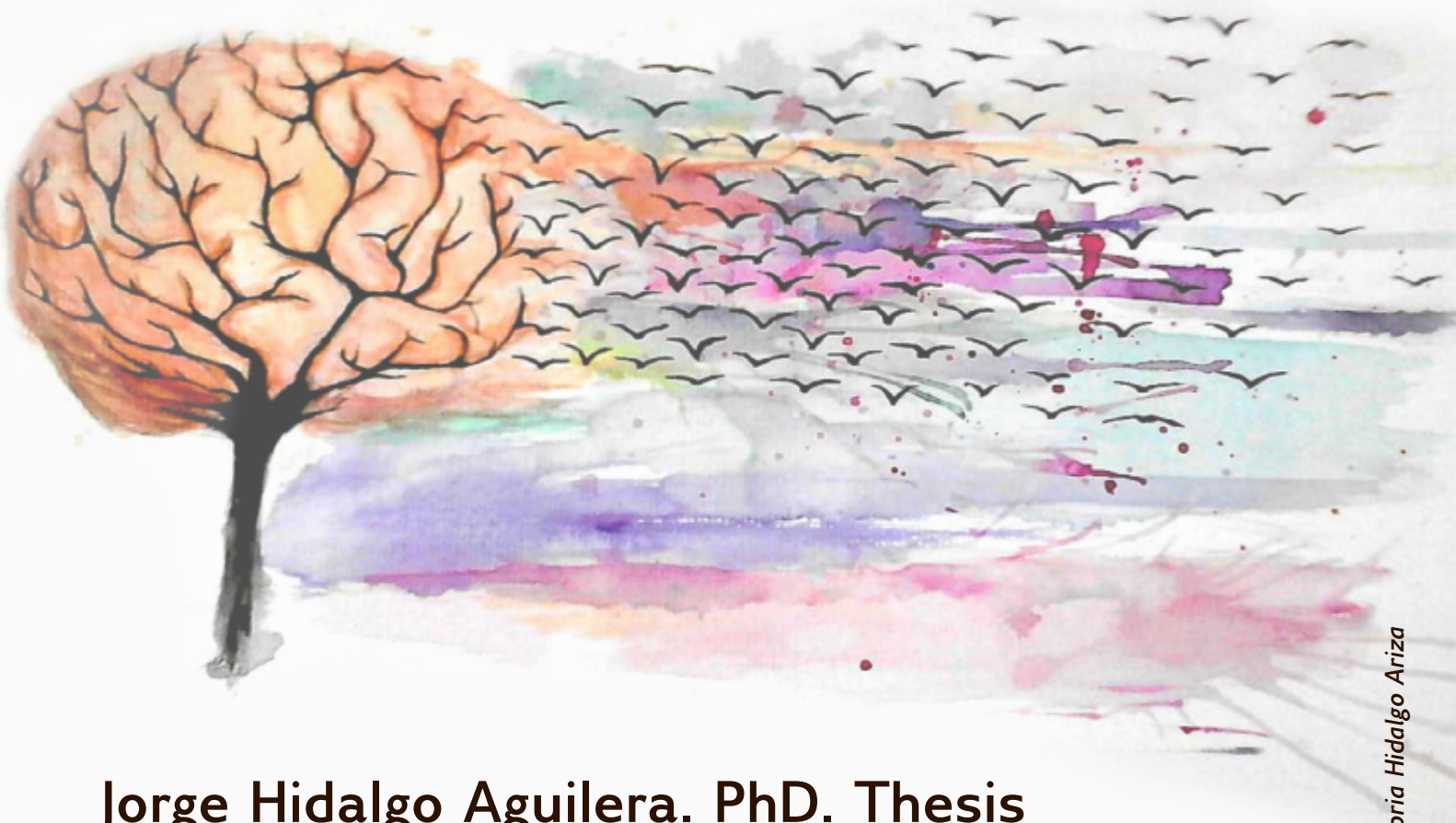
Universidad de Granada

Escuela de doctorado
Fis y Mat
Programa de posgrado en Física y Matemáticas

Novel mechanisms for phase transitions and self-organization in living systems



Jorge Hidalgo Aguilera. PhD. Thesis

Advisor: Miguel Á. Muñoz Martínez

Cover: Gloria Hidalgo Ariza

Editor: Universidad de Granada. Tesis Doctorales
Autor: Jorge Hidalgo Aguilera
ISBN: 978-84-9125-014-2
URI: http://hdl.handle.net/10481/39636

Universidad de Granada

Facultad de Ciencias

Novel mechanisms for phase transitions and self-organization in living systems

Thesis submitted by Jorge Hidalgo Aguilera
for the degree of Doctor of Philosophy

2014

Departamento de Electromagnetismo y Física de la Materia
e Instituto Carlos I de Física Teórica y Computacional

Universidad de Granada

Facultad de Ciencias

Nuevos mecanismos para transiciones de fase y auto-organización en sistemas vivos

Tesis presentada por Jorge Hidalgo Aguilera
para obtener el grado de Doctor
por el Programa de Doctorado en Fisica y Matematicas

2014

Departamento de Electromagnetismo y Física de la Materia
e Instituto Carlos I de Física Teórica y Computacional

ugr | Universidad de Granada

El doctorando Jorge Hidalgo Aguilera y el director de la tesis D. Miguel Ángel Muñoz Martínez, catedrático de Universidad,

GARANTIZAMOS: al firmar esta tesis doctoral, *Novel mechanisms for phase transitions and self-organization in living systems* ("Nuevos mecanismos para transiciones de fase y auto-organización en sistemas vivos"), que el trabajo ha sido realizado por el doctorando bajo la dirección del director de la tesis. Y hasta donde nuestro conocimiento alcanza, en la realización del trabajo se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones, así como que el doctorando ha disfrutado de una estancia en el extranjero, durante un periodo de tres meses, en la *School of Physics and Astronomy* de la Universidad de Manchester (Reino Unido).

Granada, a 29 de octubre de 2014.

Director de la tesis: Doctorando:

Fdo: Miguel Ángel Muñoz Martínez Fdo: Jorge Hidalgo Aguilera

A mis padres.

Que me han dado todo,
y los tomates, las patatas, las croquetas, el salmorejo...

Contents

1 Introduction **5**

1.1 Spotlight on phase transitions 6

1.2 Operating at a narrow line 8

1.3 Preview-summary 9

Introducción **13**

2 The emergence of criticality in living systems **21**

2.1 *'Are biological system poised at criticality?'* 21

2.2 Mathematical framework: Mapping the external world 23

Example: Ising parametrizations 26

2.2.1 Optimal mapping and Kullback-Leibler divergence 27

Numerical computation in mean-field Ising parametrizations . . 29

2.2.2 Fisher information and criticality 29

2.3 Computational models 31

2.3.1 Evolutionary model 32

Heterogeneity of the environment 36

Effective criticality of the environment 38

2.3.2 Co-evolutionary model 39

Alternative parametrizations 43

Complex internal networked topologies 43

Heterogeneity of the community 46

Multiple body interactions 47

2.3.3 Adaptive dynamics 49

Adaptive model . . . 49
Co-adaptive model . . . 50
2.4 Analytical results . . . 52
2.4.1 External heterogeneous environment . . . 53
2.4.2 Self-organized environment . . . 54
2.5 Chapter summary . . . 58

3 Bet-hedging in population dynamics 61
3.1 *'Bet hedging or not?'* . . . 61
3.2 Contact Process with hybrid dynamics . . . 62
3.3 Phases of the pure strategies . . . 65
3.4 Extinction times and Parrondo's paradox . . . 66
3.5 Dependence on external/environmental variability . . . 70
3.5.1 Analytics . . . 73
3.6 Dependence on internal/demographic variability . . . 73
3.6.1 Analytics . . . 74
3.7 Effect of temporal correlations . . . 77
3.8 Heterogeneous vs. homogeneous populations . . . 79
3.9 Chapter summary . . . 82

4 Mixed dispersal strategies 85
4.1 *'Diversity in times of adversity'* . . . 85
4.2 Simplistic model with inbreeding depression . . . 86
4.3 Phases of the pure strategies . . . 89
4.4 Benefits of the hybrid strategy . . . 89
4.4.1 Relative gain enhancement of hybrid strategies . . . 92
4.4.2 Reduction of the mean extinction time . . . 92
4.5 Evolutionary stable strategy . . . 94
4.6 Chapter summary . . . 97

5 Stochastic amplification in neural dynamics 99
5.1 *'Rythms of the brain'* . . . 99
5.1.1 Up and Down states . . . 101
5.2 Synaptic dynamics . . . 104

5.3 Models for Up and Down states 106
5.3.1 Model A: Minimal coarse-grained model 106
5.3.2 Model B: Spiking-neuron network model 109
Effective current contribution 112
Self-consistent solution 115
5.4 Stochastic amplification of fluctuations (SAF) 116
5.4.1 SAF in Model A 121
5.4.2 SAF in Model B 123
5.4.3 SAF as a collective phenomenon (Model B) 129
5.4.4 Robustness of the mechanism: effect of the synaptic facilitation 131
5.4.5 Robustness of the mechanism: SAF in a network of excitatory-inhibitory neurons 133
5.5 Chapter summary 135

6 Conclusions **137**

Conclusiones **141**

A Approach to Sanov's theorem **145**

B Fisher information and Cramérs-Rao inequality **148**

C Discrete time calculation for the Hybrid Contact Process **151**

D Conditions for stochastic amplification **153**

Bibliography **155**

Chapter 1

Introduction

As an non-experienced *complexsystemist*, any PhD. student may have been got involved, eventually, into a tense debate with other non-complexsystemist colleagues about the role of statistical physics in other disciplines, such as ecology, sociology, neuroscience or economics. The main topic focuses, usually, on the question: "why do you still call it *physics*?"

At a certain point, one develops a strong feeling about it, and generates the automatic answer that, of course, "physical systems can represent any system in nature susceptible to be modeled mathematically, in order to give an approximate description of it, as well as to make right predictions about its behavior". Consequently, complex systems in nature fit quite well into this definition. However, at a *second* certain point, another opinion arises: during many years, different scientific fields were strictly separated, but today, with the development of interdisciplinary sciences, what we are experiencing might be understood as the return – with a new look– of Natural Philosophy; or just *science*, as we used to called it in the elementary school.

In this context, it is particularly interesting how, for the same system, different "levels of knowledge" arise at each scale, that sometimes cannot be rationalized as an effective description of the precedent one. To give a hypothetical example, let imagine that computational technology would have never been developed. In this peculiar scenario, one day, an advanced civilization appears and we are granted with one single supercomputer, with no instructions about its functioning. Most probably, aimed at understanding how it works, we would start by unraveling the microscopic dynamics of

its tiny components. However, even in the case we reached to model and predict with perfect accuracy the behavior of each part of the computer, we would still be missing an additional and essential step to understand, for instance, what a program does or what software in general really means.

In a similar way, if we want to understand our brain, we need to rationalize the dynamics at the single neuron level, but this is not enough. If we want to understand the robustness of an ecological system, it is not sufficient to know which species prey upon other species. If we want to prevent a global epidemic outbreak, we need more than explaining each person-to-person infection. We need something more, because *More* is, actually, *different*.

1.1 Spotlight on phase transitions

Within the context of complex systems and collective phenomena, phase transitions have aroused a particular interest[127, 16]. Since school, we have been told that matter can be found at three different states: solid, liquid, and gas. Moreover, we learned that matter can change its state by undergoing a *phase transition*, which is actually a very common phenomenon, occurring, for instance, every time we put ice in a drink or cook a tasty soup. Roughly speaking, we say that a system undergoes a phase transition if it suddenly experiences a qualitative change on its global state. The archetypal example corresponds to the phase transition of water, but many others can be found in nature (see fig. 1.1). Specifically, the phase is characterized by the so-called *order* parameter (for instance the density), which changes by tuning the –in the statistical mechanics jargon– *control* parameter (as the temperature, pressure, etc.).

Among phase transitions, we distinguish two main types: discontinuous, or *first-order*, and continuous, also called *second-order*, phase transitions. The first kind corresponds to the case in which change is performed discontinuously and it involves an extra cost (latent heat), which is given to or received from the system. The already cited transitions of water under ordinary conditions are some examples of this. On the other hand, continuous phase transitions do not require any additional cost, and change is immediate and progressive. In this context, we find the spontaneous magnetization occurring on a piece of iron below 770 °C.

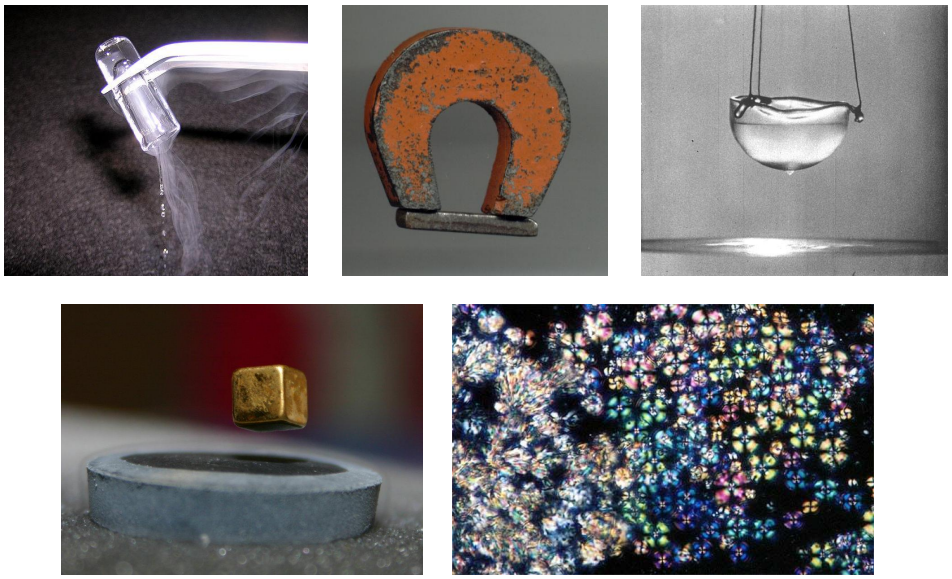


Figure 1.1: **Phase transitions are ubiquitous in nature:** From top left to bottom right: **1)** A small piece of rapidly melting argon ice simultaneously shows the transitions from solid to liquid to gas; **2)** Example of ferromagnetic material (Alnico alloy) at ordinary temperature; **3)** Liquid helium at the superfluid phase, creeping up the wall of the cup and falling down; **4)** Magnet levitating above a superconductor, cold down with liquid nitrogen; **5)** Phase transition between nematic (left) to smectic A (right) in a liquid crystal. *Source*: Wikipedia (Creative-Commons license).

The narrow line separating both phases displays many intriguing phenomena, being this one of the principal reasons why statistical physicists enjoy so much their study. For discontinuous transitions, two radically different phases meet at the separating point, therefore it is expected that both may *coexist* under certain conditions. The triple point of water constitutes an extreme case in which gas, liquid and solid water perfectly do it. In parallel, continuous phase transitions are characterized by a *critical* point separating both regimes; just at this point, systems exhibit a plethora of very interesting features [127, 16], such as scale invariance on its structure, large scale correlations and maximum susceptibility to external perturbations. The great lesson learned from statistical mechanics is that, even though the elementary constituents are very simple, criticality and scale invariance emerge as the collective behavior of a many-body system. And, what it is more stunning, such complexity only depends on just a few general aspects such as the symmetries in the system, and not on the specific details. This is

all really good news for complexsystemists, who can neglect most of the specific details about the real dynamics. As we said before, More is different, fortunately!

1.2 Operating at a narrow line

Although criticality is a very attractive phenomenon, it only constitutes a rare situation among all possible states, most of which are, actually, quite boring. For this, we have to poise the system *just* at the critical point. However, paradoxically enough, many systems in nature exhibit fingerprints of criticality without the need of any external fine-tuning. In particular, scale invariance has been evidenced in different contexts such as in earthquakes [146, 32], solar flares [67], rainfall measurements [111], chemical reactions [27], gluon-quark systems [71], turbulences in plasma [130, 117]... To be more specific, earthquakes, for instance, occur at all scales, ranging from common unperceived vibrations to, eventually, demolishing seisms. Such different magnitudes are distributed as a power-law, which is the most characteristic fingerprint of criticality.

The theory of Self-Organized Criticality (SOC) has provided a general framework to understand this enigmatic phenomenon. [4, 5, 44]. In a nutshell, such self-tuning is produced when two mechanisms dynamically operate on the control parameter: a slow *external driving* and a rapid *dissipation* with the activity (see fig. 1.2). We can illustrate the mechanism in the example of earthquakes (in a very simplistic manner): let consider two tectonic plates meeting in a fault; as a consequence of their slow motion, the stress at the contact zone is slowly incremented (external driving). At a certain point, the fault cannot take any more tension, and it releases all the accumulated energy in a very fast event (rapid dissipation). These two mechanisms, operating with different time scales, organize the system at the edge between the active and inactive phase.

However, we are forced to read the small print in the mechanism of SOC, as *it still requires the external tuning of a parameter*. If the two processes (dissipation and driving) occur with similar time scales, the dynamics gets into a complicated situation, but, if they are separated, the system self-organizes at the critical point. The interesting fact is that, contrary to any standard fine-tuning, this occurs in a more natural way, and therefore a similar self-organization is more plausible to operate in nature.

Although the theory of SOC has been enthusiastically welcomed by complexsys-

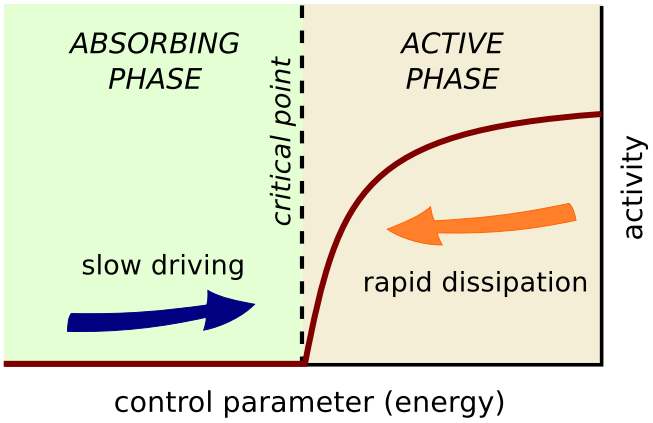


Figure 1.2: **Self-organized criticality:** a system with a continuous phase transition can be self-tuned to its critical point, as a result of two balancing forces acting on the control parameter: the slow external driving increases progressively the energy (control parameter), pushing any silent state into the active region; on the other hand, rapid energy dissipations are produced with activity. Consequently, the system self-organizes at the edged of the two phases, exhibiting critical behavior.

temists, we would like to put a warning about its generality. Specifically, power-law distributions, as the principal trace of SOC, can emerge in very different contexts which, actually, nothing have to do with criticality nor even a phase transition, as for instance in fractal geometries [87], Levy flights [87], and many other generative processes of such distributions [99]. For this reason, relating a power-law decay –that, in many cases, lacks of more than one or two scales– with SOC and universality, can be exposed to considerable and justified criticism.

1.3 Preview-summary

Up to now, we have mainly talked about phase transitions and self-organization in the context of inanimate matter. Thus the question is: How these phenomena affect living systems? Do they play any relevant role in biology? Phase transitions and self-organization are found in many biological contexts; the list is infinite, but some of them correspond to flocking behavior [136], cerebral activity [118], quorum sensing in bacterial communities [51], and so on. However, in contrast with inanimate matter, living systems are subjected to the Darwinian forces of survival, adaptation and

evolution [38]. Therefore, we expect that any *useful* mechanism relying on a phase transition could have been selected and successfully exploited by living systems in nature. The work developed during this thesis has tried to unravel different aspects of such mechanisms and their implications for the proper functioning of biological systems.

In particular, **Chapter 2** focuses on the empirical evidence that many aspect of living systems might operate at the vicinity of critical points [101]. However, given the diversity and heterogeneity between such systems, with examples ranging from brain activity [118] to flock dynamics [15], a general theory for understanding why and how living systems could dynamically tune themselves to be poised in the vicinity of a critical point is lacking. Employing tools from statistical mechanics and information theory, we show that complex adaptive or evolutionary systems can be much more efficient in coping with diverse heterogeneous environmental conditions when operating at criticality, while they remain non-critical for simple and predictable environments.

A more robust and non-expected *a priori* convergence to criticality emerges in co-evolutionary and co-adaptive set-ups in which individuals aim to represent other agents in the community with fidelity, and the environment is composed, essentially, by the community itself. While, initially, this population consists of simple individuals, complexity emerges as a global attractor of the dynamics, and the community ends to be highly heterogeneous.

