work, the phenomenon was extensively studied by using a matrix of left-eye and right-eye contrast combinations spanning the entire range from near the detection threshold to the theoretical maximum transitions. The authors aimed to get a more precise frame on the contrast dependence of dominance and transition durations and that of the occurrence of *return transitions*; *i.e.*, occasions when an eye loses and regains dominance without intervening dominance of the other eye.

In their experiments, Brascamp and co-workers used as stimulus a pair of sine-wave gratings, orthogonally oriented rightward tilted (45°) for one eye and leftward (−45°) for the other. Four subjects reported percepts by pressing and holding either of two buttons corresponding to exclusive visibility of either eye's image or releasing all keys in case of a transition. A third button was pressed when the observer was not sure about the perceived percept (*e.g.*, due to temporally unaligned eyes), because releasing the keys is also a natural response in such cases. Contrast conditions were distributed randomly over trials. Sessions consisted of four 5-min experimental trials. Each observer was presented with a 4×4 matrix of contrasts values spanning the full domain of left-eye/right-eye contrast combinations. Each subject produced an average of about 180 domi-

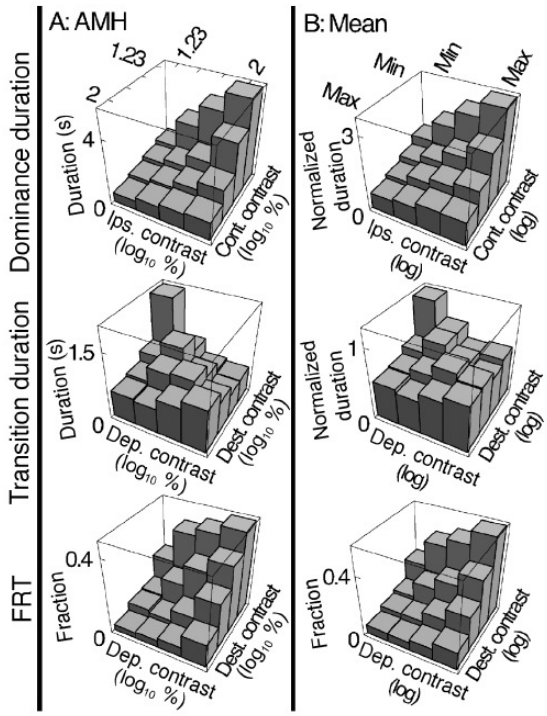

Figure 2.4: Dominance durations, transition durations, and the FRT as a function of the two eyes contrasts, for one subject (A) and averaged over all four (B). Ipsilateral (ips.) and contralateral (cont.) refer to the eye whose dominance durations are plotted and to the other eye, respectively. Contrasts were customized for each subject, with "Min" meaning near detection threshold and "Max" meaning 100%. See the text for further details. Reprinted from Brascamp et al. (2006).

nance durations per contrast condition per eye and the accompanying transition durations.

Figure 2.4 shows the results for a typical subject (Panel A) and averaged over all four (Panel B). The top and middle charts show dominance and transition durations; the bottom ones show the fraction of return transitions (FRT), that is the fraction of transition phases after which dominance returned to the previously dominant eye, remaining suppressed the image presented to the other eye. All durations in Panel B were normalized per subject relative to his/her dominance duration in the 100%/100% contrast condition; thus obtaining proportional relations between dominance and transition durations.

Since none of the observers showed a significant eye preference, the data was pooled over eyes. As a result, concerning dominance durations, contrasts are shown respect to an ipsilateral eye (the one whose dominance durations are being plotted) and its contralateral one (the other eye). Concerning transitions, the

terms "departure" and "destination" contrast refer to the eye that was dominant before that blended state and the other, respectively.

Figure 2.4 shows the surprising results that dominance and transition durations are both on the order of seconds and that the FRT varies between about 0 and as much as 0.5, depending on condition. All three quantities show a systematic dependence on contrast, which is similar for dominance durations and the FRT but different for transition durations. In the first case, both quantities increased (decreased) with the ipsilateral/departure (contralateral/destination) contrast. On the contrary, transition durations (middle) behaved differently, with the roles of departure and destination contrast being largely equivalent. Consequently, transition durations became longer when both contrasts are minimal.

Brascamp and co-workers carried out some computational experiments trying to reproduce their novel results. They used three prevailing models mainly based on adaptation and cross-inhibition to mainly reproduce the well known non-linear characteristics and stochastic properties of rivaling behavior. They used the Wilson's approach Wilson (2003) and the Kalarickal and Marshall's model Kalarickal & Marshall (2000) to account for superposition transitions and the Stollenwerk's model Stollenwerk & Bode (2003) in the attempt to reproduce picemeal-kind mixed phases, as this model present an spatial dimension (see section 2.3).

Figure 2.5 show the computational results. Apart from the contrast ranges in which Levelt's 2nd proposition is verified (white bars), the panels A-C present an important deviation from the experimental data (figure 2.4), predicting, for instance, too much short transitions and underestimating the FRT. Panel D presents the results of just one parameters region for the Kalarickal & Marshall's model in which it was able to (qualitatively) fit the original data. In this case, the noise is directly applied to the habituation variables, instead to its rates as in the original formulation.

Nevertheless, at these conditions it is verified that the model loses its capability to produce suitable Gamma-like time dominance distributions, because the dynamics is then determined by only one attractor of intermediate activity levels (figure 2.6), which is continuously driven exclusively by noise without the intervention of deterministic factors (adaptation and mutual inhibition). The authors concluded that all the present approaches exceed in the possible role that the slow adaptation variable might play during transitions, underestimating a potentially more fundamental role of fluctuations (noise).

Regarding dominance durations, Brascamp *et al.* shows that the widely accepted rule known as Levelt's second proposition is only valid in a limited contrast range; outside this range, the opposite of the proposition is true. Then, it should be replaced by the more precise statement that unilateral contrast changes mainly affect dominance durations of the eye presented with *currently* higher contrast stimulus.

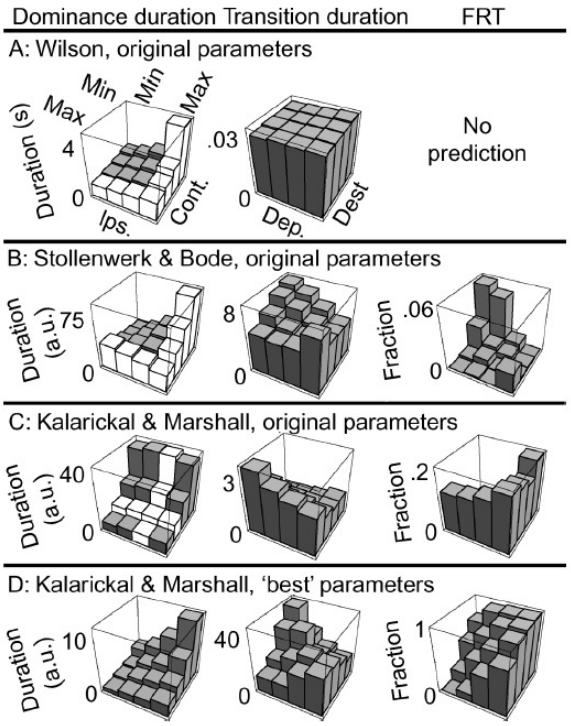

Figure 2.5: Validation of three prevailing models by using the Brascamp et al. (2006) data. From left to right the charts show the simulated dominance and transition durations and the FRT, respectively. Duration are given in arbitrary units for some models. For a limited range of inputs, the models are in agreement with the Levelt's 2nd proposition (white bars). Wilson's model is unable to produce return transition. A marked discordance with the experimental data can be seen in Panels A-C. Panel D display the behaviour of the Kalarickal & Marshall's model in a limited zone of the parameter space where a qualitative agreement with the data was do found. The nomenclature used is the same as in figure 2.4. Adapted from Brascamp et al. (2006).

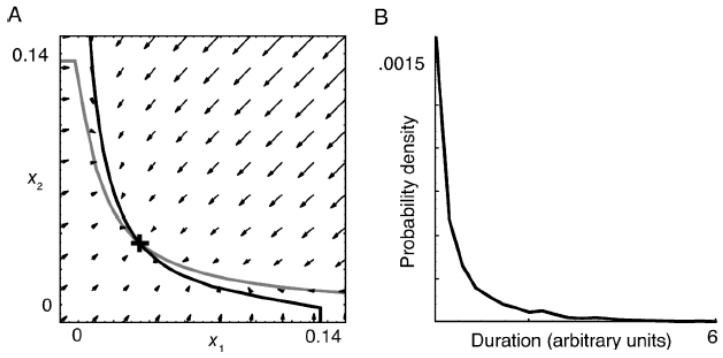

Figure 2.6: Analysis of the Kalarickal & Marshall's model in the conditions of the Panel D - figure 2.5. (A) The null-clines (black and grey lines) intersect in only one (fixed) point, to which the system tends to evolve (arrows) after eventual incursions of the noise (*i.e.*, that point is actually an attractor). (B) The model behaves purely as a stochastic system, yielding exponential-like distribution for the dominance durations, instead the characteristic Gamma functions of bistable perception. Reprinted from Brascamp et al. (2006).

History-dependence in bistable perception - Pastukhov & Braun (2011)

More recently, van Ee (2009) has reconsidered the problem of serial correlations in the perceptual time series. Positive values for this statistical measure were found by these authors, but differently to previous works that discarded the presence of such adaptation-related effects Borsellino et al. (1972); Walker (1975), impurities sources in the dominance durations such as transitions and reaction times were considered this time. Comparable results were reported for an ambiguous pattern (illusory rotating spheres) and for the classic stimulus of two orthogonal gratings as well.

Below we discuss in some detail the related work by Pastukhov & Braun (2011), who have examined the problem for the same stimuli, but in the context of the novel concept of *cumulative history*. The modeling implications of their results are considered in the next chapter of this manuscript.

Nineteen observers (including author AP) with normal or corrected-to-normal vision participated in the experiments. The subjects were shown two kind of bistable stimulus. A binocular rivalry (BR) display consisted of a pair of orthogonal (red/green) sinusoidal gratings viewed trough anaglyph glasses. The contrast of each grating was conveniently adjusted before the sessions, in order to cancel a possible innate perceptual preference for any of them. Viewing the BR display lead to a perception of clear percepts (each of the two differently oriented/colored gratings), but also the subjects reported the phenomenal appearance of two kinds of mixed percepts: ones that were 'fused' (*i.e.*, both gratings appear superimposed) and others consisting of fragments of both external stimuli

(*i.e.*, 'patchy' percepts characteristic of 'piecemeal' transitions).

The second stimulus was an illusory rotating sphere: an orthographic projection of hundred of dots distributed on a sphere surface that rotated around the vertical axis with a period of 4 s, giving an appearance of a three-dimensional structure (*i.e.*, a kinetic-depth effect (KDE)). Ambiguous perception of the sphere was created by the lack of cues (diameter and luminance of "closer" and "farther" dots) differentiating the front surface from the rear one. Such a procedure gave the impression that the structure perceived spontaneously switched, provoking that subjects reported its apparent rotation in depth as either 'front left' or 'front right'.

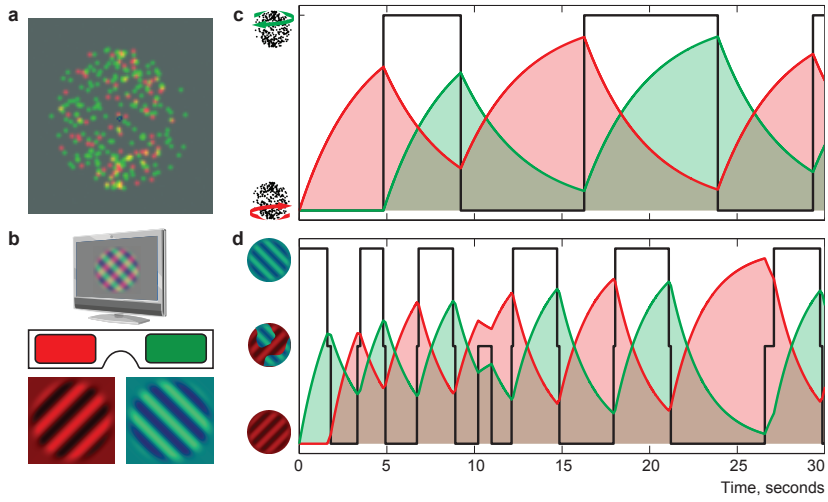

Figure 2.7: Multi-stable displays and reversals of visual appearance. (a) Observers view the two-dimensional projection of a rotating cloud of dots: phenomenally, they experience rotation in depth, with the front part moving at times to the left and at times to the right ("kinetic depth effect", KDE). (b) With each eye, observers view a grating of different color and orientation: their visual experience is dominated at times by one grating and at times by the other ("binocular rivalry", BR). Less often, a patchwork of red and green areas is experienced as well. (c), (d) Example series of reversals of visual appearances, for KDE and BR displays. The black traces indicate the reported visual appearance ('left' or 'right' for KDE; 'green', 'red', or 'patchy' for BR). The colored traces illustrate hypothetical 'cumulative histories', computed with $\tau_H = 0.5 < T_{dom} >$ (see text for details). Extracted from Pastukhov & Braun (2011).

Presentation sessions consisted of 12 blocks of 5 *min* each per observer for KDE stimulus and 18 blocks per subject for BR display. On average alternation rates showed an initial decaying transient lasting about 1 *min* before a stationary phase is reached, thus only the last four minutes of each observation block is analyzed (excluding the last incomplete percept). Average number of clear percepts

per block was 36 for KDE and 110 for BR, because spontaneous reversals were much less frequent with the KDE displays leading to longer dominance phases.

Following the standard analysis, the average dominance time $< T_{dom} >$ and the coefficient of transition C_v is then computed from the sequence of dominance periods T_i of each presentation block:

$$< T_{dom} >= \frac{1}{N} \sum_{i=1}^{N} T_i \qquad\qquad C_v = \frac{\sum_{i=1}^{N} (T_i - < T_{dom} >)^2}{\sum_{i=1}^{N} T_i} \tag{2.1}$$

Additionally, overall values for the mean and standard error of T_{dom} and C_v were obtained by combining all blocks.

More importantly, a novel concept is also used in Pastukhov & Braun (2011) to characterize the perceptual dynamics in unprecedented detail. Namely, a measure of *cumulative history* as a function of time, $H_x(t)$, is defined in terms of the previous phenomenal experience. Let $S_x(t)$ be a record of one of the possible perceptual outcomes x, defined as unity while percept x dominates, 0.5 during a mixed or patchy percept (only for BR), and zero when percept x is suppressed. The cumulative history $H_x(t)$ is then given by a linear exponential convolution of $S_x(t)$:

$$\tau_H \frac{dH_x}{dt} = -H_x(t) + S_x(t) \;\Leftrightarrow\; H_x(t) = \frac{1}{\tau_H} \int_0^t S_x(t') \exp\left[-\frac{(t-t')}{\tau_H}\right] dt', \tag{2.2}$$

where τ_H is a constant that has to be determined empirically, corresponding to the optimal time window that best explains the observed perceptual variability. This definition excludes the possible cross-influence between the competing (clear) percepts, although the combined effect of the two resulting interpretations is do accounted for. Due to the presence of the exponential-kernel, the cumulative history $H_x(t)$ reflects both how long and how recently a given percept has dominated in the past. Two representative series of dominance reports, the time courses of $S_x(t)$ and $H_x(t)$ are illustrated, respectively, by the black and the colored traces in figure 2.7 (c)-(d).

The computation of the τ_H constant is then made as followed. Let $H_l(t)$ and $H_r(t)$ be the history measures computed for the two alternative percepts produced by the given display (*e.g.*, left- or rightward rotation, left- or right-eye grating). Subsequently, linear correlation coefficients c_H between the individual dominance durations (T_l and T_r) and the values of both histories up to the onset of such experienced phases are calculated, for all the possible combinations of history and percept: $H_l \times T_l$, $H_l \times T_r$, $H_r \times T_r$, and $H_r \times T_l$.

An interval of possible values of τ_H ranging from 0.01 s to 60 s was considered. In general, negative values $c_H < 0$ are found when history and dominance periods of the same interpretation are combined (for instance, H_l and T_l), what suggest some relation of the concept of cumulative history and the adaptation of neurons representing the currently dominant percept. Indeed, this observation constitutes the first direct evidence that neural adaptation does modulate rever-

sal probabilities. Strikingly, a non-monotonic profile for the average (absolute) correlation c_H as a function of τ_H is found(see Figure 2.8).

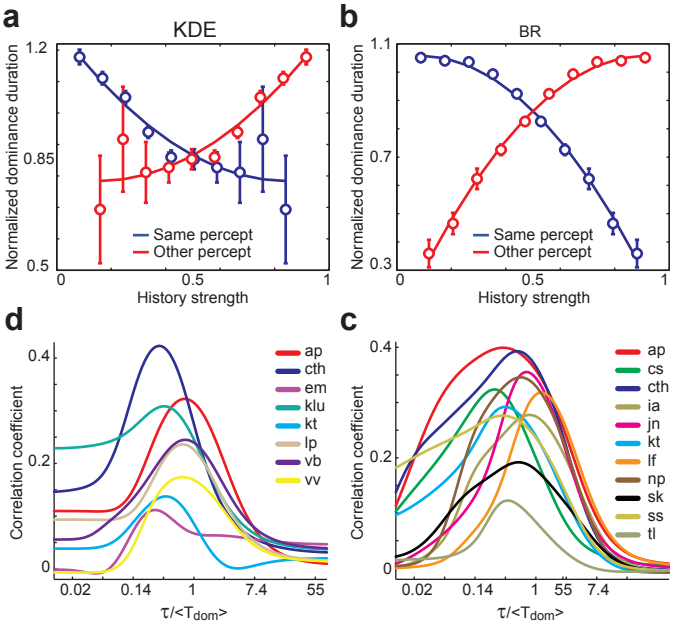

Figure 2.8: Individual dominance periods depend on dominance history. (a) and (b) Dominance periods T (normalized mean $\pm$ standard error) of a percept, as a function of the cumulative history H of the same (blue) or other (red) percept, taken at the time of the initial reversal. (c) and (d) Pearson correlation coefficients c_H (mean absolute values) between dominance duration and history, as a function of the decay constant τ_H. The colors identify different observers (see inset). For most observers, the maximal correlation is obtained when $\tau_H \approx 0.5 < T_{dom} >$. After Pastukhov & Braun (2011).

Similarly to the case of the observables T_{dom} and C_v, overall values for the mean and standard error for the maximal correlation c_H and its associated time-scale τ_H were obtained by combining all presentation blocks. In this way, it is demonstrated that optimal values of τ_H are approximately half the average dominance time $< T_{dom} >$ experienced by the subject ($\tau_H = 0.54 \pm 0.21 < T_{dom} >$ for the KDE stimulus, $\tau_H = 0.56 \pm 0.28 < T_{dom} >$ for the BR display). Significant values of maximal correlations c_H are found (see table 2.2), being $c_H = 0.24 \pm 0.10$ for the rotating sphere and $c_H = 0.30 \pm 0.08$ for the dichoptic gratings.

Based of these remarkable significant findings, Pastukhov & Braun (2011) have proposed to use such history-related quantities as complementary measures to characterize the statistics of the dynamics of bistable perception. That is, the statistics of the perceptual data can be better described by a set of four psychophysical observables: the mean $< T_{dom} >$ of the dominance durations data

and its coefficient of variation (C_v), the maximal linear correlation between perceptual history and immediate next dominance period (c_H) and its corresponding time-scale constant (τ_H). The values of the four psychophysical measures determined for each observer and display (mean and standard error, and the level of significance p associated with c_H) are listed in tables 2.1 and 2.2.

Stimulus	Subject	$< T_{dom} >$, s	C_v
KDE	ap	3.69 ± 0.10	0.46 ± 0.02
	cth	16.31 ± 1.00	0.58 ± 0.05
	em	23.85 ± 3.09	0.84 ± 0.11
	klu	9.40 ± 0.62	0.76 ± 0.05
	kt	9.49 ± 0.67	0.63 ± 0.08
	lp	8.22 ± 0.44	0.58 ± 0.05
	vb	13.79 ± 0.92	0.68 ± 0.05
	vv	5.42 ± 0.13	0.62 ± 0.02
BR	ap	2.69 ± 0.06	0.49 ± 0.02
	cs	3.27 ± 0.14	0.69 ± 0.07
	cth	3.58 ± 0.10	0.60 ± 0.02
	ia	1.26 ± 0.02	0.54 ± 0.01
	jn	2.27 ± 0.02	0.24 ± 0.01
	kt	2.68 ± 0.03	0.44 ± 0.02
	lf	1.43 ± 0.02	0.41 ± 0.01
	np	1.36 ± 0.02	0.51 ± 0.01
	sk	1.69 ± 0.02	0.56 ± 0.01
	ss	4.53 ± 0.18	0.53 ± 0.03
	tl	2.67 ± 0.03	0.32 ± 0.01

Table 2.1: Psychophysical observations obtained in the study by Pastukhov & Braun (2011). Durations-related measures are shown with their mean and standard error.

To facilitate a direct comparison of the results reported in Pastukhov & Braun (2011) with statistical analyses typically carried out in previous studies, a fifth observable is included. That is, the serial linear correlation (c_T) between successive dominance intervals T_i of the same percept:

$$c_T = \frac{\sum_{i=1}^{N}(T_i - < T_i >)(T_{i+1} - < T_{i+1} >)}{\sigma_{T_i}\sigma_{T_{i+1}}} \tag{2.3}$$

Correlations between successive dominance periods of the same percept were weak ($c_T = -0.04 \pm 0.1$ for KDE and $c_T = 0.02 \pm 0.02$ for BR) and reached significance only with the KDE display (table 2.2). Such absence of clear sequential correlations is in agreement with previous works that have considered multi-stable perception to be a stochastic phenomenon Fox & Herrmann (1967); Levelt (1967); Walker (1975); Borsellino et al. (1972).