This result has a broad range of implications for general complex adaptive systems. In particular it could apply to some bacterial communities [81] and viral populations [129] for which a huge phenotypic variability has been empirically observed. Such a large diversification can be seen as a form of *"bet hedging"*, an adaptive survival strategy analogous to stock-market portfolio management [144], which turns out to be a straightforward consequence of individuals in the community being critical. This interpretation constitutes the linking point with **Chapter 3**, which focuses on the study of bet-hedging strategies in the context of population dynamics. Here we analyze a simple model of a community of individuals reproducing by means of two different strategies: a poor but safe strategy or a better but risky (environment-dependent) one. Depending on the characteristics for each strategy, the stationary density of individuals exhibits a phase transition, changing from zero (extinction) to a sustained population (survival). In this scenario, we show that, under certain conditions, pure strategies leading irreme-

diably to extinction, when combined properly, can lead to survival. This phenomenon constitutes an instance of the so-called Parrondo's Paradox in game theory [57]. Additionally, we show that the benefits derived from bet-hedging are much enhanced for higher environmental variabilities and for small spatial dimensions, which are typical circumstances encountered by living systems in nature.

A specific case of bet-hedging corresponds to hybrid dispersal strategies developed by certain plants [68, 7, 3], which have evolved to spread their offspring by meas of two kind of seeds: with aerial seeds, that are more likely to colonize distant sites, but a the same time are much affected by fluctuations of the environment, and with terrestrial seeds, which are subjected to the negative consequences of self-reproduction (inbreeding depression), but not to external conditions. In **Chapter 4** we study a simple model of population dynamics equipped with such ingredients, and we analyze under which conditions hybrid strategies provide a significant gain respect to the pure strategies.

Finally, **Chapter 5** is dedicated to the study of neural dynamics, in particular to the so-called *Up and Down states* [147, 143], cortical oscillations in which the activity switches from intense activity intervals (Up) to quiescent periods (Down); such Up and Down states have been also related to self-organized criticality in the brain [98, 84, 17]. We focus on the experimental evidence that a class of spontaneous oscillations can emerge within the Up states, but not for Down states [128, 103, 49, 31]. By using different computational models, we show that the collective phenomenon of *"stochastic amplification of fluctuations"*, previously described in other contexts such as Ecology [94] and Epidemiology [1], explains in an elegant manner, beyond model details, this extra-rhythm emerging in the Up but not in the Down states.

Summing up, we have talked about the ubiquitousness of collective dynamics in nature and living systems, self-organization to complex behavior through simple interactions, and universality of several mechanisms underlying completely different systems. After all, a better answer for the misunderstood complexsystemist surrounded by purists, without any intention of offending Mr. Feynman, could be: "Studying complex systems is like physics: sure, it may give some practical results, but that's not why we do it."

Introducción

Como un *complexsystemista* no experimentado, cualquier estudiante de doctorado podría haberse visto envuelto, alguna vez, en un tenso debate con otros colegas no complexsystemistas sobre el papel de la física estadística en otras disciplinas, como la ecología, la sociología, la neurociencia o la economía. La discusión principal, normalmente, gira en torno a la pregunta: "¿Por qué los seguís llamando *física*?"

Llegado un cierto punto, uno desarrolla una fuerte opinión al respecto, y genera la respuesta automática de que, por supuesto, "los sistemas físicos pueden representar cualquier sistema en la naturaleza susceptible de ser modelado matemáticamente, con el objetivo de dar una descripción aproximada de él, además de hacer predicciones correctas sobre su comportamiento". De esta forma, los sistemas complejos en la naturaleza se ajustan bastante bien a esta definición. Sin embargo, llegado un *segundo* cierto punto, uno crea otra opinión sobre ello: durante muchos años, los diferentes ámbitos de la ciencia han estado fuertemente separados, pero hoy, con el desarrollo de las ciencias interdisciplinares, lo que vivimos podría ser visto como el retorno –con un cambio de *look*– de la Filosofía Natural; o "Cono", como popularmente se la llama en la escuela primaria.

En este contexto, es particularmente interesante cómo, para un mismo sistema, distintos "niveles de conocimiento" aparecen en cada escala, que muchas veces no pueden entenderse como una descripción efectiva de la precedente. Para dar un ejemplo hipotético, imaginemos que la tecnología computacional nunca se hubiera desarrollado. En este peculiar escenario, un día, una civilización avanzada aparece y nos concede un solo y único súper computador, sin más instrucciones sobre su funcionamiento. Lo más probable es que, intentando entender para qué sirve, empezaríamos por desentrañar los mecanismos de sus pequeños componentes. Sin embargo, incluso en el caso de que

alcanzáramos un nivel absoluto con el que pudiéramos predecir con total precisión el comportamiento de cada parte del ordenador, todavía nos faltaría un paso adicional y esencial hasta entender, por ejemplo, en qué consiste un programa o qué significa la palabra *software* en general.

De una forma similar, si queremos entender nuestro cerebro, necesitamos entender la dinámica individual de las neuronas, pero esto no es suficiente. Si queremos entender la estabilidad de un sistema ecológico, no basta con saber qué especies depredan otras especies. Si queremos evitar una pandemia global, necesitamos más que explicar los mecanismos de contagio de persona a persona. Necesitamos algo más, porque *Más* es, de hecho, *diferente*.

El interés de las transiciones de fase

En el contexto de los sistemas complejos y los fenómenos colectivos, las transiciones de fase han gozado de particular interés [127, 16]. Desde las escuela, nos han enseñado que la materia puede encontrarse en tres estados diferentes: sólido, líquido, o gaseoso. Además, hemos aprendido que la materia puede cambiar su estado mediante una *transición de fase*. Esto es, de hecho, un fenómeno bastante común, y ocurre cada vez que echamos hielo a nuestra bebida o preparamos una rica sopa. En términos generales, decimos que un sistema experimenta una transición de fase si, de repente, su estado global sufre un cambio cualitativo. El ejemplo más característico es el de la transiciones hielo-agua-vapor, pero muchos otros se pueden encontrar en la naturaleza (ver fig. 1.1). De forma específica, la fase se caracteriza mediante el llamado parámetro de *orden* (por ejemplo, la densidad), que cambia cuando variamos el –en términos de mecánica estadística– parámetro de *control* (como puede ser la temperatura, presión, etc.).

Entre las transiciones de fase, distinguimos dos tipos principales: discontinuas, o de *primer orden*, y continuas, también llamadas de *segundo orden*. El primer tipo corresponde al caso en el que el cambio se lleva a cabo de forma discontinua, e involucra un coste extra (calor latente), que es recibido o dado por el sistema. Las ya citadas transiciones del agua en condiciones ordinarias constituyen ejemplos de ello. Por otro lado, las transiciones de fase continuas no requieren de ningún coste adicional, y el cambio se produce de madera inmediata y progresiva. En este segundo caso encontramos la

magnetización espontánea del hierro por debajo de los 770 °.

En la estrecha línea que separa ambas fases aparece una multitud de fenómenos interesantes, siendo esta una de las principales razones por las que a los físicos estadísticos les encanta su estudio. En las transiciones discontinuas, dos fases completamente distintas convergen en el punto de separación; por tanto, cabe esperar que ambas puedan *coexistir* en determinadas circunstancias. El punto triple del agua constituye un caso extremo en el que vapor, agua y hielo lo hacen perfectamente. De forma paralela, las transiciones de fase continuas están caracterizadas por un punto *crítico* que separa ambos regímenes; justo en ese punto, los sistemas presentan una multitud de características interesantes [127, 16], como por ejemplo invariancia de escala en su estructura, correlaciones de largo alcance y máxima susceptibilidad frente a perturbaciones externas. La gran lección que nos enseña la mecánica estadística es que, incluso si los elementos que constituyen un sistema son muy simples, la criticidad e invariancia de escala emergen como un fenómeno colectivo, y, lo que es más impresionante, que dicha complejidad solo depende de algunos aspectos generales, como las simetrías del sistema, y no de los detalles específicos. Esto es una noticia estupenda para los complexsystemistas, que pueden olvidarse de la mayoría de detalles sobre la dinámica real. Como hemos dicho antes, Más es diferente, ¡afortunadamente!

Operando en un estrecho margen

Aunque la criticidad es un fenómeno muy atractivo, solo constituye una situación muy rara de todos los posibles estados en los que podemos encontrar a un sistema, la mayoría de los cuales bastante aburridos. Para ello, deberíamos *tunear* el sistema justo en el punto crítico. Sin embargo, paradójicamente, muchos sistemas en la naturaleza presentan indicios de criticidad sin la necesidad de un ajuste externo. Por ejemplo, la invariancia de escala aparece en contextos muy diferentes, como en terremotos [146, 32], erupciones solares [67], medidas de precipitaciones [111], reacciones químicas [27], sistemas de quarks-gluones [71], turbulencias en plasmas [130, 117]... Para ser más específico, los terremotos, por ejemplo, ocurren a todas las escalas, desde las más comunes vibraciones imperceptibles hasta, de forma ocasional, los seísmos demoledores. Estas magnitudes tan diferentes aparecen distribuidas en forma de ley de potencias, la huella

más característica de criticidad.

La teoría de la Criticidad Auto-Organizada (SOC, por las siglas del inglés *Self-Organized Criticality*) ha proporcionado un marco general en el que entender este hecho enigmático [4, 5, 44]. En pocas palabras, dicho auto-ajuste se produce cuando dos mecanismos actúan sobre el parámetro de control: una *carga externa* lenta, y una *disipación* rápida con la actividad (ver fig. 1.2). Podemos ilustrar este mecanismo con el ejemplo de los terremotos (siendo muy simplistas): imaginemos dos placas tectónicas que se encuentran en una falla. Como consecuencia del lento movimiento de la corteza terrestre, el estrés en la zona de contacto se incrementa poco a poco. En un momento dado, la falla no puede soportar más la tensión, y libera toda la energía acumulada en un terremoto que ocurre de manera muy rápida (disipación). Estos dos mecanismos, operando con diferentes escalas temporales, organizan al sistema al punto que separa la fase activa e inactiva.

Sin embargo, estamos obligados a leer la letra pequeña en el mecanismo de SOC, ya que *todavía requiere del ajuste externo de un parámetro.* Si los dos procesos (disipación y carga) ocurren con escalas de tiempo parecidas, la dinámica se vuelve muy complicada, pero, solo cuando dichas escalas son separadas (o, equivalentemente, cuando la carga se hace mucho más lenta que la disipación), el sistema se auto-organiza al punto crítico. El punto interesante es que, de forma contraria a cualquier *tuneo* estándar, este ajuste ocurre de una forma muy natural, y por tanto una auto-organización similar es más plausible que ocurra en la naturaleza.

Aunque la teoría de SOC ha sido acogida con gran entusiasmo por los complexsystemistas, deberíamos hacer un pequeño inciso respecto a su aplicabilidad. Las leyes de potencia, como la principal característica de SOC, pueden emergen en contextos muy diferentes que, realmente, nada tienen que ver con criticidad o ni siquiera con una transición de fase, como por ejemplo en geometrías fractales [87], vuelos de Levy [87], y muchos otros procesos capaces de generar dichas distribuciones [99]. Por esta razón, relacionar cualquier ley de potencias –que, en muchos casos carece de más de uno o dos órdenes de magnitud– con universalidad y SOC, puede ser objeto de considerable y justificada crítica.

Resumen

Hasta ahora, hemos hablado principalmente sobre transiciones de fase y auto-organización en el contexto de la materia inanimada. La siguiente pregunta es: ¿Cómo afectan estos fenómenos a los sistemas vivos? ¿Juegan algún papel relevante en biología? Las transiciones de fase y la auto-organización aparecen en muchos contextos biológicos; la lista es infinita, pero, por nombrar unos pocos, dichas transiciones aparecen en el comportamiento de manadas [136], actividad cerebral [118], *quorum sensing* en comunidades de bacterias [51], y muchos más. Sin embargo, en contraste con la materia inanimada, los sistemas vivos están sujetos a las reglas darwinianas de supervivencia, adaptación y evolución [38]. Así pues, cabe esperar que cualquier mecanismo *útil*, con base en una transición de fase, haya sido seleccionado y explotado por los sistemas biológicos en la naturaleza. El trabajo desarrollado durante esta tesis ha tratado de descifrar distintos aspectos de tales mecanismos y sus implicaciones en el buen funcionamiento de los sistemas biológicos.

En particular, el **Capítulo 2** se centra en la evidencia empírica de que muchos aspectos de los sistemas biológicos podrían estar operando en la cercanía de puntos críticos [101]. Sin embargo, dada la heterogeneidad y diversidad entre tales sistemas, con ejemplos que van desde la actividad cerebral [118] a las bandadas de pájaros [15], una teoría general con la que entender por qué y cómo los sistemas biológicos podrían organizarse de forma dinámica para situarse en un punto crítico aún no se ha llevado acabo. Mediante herramientas de mecánica estadística y teoría de la información, nosotros mostramos que los sistemas adaptativos y evolutivos complejos son mucho más eficientes haciendo frente a condiciones externas heterogéneas cuando operan en la criticidad, mientras que no lo necesitan cuando los entornos son simples y predecibles [62].

Una convergencia más robusta hacia la criticidad aparece en sistemas co-evolutivos y co-adaptativos en los cuales los individuos intentan representarse unos a otros; en este caso, el entorno está compuesto, esencialmente, por la comunidad en sí. Mientras que, inicialmente, la población puede constar únicamente de individuos simples, la complejidad emerge como el atractor global de esta dinámica, y la comunidad termina por ser muy heterogénea.

Este resultado tiene muchas implicaciones en sistemas adaptativos generales. En

particular, podría aplicarse a ciertas comunidades de bacterias [81] y poblaciones de virus [129] para las que se ha observado una alta variabilidad fenotípica. Tal diversificación puede ser entendida como una forma de "cobertura de riesgos" (en inglés conocida como "bet-hedging"), una estrategia de supervivencia análoga a la gestión del mercado de acciones en bolsa [144], y que resulta ser una consecuencia directa de individuos en una comunidad crítica.

Esta interpretación constituye el punto de conexión con el **Capítulo 3**, el cual se centra en el estudio de estrategias de bet-hedging en el contexto de dinámicas de poblaciones. Aquí, analizamos un modelo simple de una comunidad de individuos que se reproducen a través de dos estrategias distintas: una estrategia pobre pero segura, y otra muy arriesgada y dependiente de las condiciones externas. Dependiendo de las características de cada estrategia, la densidad estacionaria de los individuos muestra una transición de fase que va desde cero (extinción) a una población mantenida (supervivencia). En este escenario, nosotros mostramos que, bajo ciertas condiciones, estrategias puras que llevarían irremediablemente a la extinción, cuando se combinan adecuadamente, pueden llevar a la supervivencia. Este fenómeno constituye un ejemplo de la llamada Paradoja de Parrondo en teoría de juegos [57]. Adicionalmente, mostramos que los beneficios derivados del bet-hedging se aumentan enormemente para condiciones externas muy variables y para bajas dimensiones espaciales, que de hecho son las circunstancias típicas con las que se encuentran los sistemas vivos en la naturaleza.

Un ejemplo específico de bet-hedging corresponde a las estrategias de dispersión híbridas de algunas plantas [68, 7, 3], las cuales han evolucionado para propagar su descendencia a través de dos tipos de semilla: mediante semillas aéreas, que son mucho más propensas a colonizar sitios distantes, pero que, al mismo tiempo están muy afectadas por las condiciones y fluctuaciones del entorno; y con semillas terrestres, las cuales están sujetas a las consecuencias negativas de la intra-reproducción pero no a las condiciones externas. En el **Capítulo 4** estudiamos un modelo sencillo de dinámica de poblaciones equipado con tales ingredientes, y analizamos bajo qué condiciones las estrategias de dispersión híbridas proporcionan un beneficio significativo respecto a las estrategias puras.

Finalmente, el **Capítulo 5** está dedicado al estudio de dinámicas neuronales, en particular a los llamados estados *Up-Down* [147, 143], donde la actividad cerebral al-

terna entre periodos de intensa actividad (Up) e intervalos completamente silenciosos (Down); dicho estados Up-Down han sido también relacionados con criticidad auto-organizada [98, 84, 17]. En nuestro estudio, nos centramos en la evidencia empírica de que una serie de oscilaciones espontáneas emergen en los estados Up, mientras que esto no sucede en los Down [128, 103, 49, 31]. Usando diferentes modelos computaciones, mostramos que el fenómeno colectivo llamado *"amplificación estocástica de las fluctuaciones"*, anteriormente descrito en otros contextos como la Ecología [94] o la Epidemiología [1], explica de una forma elegante, más allá de los detalles del modelo, estas oscilaciones extras emergentes únicamente en los períodos Up, y así mismo justifica por qué no lo hacen para los Down.

En definitiva, hemos hablado sobre la ubicuidad de las dinámicas colectivas en la naturaleza y los sistemas biológicos, la auto-organización a comportamientos complejos a través de interacciones simples, y la universalidad de varios mecanismos detrás de sistemas muy diversos. Después de todo, una mejor respuesta para el complexsystemista rodeado de puristas, sin la intención de ofender al Sr. Feynman, podría ser: "El estudio de los sistemas complejos es como la física: seguramente dará alguna aplicación práctica, pero no es por eso por la que la hacemos".

Chapter 2

The emergence of criticality in living systems

2.1 *'Are biological system poised at criticality?'*

Empirical evidence has proliferated that living systems might operate at criticality [101], with examples ranging from spontaneous brain behavior [9], gene expression patterns [107], cell growth [50], morphogenesis [79], bacterial clustering [24], and flock dynamics [15]. Even if none of these examples is fully conclusive and even if the meaning of "criticality" varies across these works, the criticality hypothesis –as a general strategy for the organization of living matter– is a tantalizing idea worthy of further investigation.

Unlike models of self-organized criticality in which some inanimate systems are found to become critical in a mechanistic way [70], we focus on general adaptive or evolutionary mechanisms, specific to biological systems. In this context, the drive to criticality may arise from functional advantages of being poised in the vicinity of a critical point.

But why is a living system fitter when it is critical? Living systems need to perceive and respond to environmental cues and to interact with other similar entities. Indeed, biological systems constantly try to encapsulate the essential features of the huge variety of detailed information from their surrounding complex and changing environment into manageable internal representations, and they use these to base their actions and responses. The successful construction of these representations, which extract, summa-

rize, and integrate relevant information [45], provides a crucial competitive advantage, which can eventually make the difference between survival and extinction. We suggest here that criticality is an optimal strategy to effectively represent the intrinsically complex and variable external world in a parsimonious manner. This is in line with the hypothesis that living systems benefit from having attributes akin to criticality –either statistical or dynamical [101]– such as a large repertoire of dynamical responses, optimal transmission and storage of information, and exquisite sensitivity to environmental changes [16, 107, 25, 10, 77, 102, 122].

As conjectured long ago, the capability to perform complex computations, which turns out to be the fingerprint of living systems, is enhanced in "machines" operating near a critical point [82, 14, 75], i.e. at the border between two distinct phases: a disordered phase, in which perturbations and noise propagate unboundedly –thereby corrupting information transmission and storage – and an ordered phase where changes are rapidly erased, hindering flexibility and plasticity. The marginal, critical, situation provides a delicate compromise between these two impractical tendencies, an excellent trade-off between reproducibility and flexibility [25, 122, 10] and, on larger time scales, between robustness and evolvability [137]. Similarly, criticality has been recently shown to emerge through adaptive information processing in machine learning, where networks are trained to produce a desired output from a given input in a noisy environment; when tasks of very different complexity need to be simultaneously learned, networks adapt to a critical state to enhance their performance [55]. A specific example of this general framework are genetic regulatory networks [75, 56]. Cells ranging from those in complex organisms to single-celled microbes such as bacteria respond to signals in the environment by modifying the expression of their genes. Any given genetic regulatory network, formed by the genes (nodes) and their interactions (edges) [41]– can be tightly controlled to robustly converge to a fixed almost-deterministic attractor –i.e. a fixed "phenotype"– or it can be configured to be highly sensitive to tiny fluctuations in input signals, leading to many different attractors, i.e. to large phenotypic variability [66]. These two situations correspond to the ordered and disordered phases respectively. The optimal way for genetic regulatory networks to reconcile controllability and sensitivity to environmental cues is to operate somewhere in between the two limiting and impractical limits alluded to above [75] as has been confirmed in different experimental set-ups

[107, 79, 6]. Still, it is not clear how such tuning to criticality comes about.

Our goal here is to exploit general ideas from statistical mechanics and information theory to construct a quantitative framework showing that self-tuning to criticality is a convenient strategy adopted by living systems to effectively cope with the intrinsically complex external world. In order to provide some further intuition, we employ genetic regulatory networks as a convenient guiding example, but one could equally well consider neural networks, models for the immune response, groups of animals exhibiting collective behavior, etc., with each specific realization requiring a more detailed modeling of its special attributes.