Remarkably, the values of correlations with cumulative history were generally not only more significant than but also twice as large sequential correlations of

dominance periods. Thus, contrary to the conventional view, the perceptual trace experienced by an individual do contains a subtle but significant level of hysteresis that affects the duration of the next dominance phase. In average, a consistent trend was observed, where shorter (longer) residence times in the same percept was preceded by larger (lower) values of cumulative history values (see Figures 2.8 (a)-(b)).

However, further statistical analyses based on the concepts of entropy and mutual information, demonstrated that past history and future dominance periods, are only loosely correlated, showing that adaptive forces could not explain more than 10% of the observed variability in reversal timing Pastukhov & Braun (2011). This fact clearly suggests that adaptive processes do not trigger perceptual reversals directly, but merely modulate reversal probability.

Additionally, the inferred decay time-constant τ_H, which presumably reflects neural fatigue processes, also correlates poorly with the average dominance period $< T_{dom} >$ (Figure 2.9-a). For both displays and all observers, the linear correlation coefficient was 0.82 ($p < 0.001$) . Nevertheless, when the two kinds of stimuli are analyzed separately, this correlation value dramatically decreases, falling to 0.64 ($p = 0.09$) for KDE and to 0.36 ($p = 0.28$) for BR. Hence, contrary to the conventional view Laing & Chow (2002); Wilson (2007), dominance durations seems to be only slightly determined by adaptation rate.

Stimulus	Subject	c_H	$p(c_H)$	τ_H, s	$\frac{\tau_H}{<T_{dom}>}$	c_T	$p(c_T)$
KDE	ap	0.31 ± 0.05	0.01	4.53 ± 1.65	1.23	0.29	0.01
	cth	0.39 ± 0.08	0.01	3.57 ± 1.30	0.22	0.31	n.s.
	em	0.11 ± 0.14	n.s.	35.16 ± 31.17	1.47	0.05	n.s.
	klu	0.36 ± 0.08	0.01	1.72 ± 1.19	0.18	0.08	n.s.
	kt	0.08 ± 0.08	n.s.	6.18 ± 5.81	0.65	-0.11	n.s.
	lp	0.26 ± 0.08	0.01	3.72 ± 2.08	0.45	0.15	n.s.
	vb	0.17 ± 0.10	0.05	6.25 ± 5.75	0.45	0.07	n.s.
	vv	0.16 ± 0.04	0.01	9.58 ± 5.72	1.77	0.15	0.05
BR	ap	0.37 ± 0.05	10^{-4}	0.64 ± 0.27	0.24	0.18	0.01
	cs	0.27 ± 0.09	10^{-4}	0.94 ± 0.29	0.29	0.02	n.s.
	cth	0.36 ± 0.06	10^{-4}	0.72 ± 0.60	0.20	0.21	0.01
	ia	0.30 ± 0.03	10^{-4}	0.90 ± 0.15	0.71	0.09	n.s.
	jn	0.37 ± 0.07	10^{-4}	2.32 ± 0.55	1.02	0.16	0.01
	kt	0.30 ± 0.05	10^{-4}	0.77 ± 0.13	0.29	0.24	10^{-4}
	lf	0.40 ± 0.03	10^{-4}	2.31 ± 0.38	1.61	0.27	10^{-4}
	np	0.40 ± 0.03	10^{-4}	0.97 ± 0.16	0.71	0.22	10^{-4}
	sk	0.20 ± 0.02	10^{-4}	0.92 ± 0.15	0.54	0.10	0.05
	ss	0.48 ± 0.17	10^{-4}	0.76 ± 0.39	0.17	0.22	n.s.
	tl	0.15 ± 0.04	10^{-4}	0.84 ± 0.14	0.32	0.10	n.s.

Table 2.2: Psychophysical observations obtained in the study by Pastukhov & Braun (2011). Correlations-related measures are shown with their mean and standard error.

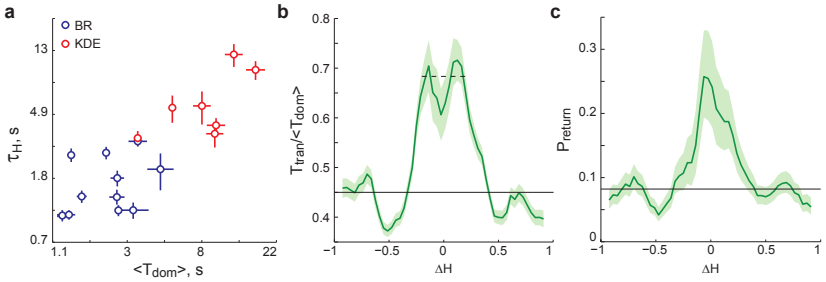

Figure 2.9: (a) Correlation between average dominance periods $< T_{dom} >$ (mean $\pm$ standard error) and decay time-constant τ_H (mean $\pm$ standard error), for 8 observers KDE and 11 observers BR. (b) Transition duration (mean $\pm$ standard error), in units of $< T_{dom} >$, as a function of history difference $\Delta H \equiv H_{red} - H_{green}$. For $\Delta H \approx 0$, transition durations increased 73% above their average value (black line). (c) Probability of return transitions (red→red or green→green) as a function of history difference ΔH. For $\Delta H \approx 0$, return probability rose 217% above its average value (black line). After Pastukhov & Braun (2011).

A possible explanation to these unexpected results can be find in the experiment with BR displays, consisting of two gratings that apart from being perceived as clear percepts, could also eventually experienced by the subject as 'patchy' patterns. Then, patchworks of red and green regions can be seen, that can not exclusively be followed by a complete transition to the other alternative interpretation, but also the subjective experience can afterwards resume to the previous uniform appearance (return or failed transition). Both kinds of transitions were profoundly affected by the balance of cumulative histories (see Figure 2.9 (c)-(d)): transitions lasted longer and the return probability was far higher, when the two cumulative histories were approximately balanced: $\Delta H \approx 0$. Thus, in the absence of any adaptive bias, perceptual transitions appear to be driven by spontaneous activity fluctuations.

In summary, the work of Pastukhov & Braun (2011) clearly demonstrates, for the first time, the presence of low but statistically significant correlation values of correlation between past perceptual history and dominance durations. Furthermore, fatigue processes do not seem to be the only factor that affects the timing of perceptual reversals, but acts together with fluctuations sources whose role turn more evident when no adaptation bias is present. Spontaneous fluctuations are able to enlarge transtion phases and even be the cause of frequent failed transitions to the competing percept, a result that strongly resembles the conclusions stated by Brascamp et al. (2006).

2.3 Current models for bistable visual perception

Existing computational models have been successful in explaining the time-averaged behavior of bistable perception. Two distinct classes of models have been proposed, according to the way in which the dominance of some percept is defined: following a *threshold* criterion or a *winner-take-all* rule. In the first case, an image becomes dominant when the activity of the neurons encoding the competing image falls until crossing certain threshold Lehky (1988).

The more recent proposals appeared in Wilson (2003); Laing & Chow (2002) and other works belong to the second kind of models. The all-or-none characteristic of perceptual switching between the competing images is implemented by comparing the activity levels of the corresponding populations.

Regarding the quantity of processing stages, hybrid models of rivalry Tong et al. (2006) have also been proposed in the last years, in order to account for the accumulating evidence for a multi-level network architecture underlying the perception of multistable patterns Lumer (1998); Dayan (1998); Grossberg et al. (2008). However, in general any model typically contains a basic modeling block, whose neurons representing the competing percepts, interacts by general mechanisms such variable (synaptically modulated) inhibition, and/or whose activity decrease by self-adaptation during dominance.

Below we present some of the prevailing models, that were partly designed to reflect the neurophysiology of the visual system and use an elaborated form of the self-adaptation and competitive inhibition. The all-or-none characteristic of perceptual switching is implemented by a winner-take-all rule: perceptual dominance is determined by the relative level of the activities of a pair of pools; *e.g.*, image 1 is perceptually dominant when the neuronal firing-rates satisfy $r_1 > r_2$.

A mean field bi-stage approach: Wilson's model

Wilson (2003) proposed a two-stage competitive neural model, to take into account the importance of monocular and binocular neurons in binocular rivalry Logothetis et al. (1996). The mathematical form of a processing stage (one population) is given by the following equations:

$$\begin{aligned}
\tau_r \frac{dr_1}{dt} &= -r_1 + \frac{R_{max}[I_1 - \beta i_2]_+^2}{(\theta + a_1)^2 + [I_1 - \beta i_2]_+^2}, \quad \tau_r = 20\ ms \\
\tau_i \frac{di_1}{dt} &= -i_1 + r_1, \quad \tau_i = 11\ ms \\
\tau_a \frac{da_1}{dt} &= -a_1 + \phi_a r_1, \quad \tau_a = 900\ ms.
\end{aligned} \tag{2.4}$$

Here r_1 is the firing rate of an excitatory neuron responding to one of the possible interpretations of the bistable stimulus, I_1, and $R_{max} = 100\ Hz$ is the maximal discharging frequency. The asymptotic firing rate is determined by a

Naka-Rushton function $f(x)$ for positive values of the argument $x = (I - \beta i)_+$, with a semi-saturation or threshold value $\theta = 10$, such that:

$$f(x = \theta + a_1) = R_{max}/2. \tag{2.5}$$

This excitatory neuron drives an inhibitory cell, i_1, which is described by a linear equation for simplicity. The same excitatory neuron receives inhibition of strength β from another similar inhibitory neuron i_2. Finally, a_1 describes the very slow self-adaptation of this neuron by an after-hyperpolarizing potential current, that contributes with a threshold increment.

The noise is directly added to the input variables. Nevertheless, Kim et al. (2006) have shown that the slow self-adaptation might be a better locus for the noise since Wilson's model simulated their data reasonably well when the noise was added to the adaptation equations.

Remarkably, Wilson's model distinguishes from others by a divisive kind of adaptation effect Shpiro et al. (2007). That is, the slope of the activation (Naka-Rushton) function, which is given by the first derivative at the (variable) threshold point $\theta + a$:

$$f'(x = \theta + a) = \frac{2R_{max}}{\theta + a}, \tag{2.6}$$

explicitly depends on the temporal evolution of the slow negative feedback process.

Laing and Chow's reduced model

Laing & Chow (2002) presented a biologically plausible model for binocular rivalry consisting of a network of Hodgkin-Huxley type neurons, in which a given percept is represented as a localized focus of active neurons. The reader interested in implementation details is referred to the appendix A of the mentioned reference. Here we present the reduced description derived from the spiking network model:

$$\begin{aligned} \frac{dr_1}{dt} &= -r_1 + f(g_1 \alpha r_1 - g_2 \beta r_2 - a_1 + I_1) \\ \tau_a \frac{da_1}{dt} &= -a_1 + \phi_a f(g_1 \alpha r_1 - g_2 \beta r_2 - a_1 + I_1) \\ \tau_d \frac{dg_1}{dt} &= 1 - g_1 - g_1 \phi_g f(g_1 \alpha r_1 - g_2 \beta r_2 - a_1 + I_1) \end{aligned} \tag{2.7}$$

r_1, r_2 represent the spatially averaged net excitatory activity of each localized population, I_1, I_2 are the input variables, a_1 and g_1, g_2 are the population adaptation and synaptic depression variables, respectively. The parameter α describes recurrent excitation, β stands for the strength of mutual inhibition and ϕ_a for adaptation strength. For simplicity the gain function f was chosen as the Heaviside step function, *i.e.*, $f(x) = 1$ for $x \geq 0$ and $f(x) = 0$ for $x < 0$. The

constants τ_a and τ_d are the time constants of the adaptation and synaptic depression, respectively. Analogous equations are formulated for the other competitive population, exchanging the variables describing the pools.

Synaptic depression was included in both the excitatory and inhibitory connections, since while depression is thought to occur only in excitatory synapses, the inhibitory neurons in the full spiking model are largely driven by the excitatory populations whose activity is influenced by such slow negative feedback.

Laing & Chow (2002) used such simplified version of the original model to obtain explicit non-linear expressions for the dependence of the dominance durations on input strengths, providing a simple explanation for Levelt's second proposition. Besides, an expression for the distribution of dominance durations was analytically calculated, a one different from the familiar description by a Gamma function (see Eq. 12 and appendix C of Laing & Chow (2002), for the analytical derivation).

Two simpler variants of Laing and Chow's model are presented in Shpiro et al. (2007):

Laing and Chow model with only adaptation (adaptation-LC)

$$\begin{aligned} \frac{dr_1}{dt} &= -r_1 + f(-\beta r_2 - \phi_a a_1 + I_1 + n_1) \\ \tau_a \frac{da_1}{dt} &= -a_1 + r_1 \end{aligned} \tag{2.8}$$

where $n(t)$ is introduced here for the fluctuations variable, being modeled as a filtered version of a white noise $\chi(t)$:

$$\frac{dn(t)}{dt} = -\frac{n(t)}{\tau_n} + \sigma_n \sqrt{\frac{2}{\tau_n}} \chi(t), \tag{2.9}$$

and $f(x)$ is a sigmoid function with threshold θ and slope proportional to $1/k$:

$$f(x) = \frac{1}{1 + \exp(-(x-\theta)/k)}. \tag{2.10}$$

Laing and Chow model with only depression (depression-LC)

$$\begin{aligned} \frac{dr_1}{dt} &= -r_1 + f(-\beta g_2 r_2 + I_1 + n_1) \\ \tau_d \frac{dg_1}{dt} &= 1 - g_1 - \phi_d g_1 r_1 \end{aligned} \tag{2.11}$$

Kalarickal & Marshall (2000) model

The authors proposed a model with *noisy* habituating reciprocal inhibitory synapses between the neurons encoding the possible interpretations of the input. Known empirical results such as Levelt's propositions and the lack of serial correlations are successfully reproduced in Kalarickal & Marshall (2000). A recent application of this model is reported in Brascamp et al. (2006) and briefly discussed in section 2.2 of this manuscript.

The equations defining the Kalarickal & Marshall (2000) model are as follows:

$$\begin{aligned} \dot{r}_i &= -r_i + (1 - r_i)W_i^+ I_i - (c_i + r_i)W_{ji}^- g_{ji} max(r_j, 0) \\ \dot{g}_{ji} &= c_2[(1 - g_{ji}) - c_3 max(r_j, 0)W_{ji}^- g_{ji}] + b(t) \end{aligned} \tag{2.12}$$

where W_i^+ denotes the synaptic efficacy of the excitatory input to neuron i, W_{ji}^- characterizes the inhibition coming from neuron j and affecting the cell i, and g_{ji} is the corresponding habituation variable ($i \neq j$). c_1, c_2 and c_3 are additional constants to be fit to empirical data.

A stochastic ingredient is given by the dynamic variable $b(t)$, that takes the values s if $r_{ij}(t) < p$ and $-s$ otherwise, where $r_{ji}(t)$ is a uniformly distributed random variable in $[0, 1)$. $p \in [0, 1]$ and s is a small positive constant ($s = 0.0025$ in the original version of the model).

Stollenwerk & Bode (2003) model

This approach is useful when considering more realistic sequences of suppression and dominance phases happening during binocular rivalry. A mixed percept can consist of parts of the images presented to the two eyes (spatial or piecemeal phases), that eventually accompanies a failed transition or a reversal to the alternative stimulus Bossink et al. (1993); Mueller & Blake (1989). The approach consists of a two-dimensionally extended model formed by basic blocks coupled with each other representing neighboring positions in the visual field.

In the basic block (see Fig. 2 of Stollenwerk & Bode (2003)), it is the output (O_1, O_2) of the Φ-neurons what is associated with the current perception, and cross-inhibition is implemented by the χ-neurons. The possibility of destabilizing the currently dominant image is based on negative feedback fed through the Ψ-neurons, characterized by a phase delay in any rivalry scenario which is quantified by $\tau_{s.I}$.

The output of $\Phi-$ and $\Psi-$ neurons in the pathway 1 of each block [4] is described by a sigmoid-type gain function

$$f(x) = \frac{1}{1 + \exp(-4\alpha(x - V))}. \tag{2.13}$$

The time evolution of Ψ cells is modeled by the following differential equation

$$\tau_{s.I}\dot{\Psi}_1 = -\Psi_1 + f(\Phi_1), \tag{2.14}$$

where $\tau_{s.I}$ indicates the time-scale of the slow "self-inhibition". Besides, $\Phi-$neurons evolve according to

$$\dot{\Phi}_1 = -\Phi_1 + I_1 + k \sum_{\substack{Neighbours \\ n}} k^n f(\Phi_1^n) - \epsilon f(\Psi_1) - \eta f(\Phi_2) + S, \tag{2.15}$$

[4] As in other models, the corresponding equation for the other pathway can easily be obtained by exchanging the indices 1 and 2

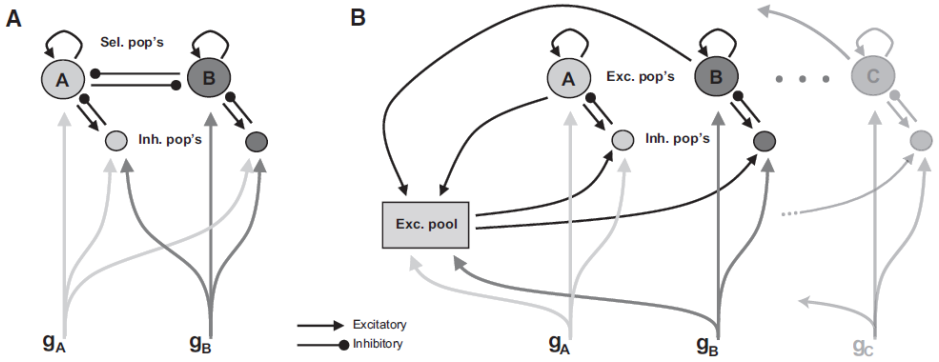

Figure 2.10: Network architecture of Moreno-Bote et al's model, 2007. Reprinted from Moreno-Bote et al. (2007).

being I_1 the external stimulus signal, k is some constant that scales the sum over the neighbours of the cell, and k^n represents a weighting factor of the output signal $f(\Phi_1^n)$ of a given neighbour, that directly depends on the inverse of its distance to the Φ_1-cell. ϵ and η indicates the self-inhibition and cross-inhibition strengths, respectively, and S describes and intrinsic activation not related with the input signal. Finally, the output of the χ-inverter neurons is set equal to the one of the Φ-neuron, thus considering its activation very fast.

Moreno-Bote et al. (2007) model

This model mainly distinguishes among all the others in two aspects. Firstly, the derivation of the rate equations is explicitly based on an energy function intended to account for some well-known properties of bistable perception such as Levelt's propositions II and IV. Secondly, it proposes a network architecture that can be easily extended to generate the typical multi-stable behavior observed when the external stimulus allows more than two competing percepts be inferred by the brain.

Two alternative architectures are proposed in Moreno-Bote et al. (2007). A simpler one consists of a pair of average populations of excitatory neurons directly connected by (inhibitory) interneurons [5], so resembling the classical view implemented in previous models. However, a significant difference is given by the fact that the response of both populations of inhibitory neurons rely on information about the *total* strength of external stimulation (see Fig. 2.10-A).

A second architecture contains more novel ingredients that made it more interesting to be included in this thesis. Specifically, the information of the total external stimulation is provided by a unique (global) excitatory subpopulation that feeds the inhibitory populations (Figure 2.10-B). These last neurons, in turn, convey local signals to the respective excitatory cells representing the com-

[5] Hence sometimes called the *DIRI* version of this model Shpiro et al. (2009).

peting percepts [6]. Finally, a feedback through the global excitatory subpopulation implements some level of mutual-inhibition between the primary excitatory populations, an idea that can make such a circuitry also suitable for modeling multi-stable perception of complex stimuli.

The basic equations of one of the populations are as follows:

$$\begin{aligned} \dot{r_1} &= -r_1 + f\Big(\alpha r_1 - \beta\Big\{\Theta\Big(\varphi(r_1 + r_2) + I_1 + I_2\Big) + \eta r_1\Big\}^2 - a_1 + I_1 + n_1\Big) \\ \dot{a_1} &= -a_1 + \phi_a r_1, \end{aligned} \tag{2.16}$$

where $\Theta(x)$ is a Heaviside step function and φ and η are constants. A non-null value of $\alpha = 0.75$ is chosen for the self-excitatory connections in Moreno-Bote et al. (2007); Shpiro et al. (2009). Differently to the previous approaches, one can note the presence of a nonlinear inhibition term, proportional to the inhibition strength β.

[6] Hence sometimes called the *GELI* version of the model Shpiro et al. (2009); Seely & Chow (2011).

3 Fitting models to the empirical data: a noise-driven regime

In this chapter, various prevailing models for bistable perception are fitted to the experimental results of Pastukhov & Braun (2011), which are summarized in section 2.2. These authors measured not only salient aspects of bistable perception (*i.e.*, mean and dispersion of dominance distributions), but also some hidden hysteresis effects ignored up to now.

As it was shown in section 2.3, most of the existing models include two basic mechanisms: (i) a slow negative feedback that precludes a population from dominating forever (*e.g.*, spike-frequency adaptation or synaptic depression) and (ii) mutual inhibition between the two populations whose activities represent the competing percepts. Such cross-inhibition avoids long periods of simultaneous dominance. Our selection of a determined model, from among other existing choices, is based on the interest to evaluate the plausibility of dissimilar neuronal mechanisms as a possible physiological basis for the phenomenon of bistable perception. That is the reason to start with two different versions of Laing and Chow's model, firstly studied in Shpiro et al. (2007): a former one whose slow feedback mechanism is just a spike-frequency adaptation and a second version with only long-term depression. In other cases, the selection is motivated by more recent studies assessing the biological plausibility of the models respect to the Levelt's propositions Curtu et al. (2008); Seely & Chow (2011). This explains that a model with a more complicated architecture like Moreno-Bote et al. (2007) is also studied, as well as a third version of the Laing and Chow's model with nonlinear adaptation Curtu et al. (2008). The model by Wilson (2003) is the fifth choice of interest, due to its peculiarity of containing an activation function with both threshold and slope being dependent on the slow evolution of adaptation. For a brief summary of such models, the interested reader is referred to section 2.3.

Extensive computational simulations of these models give strong evidence that the history-dependence effects shown by Pastukhov & Braun (2011) effectively constrain the region of the parameter space able to replicate the empirical data. Concretely, that just small regions residing inside the *bistable* or *two-attractors* region of the whole parameter space are actually adequate.

Preliminary and more elaborated results obtained with the adaptation-LC model have been already presented in several international conferences Pastukhov

et al. (2011); Rodríguez et al. (2010); Pastukhov et al. (2009a); García et al. (2009); Pastukhov et al. (2009b); García-Rodríguez et al. (2009). The same work has lead to a joint article with our collaborators at Magdeburg University, which is currently under review at the *PLoS* Computational Biology journal Pastukhov et al. (2012).

3.1 Fitting models to dominance and history measurements: general methodology

As a starting point of the analysis, the experimental data obtained by Pastukhov & Braun (2011) are fitted with one of the reduced versions of the Laing & Chow (2002) model. Only spike-frequency adaptation currents are retained in order to simulate a fatigue effect on the currently dominant population. This model, known as the adaptation-LC case, was firstly studied numerically in Shpiro et al. (2007). Subsequently, it has also been studied in a considerable numerical Shpiro et al. (2009) and analytical Curtu et al. (2008) detail as well.

The methodology followed in this thesis to fit both the adaptation-LC case and other models extends the work previously reported in Shpiro et al. (2009). In that paper, aiming at fitting a range of observable values characterizing a typical population of subjects, the authors explored a three-dimensional parameter space formed by the mutual inhibition strength (β), the common value of input ($I_1 = I_2 = I_0$) feeding both populations and the noise dispersion (σ_n). In this manuscript, two additional dimensions of the parametric space are explored, namely the ones corresponding to the strength of adaptation or depression variables (ϕ_a or ϕ_d) and the time-scale characterizing such slow dynamics (τ_a or τ_d).

The whole set of examined parameters then conforms a 5D parametric space to be explored. The time constants characterizing the firing rate evolution (τ_r) and the temporal correlation of the noise (τ_n) remain fixed across all the studies at $\tau_r = 10\ ms$ and $\tau_n = 100\ ms$, respectively. In order to guarantee that the averaged values of observables T_{dom} and C_v can be fitted by the model, within a noise range of $\sigma_n \in \{0.01, 0.22\}$, a search for the optimal values of the activation function parameters (slope $\sim 1/\kappa$ and threshold θ) always precedes the study at the mentioned 5D-hyperspace. Consequently, our study explores a total quantity of seven parameter dimensions, differently to the work of Shpiro et al. (2009).

More importantly, the experimental data collected from each observer is independently fitted, instead of a range of observable values characterizing the entire sample. Finally, we incorporate the recent data obtained by Pastukhov & Braun (2011), who have quantified for the first time the subtle hysteresis effects present in bistable perception.

In general, the dynamical regime exhibited by the free-noise system depends only on three parameters: ϕ_a (or ϕ_d), β and I_0, while it is not sensitive to parameters standing for time-scales (*e.g.*, τ_a or τ_d). Such regimes characterize the capacity of the model to alternate between different states (Figure 3.1). That is, an *oscillatory* behavior refers to the case of a continuous exchange of dominance

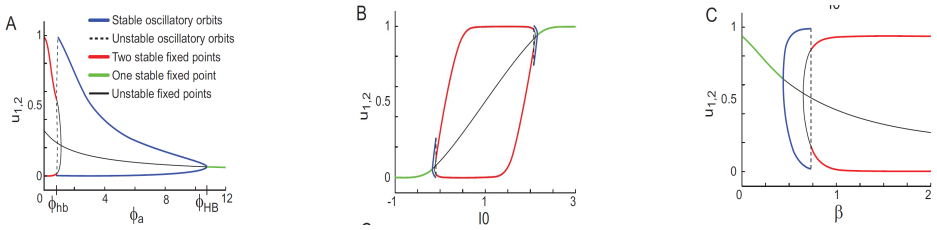

Figure 3.1: One-dimensional bifurcation diagrams of the adaptation-LC model (free-noise). The variable in the vertical axis represents fixed points or extreme values of periodic orbits, as a function of a control parameter (horizontal axis). A unique stable fixed point corresponds to a SIM-state (in green), a stable isolated periodic orbit (*i.e.*, limit cycle) corresponds to an OSC-regime (blue), while an ATT-state is associated to a pair of stable fixed points (red). Also shown are unstable fixed points (black) and unstable periodic orbits (dotted lines). (A) ϕ_a-dependence of the model, revealing bi-stability (low values of ϕ_a), oscillatory regimes (moderate ϕ_a) and SIM-states (for high values of ϕ_a). A pair of supercritical Andronov-Hopf bifurcation points are labeled as ϕ_{HB} and ϕ_{hb}. (B) Bifurcation diagram in the I_0 control parameter, showing a bistable region flanked by two oscillatory regimes. (C) Dependence on β, revealing a behavior that qualitatively consists of a mirror image of the ϕ_a diagram shown in (A): fusion states (green line) given by a unique stable fixed point at low values of β, a set of limit cycles for moderate values of the control parameter, and finally a pair of attractors (red lines) and an unstable fixed point (black) when the value of β is relatively high. Base parameter set: $I_0 = 0.5$, $\beta = 1.75$, $\phi_a = 0.25$, $\theta_a = 0.9$, $k_a = 0.14$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

between the two populations due to a presence of a limit cycle attractor for the dynamics (we will call it the *OSC*-state from here on). A *bistable* regime (sometimes also called *winner-take-all* behavior or *ATT*-state) indicates the evolution from the initial conditions to one of two attractor states where one population dominates, whereas the other one is permanently suppressed by virtue of the existing level of cross-inhibition. Finally, a third case is possible, where the rates representing the two dissimilar images/interpretations stay at approximately the same value all over the time, therefore called a *fusion* or *simultaneous* activity state (*SIM*-state). The appearance of such monostable behavior is based on the existence of a unique attractor, and has been demonstrated to happen in the context of many of the models studied throughout this thesis, for some specific set of parameters (*e.g.*, see Figure 2 of Shpiro et al. (2007)).

In order to predict the psychophysical observables reported by Pastukhov & Braun (2011), $10-30$ blocks of five minutes each were simulated with the model in a 3D grid of the relevant parameter space: (ϕ_a (or ϕ_d), β, I_0). Each of such triplets was combined with specific values of noise dispersion $\sigma_n \in \{0.01, 0.22\}$ and time-scales $\tau_a \in \{0.7, 10.0\}\ s$) (or τ_d).

Exclusive dominance was defined as in the modeling study included in Bras-

camp et al. (2006): a percept is said to dominate whenever its associated activity is 25% larger than the firing-rate of the neurons associated to the other percept. Otherwise, a mixed (transition) state is considered to occur. As it was shown by Brascamp et al. (2006), this "third percept" rarely lasts for a long time in many of the existing models, a fact likely related with the steep sigmoid functions often used in the firing-rate equations.

To compute the perceptual statistics, the first minute of each block was discarded as in experiments, in order to avoid typical transients before stationarity. Then, the obtained values for the observables $< T_{dom} >$, C_v, τ_H, and c_H were compared with the empirical values for each observer and kind of display (tables 2.1 and 2.2). If all four predictions fell within two standard-errors of the empirical values [1], the corresponding combination of model parameters was marked as a "match".

Simulations were performed on a Linux cluster (Ubuntu 10.04.3 LTS, Matlab R2010b, C++ compiler gcc 4.4.3, `http://betz.upf.edu/ganglia/`) with eight server machines each with 2 processors Quad-Core AMD Opteron^{TM} CPU 2384 @ 2.70 GHz, and other two nodes each with processors Intel(R) Xeon(R) CPU E5630 @ 2.53GHz. These nodes each have between 12 and 24 GB of RAM, although this aspect was not important while running simulations, since the resulting data was continuously saved to the hard disk. Generally, the simulation of a psychophysical experiment with 10 blocks of 5 *min* each in a $62 \times 62 \times 62$ parameter grid of a conventional model [2], takes around 20 hours of computing time.

3.2 Fitting the adaptation-LC model: a noise-driven regime

The adaptation-LC model, a simple variant of the original Laing & Chow (2002) model where the slow negative feedback is reduced to a spike-frequency adaptation current, is the first model considered (see section 2.3 for the corresponding mathematical equations). Results obtained when fitting this rate-based model to the data related with the two kind of displays used in Pastukhov & Braun (2011) (orthogonal gratings for BR and apparent rotating spheres for KDE), are shown in Figures 3.2 and 3.3, respectively. The behavior in the hyper-space $(\phi_a, \beta, I_0, \tau_a, \sigma_n)$ is illustrated at a fixed value for the common input strength $I_0 = 0.90$ and taking planar slices at distinct values of the time constant parameter τ_a. The position of each valid set of relevant parameters (ϕ_a, β) is located respect to the dynamical regimes that appear in the absence of noise (free-noise case).

Some characteristic behaviors of general applicability to other models studied in this work are visible in both figures. A weak level of mutual inhibition β gener-

[1] That means, assuming that the Central Limit theorem is valid, a 95% confidence interval in a Gaussian distribution of the observable.

[2] Differently to cases such as the Wilson (2003) model, the one by Moreno-Bote et al. (2007) studied in the subsection 3.4 contains more equations due to its more complex network architecture (section 2.3).

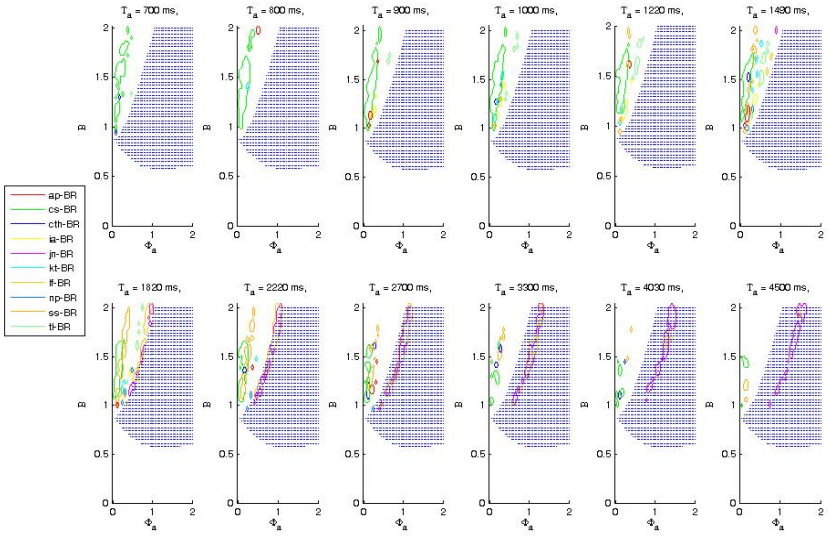

Figure 3.2: Fitting BR-experimental results with the adaptation-LC model (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is outlined by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). A few exceptions are evident for less plausible (larger) values of τ_a. The region below the blue zone is clearly unreachable, and represents monostable states (SIM-states). Base parameter set: $I_0 = 0.90$, $\theta = 0.0$, $k = 0.14$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

ally yields a simultaneous activity state in the free-noise system, with both rates remaining stable at equal values (a *SIM*-state). A sufficiently strong adaptation is, however, the cause of the emergence of an oscillatory dynamics (*OSC*-state), where the activities oscillate in counter-phase between low and high values, even if no noise is present. Finally, a combination of a high inhibition and a low adaptation results in bistable behavior between two attractors (*ATT*-state), where the dominance switches are provoked exclusively by stochastic fluctuations.

It can be seen in Figure 3.2 that most of the matching points that replicate BR data are inside the ATT-regime, just over the oscillatory one (in blue), suggesting that only a noise-induced mechanism may explain the perceptual trace statistics, characterized this time by dominance descriptors ($< T_{dom} >$ and C_v) and hysteresis measures (τ_H, and c_H) as well. However, a few exceptions are evident for less plausible (larger) values of τ_a McCormick & Williamson (1989); Sanchez-Vives et al. (2000), this pointing to an additional adaptation-driven mechanism that could produce the reversals in perception. However, a so slow negative feedback

might indicate a somehow 'frozen' system where the noise plays again a significant role in the observed switches. In contrast, for the case of the KDE display (figure 3.3), just the noise seems to be the cause of the reversals in perception that all the subjects report during the experiments.

As reported in Pastukhov & Braun (2011), an important difference between the perception of the two bistable stimuli (BR and KDE) is a high frequency of occurrence of return (failed) transitions and also the presence of long transitions (mixed states) in the BR perceptual traces (Figure 2.9 (b)-(c)), but absent when viewing illusory rotating spheres. These results correspond to the presentation of subjects with balanced inputs, but similar conclusions were previously reported in Brascamp et al. (2006), where also asymmetrical stimulation of the two eyes was studied (see section 2.2 for a summary).

In such work, the authors suggest that the presence of returned transitions in BR experiments could give strong support to a noise-driven mechanism underlying bistable perception. Brascamp et al. (2006) found no success in fitting their empirical data by many of the prevailing models, indicating that their results could unlikely be replicated by deterministic effects like the one of an adaptation dynamics. However, none of the models studied in this thesis was neither adjusted to take into account such frequent failed transitions in BR.

On the other hand, the existence of long intervals of mixed dominance as the ones reported in Pastukhov & Braun (2011) and Brascamp et al. (2006) may give support also to the presence of the counterpart of the fatiguing effects like adaptation or long-term depression, that continuously destabilize the dominance for both populations. However, another alternative mechanism of noise-induced transitions based on a model system with three stable states (attractors) could also be invoked to explain such intriguing results Brascamp et al. (2006); Martí i Ortega (2008).

Surprisingly, in most of the cases the perceptual statistics of the observers, no matter the kind of stimuli (KDE or BR), can be replicated by points lying on the ATT-region of the parameter space, thus explaining the low correlation values c_H reported in Pastukhov & Braun (2011). Moreover, such regions are generally more or less parallel to the bifurcation boundary with the oscillatory regime. This latest observation seems to support once again the importance of the additional destabilization effect produced by the adaptation currents over the dominance periods.

In contrast, in the previous work by Shpiro et al. (2009) small regions inside the OSC-regime are also predicted to be able to reproduce the bistable perception phenomena. Specifically, a balance between adaptation and noise-driven regimes is proposed there to explain the dominance data, because the matching regions contained the bifurcation boundary (see, for instance, Figure 4(a) in Shpiro et al. (2009)). Our results show that the inclusion of history-dependence measures (τ_H and c_H) rule out oscillatory behaviors as a plausible biological basis for bistable visual perception.

An additional information about the previous results can be gained by "unfolding" the volume (ϕ_a, β, I_0) associated to an intermediate value of $\tau_a = 1\ s$ onto planar subspaces at increasing values of the common input parameter I_0.

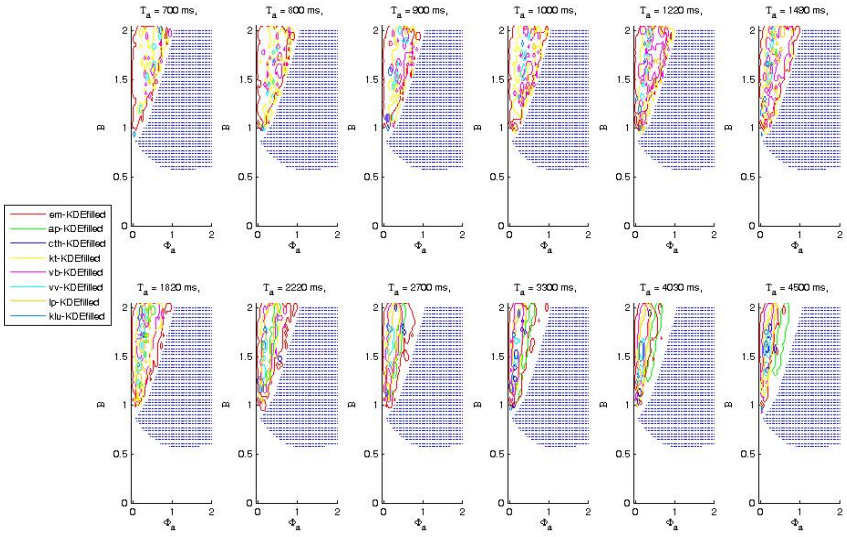

Figure 3.3: Fitting KDE-experimental results with the adaptation-LC model (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is outlined by a different color (legend on the left). The matching points are inside the ATT-regime, just over the oscillatory one (in blue). Base parameter set: $I_0 = 0.90$, $\theta = 0.0$, $k = 0.14$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

Figure 3.4 shows the projection onto such planes of the valid volumes found for each subject who was presented with BR stimuli. Note that more observers can be fitted with the adaptation-LC model when the input is fixed at a pair of values I_0, a result that will re-appear in some of the subsequent studied models.

The reason for such a general behavior can be explained by the 1D-bifurcation diagram of the free-noise model when the control parameter is I_0 (see Figure 3.1-B). Two Andronov-Hopf critical points actually exist when the common input is varied while the other parameters are kept fixed. Such a pair of values shall determine then the existence of two OSC-regions flanking the bistable region on either side, and consequently two possible available regions inside the ATT-region where to adequately fit a perceptual time-series with relatively low values of correlation c_H. Similar diagrams for this and other models can be found in other studies Shpiro et al. (2007); Curtu et al. (2008), where this characteristic dynamical behavior is shown to be deeply related with the capacity of a given model to fulfill Levelt's propositions Levelt (1965, 1966); Brascamp et al. (2006). This important issue is analyzed afterwards in section 4.2.

Below we continue the current analysis by examining other prevailing options in bistable perception modeling. The second case included in this study is the

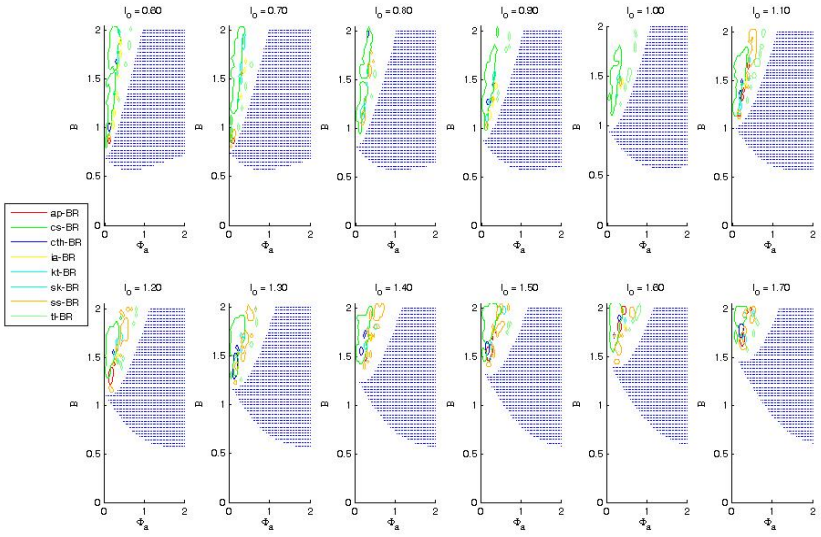

Figure 3.4: Fitting BR-experimental results with the adaptation-LC model (planar I_0-subspaces). The matching regions correspond to a fixed value of $\tau_a = 1\ s$ and various values of I_0, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.06, 0.22]$. Each observer's region is outlined by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). Base parameter set: $\tau_a = 1\ s$, $\theta = 0.0$, $k = 0.14$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

depression-LC model, *i.e.*, when the slow deterministic process that destabilizes the dominance is given by a long-term depression among neurons Shpiro et al. (2007). Then we continue with a more complex architecture that could be generalized to multi-stable stimuli with more than two possible interpretations: the Moreno-Bote's model with global excitation Moreno-Bote et al. (2007). Three-dimensional visualizations of the fitting results for these two models are provided in the following subsections, to help evaluating the validity of the conclusions derived above in the framework of the adaptation-LC model, which possesses a simpler architecture and time-independent connections between the two populations.

3.3 The depression-LC model

Next goal is assessing the robustness of our results by testing different neural mechanisms that may underly the bistable phenomenon happening when a subject is presented with ambiguous visual information. Under depression, one of these neural mechanisms, the current dominance of a given population can be cur-

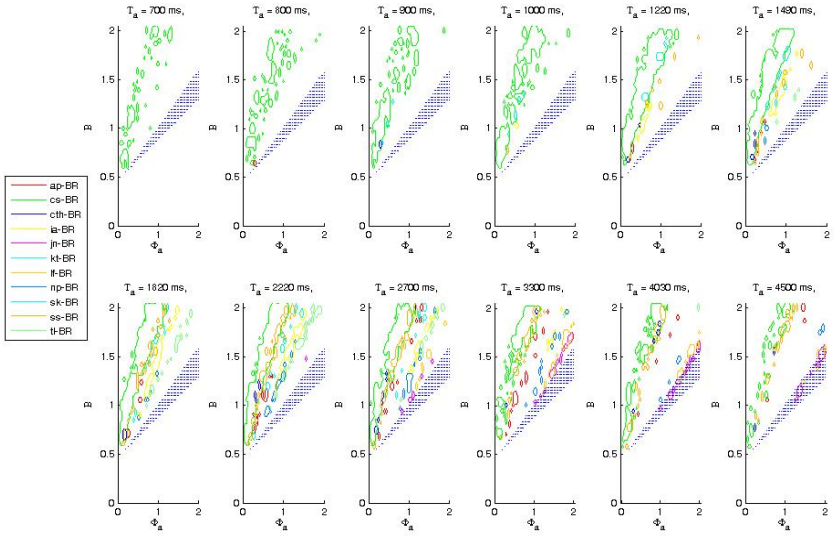

Figure 3.5: Fitting BR-experimental results with the depression-LC model (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is outlined by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). A few exceptions are evident for less plausible (larger) values of τ_a. Base parameter set: $I_0 = 0.50$, $\theta = 0.0$, $k = 0.04$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

tailed by the presence of a time-dependence in the synaptic connections between the two averaged neurons. Specifically, the cross-inhibition terms are multiplied by a long-term depressing variable described with a time-constant $\tau_d \sim 1s$. For a brief review of such a model, the reader is referred to section 2.3 and references therein.