2.2 Mathematical framework: Mapping the external world

The external environment in which living systems operate is highly variable, largely unpredictable, and describable in terms of probability distribution functions. Living systems need to modify their internal state to cope with external conditions and they do so in a probabilistic manner.

To be specific, but without loss of generality, we represent an environmental cue "perceived" and processed by a living system as a string of N (binary) variables, $\mathbf{s} = (s_1, s_2, \ \dots \ s_N)$. A specific environmental source is modeled by the probability distribution $P_{\rm src}$ with which it produces each of the 2^N possible states. This distribution is assumed to depend on a set of parameters, $\boldsymbol{\alpha} = (\alpha_1, \ \alpha_2, \dots)$, accounting for environmental variability. As $P_{\rm src}$ is a positive quantity, it is always possible to write:

$$P_{\rm src}(\mathbf{s}|\boldsymbol{\alpha}) = \frac{\exp\big(-H_{\rm src}(\mathbf{s}|\boldsymbol{\alpha})\big)}{Z_{\rm src}(\boldsymbol{\alpha})} \ , \tag{2.1}$$

which defines $H_{\rm src}$ up to a constant, independent of $\mathbf{s}$, which can be set equal to zero. The factor $Z_{\rm src}(\boldsymbol{\alpha})$ is defined by normalization condition. The last representation reminds the usual notation of statistical mechanics, suggesting to express $H_{\rm src}(\mathbf{s}|\boldsymbol{\alpha})$ as:

$$H_{\rm src}(\mathbf{s}|\boldsymbol{\alpha}) = \sum_{\mu=1}^{E} \alpha_\mu \phi^\mu_{\rm src}(\mathbf{s}) \ , \tag{2.2}$$

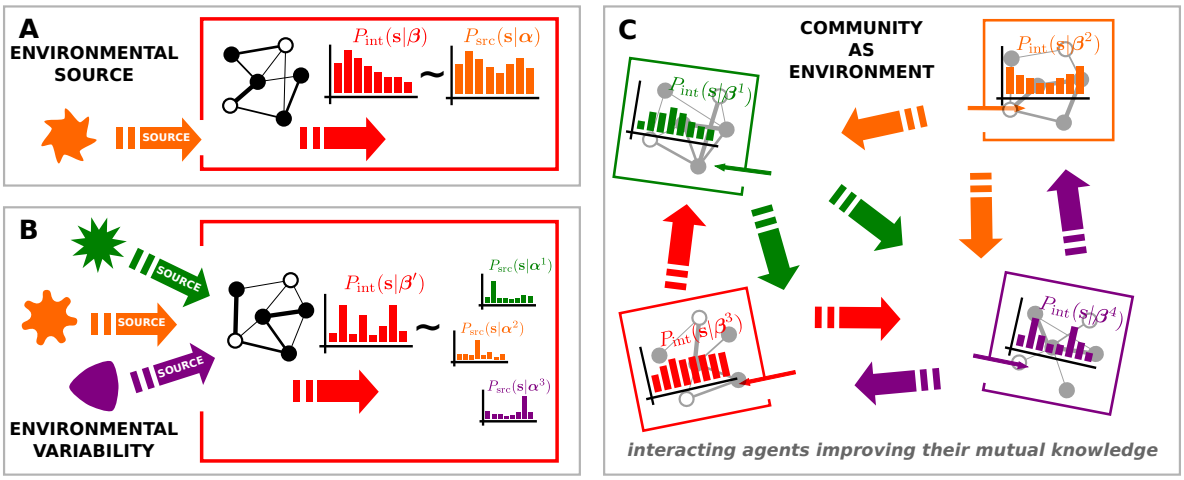


Figure 2.1: **Living systems coping with the environment**: Panel A illustrates a *living system* responding to an *environmental source* (e.g. a bacteria responding to some external conditions such as the presence/absence of some nutrients, pH concentration, or temperature). A given source, labeled by the set of parameters $\boldsymbol{\alpha}$, can only be probabilistically gauged by the system. $P_{\rm src}(\mathbf{s}|\boldsymbol{\alpha})$ is the most accurate representation that the system can potentially generate in terms of the Boolean variables (or bits) $\mathbf{s}$. However, such a representation might not be accessible to the system by merely changing its internal-state parameters, $\boldsymbol{\beta}$, and the actual internal state, $P_{\rm int}(\mathbf{s}|\boldsymbol{\beta})$, is usually an imperfect proxy for $P_{\rm src}(\mathbf{s}|\boldsymbol{\alpha})$. Panel B shows a more complex scenario, where the system has to cope with multiple and diverse sources. The internal state has to be able to accommodate each of them. In panel C, the environment is not imposed *ad hoc* but instead, it is composed of other individuals, and every agent needs to cope with ("understand") the states of the others. Each agent evolves similarly to the others in the community, trying to exhibit the same kind of state, generating in this way a self-organized environment.

where E is the number of "external" parameters specifying the source and $\phi^\mu(\mathbf{s})$ are suitable functions ("observables"). Equations 2.1 and 2.2 are a convenient parametrization of the probability distribution function, so that different values of $\boldsymbol{\alpha}$ specify distinct source distributions $P_{\rm src}$. Although this parametrization has been introduced in a heuristic way, it could constitute our starting point, i.e. we represent any external environmental source by eqs. 2.1 and 2.2, with the number of parameters E and observable functions $\boldsymbol{\phi}$ needed.

We turn now to an individual living system or *agent*, which seeks to adapt itself to cope with the perceived stimuli/signals emanating from a given environmental source.

This is accomplished by changing its internal state, encapsulated in a second probability distribution function, $P_{\rm int}$, specified by a different –smaller in principle– parameter set $\boldsymbol{\beta} = (\beta_1,\ \beta_2, \dots)$ aimed at capturing the essential features of $P_{\rm src}$ in the most efficient –though in general imperfect– way (see fig. 2.1).

In analogy with eq. 2.1, $P_{\rm int}(\mathbf{s}|\boldsymbol{\beta})$ can be written as

$$P_{\rm int}(\mathbf{s}|\boldsymbol{\beta}) = \frac{\exp\big(-H_{\rm int}(\mathbf{s}|\boldsymbol{\beta})\big)}{Z_{\rm int}(\boldsymbol{\beta})}\ , \tag{2.3}$$

where

$$H_{\rm int}(\mathbf{s}|\boldsymbol{\beta}) = \sum_{\mu=1}^{I} \beta_\mu \phi^\mu_{\rm int}(\mathbf{s}) \tag{2.4}$$

where I is the number of "internal" parameters. Henceforth we will denote the external source and its internal representation by $P_{\rm src}(\mathbf{s}|\boldsymbol{\alpha})$ and $P_{\rm int}(\mathbf{s}|\boldsymbol{\beta})$ respectively.

In our guiding example, the external cues could be, for instance, the environmental (temperature, pH,...) conditions, which are variable and can only be probabilistically gauged by a cell/bacterium. The binary vector $\mathbf{s} = (s_1, s_2,\ \dots\ s_N)$ can be thought of as the on/off state of the different N genes in its (Boolean) genetic regulatory network [75, 56, 41]. In this way, $P_{\rm src}(\mathbf{s}|\boldsymbol{\alpha})$ can be interpreted as the probability that the most convenient state aimed at by the system to cope with a given environmental condition is $\mathbf{s}$, while $P_{\rm int}(\mathbf{s}|\boldsymbol{\beta})$ is the actual probability for the genetic-network state (attractor) of a given individual –with its limitations– to be $\mathbf{s}$.

Our hypothesis is that the capacity of living systems to tune their internal states to efficiently cope with variable external conditions provides them with a strong competitive advantage. Thus, the internal state $P_{\rm int}(\mathbf{s}|\boldsymbol{\beta})$ should resemble as closely as possible the one most in accord with the environmental signal $P_{\rm src}(\mathbf{s}|\boldsymbol{\alpha})$; in other words, one seeks the distribution that the system should express in order to best respond to the external conditions. Information theory provides us with a robust measure of the "closeness" between the aimed (source) and the actual (internal) probability distribution functions [34]: the Kullback-Leibler divergence.

Example: Ising parametrizations

Throughout this chapter, we test our ideas numerically with the archetypical (mean-field) Ising model [16]:

$$H_{\text{src/int}}(\mathbf{s}|\gamma) = -\frac{1}{N}\gamma \sum_{i<j}^{N} s_i s_j = -\frac{N}{2}\gamma \left(\sum_i \frac{s_i}{N} \right)^2 + \text{constant}, \tag{2.5}$$

where γ can be either α or β. This constitutes a simple one-parameter system with a continuous phase transition and hence, a critical point. In the guiding example of genetic regulatory networks, this corresponds to an extremely simple fully-connected network in which the state of each gene is equally determined by all the other genes and, hence, the probability of a given state depends only on the total number of on/off genes, controlled by a single parameter.

Additionally, we study the two-parameters (mean-field) Ising model with four-body couplings:

$$H_{\text{src/int}}(\mathbf{s}|\boldsymbol{\gamma}) = -\frac{N}{2}\gamma_1 \left(\sum_i \frac{s_i}{N} \right)^2 - \frac{N}{4!}\gamma_2 \left(\sum_i \frac{s_i}{N} \right)^4 \tag{2.6}$$

and the case with an additional external field:

$$H_{\text{src/int}}(\mathbf{s}|\boldsymbol{\gamma}) = -\frac{N}{2}\gamma_1 \left(\sum_i \frac{s_i}{N} \right)^2 - \gamma_2 \sum_i \frac{s_i}{N} \tag{2.7}$$

Numerical factors $(-N/2)$ and $(-N/4!)$ have been introduced for convenience. In these two cases, apart of a critical point –or a critical line– in the space $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$, we find also a discontinuous phase transition, making these examples particularly interesting for our purposes. The mean-field parametrization is particularly useful when performing analytical and numerical calculations, but later we discuss the possibilities outside the mean-field.

2.2.1 Optimal mapping and Kullback-Leibler divergence

Given two general probability distributions $P(\mathbf{s})$ and $Q(\mathbf{s})$ for the set of variables $\mathbf{s}$, the Kullback-Leibler divergence (KLD from now) of $Q(\mathbf{s})$ from $P(\mathbf{s})$ is defined as

$$D(P|Q) := \sum_{\mathbf{s}} P(\mathbf{s}) \log\left(\frac{P(\mathbf{s})}{Q(\mathbf{s})}\right) , \tag{2.8}$$

and quantifies the loss of information when $Q(\mathbf{s})$ is used to approximate $P(\mathbf{s})$ [80, 34]. The KLD is non-negative and it vanishes if and only if both distributions are equal. Observe also that the KLD is not symmetric and therefore is not a properly-defined "distance". Despite of that, we will still refer to the "closeness" between distributions.

The KLD can be understood in terms of the maximum likelihood principle [34]. Consider a long sequence of T data, sampled from the distribution $P(\mathbf{s})$. In the infinite T limit, the relative frequencies of $\mathbf{s}$ converge to $P(\mathbf{s})$, but we can compute the probability $\mathcal{L}$ (or likelihood) that another model $Q(\mathbf{s})$ generates the same finite sequence of T elements. We expect that, if $P \neq Q$, this probability rapidly goes to 0, but, as illustrated in Appendix A, it decreases as

$$\mathcal{L} \sim \exp\left(-TD(P|Q)\right) \tag{2.9}$$

up to leading order. Therefore, maximizing the likelihood of a trial probability distribution function Q is equivalent to minimizing its KL divergence from the original one, P. This result is also known as Sanov's theorem [115] in the context of large deviations theory.

In the scenario where individuals capture the external information, minimizing the KLD with respect to the internal-state parameters, $\boldsymbol{\beta}$, generates the optimal, although in general imperfect, internal state aimed at representing or coping-with a given source $\boldsymbol{\alpha}$ (see Fig. 2.1A):

$$\boldsymbol{\beta}^{\text{opt}}(\boldsymbol{\alpha}) = \arg\min_{\beta'} D(P_{\text{src}}(\cdot|\boldsymbol{\alpha})|P_{\text{int}}(\cdot|\boldsymbol{\beta}')). \tag{2.10}$$

In the following, we use the simplified notation $D(P_{\text{src}}(\cdot|\boldsymbol{\alpha})|P_{\text{int}}(\cdot|\boldsymbol{\beta})) = D(\boldsymbol{\alpha}|\boldsymbol{\beta})$.

In an ever-changing world, the requirement for an individual is not just to reproduce a single source with utmost fidelity but rather to be able to successfully cope with a

group of highly different sources (see Fig. 2.1B). In order to account for broadly diverse and variable external sources, we introduce a probability density of different parameter sets $\rho_{\rm src}(\boldsymbol{\alpha})$, which describes the variability of $\boldsymbol{\alpha}$ or the probability to encounter a source with a given choice of parameters. In this more complex situation, we define the distance to the set of sources as the average of the KLDs:

$$d(\rho_{\rm src}|\boldsymbol{\beta}) = \int d\boldsymbol{\alpha}\rho_{\rm src}(\boldsymbol{\alpha})D(\boldsymbol{\alpha}|\boldsymbol{\beta}), \tag{2.11}$$

and the optimal map, as the one which minimizes the average KLD,

$$\boldsymbol{\beta}^{\rm opt}(\rho_{\rm src}) = \arg\min_{\beta'} d(\rho_{\rm src}|\boldsymbol{\beta}'). \tag{2.12}$$

This result can be interpreted in an alternative way; introducing the "averaged environment"

$$\bar{P}_{\rm src}(\mathbf{s}|\rho_{\rm src}) := \int d\boldsymbol{\alpha}\rho_{\rm src}(\boldsymbol{\alpha})P_{\rm src}(\mathbf{s}|\boldsymbol{\alpha}) \ , \tag{2.13}$$

the KLD respect to $P_{\rm int}(\mathbf{s}|\boldsymbol{\beta})$ can be rewritten as

$$D(\bar{P}_{\rm src}(\cdot|\rho_{\rm src})|P_{\rm int}(\cdot|\boldsymbol{\beta})) = d(\rho_{\rm src}|\boldsymbol{\beta}) - \int d\boldsymbol{\alpha}\rho_{\rm src}(\boldsymbol{\alpha})D(P_{\rm src}(\cdot|\boldsymbol{\alpha})|\bar{P}_{\rm src}(\cdot|\rho_{\rm src})) \ . \tag{2.14}$$

Since the last term on the right hand side does not depend on $\boldsymbol{\beta}$, the minimization of the KLD between the "averaged environment" and the internal mapping $P_{\rm src}(\mathbf{s}|\boldsymbol{\beta})$ leads to the same result as the minimization of $d(\rho_{\rm src}|\boldsymbol{\beta})$ given by eq. 2.12. In both cases, the Hessian matrix turns out to be strictly positive (see section 2.4), and therefore the eventual extrema are local minima.

A particularly interesting example of this would comprise a community of similar individuals which together strive to establish some kind of a common collective language (see Fig. 2.1C). In any of these complex situations, our working hypothesis is that an individual has a larger "fitness" when a characteristic measure, e.g. the mean, of its KLDs from the set of diverse sources is small. Therefore, fit agents have internal states close to those required by existing external conditions. This constitutes the starting point for the computational models we are going to develop, but first, in section 2.2.2 we analyze the structure of the 'information-based landscape' of internal states.

Numerical computation in mean-field Ising parametrizations

When computing the KLD, as well as the normalization constant Z of parametrizations 2.1 and 2.3, a sum over states $\mathbf{s} = (s_1, ..., s_N)$ is involved. This is in general hard to compute for large N, because the number of states grows as 2^N.

However, since the mean-field Ising parametrizations of eqs. 2.5, 2.6 and 2.7 depend on $\mathbf{s}$ only through the magnetization $m = \sum_i s_i/N$, the sums can be easily performed for such cases. Defining $\Gamma(m)$ as the number of states with $\sum_i s_i/N = m$, and using the Stirling approximation, one readily obtains:

$$\Gamma(m) = \binom{N}{\frac{N(1+m)}{2}} \underset{N\gg 1}{=} \exp\left\{-N\left(\frac{1+m}{2}\log\frac{1+m}{2} + \frac{1-m}{2}\log\frac{1-m}{2}\right)\right\}. \quad (2.15)$$

Then, we can defined the *continuous* probability density function for the magnetization m as

$$\hat{P}_{\text{src/int}}(m|\boldsymbol{\gamma}) = \Gamma(m)\frac{\exp\left(-H_{\text{src/int}}(m|\boldsymbol{\gamma})\right)}{\hat{Z}_{\text{src/int}}(\boldsymbol{\gamma})}, \quad (2.16)$$

where the normalization constant is fixed via *integrating* the previous formula in $m \in [-1, 1]$. A factor $\Delta m^{-1} = N/2$ appearing when we pass to the continuum can been re-absorbed in $\hat{Z}$.

Finally, the KLD can be computed with the integral

$$D(\boldsymbol{\alpha}|\boldsymbol{\beta}) \underset{N\gg 1}{=} \int_{-1}^{1} dm \hat{P}_{\text{src}}(m|\boldsymbol{\alpha}) \log\frac{\hat{P}_{\text{src}}(m|\boldsymbol{\alpha})}{\hat{P}_{\text{int}}(m|\boldsymbol{\beta})}. \quad (2.17)$$

2.2.2 Fisher information and criticality

We have defined the optimal map in both the cases of single and multiple sources. Now we ask: How *sensitive* is the internal state, $P_{\text{int}}(\mathbf{s}|\boldsymbol{\beta})$, to changes in the value $\boldsymbol{\beta}$? From the point of view of information theory, the answer is given by the Fisher information.

Given a probability distribution $P(\mathbf{s}|\boldsymbol{\beta})$, the Fisher information (FI) is defined as

$$\chi^{\mu\nu}(\boldsymbol{\beta}) := \left\langle \frac{\partial \log P(\cdot|\boldsymbol{\beta})}{\partial\beta_\mu}\frac{\partial \log P(\cdot|\boldsymbol{\beta})}{\partial\beta_\nu}\right\rangle_{\boldsymbol{\beta}}, \quad (2.18)$$

where μ and ν are parameter labels and the average $\langle\cdot\rangle_{\boldsymbol{\beta}}$ is performed with respect to

$P(\cdot|\boldsymbol{\beta})$. Eq. 2.18 constitutes a measure of the *amount of information* encoded in the states $\mathbf{s}$ about the parameters $\boldsymbol{\beta}$ [34]. This follows from the Cramér-Rao inequality, which states that the error made when we estimate $\boldsymbol{\beta}$ from one state $\mathbf{s}$ is, on average, greater (or at least equal) than the inverse of the Fisher information [34]. In particular, if χ happens to diverge at some point, it is possible to specify the associated parameters with maximal precision. A more detailed analysis of this interpretation is included in Appendix B.

Introducing the parametrization in eqs. 2.3 and 2.4, the FI becomes the *generalized susceptibility* in the statistical mechanics terminology:

$$\chi^{\mu\nu}(\boldsymbol{\beta}) = -\frac{\partial\langle\phi^{\mu}_{\text{int}}\rangle_{\boldsymbol{\beta}}}{\partial\beta_{\nu}} = \langle\phi^{\mu}_{\text{int}}\phi^{\nu}_{\text{int}}\rangle_{\boldsymbol{\beta}} - \langle\phi^{\mu}_{\text{int}}\rangle_{\boldsymbol{\beta}}\langle\phi^{\nu}_{\text{int}}\rangle_{\boldsymbol{\beta}} \ , \tag{2.19}$$

which measures the response of the system to parameter variations and is well-known to peak at critical points [127, 16].

To give an intuitive explanation of the Fisher information, in Fig. 2.2A we represent a two-dimensional Ising lattice at three different temperatures. From each snapshot, we can infer (in a probabilistic manner) the temperature for each panel, but the error of this estimation differs for each case. For instance, looking at the right (red) picture, we know that, most probably, it corresponds to a high temperature, being hard to distinguish the system with high temperature $T = 10^3$ or a 'very high' $T = 10^5$, because the behavior is quite similar in both cases (the dynamics is basically random). A similar argument holds for the left (blue) panel (frozen behavior). However, when we look at the central picture, the one exhibiting a fractal structure, we know that, most probably, it corresponds to a temperature close to the critical point, without much deviation from this value. In fact, the error we make in this estimation is bounded by χ^{-1}, that, in the case of the Ising model, is proportional to the inverse of the specific heat [89] (see Fig. 2.2B).