The problem is tackled first in Figures 3.5 and 3.6, where the case of the sinusoidal gratings (BR) and that of the randomly rotating sphere (KDE) displays are, respectively, examined. Similarly to Figures 3.2 and 3.3, they offer a first view of the results obtained in the larger hyper-space $(\phi_a, \beta, I_0, \tau_d, \sigma_n)$. A planar slice is made at a specific value of the common input $I_0 = 0.5$ and posteriorly unfolded in various sheets with a fixed τ_d value. First of all, note that as in the adaptation-LC model, the value of the time-scale characterizing the slow dynamics (τ_d) does not affect the location of the bifurcation line, and consequently of the dynamical regimes, of the free-noise system.

Intriguingly, despite the rather different negative feedback mechanism, the depression-LC model seems to have a behavior akin to the one already seen in the adaptation-LC model. That is, the matching dynamics that reproduces the

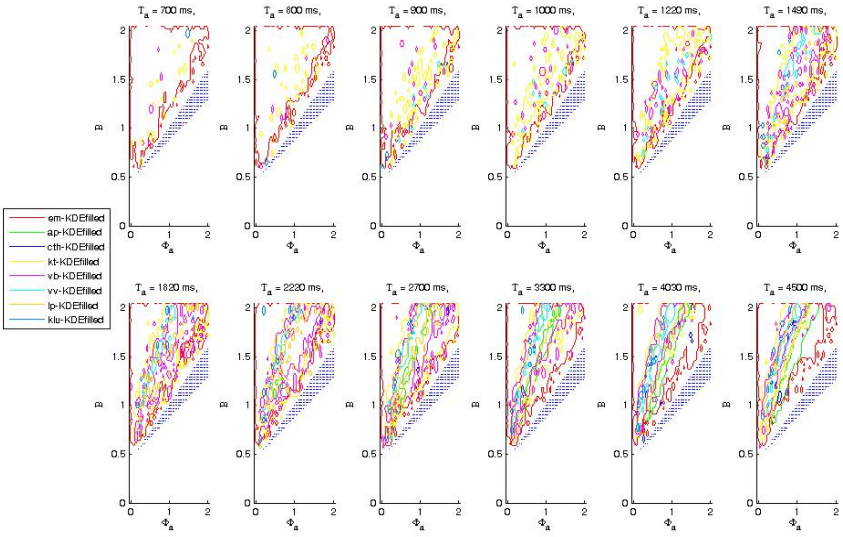

Figure 3.6: Fitting KDE-experimental results with the depression-LC model (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is outlined by a different color (legend on the left). The matching points are inside the ATT-regime, just over the oscillatory one (in blue). Base parameter set: $I_0 = 0.50$, $\theta = 0.0$, $k = 0.04$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

traditional dominance measures ($< T_{dom} >$, C_v), combined with the more subtle hysteresis factors reported in Pastukhov & Braun (2011), is largely described by a two-attractors system (*i.e.*, points in ATT-region) for the BR data, and exclusively reduces to the same kind of dynamical regime for KDE data. Thus, the noise seems to play an essential role in generating the observed reversals in perception, when a highly realistic ingredient as synaptic depression in the neural substrate is considered.

Further insight into the results obtained with the depression-LC model can be gained from Figure 3.7. Here a 3D visualization of the hyper-dimensional volume under study is presented. The right panel contains various subfigures that put into a more general context other 2D-graphs that could be shown as in the previous subsection. Some of them depict the dynamical regimes that characterize the behavior of the system in the absence of noise (first and second subfigures). Given that three dimensions are shown here, the bifurcation lines turn to be surfaces separating the fusion (green), oscillatory (blue) and bistable regions (red). The last subfigures of the left panel illustrate the location of the matching regions that mimic the original experimental data reported in Pastukhov & Braun (2011) for BR (brown) and KDE (yellow) stimuli. Similarly to the results discussed in the

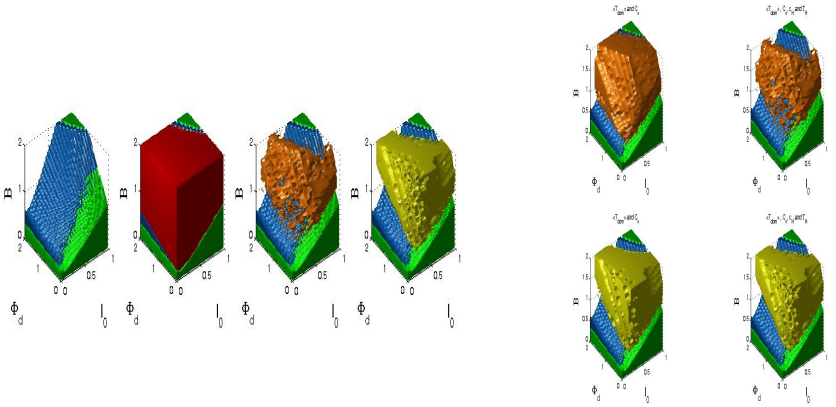

Figure 3.7: Fitting the experimental results with the depression-LC model. Left panel: (a) Bifurcation analysis of the noise-free system where a green volume indicates the parametric region that corresponds to a SIM-regime, and a blue color represents the OSC-regime. (b) The ATT-regime is shown in red color. (c) and (d) Regions matching the perceptual dynamics of human observers for BR displays (brown volume) and for KDE displays (yellow volume). The matching regions seem to lie mostly within the bistable regime and more or less parallel the oscillatory regime (blue). Right panel: (first row) The dynamic region defined by the BR distribution descriptors $< T_{dom} >$ and C_v only (left) is largely constrained when combined with the history-dependence quantifiers c_H, τ_H (right). (second row) The large experimental uncertainty in the KDE data collected, makes difficult to constrain the dynamic region defined by the KDE distribution descriptors only (left) when combined with the hysteresis measurements (right). Base parameter set: $\theta = 0.0$, $k = 0.04$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$. For more details see the text.

previous subsection, a depression-LC model should be set with parameters that visibly lie inside the bistable zone of the free-noise system.

At this point, it is useful to make a brief digression about the importance that the hysteresis phenomenon recently highlighted by Pastukhov & Braun (2011) could have to establish a modeling framework for bistable perception. In a previous study, Shpiro et al. (2009) has already made use of the most salient feature of the perceptual trace, *i.e.*, the descriptors of the distribution of dominance periods: $< T_{dom} >$ and C_v. In that paper the authors proved that these descriptors

notably reduce the regions in the parameters space that could replicate the general statistics characterizing a population of subjects. The basic result, tested in three theoretical frameworks (adaptation- and depression-LC models, apart from the Moreno-Bote et al. (2007) model with global excitation), consisted of detecting that the matching points were located in a small region near by the brink of the oscillatory regime, but spanning to either sides, thus considering both bistable and oscillatory dynamics as possible operating regimes.

An important issue that motivated the present study was to know whether the new data reported in Pastukhov & Braun (2011) could additionally constrain a given theoretical approach under study. The results presented above in subsection 3.2 give a first support to a positive answer to this question. Namely, Figures 3.3 and 3.4 show that most of the observers, no matter the kind of display presented with, can be fitted within a region that is near the bifurcation boundary between oscillatory and bistable regimes, but *inside* the ATT-region.

The second panel of figure 3.7 increases the evidence in favor of the hypothesis that the hysteresis data included in this work does constrain as well the case of the depression-LC model. The reduction of the possible regions in the parameters space is notable when the dominance distribution and history-dependence descriptors for the BR display are combined (first row). Different to this situation, the hysteresis data collected in Pastukhov & Braun (2011) does not visibly constrain the region supported by just the statistics of the dominance traces reported by the subjects viewing the illusory rotating sphere (second row). However, this fact can be explained by the large experimental uncertainty existent in the reported values of the four observables for many of the subjects (tables 2.1 and 2.2). The problem is more evident in the case of slow "switchers", *i.e.*, with larger values of average dominance time $< T_{dom} >$, due to the dramatic reduction of dominance periods that are available for analysis in the blocks of 5 min each.

3.4 An attractor network model with global excitation

Among the recent theoretical proposals to model multi-stable perception, the work by Moreno-Bote et al. (2007) occupies a privileged position as the first serious attempt to explain it through a purely attractor network approach. Needless to say that such a modeling option was slightly considered by other authors (*e.g.*, Wilson (2003)), but it seems that the complexity of the challenge *per se* over and over tilted the balance to the more intuitive image of a pseudo-deterministic periodic dynamics continuously perturbed by stochastic fluctuations. For instance, it is not self-evident that a biologically plausible bistable network, where any switch between the stable states is induced exclusively by noise, can produce the required perceptual dominance periods whose durations are well fit by Gamma or Log-normal distributions Lehky (1995); Borsellino et al. (1972); Fox & Herrmann (1967); Brascamp et al. (2005).

As it was explained in section 2.3, the Moreno-Bote et al. (2007) model mainly distinguishes among all the others in two aspects. Firstly, the derivation of the rate equations is explicitly based on an energy function intended to account for

some well-known properties of bistable perception such as the Levelt's propositions II and IV. Secondly, the basic architecture of the model network can be easily extended to generate the typical multi-stable behavior observed when the external stimulus allows more than two competing percepts be inferred by the brain.

Two alternative architectures are proposed in Moreno-Bote et al. (2007). Here we will consider only the GELI architecture that contains more novel ingredients that made it more interesting to be included in this thesis. Specifically, the information of the total external stimulation is provided by a unique (global) excitatory subpopulation that feeds the inhibitory populations. These last neurons, in turn, convey local signals to the respective excitatory cells representing the competing percepts. Finally, a feedback through the global excitatory subpopulation implements some level of mutual-inhibition between the primary excitatory populations, thus preventing the value of such a circuitry from being limited to just bistable stimuli.

For more details of the more interesting GELI-architecture of the Moreno-Bote et al. (2007) model, the reader can resort to the brief presentation made in section 2.3, or go directly to the related papers cited above. In the following, we continue by evaluating the suitability of this theoretical proposal to replicate the dynamics of the perceptual process, given the recent data reported in Pastukhov & Braun (2011).

A 3D visualization of our results for this case can be seen in figure 3.8. The left panel is similar to the one of figure 3.7 and can be useful for a rapid comparison between the behavior of the two models. The first thing that becomes glaringly obvious resides in the enormous difference between the bifurcation diagrams of their free-noise versions. In contrast to the depression-LC model, the GELI-case of the Moreno-Bote et al. (2007) model exhibits a bistable or ATT-region that is now visibly surrounded by oscillatory (blue) and simultaneous activity (green) behaviors in mostly all the planes parallel to the coordinates (ϕ_a, β) plane, when a transversal cut at some specific I_0 value is made. Moreover, when the planar slice is taken being parallel to the other coordinates planes including the I_0 axis ((ϕ_a, I_0) or (I_0, β)), the ATT-region is rarely flanked on either side by two OSC-regimes in the free-noise bifurcation diagram of the system (Figure 3.9).

As it was early pointed out in Shpiro et al. (2007), a model whose bifurcation diagram in the I_0 direction shows a slightly different behavior, can not accomplish with Levelt's propositions. Specifically, just the presence of a small oscillatory region at the left of the ATT-region in such a planar subspace (*i.e.*, at low values of I_0) can preclude the model system from fulfilling Levelt's propositions Curtu et al. (2008); Seely & Chow (2011).

Interestingly, despite the notable differences with the bifurcation diagrams of the models examined in the previous sections, once again the combined data of dominance and hysteresis measurements yield a matching region that entirely resides within the attractor regime of the free-noise system (last two subfigures of left panel in figure 3.8). Concerning the second panel, the use of the history-related observables c_H and τ_H visibly undeniably contributes to constrain the parameter region that can successfully replicate the whole set of data for BR

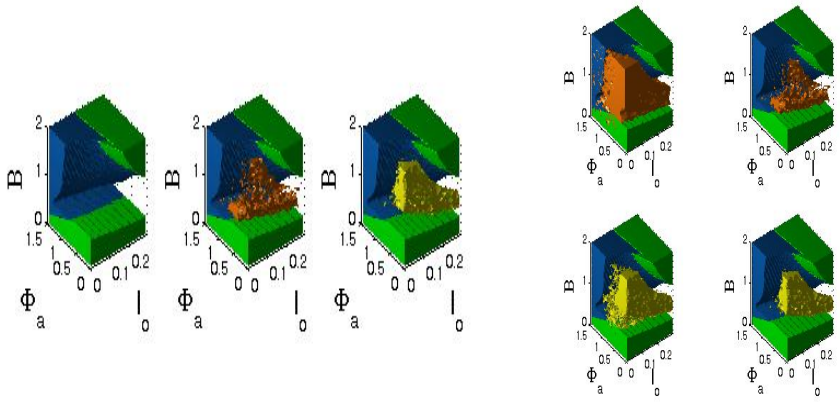

Figure 3.8: Fitting the experimental results with the Moreno-Bote (2007) model. Left panel: (a) Bifurcation analysis of the noise-free system where a green volume indicates the parametric region that corresponds to a SIM-regime, and a blue color represents the OSC-regime. The ATT-regime is not shown. (b) and (c) Regions matching the perceptual dynamics of human observers for BR displays (brown volume) and for KDE displays (yellow volume). The matching regions seem to lie mostly within the bistable regime and more or less parallel the oscillatory regime (blue). Right panel is similar to the one of Figure 3.7. Base parameter set: $\alpha = 2.00$, $\theta = 0.0$, $k = 0.03$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

display. However, they apparently do not contribute to reduce the KDE region obtained by the sole use of the observables $< T_{dom} >$ and C_v, the data used in Shpiro et al. (2009). This effect is closely related with similar results already obtained with the depression-LC model (compare to the second panel of figure 3.7), and can be explained by the larger relative errors associated to the average values collected in KDE observers, specially for slow switchers (thus providing with a rather small sample data).

Finally, note that differently to the adaptation- and depression-LC models, a strong level of excitatory recurrence α is needed to accurately simulate behavioral data [3]. Concretely, in order to reproduce most of the detailed psychophysical measurements listed in tables 2.1 and 2.2, such parameter is increased to the

[3] In fact, it is only in this case among to the rest of the models considered throughout the present manuscript, that some degree of excitatory recurrence is needed.

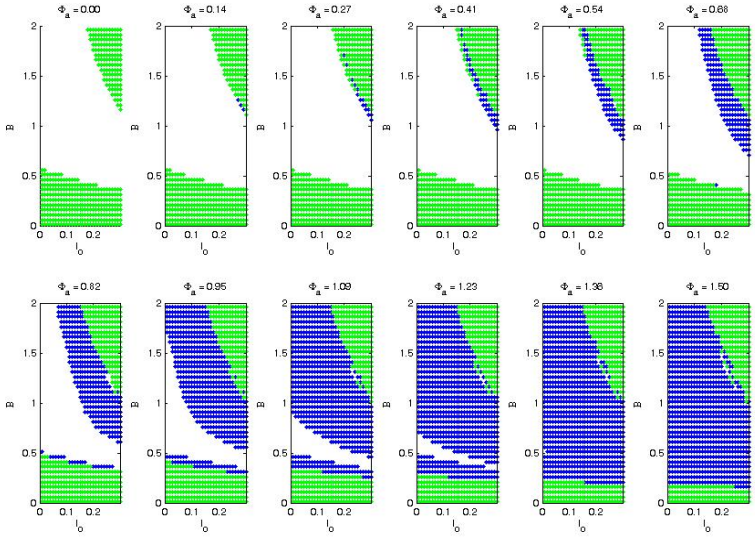

Figure 3.9: Two-dimensional bifurcation diagrams of Moreno-Bote et al's model for different ϕ_a values. Each β vs. I_0 plane shown corresponds to a particular planar slice at the 3D bifurcation diagram of Figure 3.8.

value $\alpha = 2.0$, in contrast to the smaller level $\alpha = 0.75$ published with the original paper Moreno-Bote et al. (2007), which successfully replicated the less detailed (dominance-related) data analyzed in Shpiro et al. (2009).

A possible reason why the original parameters of this model can not replicate the more complete set of data reported in Pastukhov & Braun (2011), could be the low values of correlation c_H between history and dominance periods. The study presented above (sections 3.2 and 3.3) by using two simpler versions of the Laing & Chow (2002) model, could convince the reader that such a data definitely favor the bistable region over the oscillatory one, a result that better precises the conclusions derived in Shpiro et al. (2009), which located the matching parameter's region at the bifurcation line. However, the ATT-region of the Moreno-Bote et al. (2007) model would result to be rather small for a value $\alpha = 0.75$ (see, for instance, Figure 5-(a) of Shpiro et al. (2009)), to allocate the more complete set of data shown in tables 2.1 and 2.2. Such a data set corresponds to about 20 subjects presented with two kind of stimuli and that quantifies for the first time the subtle hysteresis effects hitherto ignored but present in bistable perception.

How our fitting results obtained with the GELI-version of Moreno-Bote et al. (2007) model behave, when the time constant τ_a varies, is an issue that can be assessed by examining the figure 3.10. As in the 3D graphs, the bifurcation diagram over which the fitting data appear superimposed, clearly contrasts with the ones describing the two versions of the Laing and Chow's model studied

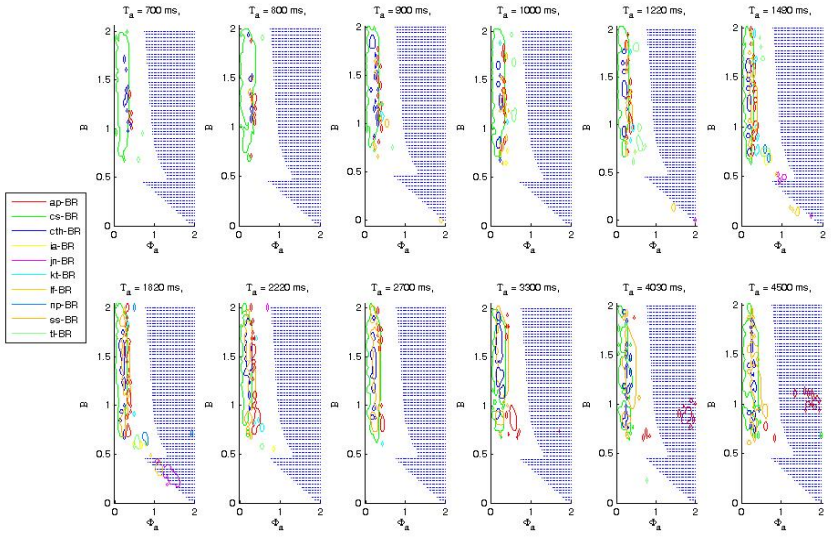

Figure 3.10: Fitting BR-experimental results with the Moreno-Bote (2007) model (GELI architecture) (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each subject's region is indicated by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). A few exceptions are evident for less plausible (larger) values of τ_a. Base parameter set: $I_0 = 0.08$, $\alpha = 2.00$, $\theta = 0.0$, $k = 0.03$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

above. Notably, these bi-dimensional diagrams are clearly in accordance with their volumetric counterpart shown in figure 3.8.

Figure 3.10 shows once again that the OSC-region is suitable to fit just a few of the huge set of behavioral data shown in tables 2.1 and 2.2. Nevertheless, such (essentially) adaptation-driven dynamics would be characterized by relatively larger values of the adaptation time-scale τ_a associated with Ca^{2+}-dependent K^+-channels, thus consequently with less biological plausibility McCormick & Williamson (1989); Sanchez-Vives et al. (2000). A similar unrealistic situation was obtained with the adaptation-LC model (*e.g.*, see figure 3.2), a result that would be emphasizing the relevance of a noise-driven mechanism in comparison to a somehow 'frozen' oscillatory system, to replicate low values of correlations c_H.

A similar 2D-view of our attempt to fit Moreno-Bote et al.'s model to the behavioral data reported by subjects viewing a KDE display is presented in figure 3.11. As in the case of the adaptation-LC model, the data associated to this kind of stimuli exclusively supports the hypothesis of a neural substrate where

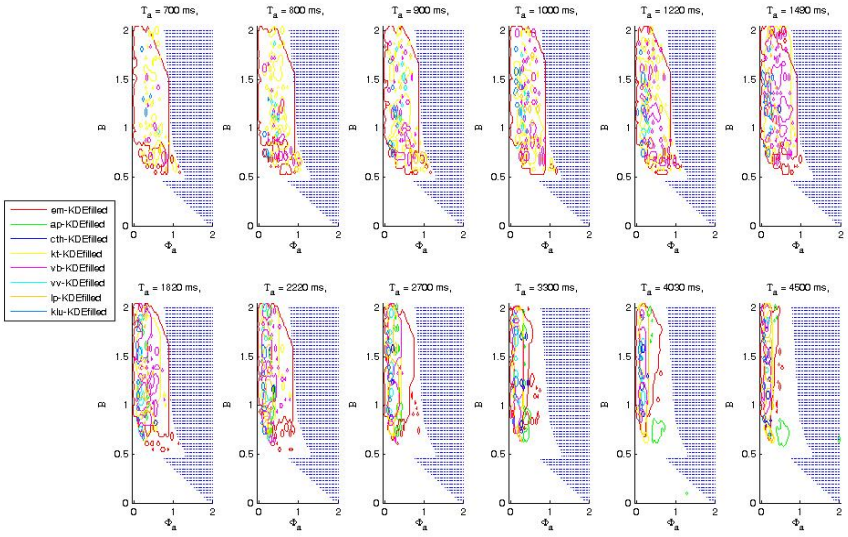

Figure 3.11: Fitting KDE-experimental results with the Moreno-Bote (2007) model (GELI architecture) (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is represented by a different color (legend on the left). The matching points are inside the ATT-regime, just over the oscillatory one (in blue). Base parameter set: $I_0 = 0.08$, $\alpha = 2.00$, $\theta = 0.0$, $k = 0.03$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

the switches would not occur in the absence of noise.

3.5 A model with a Naka-Rushton activation function

The model by Wilson (2003) can be mainly distinguished from the other ones by the kind of activation function used for the steady state in the firing-rate equations. The characteristic sigmoid-like behavior of the neuron's rate in response to a net input is described by a Naka-Rushton function. Then, as it was shown in section 2.3, an increasing adaptation elevates the threshold (located at the denominator of the function), thus exerting the expected fatigue effect over the firing-rate. However, this choice also creates a time-dependence that affects the slope of the sigmoid function at the threshold point, rather differently to most of the prevailing models used in bistable perception. Such double action of the adaptation in the Wilson (2003) model has been termed as a *divisive* effect by Shpiro et al. (2007).

This double time-dependence adds an additional handicap to fit bistable empirical data by the Wilson's model. Indeed, a pronounced sensitivity of the

firing-rate to even tiny changes in the net input is expected in the vicinity of the threshold value. Consequently, the possibility for occurring induced-noise reversals in the dominance of the corresponding neuron should be accurately coupled with such more complex dynamics created by the adapting currents. In fact, to our knowledge, no numerical simulations of the Wilson (2003) model have been published up to date, where the neural noise had been incorporated. The works by Kim et al. (2006) and Brascamp et al. (2006) appear to be rare exceptions, where the authors intended to fit their behavioral data by the first stage of Wilson's model. However, neither any methodology is presented nor explicit equations and parameters for the noise dynamics are provided. In contrast, numerical and analytical studies of the free-noise version of such a model can be found in several works Shpiro et al. (2007); Wilson (2007); Curtu et al. (2008).

To circumvent this problem, we have carried out extensive computational simulations before proceeding to fit the results reported in Pastukhov & Braun (2011) by this model, in order to test the validity of the noise range $\sigma_n \in [0.01, 0.22]$ that was used in the rest of the models. As a direct consequence, a re-calibration of the main parameters of the model was needed, in relation to the original version appeared in Wilson (2003). More precisely, the maximal rate was adjusted from $R_{max} = 100\ Hz$ to $R_{max} = 1.0$ (*i.e.*, dimensionless, as in the rest of models) and the threshold value reduced from $\theta = 10$ to an effective value of $\theta = 0.03$.

Some of our simulation results with the (first stage of) the Wilson (2003) model, are shown in figures 3.12 and 3.13. A planar slice at $I_0 = 0.5$ made at the hyper-space $(\phi_a, \beta, I_0, \tau_a, \sigma_n)$ is in turn unfolded in several planar subspaces taken at distinct τ_a values. In a quick glance, it turns clear that the Wilson's model offers less chances to fit behavioral data. Differently to the other models studied above, which permitted to find suitable parameter points for each of the 11 set of the (more precise) BR measurements, here only 9 of the 11 BR subjects could be fitted by the model. This is probably a direct consequence of the increased complexity introduced by the divisive effect of the adaptation in this model. Such a variable acts here over the saturating function's threshold and slope as well, making more difficult the adequate coupling with the noise dynamics contributing to reversals in dominance.

Nevertheless, similarly to the three models analyzed above in the previous sections, just a really small amount of the available data shown in the table 2.1 and 2.2 leave some chance that an adaptation-driven regime settled at the OSC-region could explain the empirical results from different observers and displays. Further agreement with the previous sections can be found in figure 3.14. It shows the "unfolding" of the hyper-dimensional parametric space associated to an intermediate value of $\tau_a = 1\ s$ onto planar subspaces at increasing values of the common input I_0 that feeds both populations. The projection onto such planes of the valid volumes found for each subject demonstrates that the BR-data can be fitted with the Wilson (2003) model around two different values of the input strength I_0. As it was pointed out above, this behavior is related to the existence of two oscillatory regions flanking a unique bistable region Shpiro et al. (2007); Curtu et al. (2008). Once again, the matching regions reside within the noise-driven zone.

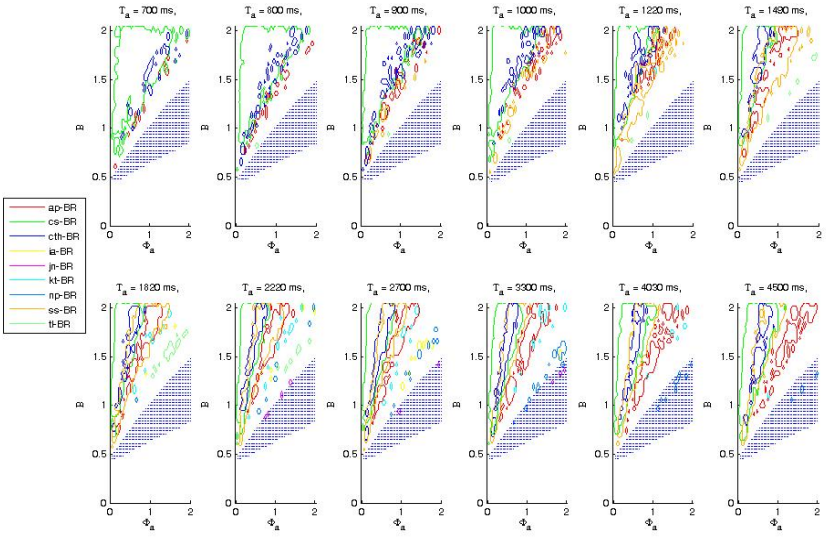

Figure 3.12: Fitting BR-experimental results with the (1st stage of) Wilson (2003) model (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is represented by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). A few exceptions are evident for less plausible (larger) values of τ_a. Base parameter set: $I_0 = 0.50$, $\theta = 0.03$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

3.6 The case of a nonlinear adaptation

In previous sections, two simplified versions of the Laing & Chow (2002) model were analyzed in the context of the detailed data recently reported by Pastukhov & Braun (2011): the adaptation-LC model (section 3.2) and the depression-LC model (section 3.3). In the first case, the slow process that conveys information about the dominant state of a given population is assumed to be given by a hyper-polarizing Ca^{2+}-dependent K^+-current that subsequently reduces the spike-frequency of the same neurons. In the case of the depression-LC model, the long-term depression present in the inhibitory synapses between the pair of competing (average) excitatory neurons acts as a slow negative feedback that eventually curtails the dominance of any of them.

To conclude our analysis of the implications of the experimental work by Pastukhov & Braun (2011) for future modeling approaches to bistable perception, in this part we study a third interesting form of the Laing and Chow's model: a one where the slow negative process is again described by an adapting current, but whose steady state is expressed by a nonlinear function, as originally pro-

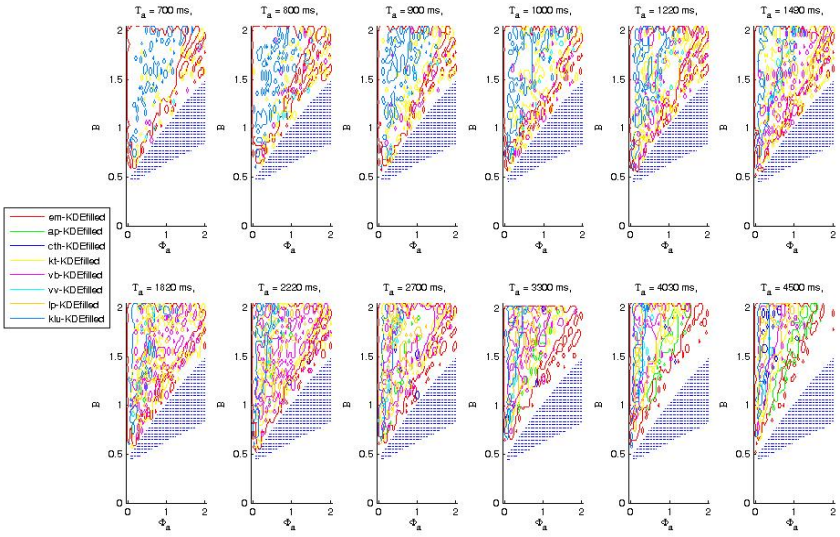

Figure 3.13: Fitting KDE-experimental results with the (1st stage of) Wilson (2003) model (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is represented by a different color (legend on the left). The matching points are inside the ATT-regime, just over the oscillatory one (in blue). Base parameter set: $I_0 = 0.50$, $\theta = 0.03$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

posed in Laing & Chow (2002) where a Heaviside function is used. In a recent work Curtu et al. (2008), such an additional saturating function is improved to a sigmoid-like one, similar to the function utilized in the firing rate equations of the adaptation- and depression-LC models. That is, a nonlinear function for the adaptation variables is used that adds two new parameters: θ_a, standing for the threshold, and κ_a related with the slope of the function at that point.

Interestingly, Curtu and co-workers argue both analytically and numerically that this variant of the Laing & Chow (2002) model can easily hold Levelt's proposition IV by just a simple manipulation of the adaptation saturating function. Choosing a value of θ_a larger than 0.5 and close to 1 would replicate the desired behavior, whereas the contrary could yield unrealistic increasing branches in the $< T_{dom} >$ plot as a function of the common value I_0. We postpone the assessment of this assertion to section 4.2; here, we limit ourselves to the fitting of such a variant of Laing and Chow's model (with a value of $\theta_a = 0.9$) to the data reported in Pastukhov & Braun (2011), which also appears tabulated in the tables 2.1 and 2.2 of the present manuscript.

Our results for the measurements related with the BR display are shown in figures 3.15 and 3.16. They correspond to planar subspaces at different values

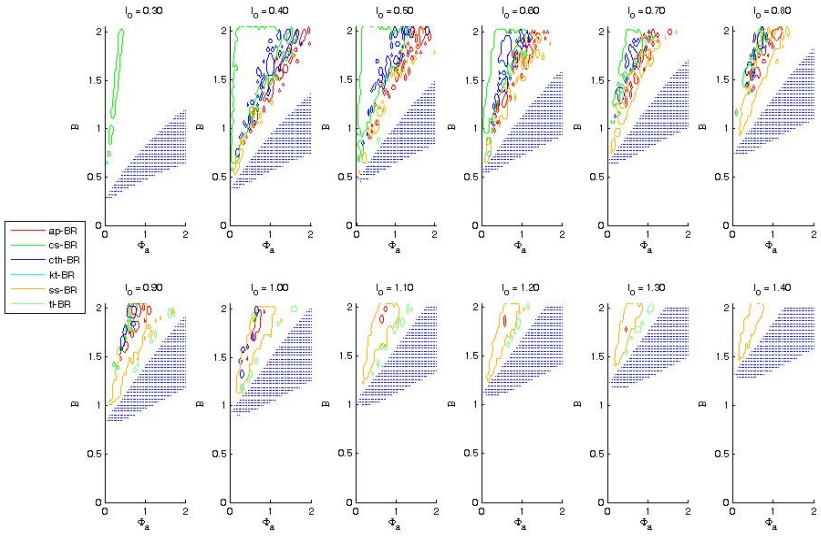

Figure 3.14: Fitting BR-experimental results with the (1st stage of) Wilson (2003) model (planar I_0-subspaces). The matching regions correspond to a fixed value of $\tau_a = 1\ s$ and various values of I_0, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is outlined by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). Base parameter set: $\tau_a = 1\ s$, $\theta = 0.03$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

of the time constant τ_a (with the I_0 value kept fixed) and at various values of I_0 (with τ_a unchanged), respectively. It can be observed that, despite the clear difference with the linear adaptation-LC model analyzed in section 3.2, the model shows results that are similar not only to the ones obtained with such a simpler variant, but also with the depression-LC version of the Laing and Chow's model (section 3.3) and Wilson's model (section 3.5), among others. The matching regions replicating the perceptual traces reported by individuals viewing the KDE display (rotating spheres) can also be fitted just with points that reside inside the bistable region (not shown).

The results presented in this chapter give strong support to the fact that the psychophysical measurements collected in Pastukhov & Braun (2011), effectively reduce the feasibility of many prevailing models for bistable perception, limiting the working regime to a one where the noise is indispensable and plays an essential role in generating the usual perceptual reversals experienced by a subject, and would be absent without it: the ATT-region. Our work refines the previous study by Shpiro et al. (2009), where a fine balance between adaptation- and noise-driven regimes is invoked to explain *averaged* dominance data (*i.e.*, described by

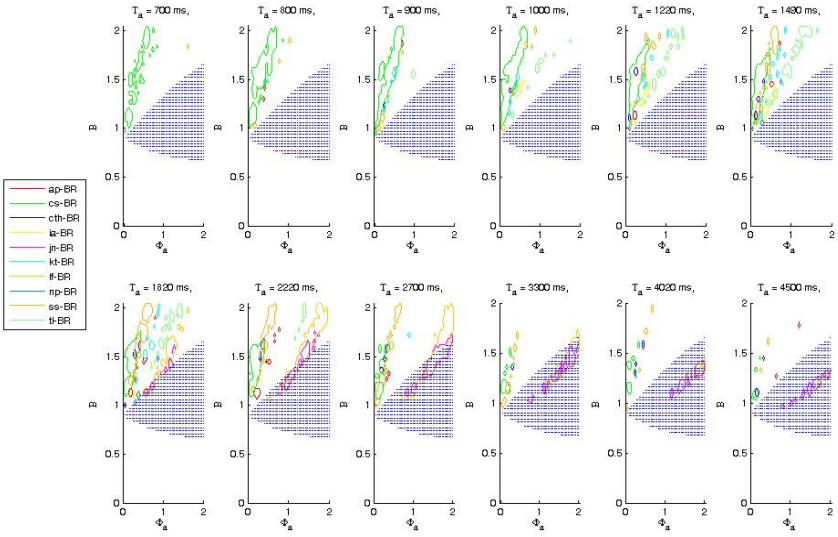

Figure 3.15: Fitting BR-experimental results with the nonlinear adaptation-LC model (planar τ_a-subspaces). The matching regions correspond to specific values of τ_a, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is outlined by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). A few exceptions are evident for less plausible (larger) values of τ_a. Base parameter set: $I_0 = 0.90$, $\theta = 0.0$, $k = 0.14$, $\theta_a = 0.9$, $\kappa_a = 0.14$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

$< T_{dom} >$ and C_v quantifiers) in bistable perception.

Indeed, the combination of such more salient data characterizing the phenomenon, but measured at the *individual* level, and the subtle *hysteresis* effects (*i.e.*, described by τ_H and (low but significant) c_H values) demonstrated by Pastukhov & Braun (2011), definitely rules out the possibility that the underlying neural system could produce the usual switches statistics purely by a fatigue (adaptation or depression) process, and where the noise role would be reduced to generate eventual deviations around a "limit cycle"-based deterministic dynamics. Such point of view has been adopted by many of the authors to explain the characteristic Gamma-like distribution widely accepted to describe the set of dominance durations of a subject, regardless the kind of display Lehky (1995); Borsellino et al. (1972); Fox & Herrmann (1967); Brascamp et al. (2005) and even of the way of stimulation Logothetis et al. (1996).

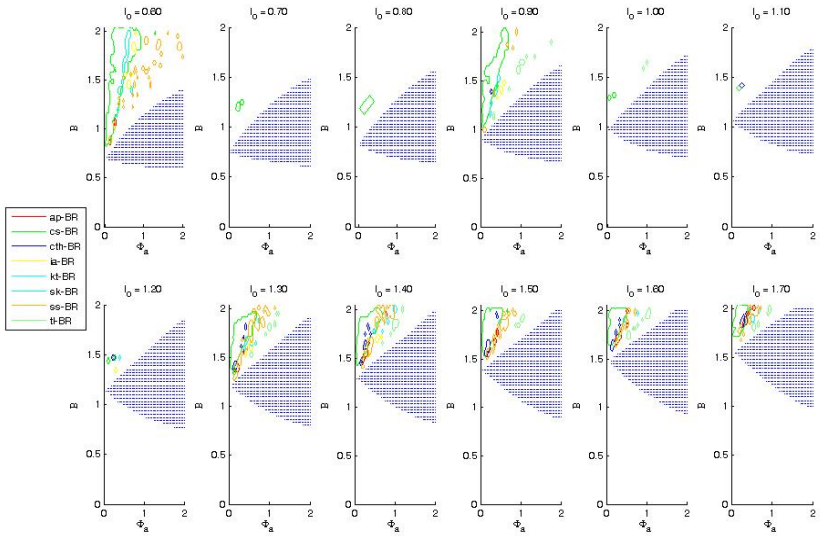

Figure 3.16: Fitting BR-experimental results with the nonlinear adaptation-LC model (planar I_0-subspaces). The matching regions correspond to a fixed value of $\tau_a = 1\ s$ and various values of I_0, after combining the valid points in a range of dispersion noise $\sigma_n \in [0.01, 0.22]$. Each observer's region is outlined by a different color (legend on the left). Most of the matching points are inside the ATT-regime, just over the oscillatory one (in blue). Base parameter set: $\tau_a = 1\ s$, $\theta = 0.0$, $k = 0.14$, $\theta_a = 0.9$, $\kappa_a = 0.14$, $\tau_n = 100\ ms$, $\tau_r = 10\ ms$.

4 Validation of psychophysical properties

In the previous chapter we have seen how the history-dependence of the perceptual process shown by Pastukhov & Braun (2011), effectively constrains the region of the parameter space able to replicate the empirical data. Concretely, that just small regions residing inside the *bistable* or *two-attractors* region of the whole parameter space are actually adequate. Here we continue the analysis by examining these constrains imposed by the experimental results of Pastukhov & Braun (2011) in the context of Levelt's propositions, a study in the same spirit of the recent work by Curtu et al. (2008); Seely & Chow (2011).

The present chapter finishes with a study about the capability of the mentioned models to reproduce more empirical data related with bistable perception. That is, the resonance effects happening when varying external *frequencies*, as shown by Kim et al. (2006) and summarized in Section 2.2 of this thesis. Importantly, a resonance respect to the noise dispersion (*i.e.*, a true *stochastic* resonance) is clearly demonstrated here for the first time, in contrast to the methodology followed in Kim et al. (2006).

4.1 Optimizing the fitting results for each observer: a weighted sum algorithm

In the previous chapter, we have exhaustively explored a hyper-dimensional parameter space of several models, in order to find any point that replicates that detailed statistics of the perceptual process that happens in individual observers presented with two different displays (BR and KDE). The data collected from each observer is independently fitted, by the application of the methodology reported in Shpiro et al. (2009) to much more parameters. Altogether, a five-dimensional parameter set is examined, formed by (1) the common value of the input strength ($I_1 = I_2 = I_0$), (2) the strength of mutual inhibition (β), (3) the strength of adaptation (ϕ_a) (or depression, (ϕ_d), in the case of the depression-LC model), (4) the associated time-scale (τ_a) for such a fatigue process (or τ_d when synaptic depression is considered) and (5) the dispersion of the noise variable (σ_n).

Additionally, a searching for effective values of the activation function pa-

rameters (slope $\sim 1/\kappa$ and threshold θ) always precedes the study of the mentioned 5D-hyperspace, in order to guarantee that average values for observables $< T_{dom} >$ and C_v can be fitted within a noise range that is kept fixed within a noise range $\sigma_n \in [0.01, 0.22]$.

However, till this point our procedure tends to give a major importance to the range of the observables reported in tables 2.1 and 2.2, and consequently assigns the same value to any set of parameters yielding a simulated value that lies within the desired interval. For instance, no distinction is made between two different sets of parametric points $(I_0^{(1)}, \beta^{(1)}, \phi_a^{(1)}, \tau_a^{(1)}, \sigma_n^{(1)})$ and $(I_0^{(2)}, \beta^{(2)}, \phi_a^{(2)}, \tau_a^{(2)}, \sigma_n^{(2)})$, regardless their outcomes when used as parameters in computational simulations of the model. That is, no greater importance is given to the former point if it generates closer values to the average data of the four observables corresponding to the subject, whereas the second one just simulates values situated at the extremes of the range permitted by the reported statistics.

Consequently, each subject has been associated to a spot in the parameter space of each model that forms a "matching" region, but no attention has been paid to the "quality" of the fitting provided by each individual point of the matching region. Besides, this approach does not allow to make model predictions for each individual separately, due to the possibility to opt for many 5D-vectors of parameters among all the ones contained into the matching zone. It is then recommendable to include some strategy in our procedure to overcome this problem. In fact, any serious attempt to fit a set of experimental data to a function or a set of equations normally includes some optimization stage aimed to minimize the (error) difference between the original data and the obtained simulated values [1].

The idea proposed here relies on the assumption that the Central Limit theorem is applicable to the measurements collected in Pastukhov & Braun (2011). More precisely, that the size of the data sample used to compute the averaged values summarized in tables 2.1 and 2.2 is large enough (*i.e.*, with 30 or a higher number of observations). In such a case, it is reasonable to assume that the possible outcomes X of a measurement of the average value of any of the four observables ($< T_{dom} >$, C_v, c_H and τ_H) can be drawn from a normal probability density which may be approximated as follows:

$$P(X) \approx \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left[-\frac{(X-\overline{X})^2}{2\sigma_X^2}\right]. \tag{4.1}$$

Here $\overline{X}$ represents the average value of the observable obtained from the finite sample of observations, and σ_X is the sample standard deviation, which are used as estimators of the corresponding parameters (mean and standard deviation) of the whole (unknown) population of the possible outcome values X.