In conclusion, if the internal states of individuals encode a continuous phase transition, the Fisher information reaches to identify the critical $\boldsymbol{\beta}_c$ in terms of the original probability $P_{\text{int}}(\mathbf{s}|\boldsymbol{\beta})$, which is where χ diverges, being maximum for finite system sizes. Additionally, for multivariate parametrizations, we take the scalar $\chi = \det(\chi^{\mu\nu})$ to account for potential divergences in any of the componentes of $\chi^{\mu\nu}$. In the following, we will refer to the critical point or maximum of the FI without distinction.

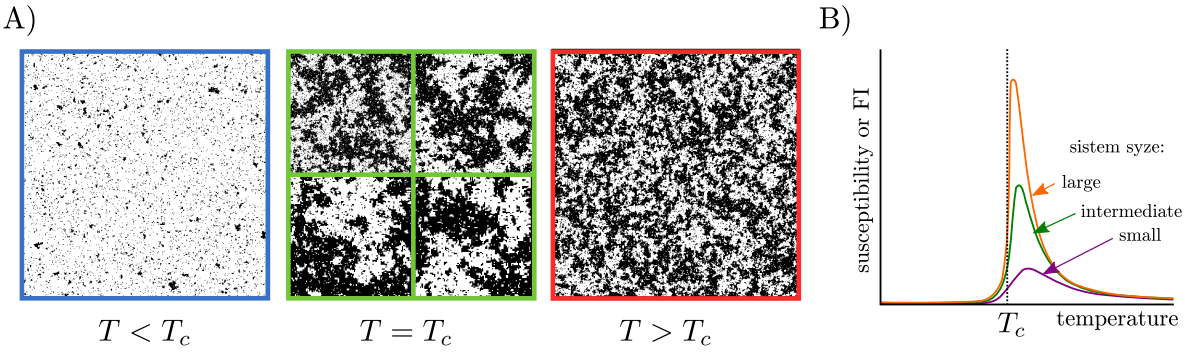


Figure 2.2: **A) Two-dimensional Ising lattice at different temperatures**: spins can point to either up (white dots) or down (black dots), and communicate with each other by neighboring interaction. Depending on the temperature, two phases emerge: order for low temperatures (left) and disorder for high temperatures (right), separated by a critical point (middle). At this temperature, the system exhibits scale invariance, as it can be seen from the four central panels, corresponding to different zooms (1x, 4x, 9x and 16x in clockwise order) of the same configuration. **B) Fisher information (or generalized susceptibility) in the Ising lattice**, for different system sizes: While it is small for low and high temperatures, the susceptibility peaks at the critical temperature, diverging for infinite system sizes.

2.3 Computational models

We have developed diverse computational evolutionary and adaptive models exploiting the ideas above. The dynamical rules employed in these models are *not* meant to, necessarily, mimic the actual dynamics of living systems, rather they are efficient ways to optimize fitness.

In the evolutionary models, inspired by the genetic algorithm [54, 56], a community of M individuals –each one characterized by its own set of internal parameters $\boldsymbol{\beta}$– evolves in time through the processes of death, birth, and mutation. Individuals with larger fitness, i.e. with a smaller mean Kullback-Leibler divergence from the rest of sources, have a larger probability to produce an offspring, which –apart from small random mutations– inherits its parameters from its ancestor. On the other hand, agents with low fitness are more likely to die and be removed from the community.

In the adaptive models, individuals can change their internal parameters if the attempted variation implies an increase of their corresponding fitnesses. These evolutionary/adaptive rules result in the ensemble of agents converging to a steady state

distribution, which we aim at characterizing.

We obtain similar results in two families of models, which differ in the way in which the environment is treated. In one case (evolutionary/adaptive), the environment is defined *ad hoc*, while, in the other (co-evolutionary/co-adaptive), the external world is self-generated by a community of co-evolving/co-adapting individuals.

2.3.1 Evolutionary model

An ensemble of M agents is exposed at each particular time to a heterogeneous complex environment consisting of S independent environmental sources, each one with a different P_{src} and thus parametrized by diverse $\boldsymbol{\alpha}$s (eqs. 2.1 and 2.2). The set of S sources is randomly extracted from a broadly distributed pool of possible sources occurring with different probabilities, $\rho_{\text{src}}(\boldsymbol{\alpha})$. The k-th agent has a representation P_{int} of the observed sources, encoded in its parameter $\boldsymbol{\beta}^k$ (eqs. 2.3 and 2.4).

Each of the agents is equipped with some initial parameter set extracted from some arbitrary distribution $p(\boldsymbol{\beta}, t=0)$, and the algorithm is run as follows:

1. At every time step, we generate S external sources, $\{\boldsymbol{\alpha}^u\}_{u=1,\dots,S}$, from the distribution $\rho_{\text{src}}(\boldsymbol{\alpha})$.

2. We compute the average KLD of every individual's internal representation to the external sources

$$d(\{\boldsymbol{\alpha}^u\}|\boldsymbol{\beta}^k) := \frac{1}{S}\sum_{u=1}^{S}\sum_{\mathbf{s}} P_{\text{src}}(\mathbf{s}|\boldsymbol{\alpha}^u)\log\frac{P_{\text{src}}(\mathbf{s}|\boldsymbol{\alpha}^u)}{P_{\text{int}}(\mathbf{s}|\boldsymbol{\beta}^k)}. \tag{2.20}$$

3. One of the individuals of the community is removed with a probability proportional to its average KLD:

$$P_{\text{kill}}(k) = \frac{d(\{\boldsymbol{\alpha}^u\}|\boldsymbol{\beta}^k)}{\sum_l d(\{\boldsymbol{\alpha}^u\}|\boldsymbol{\beta}^l)} \tag{2.21}$$

and it is replaced by a copy of another individual (offspring), which is picked randomly from the community with uniform probability.

4. The offspring inherits its parameter set from its parent or, instead, mutates with a probability ν, altering the original parameter, $\boldsymbol{\beta} \to \boldsymbol{\beta} + \boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is a random Gaussian number of zero mean and deviation $\boldsymbol{\sigma}$.

5. Time is incremented as $t \to t + 1/M$.

6. Another set of parameters $\{\boldsymbol{\alpha}^u\}_{u=1,\ldots,S}$ is generated from $\rho_{\rm src}(\boldsymbol{\alpha})$, and the process is iterated.

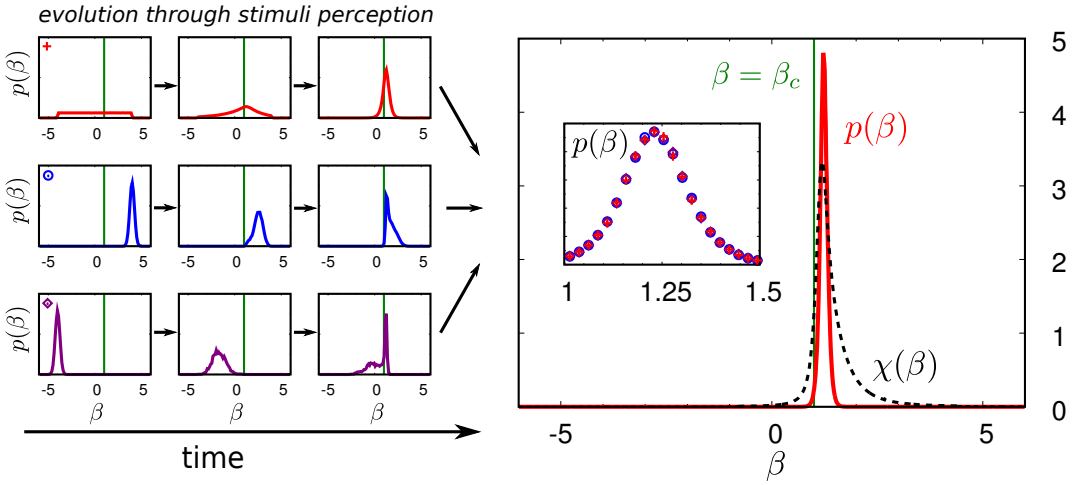


Figure 2.3: **Evolutionary model leads to near to criticality for complex environments**: Stimuli and internal states distributions correspond to mean-field Ising model parametrizations (eq. 2.5). Panels on the left represent the evolution of three different initial distributions of the parameter β in the community. In all cases the environment is described by the uniform distribution of parameters $\rho_{\rm src}(\alpha) = U([-10, 10])$. The distributions converge to the same stationary state –points and red line on the right panel–, which is peaked at the maximum of the generalized susceptibility (dashed line curve). Initial distribution are (red) uniform $U([-4, 4])$, (blue) Gaussian $N(4, 0.25)$ and (purple) $N(-4, 0.25)$. Parameters are the same as in Table 2.1, and $R = 10^4$ independent realizations.

We are interested in the stationary distribution of the individual parameters, $p(\boldsymbol{\beta}) \equiv p(\boldsymbol{\beta}, t \to \infty)$ (we start measuring at some time T_i and stop at time T_f), for which the distribution is averaged over R independent realizations of the initial distribution $p(\boldsymbol{\beta}, t = 0)$.

We have simulated the simple case in which both the external sources and internal representations correspond to the simple mean-field Ising parametrization given in eq. 2.5. Similar results can be obtained for other parametrization of sources and internal representations, for instance by considering eqs. 2.6, 2.7 with non-vanishing α_2.

Parameter	**Value**
N	100
M	100
S	10
ν	0.1
σ	0.1
$\rho_{\rm src}(\alpha_1)$ (not used in Fig. 2.5)	$U([-10, 10])$
T_i	10^4
$T_f - T_i$	10^5
R	100

Table 2.1: **Parameters of the simulation of the evolutionary model:** N is the number of spins composing each of the individuals, M is the community size, S is the number of stimuli received in every interaction with the environment, ν is the mutation probability, σ is the deviation of the mutated offspring, T_i and T_f are the initial and final time steps used for the measure and R is the number of independent realizations.

We first analyze the case of a very heterogeneous environment, which means that sources $\boldsymbol{\alpha}$ are picked from a broad distribution $\rho_{\rm src}(\boldsymbol{\alpha})$ (a more accurate definition of what this 'heterogeneity' will be given, see below). Fig. 2.3 illustrates that starting from different initial conditions $p(\boldsymbol{\beta}, 0)$ after some (sufficiently long) times the ensemble of individuals converges to a unique steady state $p(\boldsymbol{\beta}, t \to \infty)$. The resulting distribution is sharply peaked very near the critical point, at the very same location at which the Fisher information or generalized susceptibility peaks. This peak approaches the critical point $\boldsymbol{\beta} = \boldsymbol{\beta}_c$ in the limit $N \to \infty$. The set of parameters used in this simulation is listed in Table 2.1.

Fig. 2.4 illustrates the dependence of the results on parameters. The main conclusions are:

- The system becomes closer and closer to the critical point as the system-size N is enlarged.

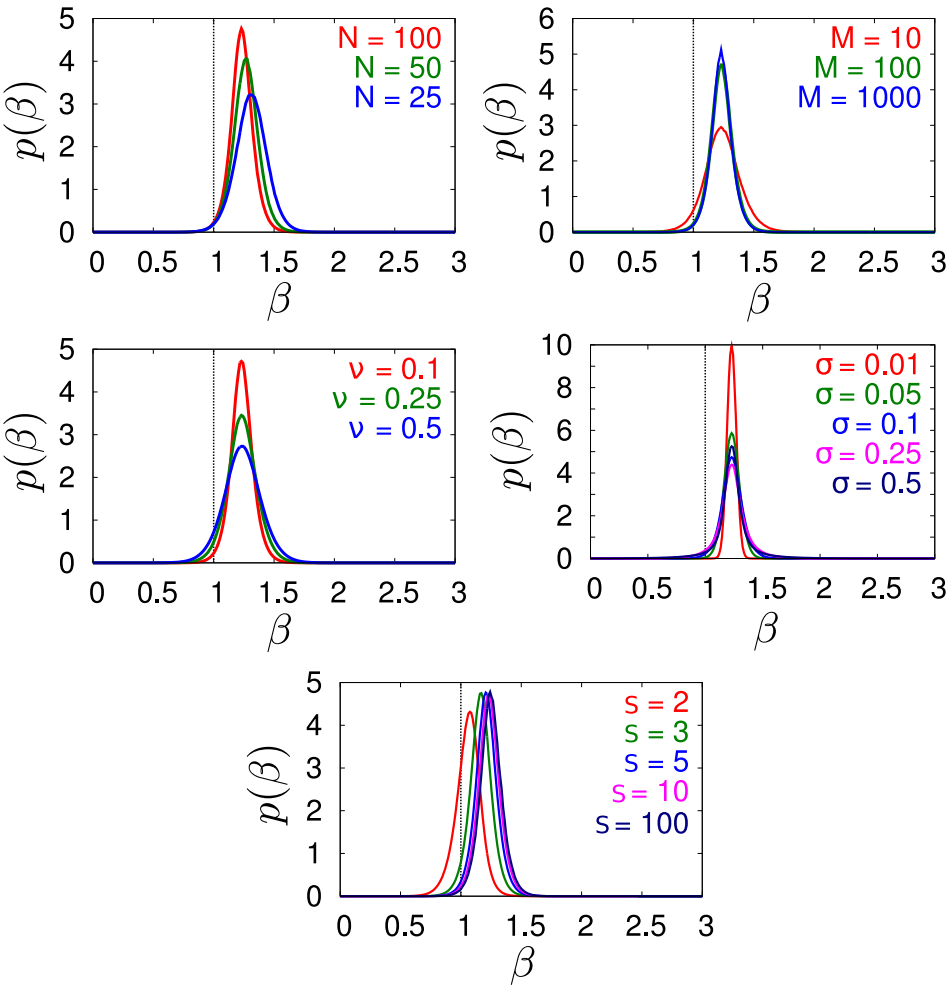


Figure 2.4: **Dependence on parameters in the evolutionary model**: Stationary distribution $p(\beta)$ as a function of diverse parameters; different colors in each plot stand for different values of (from top to bottom and from left to right): number of spins N, community size M, mutation probability ν, mutation deviation σ, and number of external sources S. Unless otherwise stated, other parameters take the same values as in Table 2.1. The dashed lines indicate the critical point location (in the limit $N \to \infty$).

- The distribution reaches an asymptotic shape as the ensemble size grows.
- The distribution becomes sharper for smaller mutation rates.
- The distribution becomes much sharper for small mutation variances.
- The distribution reaches an asymptotic shape as the number of external sources

is increased.

At this point, we study the role of the environment in leading to criticality. Computer simulations show that in the case of very homogeneous environments, occurring when all sources in the pool are similar –$\rho_{\text{src}}(\boldsymbol{\alpha})$ sufficiently narrow– the optimal $\boldsymbol{\beta}$ strongly depends on the specific sources, resulting in detail-specific internal states (see bottom panels in Fig. 2.5). On the other hand, if the external world is sufficiently heterogeneous –broad $\rho_{\text{src}}(\boldsymbol{\alpha})$– the optimal internal state becomes peaked near the critical point (upper panels in Fig. 2.5).

This result is the key finding of this model: *only when the environment is sufficiently complex, individuals evolve to the vicinity of the critical point, while they do not if the external world is very specific and predictable.*

We note that there is some still dependence of the choice of the environment in the resulting criticality (i.e. the mean β of the stationary $p(\beta)$), so the approach to criticality is less precise in this model than in the co-evolutionary one, that we discuss later, in which the environment changes progressively as the agents co-evolve, allowing the system to systematically approach the critical point with precision.

Finally, all the results of the evolutionary model hold for the analogous "adaptive model", illustrated in section 2.3.3.

Heterogeneity of the environment

Up to now, we have referred to "homogeneity" and "heterogeneity" of the environment in a rather vague –but intuitive– manner. To objectively measure the level of heterogeneity of the environment, we can use the (continuous) Shannon entropy of the distribution of moments (observables) in the external distributions. Given the environment characterized by its distribution of sources $\rho_{\text{src}}(\boldsymbol{\alpha})$, we compute its associated distribution of moments $\varrho_{\text{src}}(\langle\boldsymbol{\phi}\rangle)$:

$$\varrho_{\text{src}}(\langle\boldsymbol{\phi}\rangle) = \rho_{\text{src}}(\boldsymbol{\alpha}) \left| \frac{\partial\langle\boldsymbol{\phi}\rangle}{\partial\boldsymbol{\alpha}} \right|^{-1}. \tag{2.22}$$

Then, the Shannon entropy can be computed as

$$S[\varrho_{\text{src}}] = -\int d\langle\boldsymbol{\phi}\rangle\, \varrho_{\text{src}}(\langle\boldsymbol{\phi}\rangle) \log \varrho_{\text{src}}(\langle\boldsymbol{\phi}\rangle). \tag{2.23}$$

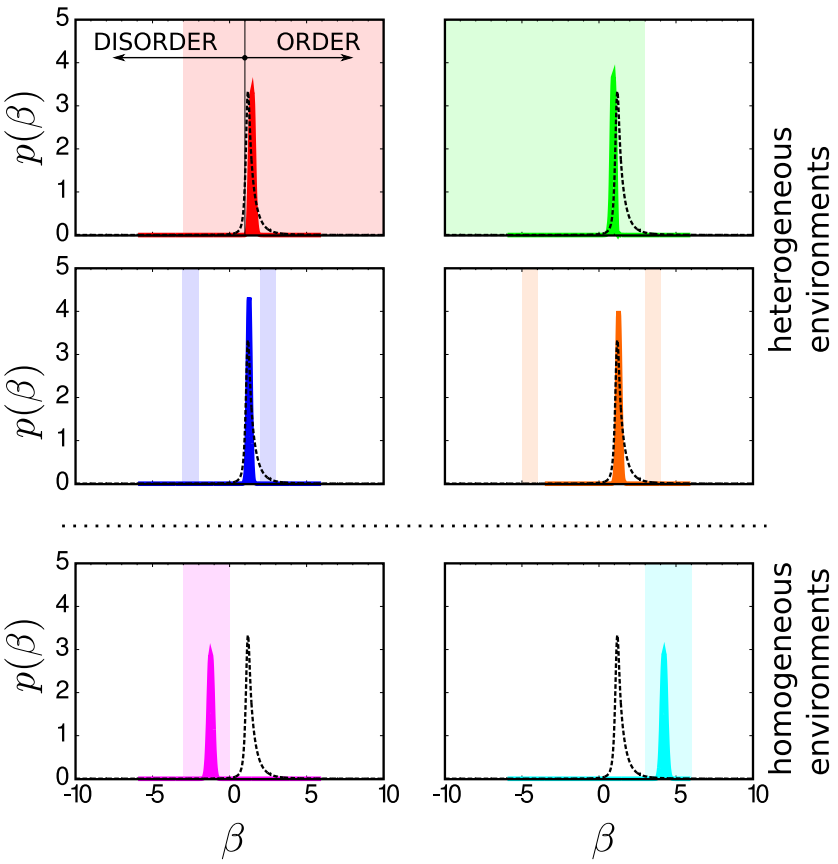


Figure 2.5: **Individuals become critical only for heterogeneous environments**: Using the parametrization of eq. 2.5 for both the sources and the internal states, we run the evolutionary model and the community eventually reaches the steady state distribution of parameters, $p(\beta)$. The six panels in the figure correspond to different supports (colored regions) for uniform source distributions, $\rho_{\rm src}(\alpha)$. The dashed line is the FI of the internal probability distribution, which exhibits a peak at the critical point separating an ordered from a disordered phase. Heterogeneous source pools (Top and Middle) lead to distributions peaked at criticality, whereas for homogeneous sources (Bottom), the communities are not critical but specialized.

Note that, as it is defined in the continuum limit, the entropy can be negative [34].

To check the validity of this measure, we randomly generate different distributions $\rho_{\rm src}(\alpha)$ for the one-parameter case, where the parameter α refers to the mean-field Ising model, eq. 2.5. For each distribution $\rho_{\rm src}(\alpha)$, we compute its optimal representation β as well as the entropy of $\varrho_{\rm src}(\langle\phi\rangle)$. The result is plotted in Fig. 2.6 and it is compared with the corresponding Fisher information $\chi(\beta)$. We see that, when the entropy is low

(homogeneous environments), the optimal mapping β can lie at any value, but, when the entropy increases (heterogeneous environments), β is gradually confined into the critical region. Thus, even though we have not been able to specify a quantitative distinction between heterogeneous and homogeneous environments, the entropy of the distribution of moments does constitute a good proxy for the environmental heterogeneity.