Note that the exponent in the Gaussian distribution of equation (4.1) is nothing else than the quadratic difference between the outcome value X and the expected average value $\overline{X}$, relative to the square of the measurement error σ_X^2. The equation therefore allows for the calculation of probabilities of outcomes of

[1] For instance, the Least Squares method, the Steepest Descent algorithm and Newton's method, among others.

a concrete set of perceptual traces, by combining in a single number the contrary effects given by the absolute difference between X and the population mean (numerator) and the one coming from the unavoidable variability of the measurement process (denominator).

Since the possible outcome value of a particular measurement of any of the four observables $< T_{dom} >$, C_v, c_H and τ_H, obtained from a set of perceptual traces, constitutes a statistical independent event, the joint probability describing the complete set of observables is given by the product of the corresponding four marginal probabilities [2]:

$$P(< T_{dom} >, C_v, c_H, \tau_H) \approx \prod_{X \in \mathcal{X}} \frac{1}{\sigma_X \sqrt{2\pi}} \exp - \frac{(X - \overline{X})^2}{2\sigma_X^2}, \tag{4.2}$$

where $\mathcal{X} = \{< T_{dom} >, C_v, c_H, \tau_H\}$. This equation gives the frequency of occurrence of a combined set of *particular* (*i.e.*, before averaged) values that might be used later to obtain the averaged values collected in tables 2.1 and 2.2. Consequently, it can be used to evaluate the relevance of a specific 5D vector of parameters $v = (I_0, \beta, \phi_a, \tau_a, \sigma_n)$ (or $v = (I_0, \beta, \phi_d, \tau_d, \sigma_n)$ when depression is considered) contained in a subject's matching region. To that end, a *weight* factor may be assigned to the vector v according to its simulated values X for each observable, and its relative difference to the empirically found average $\overline{X}$:

$$w_v(X) = \exp\left[-\frac{(X - \overline{X})^2}{2\sigma_X^2}\right], \qquad X \in \mathcal{X} = \{< T_{dom} >, C_v, c_H, \tau_H\}. \tag{4.3}$$

Subsequently, an overall weighting factor characterizing the current vector v can be constructed by taking the product of these four numbers:

$$W_v = W_v(< T_{dom} >, C_v, c_H, \tau_H) = w_v(< T_{dom} >) w_v(C_v) w_v(c_H) w_v(\tau_H). \tag{4.4}$$

The usefulness of the last definitions can be easily explained as follows. Let

$$\begin{aligned} v^{(1)} &= (I_0^{(1)}, \beta^{(1)}, \phi_a^{(1)}, \tau_a^{(1)}, \sigma_n^{(1)}), \\ v^{(2)} &= (I_0^{(2)}, \beta^{(2)}, \phi_a^{(2)}, \tau_a^{(2)}, \sigma_n^{(2)}), \\ &\vdots \\ v^{(m)} &= (I_0^{(m)}, \beta^{(m)}, \phi_a^{(m)}, \tau_a^{(m)}, \sigma_n^{(m)}) \end{aligned} \tag{4.5}$$

be the m points in the parameter space that constitute the matching region of a given observer, and $W_{v^{(i)}}$, $i = 1, 2, \ldots m$ their associated weights. Then, the relevance of each of the m available 5D points can then be easily quantified by

[2]Note that it is not argued an statistical independence between the *average* observable outcomes of different perceptual traces. In fact, the average values of τ_H and $< T_{dom} >$ seem to be correlated, as clearly pointed out in Pastukhov & Braun (2011) (see Figure 3-C).

using the expression (4.4) and finally compute an overall *weighted* average that reduces the whole matching region to a unique point $\overline{v}$

$$\overline{v} = \frac{W_{v^{(1)}} v^{(1)} + W_{v^{(2)}} v^{(2)} + \ldots + W_{v^{(m)}} v^{(m)}}{W_{v^{(1)}} + W_{v^{(2)}} + \ldots + W_{v^{(m)}}}, \tag{4.6}$$

with its corresponding variance:

$$\sigma_v^2 = \frac{W_{v^{(1)}} (v^{(1)} - \overline{v})^2 + W_{v^{(2)}} (v^{(2)} - \overline{v})^2 + \ldots + W_{v^{(m)}} (v^{(m)} - \overline{v})^2}{W_{v^{(1)}} + W_{v^{(2)}} + \ldots + W_{v^{(m)}}}. \tag{4.7}$$

Similar averages across the matching zone of the subject can be computed over the predicted values for the observables $X \in \{< T_{dom} >^{(i)}, C_v^{(i)}, c_H^{(i)}, \tau_H^{(i)}\}$

$$\overline{X} = \frac{W_{v^{(1)}} X^{(1)} + W_{v^{(2)}} X^{(2)} + \ldots + W_{v^{(m)}} X^{(m)}}{W_{v^{(1)}} + W_{v^{(2)}} + \ldots + W_{v^{(m)}}}, \tag{4.8}$$

with its corresponding variance:

$$\sigma_X^2 = \frac{W_{v^{(1)}} (X^{(1)} - \overline{X})^2 + W_{v^{(2)}} (X^{(2)} - \overline{X})^2 + \ldots + W_{v^{(m)}} (X^{(m)} - \overline{X})^2}{W_{v^{(1)}} + W_{v^{(2)}} + \ldots + W_{v^{(m)}}}. \tag{4.9}$$

To further facilitate the understanding of the methodology explained above, such averages were computed taking the adaptation-LC model as example. These values obtained from the modeling study can then be compared with the averaged values computed from the experimental data (tables 2.1 and 2.2). The results are shown in Figures 4.1

Importantly, the unique parameter point $\overline{v}$ defined by expressions (4.6) and (4.7) and the average vector of observables $\overline{X}$ determined by equations (4.8) and (4.9), now characterize the given subject in the scope of the model under analysis. This fact allows to run additional computational experiments to make new predictions subject by subject. We make use of this advantage in the next sections, making possible to study the classical problem of Levelt's propositions (Section 4.2). Lastly, in Section 4.3 we will be able to use our fitting results to study the challenging topic of stochastic resonance, and evaluate some interesting conclusions stated in the experimental study by Kim et al. (2006).

4.2 Testing the fitting: Levelt's propositions

The propositions stated by Levelt (1965, 1966) form a set of well established psychophysical laws in bistable visual perception, a topic where the experimental research have continued up to date Brascamp et al. (2006); Klink et al. (2008); Moreno-Bote et al. (2010). Particularly, a compact version of the first three propositions enunciated by Levelt has been the object of investigation in many different conditions: that the effect of decreasing the stimulus strength in the image presented to one eye is limited to increase the dominance duration of the *competing* image, while (intriguingly) no change is detected in the perception of the unchanged stimulus. Such surprising effect has been demonstrated not only in

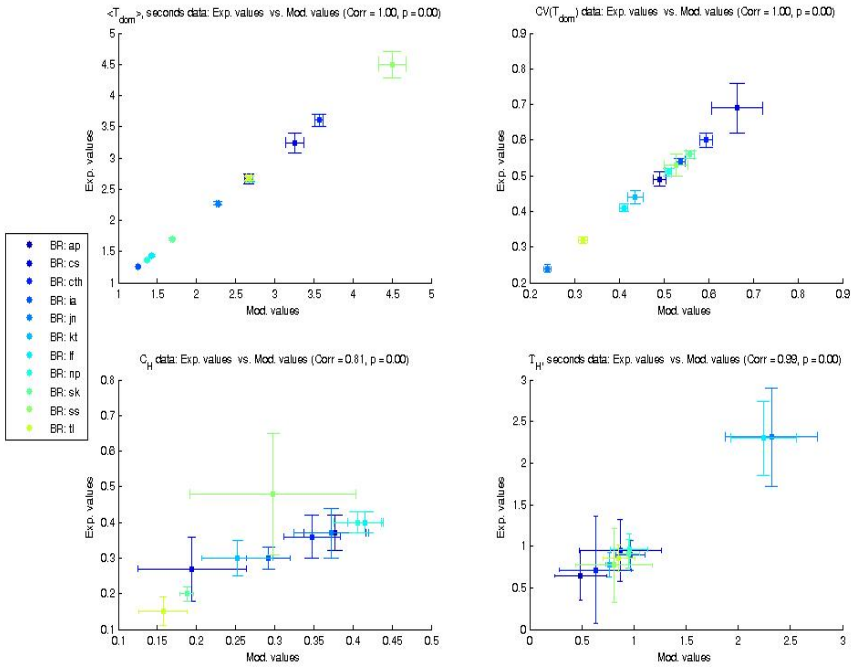

Figure 4.1: An example of the application of the weighted sum algorithm defined by equations (4.8) and (4.9). A reasonable agreement between empirical and simulated data is typically found.

(conventional) experimental designs where the ambiguous stimulus is presented continuously, but also when the pair of incongruent images are discontinuously presented with interleaved blank periods and even swapped between the eyes Logothetis et al. (1996).

Theoretical work has frequently tackled the problem of evaluating a prevailing model in bistable perception, in order to determine its capability to replicate the related Levelt's results Shpiro et al. (2007); Seely & Chow (2011); Moreno-Bote et al. (2010). For instance, recent attempts have been made to replicate the more general behavior shown in Brascamp et al. (2006); Klink et al. (2008) respect to the compact version of Levelt's law II, namely: when an unilateral contrast change takes place, it mainly affects dominance durations of the eye presented with *currently* higher contrast stimulus. Specifically, Moreno-Bote et al. (2010) has hypothesized the normalization of the inputs as a possible explanation of this strange phenomenon.

As part of this effort, here we evaluate the plausibility of our findings presented in section 3.1, in the context of the Levelt's propositions. We make use once again of the parameter points computed for each subject by using the equation 4.6,

limiting ourselves to the observers presented with the pair of orthogonal sinusoidal gratings (BR display). KDE subjects are not taken into account due to the larger uncertainty present in their reported measurements. Needless to say, nevertheless, that such type of stimuli indeed admit the manipulation of the contained clues that influence the possible interpretation outcome Klink et al. (2008). These authors clearly demonstrate that the validity of Levelt's propositions extends to the rather different structure-from-motion patterns like the rotating spheres (KDE display) used in Pastukhov & Braun (2011).

Levelt's fourth proposition

The possibility to replicate Levelt's fourth proposition in the context of the non-linear adaptation-LC model is illustrated in figure 4.2. The profile of the average dominance duration $< T_{dom} >$ as a function of the common input $I_1 = I_2 = I_0$ is shown, for each BR subject. In contrast to the expected decreasing curve predicted in Curtu et al. (2008) for a threshold value $\theta_a = 0.9$, it is primarily the global presence of a characteristic non-monotonic behavior that jumps out at the reader. Similar non-monotonic profiles in the context of several models (including the adaptation-LC variant) have been reported in a previous modeling study Shpiro et al. (2007). The authors verify the occurrence of such a behavior in computational simulations of the noisy systems (as an example, see Figure 6A of Shpiro et al. (2007), for the depression-LC model), and in the less realistic free-noise cases as well (figures 3C, 4D and 5B in the same reference).

The explanation of the undesired non-monotonic profile of $< T_{dom} >$ can be found in the behavior of the associated free-noise system. According to Shpiro et al. (2007), such profile can be observed for high values of the inhibition strength parameter β, a situation that favors the presence of a bistable region in the bifurcation diagram of the free-noise system. Additionally, such region of attractors should be flanked by two oscillatory regimes. As we have discussed in subsection 3.6, this is precisely what one can see in the 1D bifurcation diagram of the adaptation-LC model, obtained when the common input I_0 is varied while the other parameters remain unchanged (Figure 3.1-b): two Andronov-Hopf critical points exist, giving rise to a pair of limit cycles on either side of the ATT-region.

Before continuing, a brief digression is necessary. The work by Shpiro et al. (2007) clearly warns on an important difference between the noisy and the unperturbed system. That is, that the appearance of a non-monotonic behavior in the $< T_{dom} >$ *vs.* I_0 graph of the free-noise system is limited to intermediate values of the inhibition parameter β, but below some critical value β_c. This is because higher values of inhibition make the ATT-region in the middle unreachable by definition [3]. On the contrary, the presence of noise may certainly generate non-monotonic profiles of $< T_{dom} >$ for values of $\beta \leq \beta_c$, as is shown in figure 6A of Shpiro et al. (2007). The influence of some "ghost" presence of the ATT-regime for values of $\beta \sim \beta_c$ in the free-noise regime is then invoked by the authors.

[3] Actually, the situation demands the consideration of the other dimension of the parameter space given by the input strength I_0

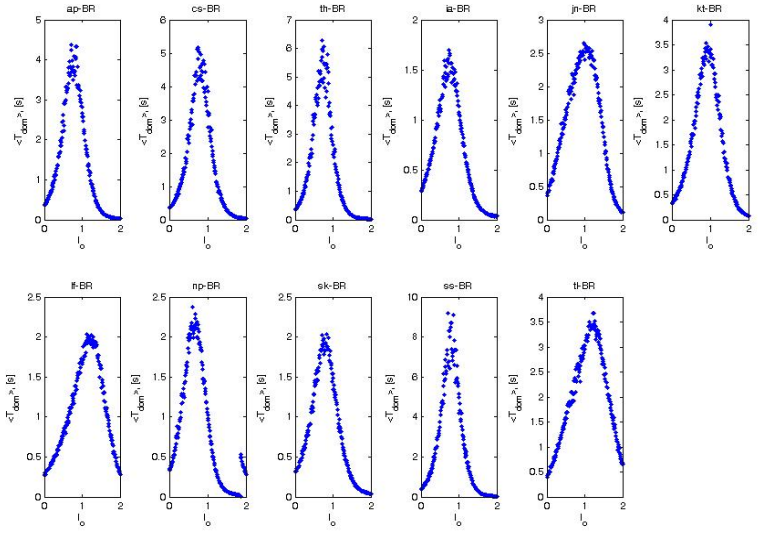

Figure 4.2: The average dominance time $< T_{dom} >$ for BR observers being modeled as an nonlinear adaptation-LC system each, is plotted as a function of the common input strength I_0. A non-monotonic profile is typically found, in spite of the characteristic decreasing curve found empirically (Levelt's proposition IV). Each observer is described by a unique 4D vector of parameters $(\overline{\beta}, \overline{\phi_a}, \overline{\tau_a}, \overline{\sigma_n})$, which is obtained by a (weighted) average (Eq. (4.6)) over the whole matching region found in subsection 3.6.

As it was commented in the mentioned subsection, the existence of two bifurcation boundaries delimiting the bistable region from the oscillatory ones, opened a double possibility to fit the model to the behavioral results of Pastukhov & Braun (2011) (see comments related with figure 3.4, page 52). This fact suggests that an alternative (mathematical) explanation for the non-monotonic behavior observed in the $< T_{dom} >$ profile could be in the I_0 component of the average vector $\overline{v}$ of parameters obtained from equation 4.6. Namely, that the average $\overline{I_0}$ value necessarily lies in the middle of the bistable region, at approximately the same distance from both bifurcation boundaries.

Note that the average indicated by Equation (4.6) to obtain a 5D vector $\overline{v}$ of parameters defining each subject in the parameter space $(\beta, \phi_a, \tau_a, \sigma_n)$, is *not* actually independent in each parameter direction. In fact, only the simulated values of the observables ($< T_{dom} >$, C_v, τ_H and c_H) matter in the computation of the weights given by the expression (4.4). Nevertheless, the average in any parameter dimension along which the matching region is divided in two feasible sets of points, will contribute with equal weights in the final averaged parameter.

Another model of interest to be analyzed in relation with Levelt's proposi-

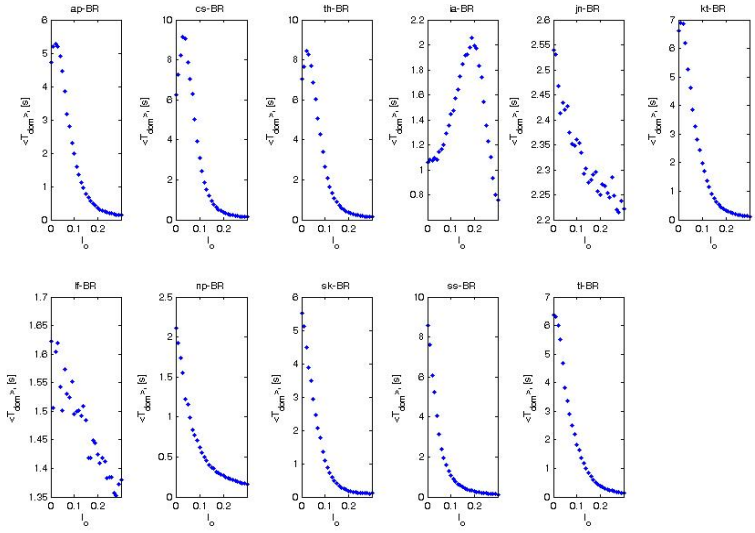

Figure 4.3: The average dominance time $< T_{dom} >$ for BR observers being modeled as a Moreno-Bote(2007) system each, is plotted as a function of the common input strength I_0. A non-monotonic profile is rarely found, in agreement with the characteristic decreasing curve found empirically (Levelt's proposition IV). Each observer is described by a unique 4D vector of parameters $(\overline{\beta}, \overline{\phi_a}, \overline{\tau_a}, \overline{\sigma_n})$, which is obtained by a (weighted) average (4.6) over the whole matching region found in subsection 3.4.

tions is the Moreno-Bote (2007) model. As it was mentioned in section 3.4, it distinguishes from the others because its dynamical equations are explicitly based on an energy function intended to account for Levelt's propositions (see, for instance, figures 2 and 10 of Moreno-Bote et al. (2007)). The numerical results of our attempts to reproduce Levelt's fourth proposition by the Moreno-Bote (2007) model are shown in Figure 4.3. It can be seen that in most of the cases, such a model does hold this empirical law for bistable perception.

Levelt's second proposition

In contrast to the careful scrutiny that in many modeling studies has usually received the fourth law early enunciated by Levelt, his second proposition has just started to be examined in recent works Seely & Chow (2011); Moreno-Bote et al. (2010). To our knowledge, neither a rigorous theoretical study about the suitability of prevailing models respect to such empirical behavior exists to date, comparable to the excellent analytical study reported in Curtu et al. (2008) with respect to the Levelt's law IV.

As it has been mentioned elsewhere, Levelt's second proposition describes the

effect of the monocular manipulation of the images on the subjective experience of a subject presented with two incongruent sensory inputs (chapter 1). It states that the effect of decreasing the stimulus strength in the image presented to one eye is limited to the increment of the dominance duration of the *competing* image, while no change is detected in the perception of the unchanged stimulus Levelt (1965, 1966). However, it is already known that the validity of this statement is limited to relatively low values of the input strength.

A more rigorous psychophysical study of binocular rivalry by Brascamp et al. (2006), where the contrasts values examined spanned the entire range from near the detection threshold to the theoretical maximum, has shown that the original version of Levelt's second proposition is only valid in a limited contrast range (the mentioned work is summarized in section 2.2). Outside this range, the opposite is true. Thus, it is necessary to replace it by the more precise statement that unilateral contrast changes mainly affect dominance durations of the eye presented with *currently* higher contrast stimulus.

Below we discuss our modeling results with regard this other Levelt's proposition. Note, that the updated version of Levelt's second law leads to a more complete picture about the behavior of the average dominance time $< T_{dom} >$ of each percept as a function of the unilateral change of the input strength, let's say I_2: a pair of monotonic curves that intercept at the intermediate value of no input bias $I_1 = I_2$.

Figures 4.4 and 4.5 show the simulated results with the Wilson's model and the non-linear version of the LC-adaptation model, respectively for the BR subjects. A reasonable agreement with the lasted version of the proposition can be seen in both cases, with most of the subjects being described by two monotonic $< T_{dom} >$ profiles: a one with increasingly long dominant phases with mean $< T^{(2)}_{dom} >$ (green) that correspond to the variable stimulus, and another decreasing curve describing the behavior of the average dominance durations $< T^{(1)}_{dom} >$ (blue) of the unchanged monocular stimulus.

A different situation occurs when the subject's perceptual experience is simulated by a depression-LC model (Figure 4.6). Then a pair of monotonic curves intercepting when no input bias exists ($I_1 = I_2$) is rarely found, in clear contradiction with Levelt's second proposition. Specifically, the average value of the intervals when the variable stimulus (with input strength I_2) dominates shows an unexpected non-monotonic profile (green). These numerical results confirm the necessity of additional (analytical/numerical) studies concerning the suitability of the depression-LC model, in the same spirit of the existing work that has been done in the scope of the fourth proposition of Levelt Seely & Chow (2011); Curtu et al. (2008). Obtaining the corresponding bifurcation diagrams of the free-noise system when only one input strength changes while the other remains constant may be an interesting starting point, similarly to the analyses made when the two inputs variables simultaneously vary ($I_1 = I_2$).

To end this section, we would like to remark also the case of Moreno-Bote et al.'s model, similarly as it was done above for Levelt's proposition IV. A first inspection of figure 4.7 clearly suggests that this model fails to replicate the sec-

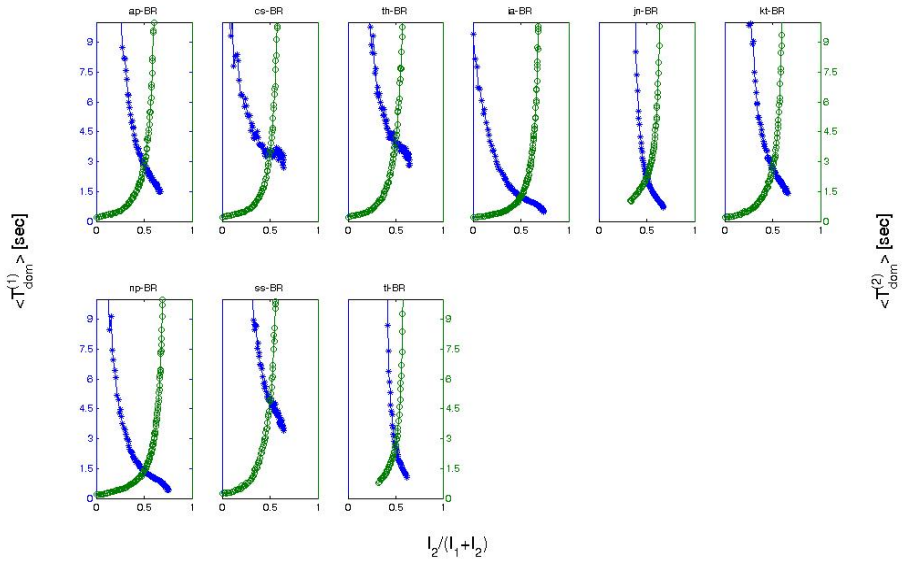

Figure 4.4: The average dominance time $< T_{dom} >$ for the BR percepts by using the Wilson (2003) model, is plotted as a function of the variable input strength I_2 (I_1 is kept fixed). A pair monotonic curves is found in most of the observers intercepting when no input bias exists, in agreement with Levelt's second proposition. Each observer is described by a unique 4D vector of parameters $(\overline{\beta}, \overline{\phi_a}, \overline{\tau_a}, \overline{\sigma_n})$, which is obtained by a (weighted) average 4.6 over the whole matching region found in subsection 3.4.

ond law enunciated by Levelt, for the BR subjects. Contrary to the expected behavior, the dominance of the two stimuli are almost equally affected by the unilateral manipulation of just one of them, being described by curves with the same trend, *i.e.*, a decreasing profile of the mean duration of their presence intervals on awareness when the variable input strength increases.

Unlike the case of the depression-LC model, for which no analytical or numerical study exists to date suggesting a possible solution for the unexpected behavior shown in figure 4.6, Moreno-Bote et al. (2010) has proposed a minor modification to the original model that may resolve the conflict between theory and experiment shown in the previous figure. Concretely, a preprocessing stage in their external inputs I_1 and I_2 of the two populations can be added, by replacing them by a *normalized* value:

$$I_i = s\frac{I_i}{I_i + I_j + I_{bg}}, \qquad i \neq j \tag{4.10}$$

where I_{bg} denotes a background or baseline activity, and s is a scaling coefficient introduced in order to produce similar alternation rates for the models with

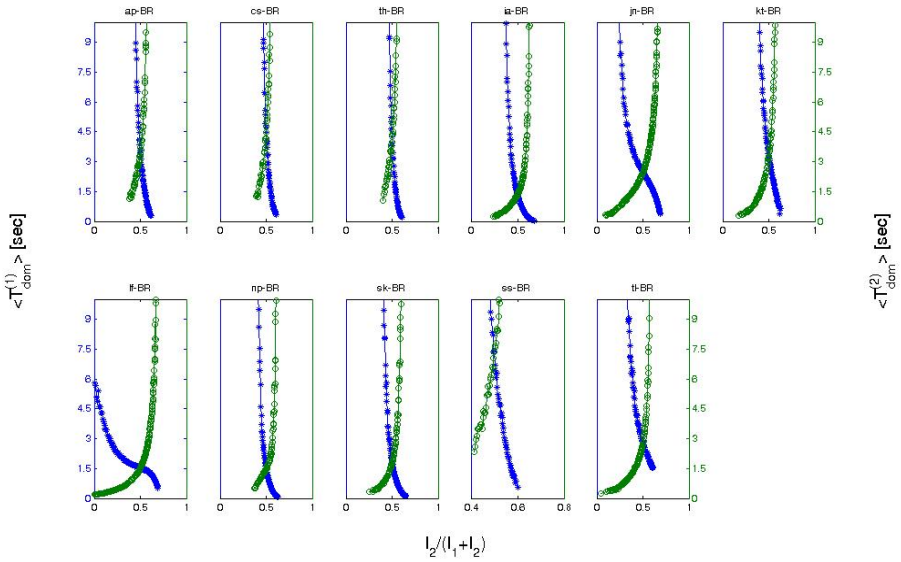

Figure 4.5: The average dominance time $< T_{dom} >$ for the BR percepts by using the non-linear LC-adaptation model, is plotted as a function of the variable input strength I_2 (I_1 is kept fixed). A pair of monotonic curves is found in most of the observers intercepting when no input bias exists ($I_1 = I_2$), in agreement with Levelt's second proposition. See text for more details.

and without input gain normalization. Consequently, the value of this scaling parameter should be irrelevant to the qualitative behavior of the model.

The numerical results of our second attempt to replicate Levelt's law II are shown in figure 4.8. Following Moreno-Bote et al. (2010), the values $s = 0.182$ and $I_{bg} = 0.01$ have been used for the scaling coefficient and the background current, respectively. It can be observed that now approximately six of eleven of the simulated results reasonably match the expected behavior. However, as in the case of the LC-depression model, further analysis is needed to uncover the dynamical factors determining the notable disagreement with the empirical behavior contained in the second proposition of Levelt.

4.3 Stochastic resonance in bistable perception

As it was shown in section 2.2, the experiments carried out by Pastukhov & Braun (2011) precise, in unprecedented detail, the statistics inherent to the phenomenon of bistable perception. The authors demonstrated the existence of subtle hysteresis effects in the perceptual traces of individuals, that were largely ignored in previous experimental studies. A couple of new observables, the correlation between a cumulative history and the dominance periods (c_H) and the time-scale

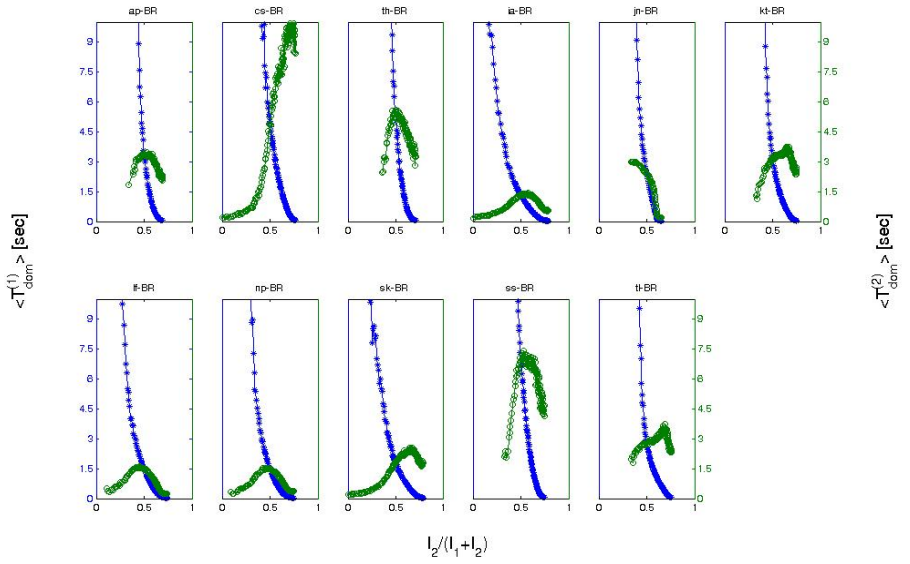

Figure 4.6: The average dominance time $< T_{dom} >$ for the BR percepts by using the LC-depression model, is plotted as a function of the variable input strength I_2 (I_1 is kept fixed). A pair of monotonic curves intercepting when no input bias exists ($I_1 = I_2$) is rarely found, in clear contradiction with Levelt's second proposition. The non-monotonic profile (green) obtained for the mean dominance duration $T^(2)_{dom}$ of the changing stimulus is clearly incorrect. See text for more details.

(τ_H) of such history, are proposed to described such history-dependence which has been hitherto unnoticed until today.

On the other hand, in section 3.1 of this thesis we have shown in the context of several prevailing models, that the data reported in Pastukhov & Braun (2011) strongly support the existence of some kind of noise-driven bistable system that could serve as neural substrate for the overall process. This contradicts the widely extended opinion that a limit cycle-based behavior, contaminated with noise Laing & Chow (2002); Wilson (2003), or some balance between fatiguing feedbacks (adaptation, depression) and noise Shpiro et al. (2007), may explain the phenomenon.

In this section we further stress the above idea, in the scope of the phenomenon of stochastic resonance Gammaitoni et al. (1998). A purely noise-driven bistable system typically shows a switching time given by the inverse of the escaping-rate r_k as stated by Kramer's formula (valid in a weak noise regime Kramers (1940)):

$$r_k = \frac{1}{\sqrt{2\pi}} \exp(-\Delta V/D), \tag{4.11}$$

where ΔV stands for the potential barrier that defines the currently system state

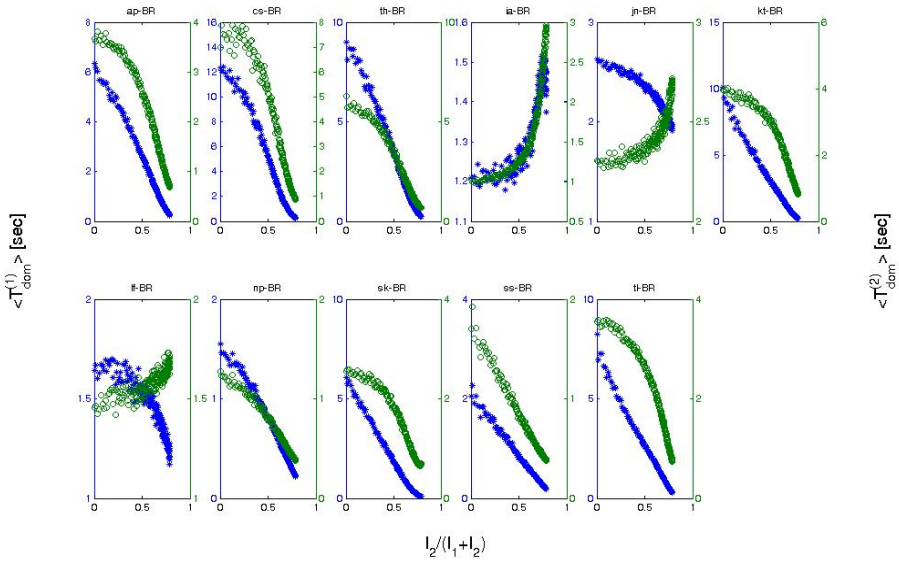

Figure 4.7: The average dominance time $< T_{dom} >$ for BR observers being modeled as a Moreno-Bote(2007) system each, is plotted as a function of the common input strength I_0. Two decreasing curves are typically found, in clear contradiction with Levelt's second proposition. See text for more details.

(*i.e.*, the well or basin of attraction of the occupied fixed point) and that may be overcome by the intervention of fluctuations with dispersion D. Then, a high degree of synchrony between the external modulation and the system dynamics is typically found when modulating frequency matches the value of the reversal rate given by equation (4.11).

In the models studied in this thesis, however, the effect of the noise is combined with the one of the deterministic factors (adaptation, depression, cross-inhibition) that dynamically change the associated energy landscape. Hence, equation (4.11) is not longer applicable. They could facilitate a transition by raising the currently active well until some saturation level is reached, but without the complete annihilation of the intrinsic bi-stability, as it was demonstrated in section 3.1. This must lead to the reduction of the potential barrier ΔV needed to hop to the still empty well, although just the noise would be able to complete a new transition Moreno-Bote et al. (2007).

Following the methodology reported by Kim et al. (2006) to investigate the presence of stochastic resonance in visual bistable perception, the two inputs parameters values, I_1 and I_2, were modulated in anti-phase by adding a square-wave signal of variable frequency value. As it was explained in subsection 2.2, this empirical procedure is expected to dynamically modify the depth of the two-well energy landscape characteristic of a bistable system, an analysis based on the

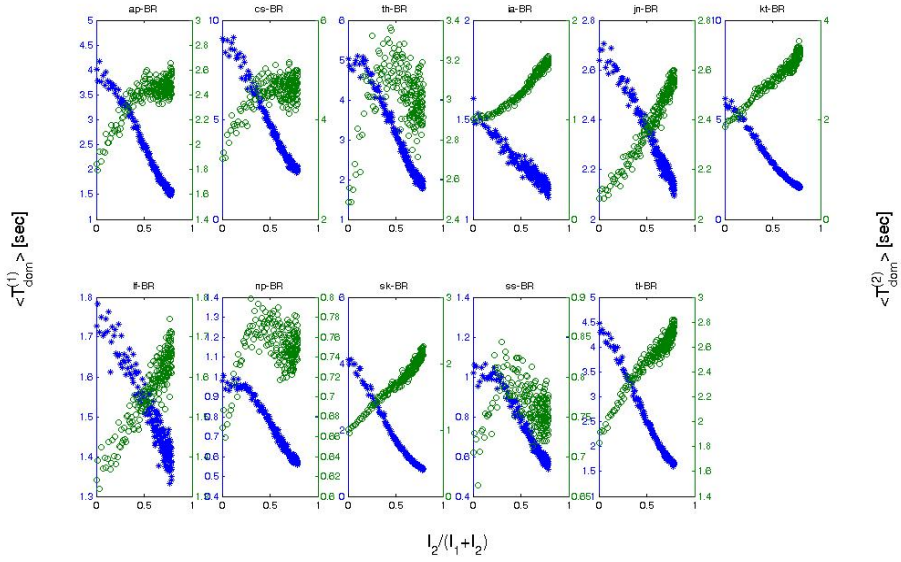

Figure 4.8: The average dominance time $< T_{dom} >$ for the BR percepts by using the Moreno-Bote et al. (2007) model with gain normalized currents (eq. (4.10)), is plotted as a function of the variable input strength I_2 (I_1 is kept fixed). A pair of monotonic curves is found in most of the observers intercepting when no input bias exists ($I_1 = I_2$), in agreement with Levelt's second proposition. See text for more details.

Levelt's second proposition. A high degree of synchrony between the perturbed system and the external signal could turn to be visible in a notable change in the probability distribution of dominance durations, when the frequency of the periodic perturbation matches the natural alternation rate (see Fig. 2.2). Then, a large concentration of the probability around the modulation half-period (HP) should be observed Gammaitoni et al. (1998).

Equivalently, such increment in synchrony between the system and the external signal should be visible in the coefficient of variation , C_v, of the new perceptual trace. Figures 4.9 and 4.10 show the results obtained with Wilson's and the LC-depression models, respectively. In both cases, a noise-dependence of the degree of synchrony between the external signal and the modulated system is observed. The effect can be seen in more detail by taking some planar slices at different modulation values ΔI (Figure 4.11).

Importantly, it is shown that a true stochastic resonance can be obtained in this kind of models by varying the noise dispersion instead of the modulating frequency as done in Kim et al. (2006). Moreover, the effect is not limited to a narrow range of modulation percent ΔI, but it exists for even very weak signals to which the system develops an extreme sensitivity. These results are in disagreement with the predictions made in the same work, based on the empirical

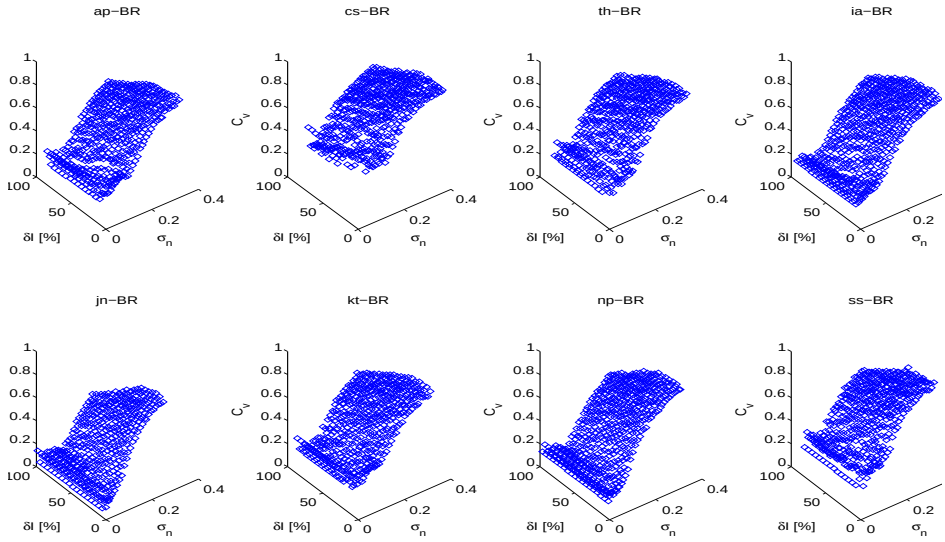

Figure 4.9: Variation coefficient (C_V) of T_{dom} variable when the system is perturbed by an external periodic signal: Wilson (2003) model. The presence of stochastic resonance is shown by a dip in the C_V surface along the σ_n dispersion axis. Notably, even for very weak modulation percents the phenomenon is observed.

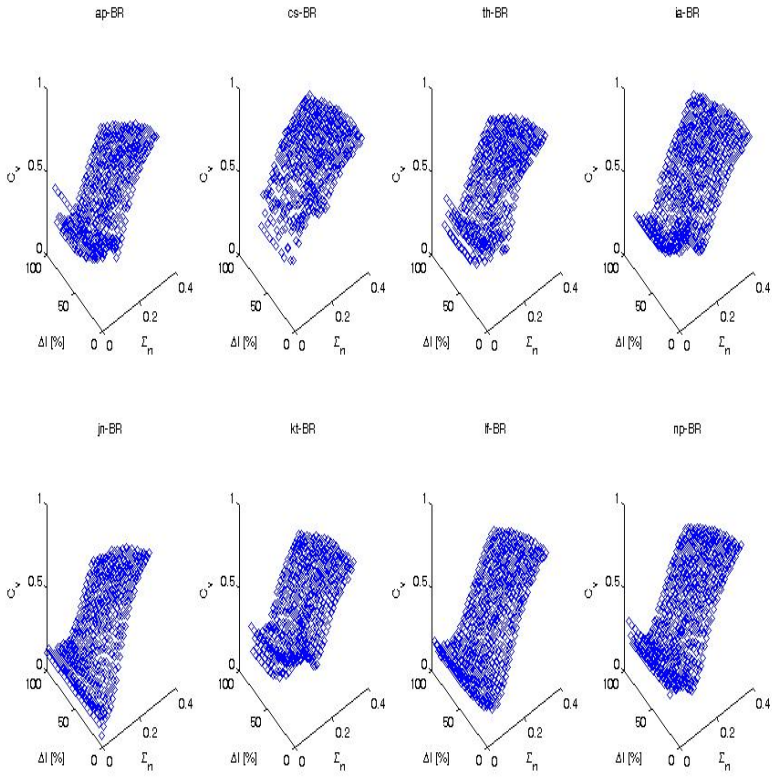

Figure 4.10: Variation coefficient (C_V) of T_{dom} variable when the system is perturbed by an external periodic signal: Depression-LC model. The presence of stochastic resonance is shown by a dip in the C_V surface along the σ_n dispersion axis. Remarkably, even for very weak modulation percents the phenomenon is observed.

finding that resonance can be obtained for $20\% - 30\%$ of external modulation when the frequency is used as control parameter.

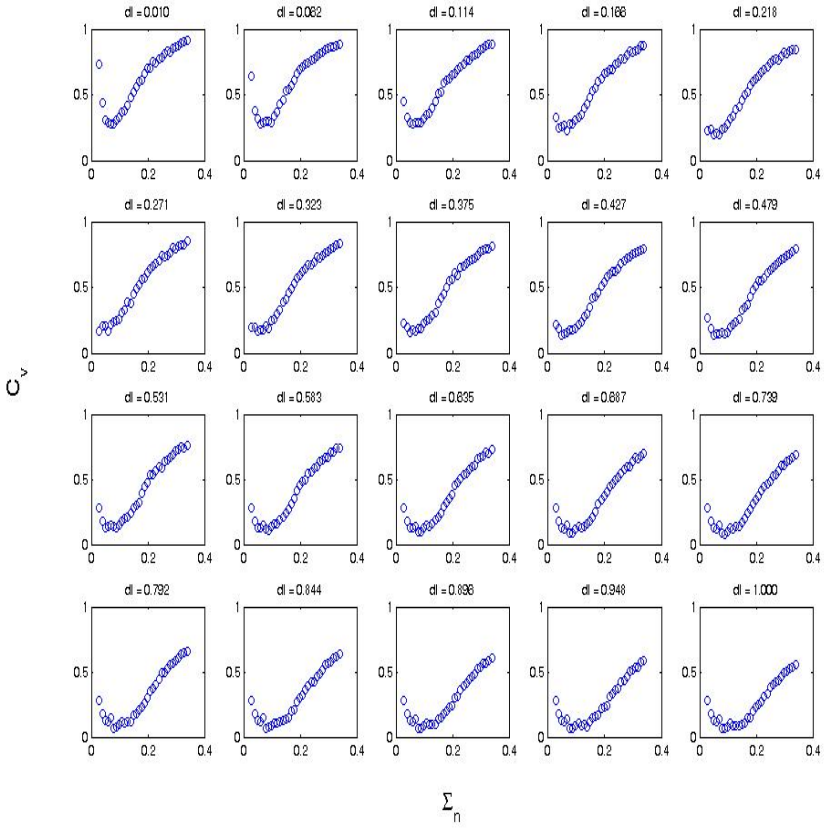

Figure 4.11: The profile of the variation coefficient (C_V) of T_{dom} at different modulation values ΔI, corresponding to the observer *np* of Figure 4.10. Notably, even for very weak modulation percents the phenomenon of stochastic resonance is observed.