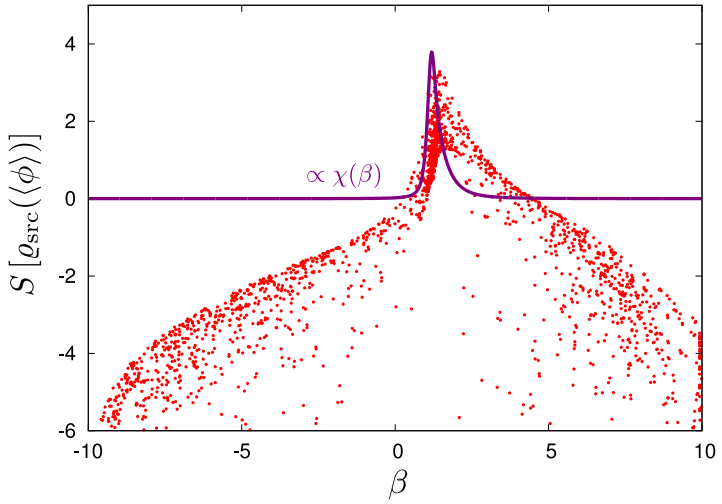


Figure 2.6: **Measuring the heterogeneity of the environment**: Each dot corresponds to a randomly generated distribution of sources $\rho_{\rm src}(\alpha)$, i.e. an environment. For each of them, we compute the associated optimal representation β as well as the continuous entropy of its moment distribution, $S[\varrho_{\rm src}(\langle\phi\rangle)]$. Because the space of all possible distributions $\rho_{\rm src}(\alpha)$ can be sampled in many different ways, we restrict ourselves to the case of uniform distributions with compact support in the range $\alpha \in [-10, 10]$. The purple line is the Fisher information (rescaled for visual comparing). We can see that, as the entropy increases, the mapping is confined to the critical region, whereas it can be at anywhere for low entropies. The parameters α, β refer to the mean-field Ising, eq. 2.5, with $N = 100$.

Effective criticality of the environment

As we have shown in section 2.2.1, minimizing the KLD to the "averaged environment" –i.e. the distribution of sources resulting by averaging over different environmental parameters– leads to the same result as minimizing the mean KLD to the sources, which is what we implement in the simulations. With this equivalence, our results show that agents seeing a complex "averaged environment" tend to become critical.

One might wonder whether the resulting closeness to criticality in these models is not just a byproduct of the environment itself being critical in some way. In fact, it has been recently shown that complex environments, when hidden variables are averaged out, can be effectively described by the Zipf's law [119], a signature of criticality [101]. Thus it may not be surprising that individuals tune their parameters near the criticality to minimize the KLD with respect to such a critical environment.

We show that our results cannot be generically explained in terms of this phenomenon. For the cases in which the individuals end up near criticality, the averaged environment is not necessarily critical. To illustrate this, we consider each of the pools of sources used in Fig. 2.5 to illustrate the evolution of agents in the presence of complex environments. As it is not possible to identify the criticality of the "average environment" by looking at the peak of a susceptibility, following the seminal paper by Mora and Bialek[101] we say that a particular distribution is critical if it obeys Zipf's law. To check this, for each particular environment, we plot the probability of states ordered by their rank and the energy as a function of the entropy (see [101]). In critical cases, the energy should be a linear function of the entropy, and the rank ordering should obey the Zipf's law [101, 119].

The results are shown in Fig. 2.7, where we have kept the same relative position and color code as in Fig. 2.5. Only two of the six averaged environments presented in Fig. 2.5 turn out to be critical in the sense of Zipf's law (upper panels in Fig. 2.7). In the two central cases, the averaged environment is not critical, but the optimal internal distribution peaks around criticality (central panels in Fig. 2.7). This case corresponds to an environment composed, essentially, of two very different type of sources, and individuals have to accommodate to the critical point to respond to both of them efficiently. This demonstrates that our approach works in a general scenario of heterogeneous sources, without requiring the environment to be Zipfian.

2.3.2 Co-evolutionary model

Here we discuss the type of evolutionary model in which every individual receives stimuli from its surrounding world, which is nothing but the set of the other individuals in the community. Hence, the environment perceived by each individual consists of the other systems in the community, which it aims at "understanding" and coping with.

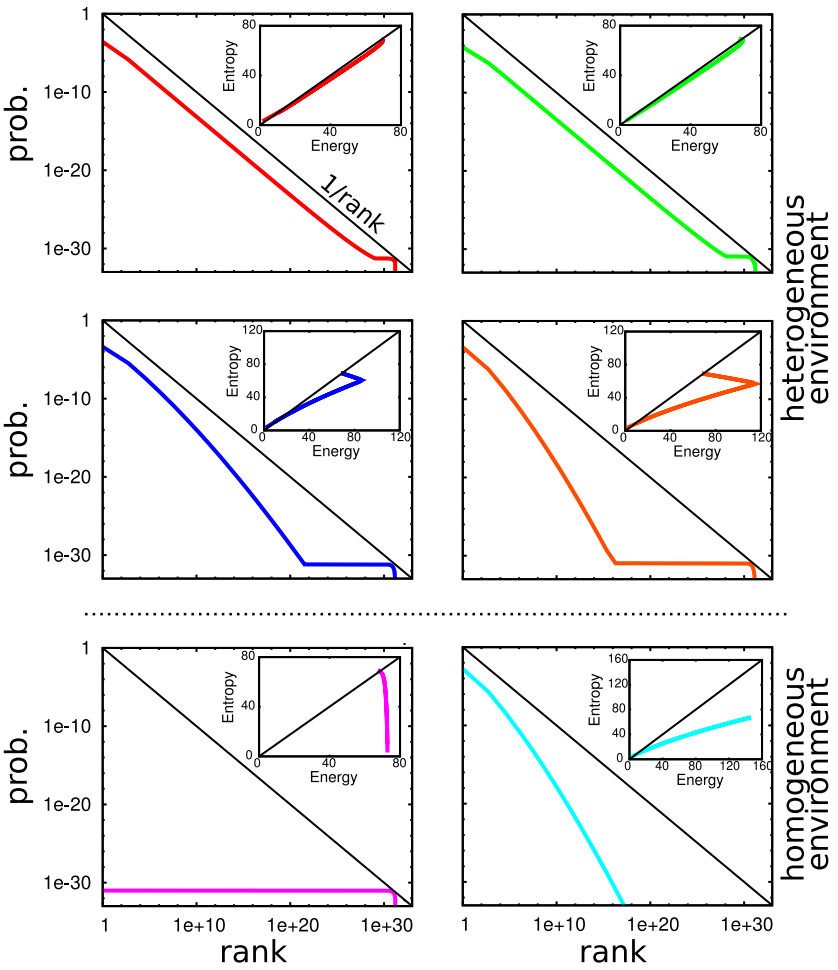


Figure 2.7: **Heterogeneity and criticality of the averaged environment**: Each panel of the figure refers to the corresponding panel, with the same position and color code, in Fig. 2.5. In each one, the main plot shows the probability of states in the "averaged environment" $\bar{P}_{\rm src}(\mathbf{s}|\rho_{\rm src})$ (eq. 2.13) with the states $\mathbf{s}$ ranked in order of decreasing probability. The inset shows the energy associated to $\bar{P}_{\rm src}$ as a function of the entropy. The black lines define the expected linear behavior in the critical case [101]. Only the red and green settings correspond to a critical distribution obeying also Zipf's law $(1/x)$. The more interesting cases are the blue and the orange ones: for these, the internal distribution is critical even though the average environment is not.

In the simplest computational implementation of this idea, a pair of individual agents is randomly selected from the community at each time step and each of these two individuals constitutes the environmental source for the other (a more general

implementation with multiple interactions at the same time is discussed later). Given that the KLD is not symmetric, one of the two agents has a larger fitness and thus a greater probability of generating progeny, while the less fit system is more likely to die.

More specifically, let call $P_{\text{int}}(\mathbf{s}|\boldsymbol{\beta}^k) \propto \exp\left(-H_{\text{int}}(\mathbf{s}|\boldsymbol{\beta}^k)\right)$ the probability distribution describing the k-th agent of the community (we use the same parametrization for all the individuals, with different parameter $\boldsymbol{\beta}^k$). Starting with an ensemble of M individuals whose coupling parameters are extracted from an arbitrary distribution, $p(\boldsymbol{\beta}, t = 0)$, the evolutionary dynamics proceeds as follows:

1. At each time step, two individuals, i and j, are randomly selected.

2. Each individual has a probability to be removed proportional to its relative KLD respect to the other:

$$P_{\text{kill}}(i) = \frac{D(\boldsymbol{\beta}^j|\boldsymbol{\beta}^i)}{D(\boldsymbol{\beta}^j|\boldsymbol{\beta}^i) + D(\boldsymbol{\beta}^i|\boldsymbol{\beta}^j)}, \quad P_{\text{kill}}(j) = 1 - P_{\text{kill}}(i), \tag{2.24}$$

   where –as the KLD is not symmetric– $P_{\text{kill}}(i) \neq P_{\text{kill}}(j)$ unless $\boldsymbol{\beta}^i = \boldsymbol{\beta}^j$. One of the two individuals is removed from the community, while the other creates and offspring.

3. Offspring mutate with a probability ν, modifying its parameters from $\boldsymbol{\beta}$ to $\boldsymbol{\beta} \to \boldsymbol{\beta} + \boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is randomly chosen from a multivariate Gaussian distribution with zero mean and deviation $\boldsymbol{\sigma}$.

4. Time is updated to $t \to t + 1/M$.

5. Another couple of individuals i' and j' is picked, and the process is iterated.

To compute the stationary distribution of parameters $p(\boldsymbol{\beta}) \equiv p(\boldsymbol{\beta}, t \to \infty)$, we iterate T_i time steps and then perform measurements during $T_f - T_i$ steps. Results are averaged over R realizations of the evolutionary process.

As for the evolutionary model, we first study the stationarity of the final distribution of parameters $\boldsymbol{\beta}$ in the community. In Fig. 2.8, the co-evolution of $M = 100$ agents –which on their turn are sources– leads to a very robust evolutionarily-stable steady-state distribution (in this case, we illustrate our results with the Ising parametrization

with an external field, eq. 2.7, but after we discuss other parametrizations). Indeed, different panels on the left show that, for three substantially different initial parameter distributions, the community co-evolves in time to a unique localized steady state distribution, which turns out to be peaked at the critical point (i.e. where the Fisher information peaks, see Fig. 2.8 right panel). This conclusion is robust against model details, different parametrizations and computational implementations: the solution peaked at criticality is an evolutionary stable attractor of the dynamics. The same conclusions hold for an analogous "co-adaptive model" in which the systems adapt rather than dying and replicating (see section 2.3.3).

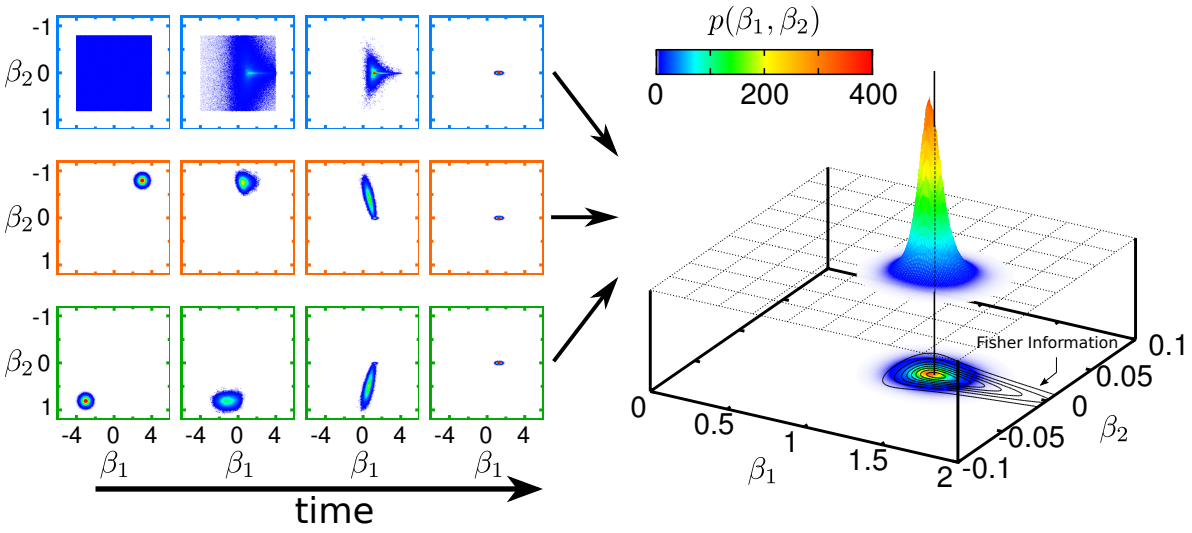


Figure 2.8: **Co-evolutionary model leads self-consistently to criticality:** In this case, each agent i ($i = 1, ..., M$) is characterized by a 2-parameter internal state distribution of eq. 2.7 (mean-field Ising with an external field), and the rest of the community acts as the external environment it has to cope with, i.e. the agents try to "understand" each other. (Left) The co-evolutionary rules drive the community very close to a unique localized steady state. As shown (Right), this is localized precisely at the critical point, i.e. where the FI of the internal state distribution exhibits a sharp peak (as shown by the contour plots and heat maps).

Numerical parameters are summarized in Table 2.2.

We study also the dependence on the parameters, in particular, on number of spins, N, number of individuals, M, the mutation probability, ν, and the deviation of the mutations, $\boldsymbol{\sigma}$. Results are summarized in Fig. 2.9 for the same parametrization used

Parameter	Value
N	100
M	100
ν	0.1
σ	0.1
σ_1, σ_2	$1\sigma, 0.1\sigma$
T_i	10^4
$T_f - T_i$	10^5
Init. Distribution	$N(-3, 0.25) \cdot N(-0.25, 0.05)$
	$N(3, 0.25) \cdot N(0.25, 0.05)$
	$U([-4, 4] \times [-0.8, 0.8])$
R	100 (10000 for transients)

Table 2.2: **Parameter values in the simulation of the co-evolutionary model**. See also Table 2.1.

in Fig. 2.8 (eq. 2.7)

Alternative parametrizations

We show that the convergence to the critical point is robust for alternative parametrizations of the internal states $P_{\text{int}}(\mathbf{s}|\boldsymbol{\beta})$. Figs. 2.10 and 2.11 correspond to stationary distribution $p(\boldsymbol{\beta})$ in the one-parameter case (eq. 2.5) and the quadratic-quartic case (eq. 2.6), respectively, together with the determinant of the FI.

Complex internal networked topologies

We now scrutinize a different parametrization of the model in which the internal probability distribution of each individual/agent is not a "mean-field" one, in the sense that every s_i variable is coupled to all others, but instead, possible interactions are encoded in a network, such that each s_i interacts only with other s_j directly connected to it, i.e. for which the *adjacency matrix*, element $a_{ij} \neq 0$. We have restricted this analysis to the case of one parameter β.

The co-evolutionary dynamics is as before, with the only difference that now the structure of the probability characterizing each individual is as follows:

- Given N spin variables, we generate a fixed adjacency matrix of interactions

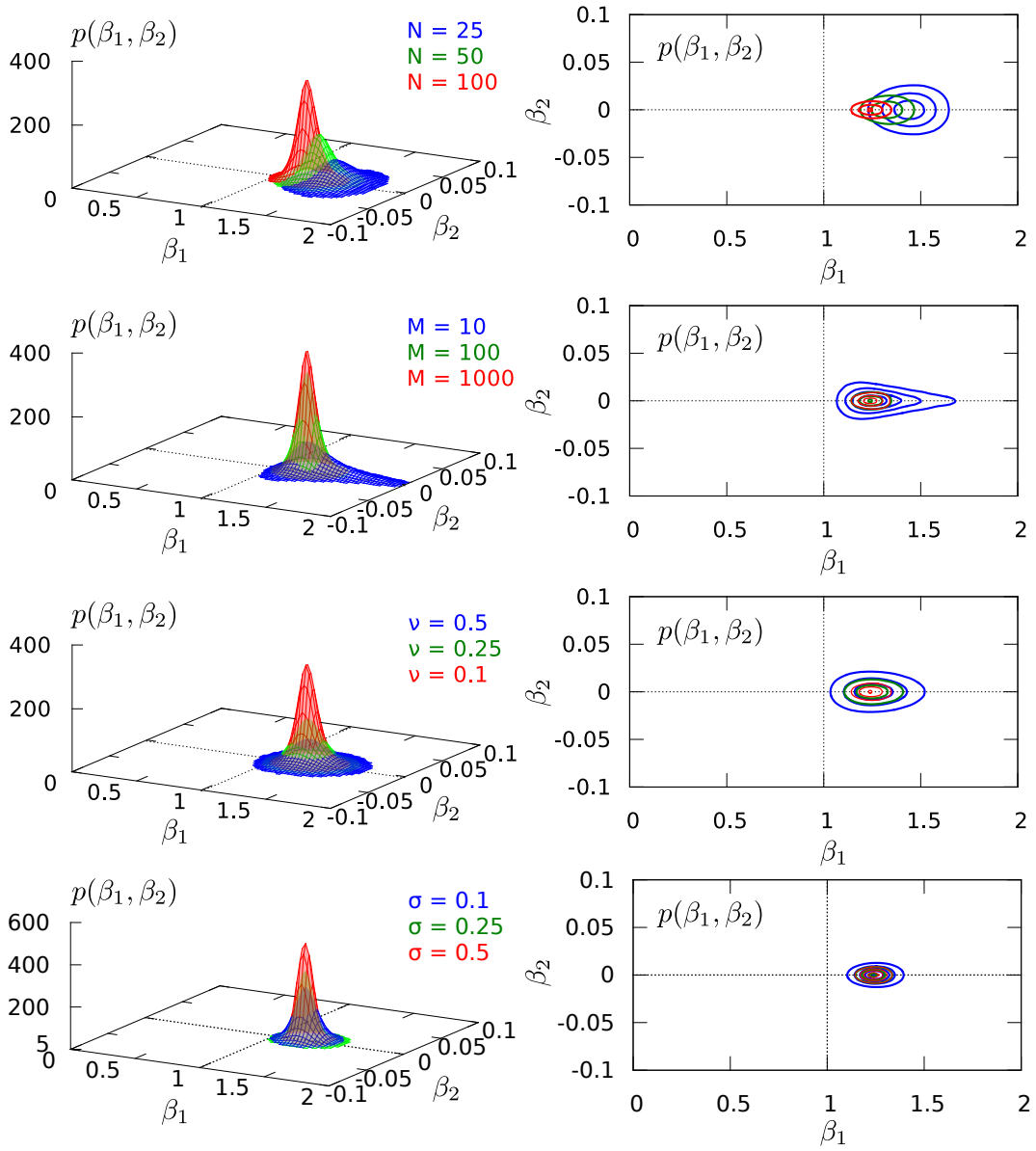


Figure 2.9: **Parameter dependence in the co-evolutionary model:** we plot the stationary distribution $p(\beta_1, \beta_2)$ as a function of parameters (for the linear-quadratic case, eq. 2.7). Different colored lines in each plot correspond to different values of N, community sizes M, mutation parameters ν, and σ. For larger communities sizes the stationary distribution becomes sharper. Parameter values are listed in Table 2.2 (unless otherwise specified).

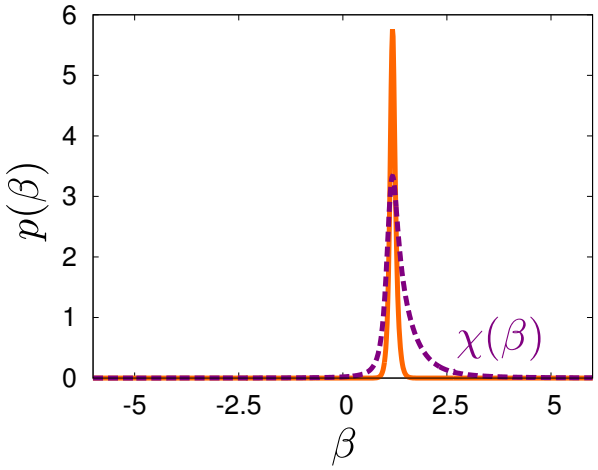


Figure 2.10: **Stationary distribution $p(\beta)$ in the co-evolutionary model in the 1-parameter case, eq. 2.5**: we compare the stationary distribution (orange line) with the generalized susceptibility (purple dashed line). As in Fig. 2.8, the individual parameters converge to the peak of the FI. Parameters are set to $N = 100$, $M = 100$, $\nu = 0.1$, $\sigma = 0.1$.