5 Conclusions and future work

In this thesis, a set of some prevailing rate-based models for bistable perception have been considered in order to find the implications of the novel results reported in Pastukhov & Braun (2011). These authors have quantified not only salient aspects of bistable perception (mean and dispersion of dominance distributions), but also some hidden hysteresis effects ignored up to now. Their experimental report includes a classic binocular rivalry (BR) display (two orthogonal gratings) together with illusory rotating spheres created by a structure-from-motion or kinetic-depth-effect (KDE). The different characteristics of the models considered here allow to evaluate the biological plausibility of distinct explanations for the neural substrate underlying the phenomenon.

As a common feature, the models treated in the present work include two basic mechanisms: (i) a slow negative feedback that precludes a population from dominating forever (*e.g.*, spike-frequency adaptation or synaptic depression) and (ii) mutual inhibition between the two populations whose activities represent the competing percepts, that permit exclusive dominance. However, various new situations are considered, such as the presence of a slow fatigue process given by linear Shpiro et al. (2007) and nonlinear Curtu et al. (2008) functions, complex architectures Moreno-Bote et al. (2007) and a non-symmetric neuronal activation functions Wilson (2003).

Extensive computational simulations of these models rigorously demonstrate that the history-dependence of the perceptual process shown by Pastukhov & Braun (2011), effectively constrains the region of the parameter space able to replicate the empirical data. More precisely, that just small regions residing inside a *bistable* or *two-attractor* region of the whole parameter space are actually suitable to reproduce the experimental results, both for BR and KDE displays. Remarkably, the results remain valid across all the different classes of models considered, regardless the details of the neuronal implementation.

Such dynamical regime is typically defined by a combination of a high inhibition and a low adaptation level, where the dominance switches are provoked exclusively by stochastic fluctuations, and would be absent without it. Our results strongly support the novel hypothesis that neural noise could play an essential role in bistable perception Brascamp et al. (2006), and clarifies the previous modeling study by Shpiro et al. (2009) who proposed a balance between deterministic (fatigue) and stochastic forces to explain the observed timing of perceptual reversals.

The biological plausibility of the parameter region found for each of the models considered, is further tested with respect to the widely known Levelt's propositions, a study in the same spirit of the recent work by Curtu et al. (2008); Seely & Chow (2011). To that end, we make use of weighted averages across the parameter regions computed for each subject in the first part of this thesis. The mentioned weighted sum algorithm proposed here accounts for the relevance of each parameter point, according to its capability to replicate the whole set of experimental data reported by Pastukhov & Braun (2011), and constitutes an important improvement to the methodology proposed by Shpiro et al. (2007).

Our analysis of Levelt's propositions is limited to the observers presented with the pair of orthogonal sinusoidal gratings (BR display). KDE subjects are not taken into account, due to the larger uncertainty present in their reported measurements. Computational simulations of the rate-based models considered allow to discover how different neuronal mechanisms clearly differ in their suitability to replicate additional behavioral data. For instance, models with a slow fatiguing process given by spike-frequency adaptation Wilson (2003); Shpiro et al. (2007), no matter if they are being described by linear Shpiro et al. (2007) or nonlinear Curtu et al. (2008)) functions of the activity, replicate quite well Levelt's second law.

Conversely, a notable discrepancy between model and empirical results is found when such negative feedback is described as a long-term depression affecting the synapses between the competing neurons representing the two alternative interpretations Laing & Chow (2002); Shpiro et al. (2007). Then, only realistic (increasing) profiles for the mean dominance duration as a function of an unilateral manipulation is obtained for the unchanged stimulus, while unexpected non-monotonic curves corresponded to the varying stimulus.

Future directions of research could include an analytical treatment of Levelt's second law aimed to uncover the dynamical factors that determine an unexpected non-monotonic profile in some of the models treated in this thesis. Such a study should concern the suitability of the depression-LC model, in the same spirit of the existing work that has been done in the scope of the fourth proposition of Levelt Seely & Chow (2011); Curtu et al. (2008). Obtaining the corresponding bifurcation diagrams of the free-noise system when only one input strength changes while the other remains constant may be an interesting starting-point, similarly to the analyses made when the two inputs variables simultaneously vary (Levelt's fourth law).

The interesting proposal by Moreno-Bote et al. (2007) is also analyzed with respect to Levelt's law II, a study that confirmed that the absence of a preprocessing of the inputs could be a shortcoming in the original version of this model. Their simple replacement by gain normalized currents Moreno-Bote et al. (2010) easily makes the model show the expected behavior. Concerning Levelt's fourth proposition, however, this theoretical proposal (even in its original form) was the only one among the five models treated that showed realistic behaviors. Specifically, all but one of eleven subjects whose behavioral data reported by Pastukhov & Braun (2011) were successfully fitted by this model, showed a decreasing profile of the average dominance interval, when the input strength of the two inputs

were simultaneously varied. In contrast, a non-monotonic curve was found in all the other cases, as previously predicted for noise-driven regimes in the excellent study of Shpiro et al. (2007).

The present work finishes with a study about the capability of the mentioned models to reproduce more empirical data related with bistable perception. That is, the resonance effects happening when varying external *frequencies*, as shown by Kim et al. (2006). Importantly, a resonance respect to the noise dispersion (*i.e.*, a true *stochastic* resonance) is clearly demonstrated here for the first time, in contrast to the methodology followed in Kim et al. (2006) where a maximal degree of synchrony is shown as a function of the modulating frequency. It is also considered their estimations of noise dispersion ($20 - 30\%$ of the input) and its locus (adaptation variables), by demonstrating that increased sensitivity to even weak signals of the order of less than 10% can be obtained with the models considered, with the noise variable simply entering as part of the net input feeding the neuron.

Interesting lines of research regarding the work by Kim et al. (2006) could include an analytical treatment of the problem, based on the reduction of dimensions (usually four dynamical variables, plus two time-dependent noise sources). Besides, more numerical work may be addressed to study the role of the time-correlation of the colored noise in stochastic resonance, in the scope of bistable perception models working in a noise-driven regime. Furthermore, it is important to understand how our numerical results depend on the temporal structure of the fluctuations. Namely, different degree of noise color (correlation) could create further history dependence in the simulated perceptual trace, and finally change our conclusions respect to the fit of the behavioral data reported in Pastukhov & Braun (2011).

Bibliography

Andrews, T. J. & Purves, D. (1997). Similarities in normal and binocularly rivalrous viewing. *Proceedings of the National Academy of Sciences of the United States of America*, 94(18):9905–9908.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychological review*, 96(1):145–167.

Blake, R. & Fox, R. (1974). Adaptation to invisible gratings and the site of binocular rivalry suppression. *Nature*, 249(456):488–490.

Blake, R., Westendorf, D. H., & Overton, R. (1980). What is suppressed during binocular rivalry? *Perception*, 9(2):223–231.

Borsellino, A., De Marco, A., Allazetta, A., Rinesi, S., & Bartolini, B. (1972). Reversal time distribution in the perception of visual ambiguous stimuli. 10(3):139–144.

Bossink, C. J., Stalmeier, P. F., & Weert, C. M. D. (1993). A test of levelt's second proposition for binocular rivalry. *Vision Research*, 33:1413–1419.

Brascamp, J. W., Ee, R. V., Noest, A. J., Jacobs, R. H. A. H., & van den Berg, A. V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *Journal of Vision*, 6(11):1244–1256.

Brascamp, J. W., Knapen, T. H., Kanai, R., Noest, A. J., van Ee, R., & van den Berg, A. V. (2008). Multi-timescale perceptual history resolves visual ambiguity. *PLoS ONE*, 3(1):e1497.

Brascamp, J. W., Knapen, T. H. J., Kanai, R., van Ee, R., & van den Berg, A. V. (2007). Flash suppression and flash facilitation in binocular rivalry. *Journal of Vision*, 7(12):1–12.

Brascamp, J. W., van Ee, R., Pestman, W. R., & van Den Berg, A. V. (2005). Distributions of alternation rates in various forms of bistable perception. *Journal of Vision*, 5(4):287–298.

Brown, R. J. & Norcia, A. M. (1997). A method for investigating binocular rivalry in real-time with the steady-state vep. *Vision Research*, 37:2401–2408.

Carter, O. L. & Pettigrew, J. D. (2003). A common oscillator for perceptual rivalries? *Perception*, 32(3):295–305.

Cobb, W. A., Morton, H. B., & Egglinger, G. (1967). Cerebral potentials evoked by pattern reversal and their suppression in visual rivalry. *Nature*, 216:1123–1125.

Cosmelli, D. & Thompson, E. (2007). Mountains and valleys: binocular rivalry and the flow of experience. *Consciousness and Cognition*, 16(3):623–641.

Crick, F. (1996). Visual perception: rivalry and consciousness. *Nature*, 379:485–486.

Crick, F. & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2:263–275.

Crick, F. & Koch, C. (1998). Consciousness and neuroscience. *Cereb. Cortex*, 8:97–107.

Curtu, R., Shpiro, A., Rubin, N., & Rinzel, J. (2008). Mechanisms for frequency control in neuronal competition models. *SIAM journal on applied dynamical systems*, (2):609–649. ISSN 1536-0040.

Dayan, P. (1998). A hierarchical model of binocular rivalry. *Neural Computation*, 10:1119–1135.

DuTour, E. F. (1760). Discussion d'une question d'optique. *Mémoire de mathématique et de physique présentés par Divers Savants*, 3:514–530.

Fox, R. & Check, R. (1968). Detection of motion during binocular rivalry suppression. *Journal of Experimental Psychology*, 78:388–395.

Fox, R. & Check, R. (1972). Independence between binocular rivalry suppression duration and magnitude of suppression. *Journal of Experimental Psychology*, 93(2):238–289.

Fox, R. & Herrmann, J. (1967). Stochastic properties of binocular rivalry alternations. *Percept. Psychophys.*, 2(9):432–446.

Gammaitoni, L., Hanggi, P., Jung, P., & Marchesoni, F. (1998). Stochastic resonance. *Rev. Mod. Phys.*, 70:223–287.

García, P., Deco, G., Pastukhov, A., Braun, J., & Guillamon, T. (2009). Dynamical insights on the history-dependence during continuous presentation of rivaling stimuli. Dins *BMC Neuroscience*, volum 10, p. P364. Eighteenth Annual Computational Neuroscience Meeting. Conference abstract, http://www.biomedcentral.com/1471-2202/10/S1/P364.

García-Rodríguez, P. E., Deco, G., Pastukhov, A., Braun, J., & Guillamon, T. (2009). History-dependence in bistable perception: dynamical insights. Dins *Workshop and Advanced Course on Deterministic and Stochastic Modeling in Computational Neuroscience and other Biological topics*, ps. 138–139.

Grossberg, S., Yazdanbakhsh, A., Cao, Y., & Swaminathan, G. (2008). How does binocular rivalry emerge from cortical mechanisms of 3-d vision? *Vision Research*, 48:2232–2250.

J. Hertz, A. K. & Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Adison-Wesley.

James, W. (1981). *The principles of psychology*. Cambridge, MA: Harvard University Press.

Kalarickal & Marshall (2000). Neural model of temporal and stochastic properties of binocular rivalry. *Neurocomputing*, 32–33:843–853.

Kim, Y.-J., Grabowecky, M., & Suzuki, S. (2006). Stochastic resonance in binocular rivalry. *Vision Research.*, 46(3):392–406.

Klink, P., van Ee, R., & van Wezel, R. (2008). General validity of levelt's propositions reveals common computational mechanisms for visual rivalry. *PLoS ONE*, 3(10):e3473.

Kohler, W. & Wallach, H. (1944). Figural after-effects. an investigation of visual processes. *Proceedings of the American Philosophical Society*, 88(4):269–357.

Kovacs, I., Papathomas, T. V., Yang, M., & Feher, A. (1996). When the brain changes its mind: Interocular grouping during binocular rivalry. *Proc. Natl. Acad. Sci. USA*, 93:15508–15511.

Kramers, H. A. (1940). Brownian motion in a field of force and diffusion model of chemical reactions. *Physica*, 7:284.

Laing, C. R. & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *J. Comput. Neurosci.*, 12:39–53.

Lansing, R. W. (1964). Electroencephalographic correlates of binocular rivalry in man. *Science*, 146:1325–1327.

Lee, S.-H. & Blake, R. (1999). Rival ideas about binocular rivalry. *Vision Research*, 39:1447–1454.

Lehky, S. R. (1988). An astable multivibrator model of binocular rivalry. *Perception*, 17:215–228.

Lehky, S. R. (1995). Binocular rivalry is not chaotic. *Proc. R. Soc. Lond. Ser. B*, 259:71–76.

Lehky, S. R. & Maunsell, H. R. (1996). No binocular rivalry in the lgn of alert macaque monkeys. *Vision Research*, 36(9):1225–1234.

Lehmkuhle, S. W. & Fox, R. (1975). Effect of binocular rivalry suppression on the motion aftereffect. *Vision Research*, 15(7):855–859.

Leopold, D. A. & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkey's percepts during binocular rivalry. *Nature*, 379:549–553.

Leopold, D. A. & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences*, 3(7):254–264.

Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature neuroscience*, 5(6):605–609.

Levelt, W. J. M. (1965). *On binocular rivalry*. Institute for Perception RVO-TNO: Soesterberg (The Netherlands).

Levelt, W. J. M. (1966). The alternation process in binocular rivalry. *British J. Psychol.*, 57(3-4):225–238.

Levelt, W. J. M. (1967). Note on the distribution of dominance times in binocular rivalry. *British J. Psychol.*, 58(1):143–145.

Liu, L., Tyler, C. W., & Schor, C. M. (1992). Failure of rivalry at low contrast: evidence of a suprathreshold binocular summation process. *Vision Research*, 32(8):1471–1479.

Logothetis, N. K. (1998). Single units and conscious vision. *Phil. Trans. R. Soc. London B: Biol. Sciences*, 353:1801–1818.

Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature*, 380:621–624.

Logothetis, N. K. & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science*, 245:761–763.

Long, G. M. & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychological Bulletin*, 130(5):748–768.

Lumer, E. D. (1998). A neural model of binocular integration and rivalry based on the coordination of action-potential timing in primary visual cortex. *Cerebral Cortex*, 8:553–561.

Lumer, E. D., Friston, K., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, 280:1930–1934.

Maier, A., Wilke, M., Logothetis, N. K., & Leopold, D. A. (2003). Perception of temporally interleaved ambiguous patterns. *Nature neuroscience*, 13(13):1076–1085.

Martí i Ortega, D. (2008). *Neural stochastic dynamics of perceptual decision making*. Tesi Doctoral, Universitat Pompeu Fabra.

McCormick, D. A. & Williamson, A. (1989). Convergence and divergence of neurotransmitter action in human cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 86(20):8098–102.

Melnikov, V. I. (1993). Schmitt trigger: A solvable model of stochastic resonance. *Physical Review E*, 48:2481–2489.

Moreno-Bote, R., Rinzel, J., & Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *J Neurophysiol*, 98(3):1125–39.

Moreno-Bote, R., Shpiro, A., Rinzel, J., & Rubin, N. (2010). Alternation rate in perceptual bistability is maximal at and symmetric around equi-dominance. *Journal of Vision*, 10–11(1):1–18.

Mueller, T. J. (1990). A physiological model of binocular rivalry. *Visual Neuroscience*, 4:63–73.

Mueller, T. J. & Blake, R. (1989). A fresh look at the temporal dynamics of binocular rivalry. *Biol. Cybern.*, 61:223–232.

O'Shea, R. P. (1999). Translation of DuTour (1760). Available at `http://psy.otago.ac.nz:800/r_oshea/dutour60.html`.

O'Shea, R. P. & Crassini, B. (1981). Interocular transfer of the motion after-effect is not reduced by binocular rivalry. *Vision Research*, 21(6):801–804.

Pastukhov, A. & Braun, J. (2011). Cumulative history quantifies the role of neural adaptation in multistable perception. *Journal of Vision*, 11(10)(12):1–10.

Pastukhov, A., Braun, J., García, P., & Deco, G. (2009a). Multistable perception is never memoryless. Dins *Perception*, volum 38, p. 184. Thirty-second European Conference on Visual Perception. Conference abstract, `http://www.perceptionweb.com/abstract.cgi?id=v090827`.

Pastukhov, A., Braun, J., García-Rodríguez, P. E., & Deco, G. (2009b). History dependence in multistable perception highlights the role of noise in visual perception. Dins *13th Annual Meeting of the Association for the Scientific Study of Consciousness*, ps. 212–213. Conference abstract, `http://kobi.nat.uni-magdeburg.de/node/423`.

Pastukhov, A., García-Rodríguez, P., Haenicke, J., Guillamon, A., Deco, G., & Braun, J. (2011). Stable but sensitive: multi-stable perception arbitrates to the exploitation-exploration dilemma. Dins *Frontiers in Computational Neuroscience*. Computational Neuroscience & Neurotechnology Bernstein Conference & Neurex Annual Meeting. Conference abstract, `http://www.frontiersin.org/10.3389/conf.fncom.2011.53.00145/event_abstract`.

Pastukhov, A., García-Rodríguez, P. E., Haenicke, J., Guillamon, A., Deco, G., & Braun, J. (2012). Multi-stable perception balances stability and sensitivity. Under review at *PLoS Computational Biology*.

Pearson, J. & Clifford, C. G. (2004). Determinants of visual awareness following interruptions during rivalry. *Journal of Vision*, 4(3):196–202.

Pikovsky, A. S. & Kurths, J. (1997). Coherence resonance in a noise-driven excitable system. *Physical Review Letters*, 78(5):775–778.

Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience*, 3(11):1153–1159.

Rodríguez, P. G., Pastukhov, A., Deco, G., Braun, J., Guillamon, A., & Haenicke, J. (2010). Variability of dominance phases in bistable perception: history-dependence and noisy dynamics. Dins *Frontiers in Computational Neuroscience*. Bernstein Conference on Computational Neuroscience. Conference abstract, http://www.frontiersin.org/10.3389/conf.fncom.2010.51.00083/event_abstract.

Rubin, N. & Hupé, J. M. (2004). Dynamics of perceptual bistability: plaids and binocular rivalry compared. In: Binocular Rivalry, edited by Alais D., Blake R. Cambridge, MA: MIT Press.

Sanchez-Vives, M. V., Nowak, L. G., & McCormick, D. A. (2000). Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. *J Neurosci*, 20(11):4286–4299.

Seely, J. & Chow, C. C. (2011). The role of mutual inhibition in binocular rivalry. *J Neurophysiol.*, 106(5):2136–2150.

Sengpiel, F. (1997). Binocular rivalry: ambiguities resolved. *Current Biology*, 7:R447–R450.

Sengpiel, F., Blakemore, C., & Harrad, R. (1995). Interocular suppression in the primary visual cortex: a possible neural basis of binocular rivalry. *Vision Research*, 35(2):179–195.

Sheinberg, D. L. & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proc. Natl. Acad. Sci. USA*, 94:3408–3413.

Shpiro, A., Curtu, R., Rinzel, J., & Rubin, N. (2007). Dynamical characteristics common to neuronal competition models. *Journal of neurophysiology*, (1):462–73.

Shpiro, A., Moreno-Bote, R., Rubin, N., & Rinzel, J. (2009). Balance between noise and adaptation in competition models of perceptual bistability. *Journal of computational neuroscience*, 27(1):37–54.

Srinivasan, R., Russell, D. P., Edelman, G. M., & Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *J. Neurosci.*, 19(13):5435–5448.

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, 13(7):310–318.

Stollenwerk, L. & Bode, M. (2003). Lateral neural model of binocular rivalry. *Neural Computation*, 15:2863–2882.

Sugie, N. (1982). Neural models of brightness perception and retinal rivalry in binocular vision. *Biological Cybernetics*, 1:13–21.

Tong, F. (2001). Competing theories of binocular rivalry: a possible resolution. *Brain and Mind*, 2:55–83.

Tong, F. & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, 411:195–199.

Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends Cogn. Neurosci.*, 10(11):502–511.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, 21:753–759.

Tononi, G., Srinivasan, R., Russell, D. P., & Edelman, G. M. (1998). Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *Proc. Natl. Acad. Sci. USA*, 95:3198–3203.

Toppino, T. C. & Long, G. M. (1987). Selective adaptation with reversible figures: Don't change that channel. *Perception and Psychophysics*, 42:37–48.

van Boxtel, J. J. A., van Ee, R., & Erkelens, C. J. (2007). Dichoptic masking and binocular rivalry share common perceptual dynamics. *Journal of Vision*, 7(14)(3):1–11.

van Ee, R. (2009). Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: evidence from serial correlations in perceptual bistability. *Journal of the Optical Society of America A*, 26(12):2612–2622.

Wade, N. J. & Wenderoth, P. (1978). The influence of colour and contour rivalry on the magnitude of the tilt after-effect. *Vision Research*, 18(7):827–835.

Wales, R. & Fox, R. (1970). Increment detection thresholds during binocular rivalry suppression. *Perception and Psychophysics*, 8(2):90–94.

Walker, P. (1975). Stochastic properties of binocular rivalry alternations. *Perception and Psychophysics*, 18(6):467–473.

Wheatstone (1838). On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Phil. Trans. R. Soc. London*, 128:371–394.

Wilson, H. R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proc. Natl. Acad. Sci. USA*, 100:14499–14503.

Wilson, H. R. (2007). Minimal physiological conditions for binocular rivalry and rivalry memory. *Vision Research*, 47:2741–2750.

Zhou, Y. H., Gao, J. B., White, K. D., Merk, I., & Yao, K. (2004). Perceptual dominance time distributions in multistable visual perception. *Biological Cybernetics*, 90(4):256–263.

Colofo

Aquesta tesi s'ha escrit en LaTeX fent servir l'editor `kile`.