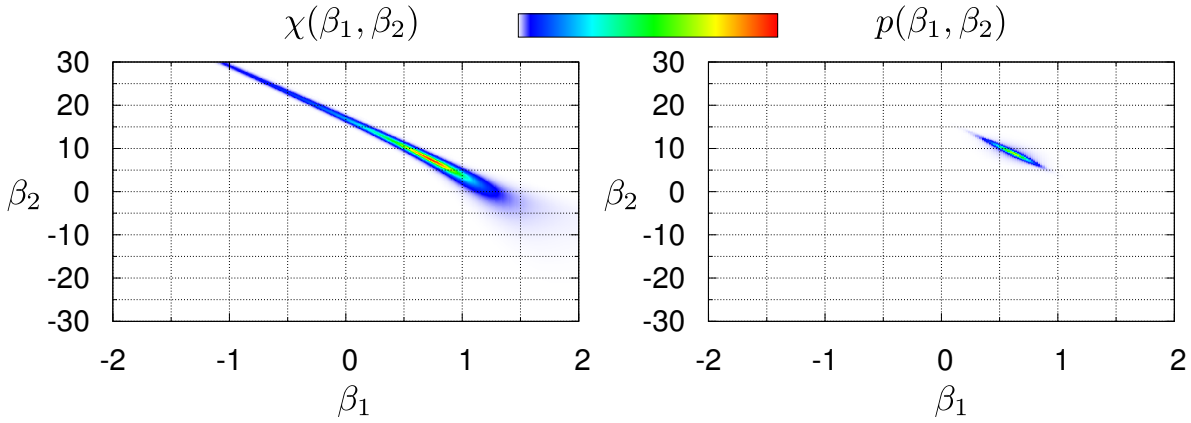


Figure 2.11: **Stationary distribution $p(\beta_1, \beta_2)$ in the co-evolutionary Model with the quadratic-quartic parametrization, eq. 2.6**: as above, we compare the stationary distribution (right panel) with the generalized susceptibility (left panel). Again, the community evolves toward the global maximum of the FI. Parameter values: $N = 100$, $M = 100$, $\nu = 0.1$, and $\sigma_1 = \sigma_2 = 0.1$.

$\widehat{a}$. The probability to find a certain configuration $\mathbf{s}$ in the k-th individual is $P^{\widehat{a}}_{\text{int}}(\mathbf{s}|\beta^k) \propto \exp\{-H^{\widehat{a}}_{\text{int}}(\mathbf{s}|\beta^k)\}$ with

$$H^{\widehat{a}}_{\text{int}}(\mathbf{s}|\beta^k) = -\beta^k \frac{1}{N} \sum_{i,j>i}^{N} a_{ij} s_i s_j. \tag{2.25}$$

- The system is iterated as in the co-evolutionary model, leaving $\widehat{a}$ fixed in time and identical for all individuals.

As the structure of $\widehat{a}$ is an arbitrary one, the calculation of the distances between distributions needs to be explicitly computed by summing over the 2^N possible states (which severely limits the maximum size in computer simulations; $N \sim 20$).

Results for the stationary distribution of parameters $p(\beta)$ and different types of network architectures, together with the corresponding curves of generalized susceptibilities computed as

$$\chi^{\widehat{a}}(\beta) = \left\langle \left(\frac{1}{N} \sum_{i,j>i}^{N} a_{ij} s_i s_j \right)^2 \right\rangle_{P^{\widehat{a}}(\mathbf{s}|\beta)} - \left\langle \frac{1}{N} \sum_{i,j>i}^{N} a_{ij} s_i s_j \right\rangle^2_{P^{\widehat{a}}(\mathbf{s}|\beta)} \tag{2.26}$$

are shown in Fig. 2.12. In all cases, the main result of this chapter holds: the resulting internal parameters distributions peak around the critical point.

Heterogeneity of the community

We find an interesting observation in Fig. 2.9 about the dependence on the mutations variance σ^2 of the resulting distribution: when this deviation increases, distributions generally become more peaked around the critical point. In principle, this conclusion could be confusing and paradoxical, but tells us that, when the progeny tends to diversify (i.e. more heterogeneous), the community shrinks to the critical point.

To analyze this effect, we focus in the simple one-parameter case (eq. 2.5). We study three different aspects of the stationary distribution of $p(\beta)$ depending on σ: *i)* The standard deviation of the stationary distribution, *ii)* the kurtosis, which gives a

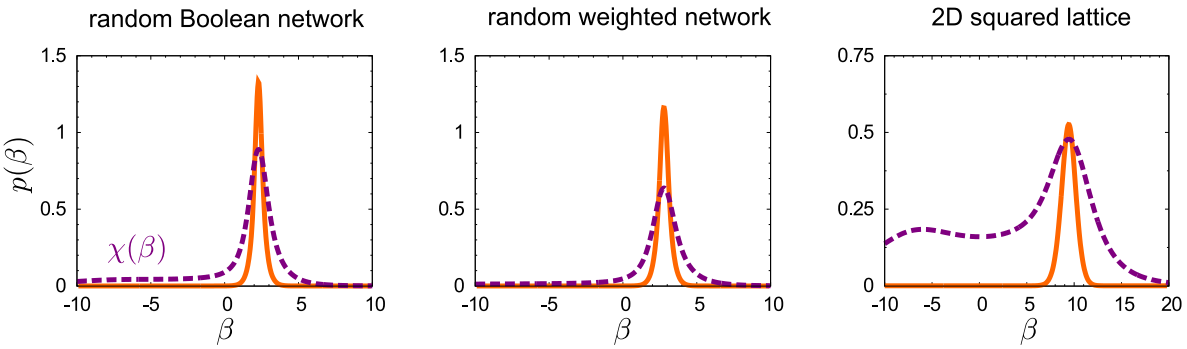


Figure 2.12: **Stationary parameter distributions for three different network structures $\widehat{a}$ of N nodes**: Three cases are studied: (left) Random Boolean network: connections are not weighted, i.e. $a_{ij} = a_{ji} = \{0, 1\}$, with mean connectivity $N/2$, (center) Random weighted network: in this case, $a_{ij} = a_{ji} = \eta$, where η is a random number between 0 and 1 (right) Regular 2D lattice with periodic bounding conditions. In all cases, parameters have been set to $\nu = 0.1$, $\sigma = 0.5$, $M = 100$, and $N = 20$ for the two left and central panesl, and $N = 25$ for right one. In dashed line, the FI has been plotted and re-scaled for visual comparison.

measurement of the "peakedness" of the given distribution, which is defined as

$$\text{kurtosis} = \frac{\langle \beta^4 \rangle - \langle \beta \rangle^4}{\langle \beta^2 \rangle - \langle \beta \rangle^2} - 3, \tag{2.27}$$

being null for a Gaussian distribution, positive if it is more peaked, and negative if it is flatter, and finally *iii)* the position of the 25th, 50th, 75th and 90th percentiles.

Results are summarized in Fig. 2.13. In particular, we see that most of the population –but not all– concentrate in the critical point when increasing σ, which can be seen for instance in the deviation where we get the 75% of the total population (75th percentile). Additionally, the distribution becomes sharper (as it is seen from the kurtosis). However, the standard deviation also increases because of highly mutated offspring.

Multiple body interactions

The co-evolutionary model has been introduced taking only two individuals at each time step. Are the previous results affected if a larger number of individuals is allowed to interact at the same time?

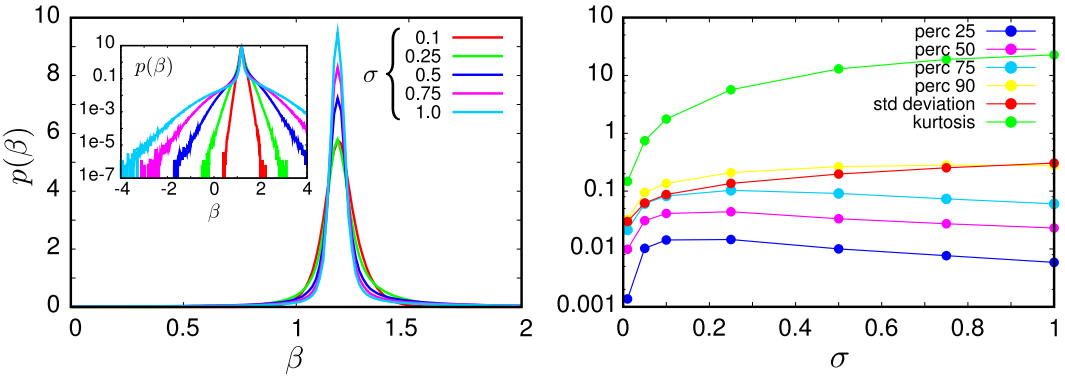


Figure 2.13: **Mutation variability enhances the approach to criticality in the co-evolutionary model** (results shown for the 1-parameter distribution of eq. 2.5): Most individuals concentrate in the critical point when increasing the mutation variance σ^2. This can be seen from the deviation where it is confined the 25%, 50% and 75% of the total population (dark blue, magenta and light blue curves), which shrinks for high values of σ. However, this variability produces also long tails (inset panel) due to the highly mutated offspring, and therefore, increasing the kurtosis.

To answer this question, here we study a variant of the co-evolutionary model in which K individuals $(i_1, ..., i_K)$ are randomly picked at each time step, and they compete among themselves; the probability to die is proportional to its (normalized) average KLD to the remaining ones, i.e.

$$P_{\text{kill}}(i_k) = \frac{\sum_{l=1}^{K} D(\boldsymbol{\beta}^{i_l}|\boldsymbol{\beta}^{i_k})}{\sum_{m=1}^{K}\sum_{l=1}^{K} D(\boldsymbol{\beta}^{i_l}|\boldsymbol{\beta}^{i_m})}, \qquad k = 1, ..., K, \tag{2.28}$$

and then it is replaced by a copy of one of the remaining $K-1$ individuals (and mutations are introduced with probability ν).

We have implemented the simulation with the simple parametrization of eq. 2.5 (Ising with one parameter). Results are summarized in Fig. 2.14 where we plot the mean value of the parameter over the entire community and 10^3 realizations of the same initial condition. It can be seen that the time to reach stationarity increases with K. When K increases, the drift which moves the system towards the criticality is lower. This is related to the fact that, by averaging over more and more individuals, the source

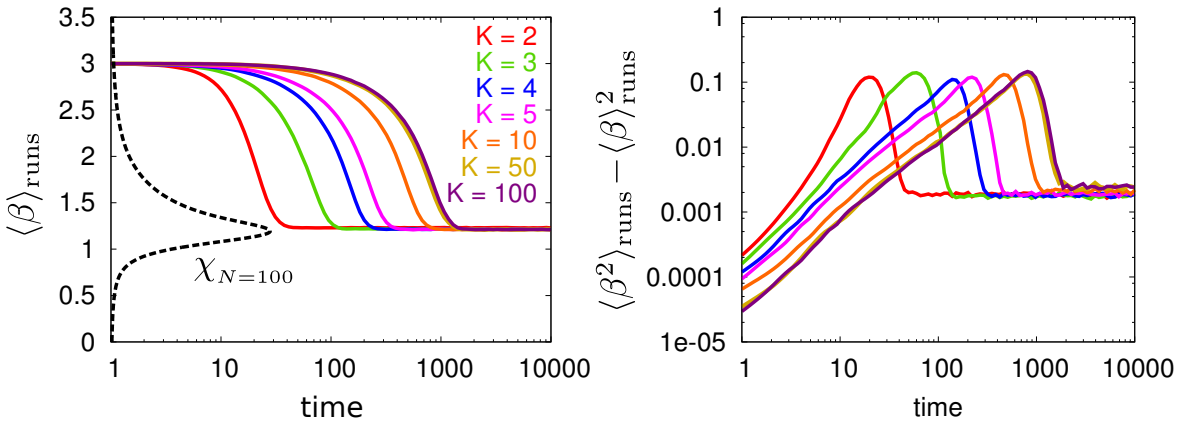


Figure 2.14: **Time evolution in the co-evolutionary model with K-body interactions** (with parametrization of eq. 2.5): Solid lines represent the time evolution of mean values $\langle\beta\rangle := \int d\beta\ p(\beta,t)\beta$. The relaxation to the stationary state depends on the effective number of individuals K with which each single agent interacts. The larger the value of K the larger the relaxation time. The community evolves very close to the maximum of the generalized susceptibility (plotted with dashed lines). The initial condition is $\beta = 3$ for all of the individuals and parameters are $N = 100$, $M = 100$, $\nu = 1$ and $\sigma = 0.1$.

effectively becomes more and more homogeneous.

2.3.3 Adaptive dynamics

We briefly show that our results are robust in the sense that criticality is also obtained in a adaptive –rather than evolutionary– type of dynamics. For this purpose, analogously to the previous evolutionary and co-evolutionary implementations, we introduce the "adaptive" and "co-adaptive" model.

Adaptive model

The adaptive model, in which a agents in a community of M individuals change their parameters to accommodate to an external environment, proceeds as follows:

1. As in the evolutionary models, every source is parametrized by eqs. 2.1 and 2.2 and the individuals by eqs. 2.3 and 2.4.

2. We start with M individuals whose parameters are distributed as $p(\boldsymbol{\beta}, t = 0)$.

Note that, in this model the individuals are completely independent, and having a big community is only useful in terms of the statistics.

3. At every time step, we generate S external sources from the pool $\rho_{\rm src}(\boldsymbol{\alpha})$.

4. At every time step, every individual can change its parameters by a small jump in one of them; among all the possibilities it chooses the one which minimizes the mean KLD to the sources:

$$\boldsymbol{\beta}^i \longrightarrow \arg\min_{\boldsymbol{\epsilon}\in\mathcal{E}} \frac{1}{S}\sum_{u=1}^{S} D(\boldsymbol{\alpha}^u|\boldsymbol{\beta}^i+\boldsymbol{\epsilon}), \quad \mathcal{E}=\{(\pm\epsilon_1,0,...0),(0,\pm\epsilon_2,0,...),...\} \quad (2.29)$$

5. With probability ν, we introduce some noise in the adaptation, and every individual can change its parameters with a Gaussian random fluctuation $\boldsymbol{\xi}$, with mean zero and variance $\boldsymbol{\sigma}$.

6. Time is incremented as $t \to t+1$. Another set of sources is generated from $\rho_{\rm src}(\boldsymbol{\alpha})$ and the process is iterated.

Fig. 2.15 shows the dynamics of the community with three different environments (the same ones as in the evolutionary simulation, Fig. 2.5, and similar parametrization). As in the analogous evolutionary model, heterogeneous environments lead to a tuning of parameters to criticality (top and middle panels), while very specific sources (bottom panels) do not.

Co-adaptive model

We have also developed the following co-adaptive model (counterpart of the co-evolutionary model):

1. We start with M individuals with parameters distributed as $p(\boldsymbol{\beta}, t=0)$. Here t refers to time.

2. At every time step, every individual can change its parameters by a small jump in one of them. Among all the possibilities, it chooses the one which minimizes

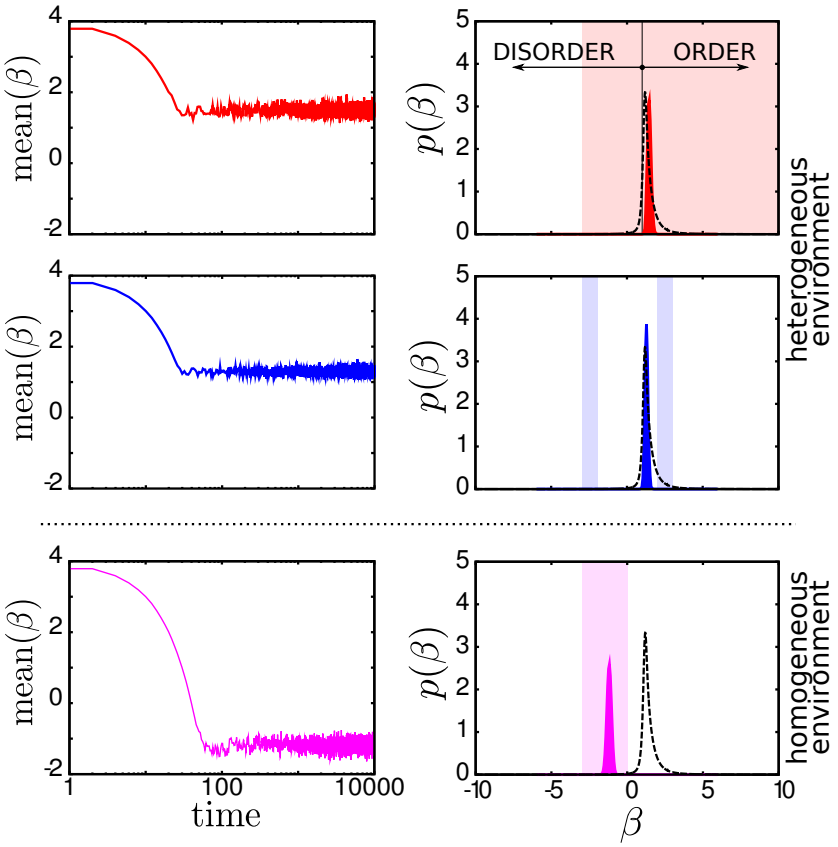


Figure 2.15: **Adaptive dynamics in the presence of varying external environments**: The figure shows the average of the internal parameters in the community and the stationary distributions for different environments. As obtained for the analogous evolutionary model, the stationary solution peaks at the critical point if the external environment is sufficiently heterogeneous, while it does not for simpler environments. Parameters, probability parametrization and initial conditions are similar to the ones used in Fig. 2.5 (same color code), and $\epsilon = 0.1$, $\nu = 0.3$.

the mean KLD to the rest of the community:

$$\boldsymbol{\beta}^i \longrightarrow \arg\min_{\boldsymbol{\epsilon}\in\mathcal{E}} \frac{1}{M-1}\sum_{j\neq i}^{M} D(\boldsymbol{\beta}^j|\boldsymbol{\beta}^i+\boldsymbol{\epsilon}), \quad \mathcal{E}=\{(\pm\epsilon_1,0,...0),(0,\pm\epsilon_2,0,...),...\} \tag{2.30}$$

3. We introduce some noise in the adaptation, and with probability ν every individual can change its parameter with a small random fluctuation $\boldsymbol{\xi}$, Gaussian

distributed with zero mean and deviation $\boldsymbol{\sigma}$.

4. Time is incremented as $t \to t+1$ and the process is iterated.

In Fig. 2.16 we show the evolution of the mean and variance of $\boldsymbol{\beta}$ in the community with three different initial conditions (same as in the co-evolutionary simulation, Fig. 2.8), together with the stationary distribution. Individuals cluster again in the vicinity of the critical point, and the result is independent of the initial conditions.

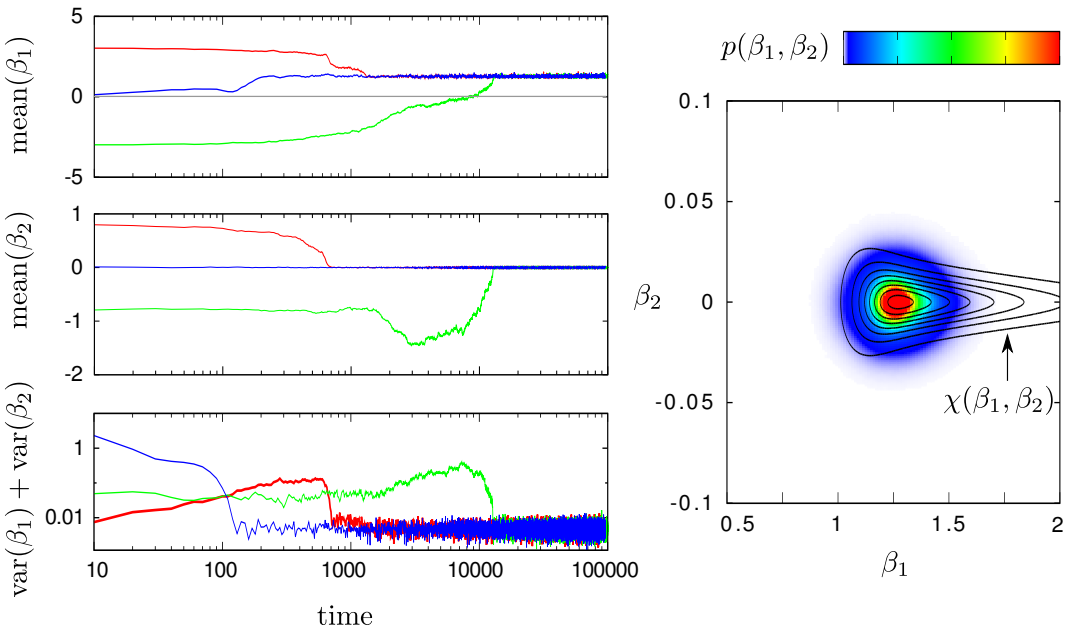


Figure 2.16: **Co-adaptive dynamics**. (Left) Time dependence of the first moments of the distribution of internal parameters. Regardless of the initial conditions, the community moves toward the critical point ($\beta_1 = 1$ and $\beta_2 = 0$) and reaches the same stationary solution as the corresponding co-evolutionary model. The stationary probability, on the right panel, shows a peak where the generalized susceptibility or FI peaks (see contour plots). Parameters, probability distribution parametrization, and initial conditions are similar to the ones used in Fig. 2.8. $\epsilon_1 = 0.1$, $\epsilon_2 = 0.01$, $\nu = 0.3$

2.4 Analytical results

In the light of the results obtained, we present different mathematical approaches to understand the previous findings. In general, our goal is to explicitly calculate the

optimal representation for the cases in which individuals cope with an heterogeneous environment or with themselves.

Given the external source $\boldsymbol{\alpha}$ (which could also represent another individual), the optimal map can be calculated by introducing the parametrization of P_{src} and P_{int} (eqs. 2.1, 2.2 and 2.3, 2.4, respectively) into the definition of the KLD (eq. 2.8); then, deriving in $\boldsymbol{\beta}$ to find the minimum:

$$0 = \frac{\partial}{\partial \beta_\mu} D(\boldsymbol{\alpha}|\boldsymbol{\beta}) = -\langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\beta}} + \langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\alpha}} \Longrightarrow \langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\beta}} = \langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\alpha}}, \tag{2.31}$$

where the index μ runs for different internal parameters ($\mu = 1, ..., I$). This equation, which implicitly defines the optimal internal parameters, has an intuitive interpretation: the minimum of the KLD is obtained when the first I moments of the distribution $P_{\text{int}}(\mathbf{s}|\boldsymbol{\beta})$ exactly match those measured from $P_{\text{src}}(\mathbf{s}|\boldsymbol{\alpha})$.

To check whether extrema are minima or maxima, one needs to evaluate the Hessian matrix at the stationary point,

$$\frac{\partial^2}{\partial \beta_\mu \partial \beta_\nu} D(\boldsymbol{\alpha}|\boldsymbol{\beta}) = -\frac{\partial}{\partial \beta_\nu} \langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\beta}} = \langle \phi^\mu_{\text{int}} \phi^\nu_{\text{int}} \rangle_{\boldsymbol{\beta}} - \langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\beta}} \langle \phi^\nu_{\text{int}} \rangle_{\boldsymbol{\beta}} = \chi^{\mu\nu}(\boldsymbol{\beta}), \tag{2.32}$$

which coincides with the generalized susceptibility defined in eq. 2.19. As the FI matrix is a positive-definite matrix (excluding the trivial case of a factorized $P_{\text{int}}(\mathbf{s}|\boldsymbol{\beta})$ for which it vanishes), if the solution of eq. 2.31 exists, it corresponds to a minimum of the KLD.

Similarly for the case of multiple sources, we can optimize $d(\rho_{\text{src}}|\boldsymbol{\beta})$, and the equation is:

$$\langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\beta}} = \int d\boldsymbol{\alpha}\ \rho_{\text{src}}(\boldsymbol{\alpha})\ \langle \phi^\mu_{\text{int}} \rangle_{\boldsymbol{\alpha}}. \tag{2.33}$$

Here the system adopts the single value of $\boldsymbol{\beta}$ that better describes *on average* the varying environment.

2.4.1 External heterogeneous environment

We consider the simple example in which both the sources and the system are characterized by a single parameter ($E = I = 1$) and same parametrization ($\phi_{\text{src}} = \phi_{\text{int}} = \phi$), and we assume that a phase transition occurs at some parameter value $\alpha = \alpha_c$, i.e. the

observable $\langle\phi\rangle_\alpha$ has a sigmoid shape (which becomes steeper as N increases) with an inflection point at $\alpha = \alpha_c$ (our analysis can be extended to more general cases where there is no built-in phase transition in the source distributions but they are merely sufficiently heterogeneous). The two plateaus of the sigmoid function correspond to the so-called *disordered* and *ordered* phases, respectively. When $\rho_{\rm src}(\alpha)$ has support on both sides of the sigmoid function, the distribution of moments $\varrho_{\rm src}(\langle\phi\rangle)$ (eq. 2.22) is broad, and then we say that the sources are "heterogeneous".

By solving the equation for the optimal β (eq. 2.33), i.e. the one with $\langle\phi\rangle_\beta$ equal to the average moment on the sources, it is obvious that the moment to be reproduced lies somewhere in between the two asymptotic values of the sigmoid with the values of β for which intermediate moments are concentrated near the inflection or critical point, α_c. Indeed, as $\chi = -\frac{d<\phi>_\beta}{d\beta}$, the critical region, where the generalized susceptibility χ has a peak, is the region of maximal variability in which different complex sources can be best accounted for through small parameter changes, in agreement with the finding that many more distinguishable outputs can be reproduced by models poised close to criticality [89].

We illustrate this argument with a particular example in Fig. 2.17. In this case, both the sources and the internal representations are modeled by eq. 2.5, each one containing only one parameter. We considered $\rho_{\rm src}(\alpha)$ uniform in the range $(-10, 10)$. We see that the minimum average KLD to the sources appears where the FI maximizes, i.e. at the critical point.

2.4.2 Self-organized environment

Let consider two individual agents A and B, where the source for A is B and vice versa. The relative fitnesses of A and B are determined by how well the set of cues (described by the probability distribution $P_{\rm src}$) of one individual is captured by the other with minimum information loss, and vice versa. For simplicity, let start by the simplest case of single-parameter distributions. If $\beta_A = \beta_B$, the two distributions would be identical and the KLD would vanish. However, this is not a stable solution. Indeed, if the two parameters are not identical but close ($\beta_A \simeq \beta_B$), as the KLD is not symmetric, one of the individuals has a larger fitness. Actually, introducing the general parametrizations

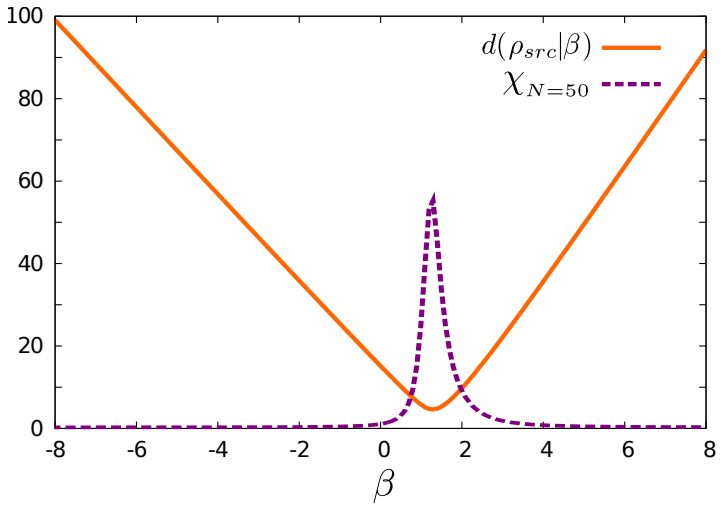


Figure 2.17: **Average KLD to the sources and optimal representation**. The orange line is the averaged KLD, $d(\rho_{\text{src}}|\beta)$, of the internal parameters β (eq. 2.11). The dashed line is proportional to the FI (eq. 2.19). The minimum of the distance is located close to the critical point (characterized by the peak of the FI). In this case both the sources and the internal representations are modeled by eq. 2.5 (one-parameter Ising) with $N = 50$,source parameters are uniformly distributed in the range $[-10, 10]$.

of eqs. 2.3, 2.4 in the definition of the KLD, it is possible to compute the difference

$$D(\beta_A + \delta\beta|\beta_A) - D(\beta_A|\beta_A + \delta\beta) = \frac{1}{6}\nabla\chi(\beta_A)\delta\beta^3 + \mathcal{O}(\delta\beta^4) \tag{2.34}$$

where $\nabla\chi$ is the gradient of the generalized susceptibility and $\delta\beta = \beta_B - \beta_A$. This implies that the individual whose parameter corresponds to the state with larger χ has a smaller KLD and is thus fitter. As χ peaks at the critical point, thus our key finding is that, for a family of individuals with similar parameters, the fittest possible agent sits exactly at criticality, and it is best able to encapsulate a wide variety of distributions.

We can illustrate the mechanism with a different approach. We consider an individual, characterized by the parameters $\boldsymbol{\beta}$, coping with similar –but not equal– individuals playing the role of the sources, characterized by the distribution $\rho_{\text{src}}(\boldsymbol{\alpha})$. Then we study how the distance to the critical point is modified by the internal representation. That is, we take a narrow distribution of $\boldsymbol{\alpha}$ values, centered at some average value $\bar{\boldsymbol{\alpha}}$ at some distance to criticality, and we wonder how does the distance to criticality change when the optimal internal representation is constructed. Does it grow or does it generically

diminish?

For this purpose, we take a distribution $\rho_{\rm src}(\boldsymbol{\alpha})$ which is different from zero only in a small region U. We can therefore expand both of the left and the right hand sides of eq. 2.33 around the mean value $\bar{\boldsymbol{\alpha}} := \int_U d\boldsymbol{\alpha}\, \boldsymbol{\alpha}\rho_{\rm src}(\boldsymbol{\alpha})$, and then we obtain

$$(\beta_\mu - \bar{\alpha}_\mu) = \left(\chi^{-1}\right)^{\mu\nu}(\bar{\boldsymbol{\alpha}})\frac{1}{2}\,\frac{\partial}{\partial\alpha_\nu}\chi^{\gamma\delta}(\boldsymbol{\alpha})\bigg|_{\bar{\boldsymbol{\alpha}}}\int_U d\boldsymbol{\alpha}'\rho_{\rm src}(\boldsymbol{\alpha}')(\alpha'_\gamma - \bar{\alpha}_\gamma)(\alpha'_\delta - \bar{\alpha}_\delta)\ , \tag{2.35}$$

where $\chi^{\mu\nu}(\boldsymbol{\alpha})$ are the elements of the Fisher information matrix (eq. 2.19). This is an equation for the deviation of the internal parameter $\boldsymbol{\beta}$ respect to $\bar{\boldsymbol{\alpha}}$. To understand its relation with the critical point, we rewrite eq. 2.35 as

$$(\boldsymbol{\beta} - \bar{\boldsymbol{\alpha}}) = \chi^{-1}(\bar{\boldsymbol{\alpha}})\nabla\Omega(\bar{\boldsymbol{\alpha}}), \tag{2.36}$$

where we have defined the scalar field $\Omega(\boldsymbol{\alpha})$ as

$$\Omega(\boldsymbol{\alpha}) := \frac{1}{2}\int_U d\boldsymbol{\alpha}'\ \rho_{\rm src}(\boldsymbol{\alpha}')(\boldsymbol{\alpha}' - \bar{\boldsymbol{\alpha}})^T\cdot\chi(\boldsymbol{\alpha})\cdot(\boldsymbol{\alpha}' - \bar{\boldsymbol{\alpha}})\ . \tag{2.37}$$

As $\chi(\boldsymbol{\alpha})$ is a positive-definite matrix, from eq. 2.37 we see that $\Omega(\boldsymbol{\alpha})$ is a positive quantity. If we introduce a base of eigenvectors of $\chi(\boldsymbol{\alpha})$, $\mathbf{v}^\gamma(\boldsymbol{\alpha})$, with eigenvalues $\lambda^\gamma(\boldsymbol{\alpha})$, we get

$$\Omega(\boldsymbol{\alpha}) = \frac{1}{2}\lambda^\gamma(\boldsymbol{\alpha})\int_U d\boldsymbol{\alpha}'\ \rho_{\rm src}(\boldsymbol{\alpha}')\left[\mathbf{v}^\gamma(\boldsymbol{\alpha})^T\cdot(\boldsymbol{\alpha}' - \bar{\boldsymbol{\alpha}})\right]^2\ . \tag{2.38}$$

At the critical point, at least one $\lambda^\gamma(\boldsymbol{\alpha})$ diverges (in the thermodynamic limit), so $\Omega(\boldsymbol{\alpha})$ has a maximum at the critical point. Note that $\mathbf{v}^\gamma(\boldsymbol{\alpha})$ cannot vanish because it is an unitary vector. Therefore, the gradient of Ω points to the critical point, at least if $\bar{\boldsymbol{\alpha}}$ is not too far from it in a such a way that there are not other local maxima to which the gradient could be pointing.

Finally, we project both sides of eq. 2.36 over the gradient of $\Omega(\bar{\boldsymbol{\alpha}})$:

$$(\nabla\Omega(\bar{\boldsymbol{\alpha}}))^T\cdot(\boldsymbol{\beta} - \bar{\boldsymbol{\alpha}}) = (\nabla\Omega(\bar{\boldsymbol{\alpha}}))^T\cdot\chi^{-1}(\bar{\boldsymbol{\alpha}})\cdot\nabla\Omega(\bar{\boldsymbol{\alpha}})\ . \tag{2.39}$$

As χ is positive-definite, also its inverse χ^{-1} is a positive-definite matrix, and the projection of $(\boldsymbol{\beta} - \bar{\boldsymbol{\alpha}})$ on the gradient of $\Omega(\bar{\boldsymbol{\alpha}})$ (which points to the critical point) is

also positive.

Consequently, *the internal map* $\rho_{\text{src}}(\boldsymbol{\alpha}) \longrightarrow \boldsymbol{\beta}$ *is closer to the critical point than* $\bar{\boldsymbol{\alpha}}$, indicating that there is an overall drift towards parameter regions with larger Fisher information. After iterating the same mapping for all the systems in the community, they converge to the critical point.

To give and example, we can consider a system characterized by only one parameter. As above, we want to find the optimal mapping β which solves eq. 2.33. We consider a general functional form for $\langle\phi\rangle_\alpha$ (omitting the sub-indexes in $\phi_{\text{src/int}}$) with a phase transition in $\alpha = \alpha_c$: at a finite size, $\left\langle\phi\right\rangle_\alpha$ has a sigmoid shape, with the inflection point (i.e. the local maximum/minimum of its derivative) corresponding to the transition point α_c. Suppose that $\rho_{\text{src}}(\alpha)$ is uniform in an interval between $\bar{\alpha} - a$ and $\bar{\alpha} + a$; we want to compare the distance $|\bar{\alpha} - \alpha_c|$ with $|\beta - \alpha_c|$, where β is the optimal value obtained with eq. 2.33. In the limit of small interval a, we obtain

$$\left\langle\phi\right\rangle_\beta = \int_{\bar{\alpha}-a}^{\bar{\alpha}+a} d\alpha \frac{1}{2a}\left\langle\phi\right\rangle_\alpha \sim \left\langle\phi\right\rangle_{\bar{\alpha}} + \left\langle\phi\right\rangle_{\bar{\alpha}}^{''} \frac{a^2}{6} \ , \tag{2.40}$$

where $\left\langle\phi\right\rangle_{\bar{\alpha}}^{''}$ is the second derivative of $\left\langle\phi\right\rangle_\alpha$ evaluated in $\bar{\alpha}$. Expanding the left hand side of the equation around $\bar{\alpha}$ we obtain

$$\beta - \bar{\alpha} \sim \frac{\left\langle\phi\right\rangle_{\bar{\alpha}}^{''}}{\left\langle\phi\right\rangle_{\bar{\alpha}}^{'}} \frac{a^2}{6} \ . \tag{2.41}$$

Suppose now that the sigmoid has a positive (negative) derivative, then the second derivative will be positive (negative) below the critical point and negative (positive) above, giving β always closer to the inflection point (i.e. the critical point). Note that the larger the value of a (i.e. more heterogeneous the distribution $\rho_{src}(\alpha)$), the closer β to the critical point.

In conclusion, both mathematical approaches lead to the same result: in the scenario where individuals play the role of the sources for the others and co-evolve to improve their mutual knowledge, the community dynamically self-organizes around the critical point. Consequently, the critical point arises as the stable attractor of such information-based interaction.

2.5 Chapter summary

Summing up, in this chapter we have focused on the hypothesis that some aspect of living systems may operate at critical points. We have addressed this question by employing different analytical and computational tools, relying on concepts from statistical mechanics and information theory. Under the mild assumption that living systems need to construct good, although approximate, internal representations of the complex world, and that such good maps provide a crucial competitive advantage, we have shown for different scenarios that individuals evolve/adapt to a critical point –if there is the possibility– in order to efficiently cope with a complex environment [62]. In particular:

- We have introduced the mathematical framework: sources can only be probabilistically perceived by the individuals, but such representations might not be accessible to the systems by tuning their internal structure. The optimal map is given by the minimization of Kullback-Leibler divergence between the internal and source distributions. We have also characterized the critical point in the space of possible representations with the maximum of the Fisher information.

- We have implement an evolutionary model in which fitter individuals, i.e. those with better maps, are selected to evolve in a genetic algorithm. At each time, individuals are exposed to randomly generated sources. We show that, when the the pool of sources is highly heterogeneous, individuals evolve around the critical point, while the do not when the environment is very predictable and specific.

- We have given a more accurate definition of the heterogeneity/homogeneity of the environment. Moreover, we have checked that this tuning to criticality in the community is not, in general, a byproduct of the environment itself being critical.

- A more robust convergence to criticality is obtained for a co-evolutionary model, in which the environment is not externally imposed, but it is composed of the community itself, i.e. individuals play the role of sources for the others. Surprisingly enough, even if the only goal of individuals is to "understand" each other with precision, the community turns out to be critical, thus becoming highly complex.

- We have checked that our results are robust against different parametrizations of the internal systems presenting a critical point, as well as in a variant of the model in which many individuals interact simultaneously at each time.

- Similar results, for both externally and self-consistently introduced environments, are obtained with analogous adaptive and co-adaptive models, in which individuals, rather than following a genetic algorithm, dynamically tune their parameters to improve their fitness.

- We have developed several mathematical analyses that illustrate the previous mechanisms. In particular for the co-evolutionary/co-adaptive dynamics, the critical point emerges as the evolutionarily stable strategy, where the non-symmetric nature of interactions turns out to be a key ingredient.

Our results can be applied for general complex adaptive systems, but in particular, they could be related to populations in which a huge phenotypic variability has been empirically observed, for instance in some bacterial communities [81] and viral populations [129]. Such diversification represents and adaptive survival strategy when the environment is highly unpredictable, in order to minimize the long-term risk of extinction [144, 145]. In our framework, it constitutes a direct consequence of a community being critical. In the following chapters, we study this mechanism in the context of population dynamics, popularly called "bet-hedging" in analogy with stock-market portfolio management.

Chapter 3

Bet-hedging in population dynamics

3.1 *'Bet hedging or not?'*

In natural environments, individuals have to choose between a variety of different evolutionary strategies, characterized by different time-dependent payoffs and risks. Particularly relevant is the case in which the choice is between a relatively safe strategy, with low but stable payoff, and a variable one, potentially very productive, but risky. An example of this are micro-organisms able to metabolize two different resources [100, 134, 40, 125]: one consistently available at a fixed though low level, and a second, more abundant on average but fluctuating in time.

The lack of knowledge about environmental conditions forces individuals to make a blind decision on whether to specialize into exploiting one or the other, or instead to develop a hybrid "bet-hedging" strategy, alternating both routes, either stochastically or following information from sensory systems [120, 81]. Hybrid strategies can be also exploited at a community level, for example by means of phenotypic variability [81, 125].

The concept of bet-hedging was first formalized in the context of information theory [76] and portfolio management [46]. Later, it was conjectured that living organisms may decrease their risk in unpredictable environments by developing bet-hedging strategies [120, 140, 30, 37, 69], as it has been empirically confirmed in bacterial and viral communities [129, 144, 145, 134, 40, 125, 58], in insects [65], and in a wealth of other examples in population ecology, microbiology, and evolutionary biology [140].

Given their ubiquity, bet-hedging strategies have attracted a lot of interest from the

perspective of evolutionary game theory [124, 106]. An interesting and non-trivial result in this context is the so-called Parrondo's paradox [57, 110] in which an alternation of two distinct losing strategies can lead to a winning one. Most of these studies, including application in population dynamics [140], rely on mean-field analyses describing well-mixed communities.

In this chapter, we study optimal bet-hedging strategies in the search of general patterns and conditions. For this, we employ a parsimonious approach relying on simple individual-based stochastic modeling.

3.2 Contact Process with hybrid dynamics

Our starting point is the simplest possible birth-death stochastic model on a lattice, i.e. the contact process (CP) [88, 59]. Individuals occupy a square lattice or network, with at most one individual per site (see fig. 3.1, left), and at every discrete time step, each individual can either:

- With probability p, produce an offspring at a randomly chosen neighboring site, provided it was empty,
- or, complementary with probability $1 - p$, die and be removed from the system.

This simple dynamics can either generate an active phase characterized by a non-vanishing density of individuals or, alternatively, lead ineluctably to the absorbing state in which the population is extinct, depending on the value of p. A critical point, p_c, separates these two different phases [88, 59] (see fig. 3.2).

We consider a variant in which individuals can choose between two strategies (see fig. 3.1, right):

- A *conservative* one, corresponding to exploitation of a constantly available resource.
- A *risky* one, exploiting a variable/unpredictable resource.

These two alternatives affect the value of p experienced by each individual. The conservative strategy corresponds to a CP in which p is kept constant at a relatively low value,

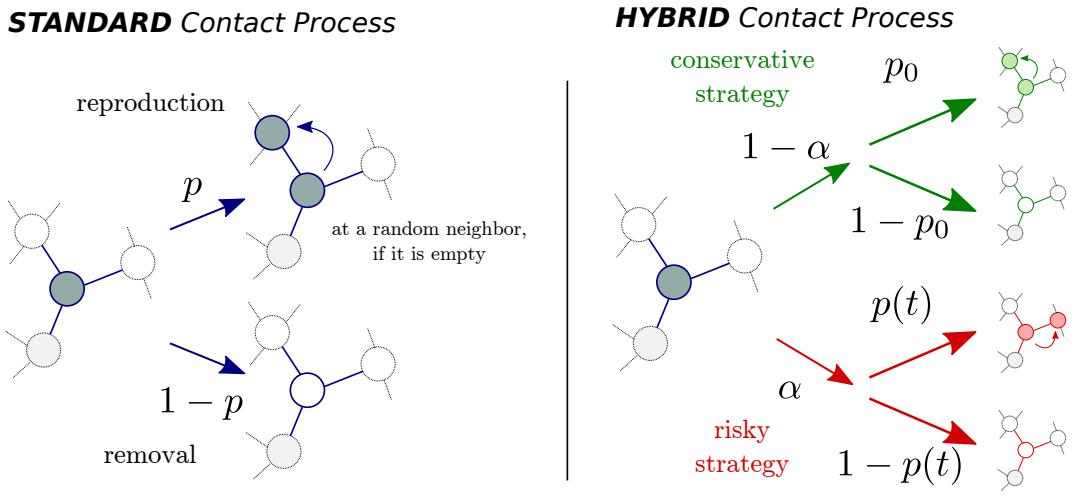


Figure 3.1: **(Left) Dynamics in a standard Contact Process**: each particle in the lattice either i) produces a new one, with prob. p, on a random neighboring site (provided it was empty) or ii) is removed from the lattice, with prob $1-p$. **(Right) Contact Process with a hybrid dynamics**: at each time, every particle chooses between the conservative (with prob. $1-\alpha$) and the risky (with prob. $1-\alpha$) spreading strategy. The conservative dynamics is characterized by a constant, relatively low spreading prob. p_0, while the risky one has a variable and unpredictable prob. $p(t)$.

p_0. In the risky strategy, demographic probabilities depend upon variable external conditions, i.e. $p = p(t)$, where $p(t)$ is a random noise –common to all individuals choosing the risky strategy in the community– drawn at every time-step. Individuals can hedge their bets by randomly choosing between the conservative and the risky strategy at each time step.

This choice is controlled by a *risk parameter* α which acts as a control parameter: with probability α each individual chooses the risky strategy and with probability $1-\alpha$ the conservative one. In the language of game theory, $\alpha = 0$ and $\alpha = 1$ are "pure strategies" and the range $0 < \alpha < 1$ describes a set of "hybrid strategies".

We implement the dynamics with *synchronous* updating[1] [59] in different network topologies: one- two and three-dimensional regular lattices and fully-connected (FC) network; after four dimensions, the CP is known to present similar properties to the FC network [88], constituting the so-called critical dimension. While otherwise stated,

[1]In the literature, however, the name "Contact Process" is usually identified with an *asynchronous* updating rule, as an instance of the SIS model.

α is kept as a fixed control parameter. Simulations are initialized with a fully occupied configuration; the dynamics proceeds as follows:

1. At every step, a fresh value of $p(t)$ is extracted from a Gaussian distribution $N(\bar{p}, \sigma^2)$ (this choice will permit, later, an easier treatment in a continuous stochastic theory). For large variance σ^2, it can happen that $p(t) < 0$ or $p(t) > 1$; in this case we simply enforce $p(t) \to 0, 1$, respectively (actually, we make $p(t) < 0 \to p(t) = \epsilon$, as we explain later).

2. The network/lattice is updated synchronously; with probabilities α and $1 - \alpha$, each individual selects the risky or the conservative strategy respectively.

3. Each individual either reproduces or dies with the corresponding probabilities; all dying individuals are removed from the system and afterward offspring are placed at random neighbors of their corresponding parents, keeping the constraint of up to one individual per site.

4. Finally, time is incremented in one unit and the process is iterated.

Before going the next section, let us clarify some technicalities about the computational implementation:

- The synchronous updating presents some ambiguities; for instance, if a particle occupying node 1 tries to produce an offspring on node 2, but, parallely, there is a particle at node 2 which has decided to die, what is the state of node 2 at the next time step? We choose to favor the activity (node 2 would stay active at the next time step in the previous example), but the general features of the model are insensitive to this choice. We do not use the standard *asynchronous* updating, more common in simulations of the CP, because within this choice, as the spreading probability depends on time ($p = p(t)$), one has to implement the Gillespie's algorithm [53] with time-dependent rates, becoming a cumbersome task [2].

- Additionally, we introduce a cut-off in $p(t) \geq \epsilon$. This numerical trick only changes slightly the kind of distribution we are using, and avoids deterministic global extinctions (because the removal prob. is $1 - p(t)$) after a characteristic number

of steps $\sim \text{Prob}(p(t) \leq 0)^{-1}$. Note that this undesirable effect would appear also in the "active" phase at the thermodynamic limit $N \to \infty$, and therefore the introduction of a cut-off becomes essential in order to compute mean extinction times when the Gaussian variance is large.

- Finally, note that, unless we are interested in measuring the mean extinction time, there always exists the possibility to fall into the absorbing state for any finite system (even for the active phase and after the introduction of the cut-off ϵ). This probability decreases exponentially for larger systems size, but it can happen, and it actually do. Therefore, any finite system has to be treated under the quasi-stationary approximation, which can be implemented in several manners [52, 43, 104]: either i) we restrict our measurements to survival runs of the dynamics, or, more easily, ii) we reactivate the system every time the activity dies by introducing just one particle (equivalent to include a low external drift). The last solution can be modified to provide exact solutions [42], but it is more than enough for many purposes.

3.3 Phases of the pure strategies

We first analyze the phases of the CP with one of the pure strategies, either $\alpha = 0$ or $\alpha = 1$. The phase diagram for the risky strategy ($\alpha = 1$) is shown in fig. 3.2 for different network dimensions as a function of the parameters $(\bar{p}, \sigma^2)$; note that $\alpha = 0$ corresponds to a risky strategy with $\sigma^2 = 0$. A critical line separates the absorbing phase (zero density) and the active phase (positive density) in each dimension. Let us remark that:

- For each network dimension, and fixing the mean reproducing probability $\bar{p}$, external noise variance tends to destroy the activity, leading to the absorbing phase for high enough values of σ^2 [133].

- For a fixed external variability σ^2, the critical point $\bar{p}_c$ is shifted to higher values of $\bar{p}$ for lower dimensions, coinciding with the common idea that *intrinsic* noise, larger in low dimensions, constitutes a negative effect for the activity.

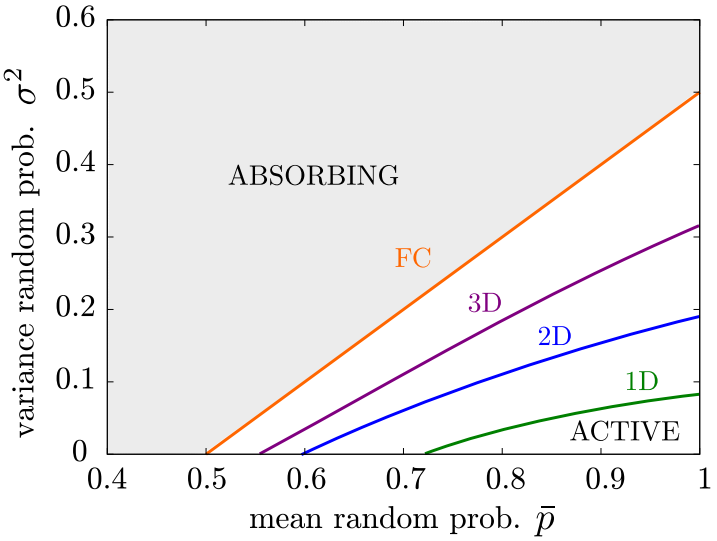


Figure 3.2: **Phase diagram of the CP with time-dependent rates:** the spreading probability per site, $p(t)$, is a random, non-correlated in time variable distributed as a Gaussian $N(\bar{p}, \sigma^2)$; complementary, the probability of a particle to be removed is $1-p(t)$. The standard CP, $p(t) = p_0$, constitutes the particular case with zero variance (Bottom axis). A critical line separates the active phase (non vanishing density) from the absorbing phase (null density), being this region larger for lower dimensions. System sizes are $N = 10000$ for 1D, 2D and FC lattices and $N = 10648$ in the 3D case.

Recent studies [133] has noted that, within the active region, a special phase, called *temporal Griffiths phase*, emerges when temporal disorder ($\sigma^2 > 0$) is introduced in the parameters; there, the system presents scaling properties for a whole range of parameters, as if the critical point were stretched. We have made no attempt to distinguish such a phase from the active one in our analysis.

3.4 Extinction times and Parrondo's paradox

In well-mixed systems it is known that the alternation of two bad strategies can outperform a better one (see e.g. [76, 69, 140]). We describe the underlying mechanism with a simple continuous-time mean-field calculation [52, 88, 59]. In the literature, however, this effect is typically explained using discrete-time multiplicative processes; its discrete analogy is discussed in Appendix C.

We consider the case $p(t) = \bar{p} + \sigma\xi(t)$ where $\xi(t)$ is a delta-correlated Gaussian white

noise. Given the presence of an external global noise, even in the large system-size limit ($N \to \infty$) the mean-field dynamics needs to be described by a stochastic equation, and the density of individuals, $\rho(t)$, obeys the following Langevin equation:

$$\begin{aligned}\dot{\rho}(t) &= \alpha\left[(\bar{p}+\sigma\xi(t))\,\rho\,(1-\rho)-(1-\bar{p}-\sigma\xi(t))\,\rho\right]+\\ &\quad (1-\alpha)\left[p_0\rho\,(1-\rho)-(1-p_0)\,\rho\right].\end{aligned} \tag{3.1}$$

Defining the *average spreading probability*,

$$p_{\rm av}(\alpha) = \alpha\bar{p} + (1-\alpha)p_0, \tag{3.2}$$

the equation is simply written as

$$\dot{\rho}(t) = (2p_{\rm av}(\alpha)-1)\,\rho - p_{\rm av}(\alpha)\rho^2 + 2\alpha\sigma\rho\left(1-\frac{\rho}{2}\right)\xi(t). \tag{3.3}$$

Up to leading order, we can neglect the quadratic saturation effects and higher order environmental-noise terms (valid for $\rho \ll 1$), obtaining, the linear Langevin equation

$$\dot{\rho} \approx (2p_{\rm av}(\alpha)-1)\,\rho + 2\sigma\alpha\rho\xi(t). \tag{3.4}$$

Changing variables (using Itô calculus) to $y = \log(\rho)$ and averaging over realizations $\langle\cdot\rangle$, eq. 3.4 becomes $\frac{d}{dt}\langle\log\rho\rangle = G(\alpha)$ with the *exponential growth rate*

$$G(\alpha) = -2\sigma^2\alpha^2 + 2p_{\rm av}(\alpha) - 1, \tag{3.5}$$

whose sign determines whether the population tends to grow or to shrink, allowing to distinguish between the active and the absorbing phase, respectively.

The growing rate is a parabola in α with negative quadratic coefficient. Therefore there exists an optimal strategy α^*, that can be found by maximizing $G(\alpha)$ in the range $[0,1]$; the result is either a pure or a hybrid one depending on parameter values:

$$\alpha^* = \begin{cases} 0 & \bar{p} < p_0 \\ \dfrac{\bar{p}-p_0}{2\sigma^2} & p_0 < \bar{p} < p_0 + 2\sigma^2 \\ 1 & \bar{p} > p_0 + 2\sigma^2. \end{cases} \tag{3.6}$$

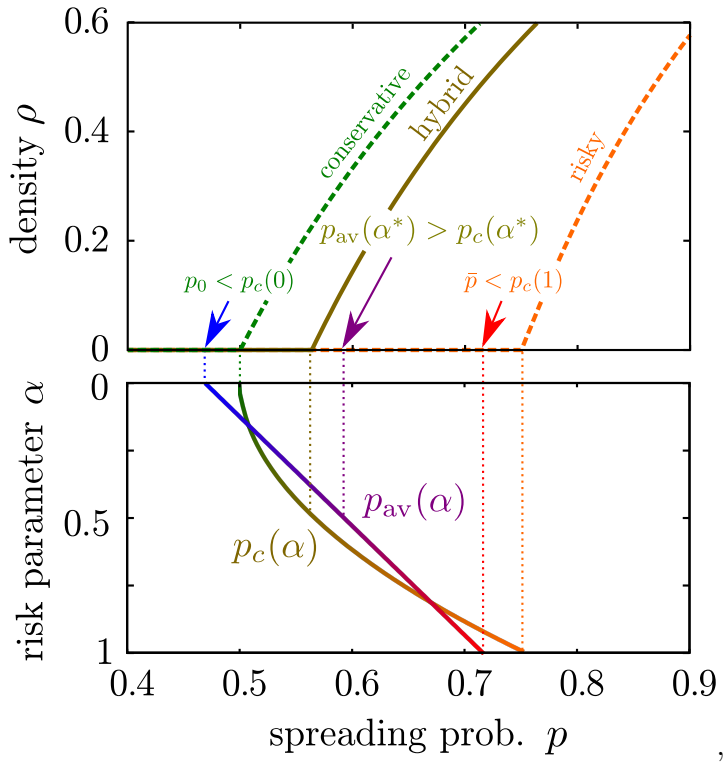


Figure 3.3: Upper panel: phase diagram for risky, conservative, and hybrid strategies. The critical points for the pure strategies lie at $p_c(0) = 1/2$ and $p_c(1) = 1/2 + \sigma^2$, respectively. (Lower panel) The average value of p for a hybrid strategy, defined by the risk parameter α, interpolates linearly between p_0 and $\bar{p}$ (blue/red line), while its critical point is given by a quadratic interpolation (green/orange line). Consequently, combining two subcritical values of p_0 and $\bar{p}$ (marked by arrows) with an intermediate value of α between the two line intersections, the resulting hybrid strategy can be supercritical. A similar enhancement of the stationary density is obtained generically, independently of whether p_0 and $\bar{p}$ are subcritical.

As the critical point occurs for $G(\alpha) = 0$, then

$$p_c(\alpha) = \frac{1}{2} + \sigma^2\alpha^2, \tag{3.7}$$

interpolating quadratically between the critical points of the pure strategies ($p_c(0) = \frac{1}{2}$ and $p_c(1) = \frac{1}{2} + \sigma^2$). On the other hand, the average spreading probability is a linear-in-α interpolation between the two limiting pure values (eq. 3.2) . As illustrated in Fig. 3.3, these two different types of behavior open the possibility for supercritical behavior, $p_{\rm av}(\alpha) > p_c(\alpha)$, even when both pure strategies, p_0 and $\bar{p}$, are subcritical. This

Condition	**Constraint**
$G(0) < 0$	$p_0 < 1/2$
$G(1) < 0$	$\bar{p} < 1/2 + \sigma^2$
$G(\alpha^*) > 0$	$\sigma^2 < \dfrac{(\bar{p} - p_0)^2}{2(1 - 2p_0)}$
$0 \leq \alpha^* \leq 1$	$p_0 \leq \bar{p} \leq p_0 + 2\sigma^2$

Table 3.1: **Parameter constraints for the Parrodox's paradox** obtained from the mean-field analysis.

paradoxical effect represents an instance of the Parrodox's paradox [57], who illustrated this mechanism in the study of Brownian ratchets. In a nutshell, the Parrondo's paradox states that *there exist pairs of games, each with a higher probability of losing than winning, for which it is possible to construct a winning strategy by playing the games alternately.* The set of constraints required for such a situation is summarized in table 3.1, and fig. 3.4 illustrates the region in the space of parameters.

In the case in which the two pure strategies are active, the same argument shows that a larger density can be achieved by hybrid strategies. In the following we consider both types of pure strategies, absorbing and active, and try to quantify the gain induced by bet-hedging.

Fixing the parameters $(p_0, \bar{p}, \sigma)$ to poise the respective pure strategies with $\alpha = 0, 1$ in the absorbing phase, we look for values of α for which hybrid strategies are active. For this, we measure the mean extinction time, τ, as a function of the system size N, for the 1D, 2D, 3D lattices and the FC network. Observe that, owing to fluctuations, any finite system is condemned to end up in the absorbing state. However, its mean lifetime increases exponentially with N, $\tau \sim \exp(N)$ in the active phase ([52, 133]; in the presence of variable parameters, there can also be a slower-than-exponential power-law scaling of $\tau(N)$ within the active temporal Griffiths phase [133]), making the system stable in the large-N limit. On the other hand, a logarithmic increase, $\tau \sim \log(N)$ is expected in the absorbing phase [52, 133].

As shown in fig. 3.5 for different values of α, τ grows with system size very slowly (logarithmically) for the two different pure dynamics ($\alpha = 0, 1$), while it increases exponentially for some hybrid strategies. We have made no attempt here to accurately determine the values of α delimiting the active phase for each dimension, but just

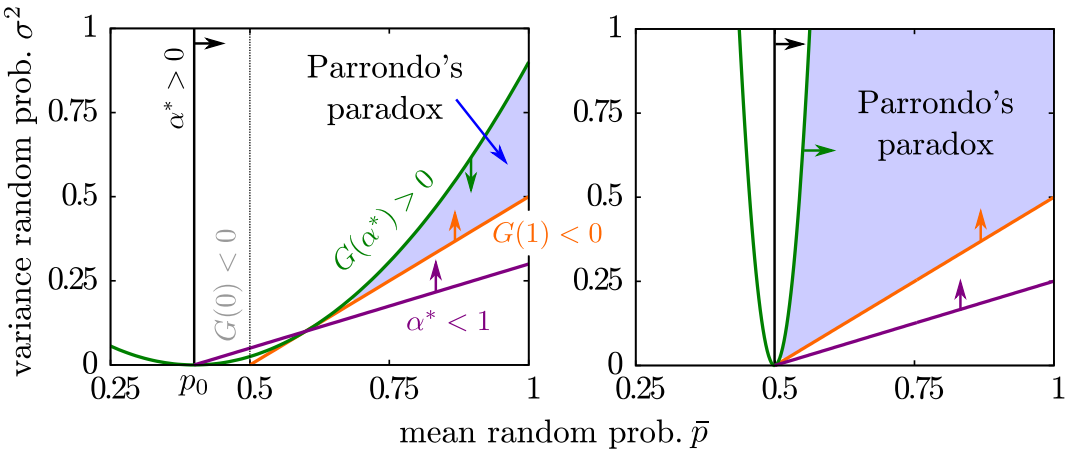


Figure 3.4: **Region of parameters exhibiting the Parrondo's paradox** for different values of the conservative spreading probability ($p_0 = 0.3$ on the left, and $p_0 = 0.499$ on the right): the blue shaded region represents a configuration of parameters where both pure strategies lead to a deterministic extinction ($G(\alpha = 0, 1) < 0$), while a mixture of them can produce a persistent activity ($G(\alpha^*) > 0$). The orange straight line separates the absorbing phase (above the curve) of the active phase (below the curve) for the risky strategy. When the conservative spreading probability approaches its the critical value $p_0 \to 1/2$, any combination with a loosing risky strategy becomes a winning one.

verified the stabilizing effect of hybrid strategies. Notice how the advantageous consequences of bet-hedging are not limited to the mean-field case –for which analytical understanding is available– but are important also in low-dimensional systems where demographic fluctuations play a key role.

In what follows, we analyze how the benefits of bet-hedging depend on the level of stochasticity, both external (environmental) and intrinsic (demographic).

3.5 Dependence on external/environmental variability

First, we study the dependence on environmental variability σ^2. For this, we tune the two pure strategies to have the same stationary density $\langle \rho(\alpha = 0) \rangle = \langle \rho(\alpha = 1) \rangle = 0.3$ and analyze how the steady-state density, ρ, depends on α for different values of σ^2. We have chosen the low value 0.3 because it still permits the activity to grow up by

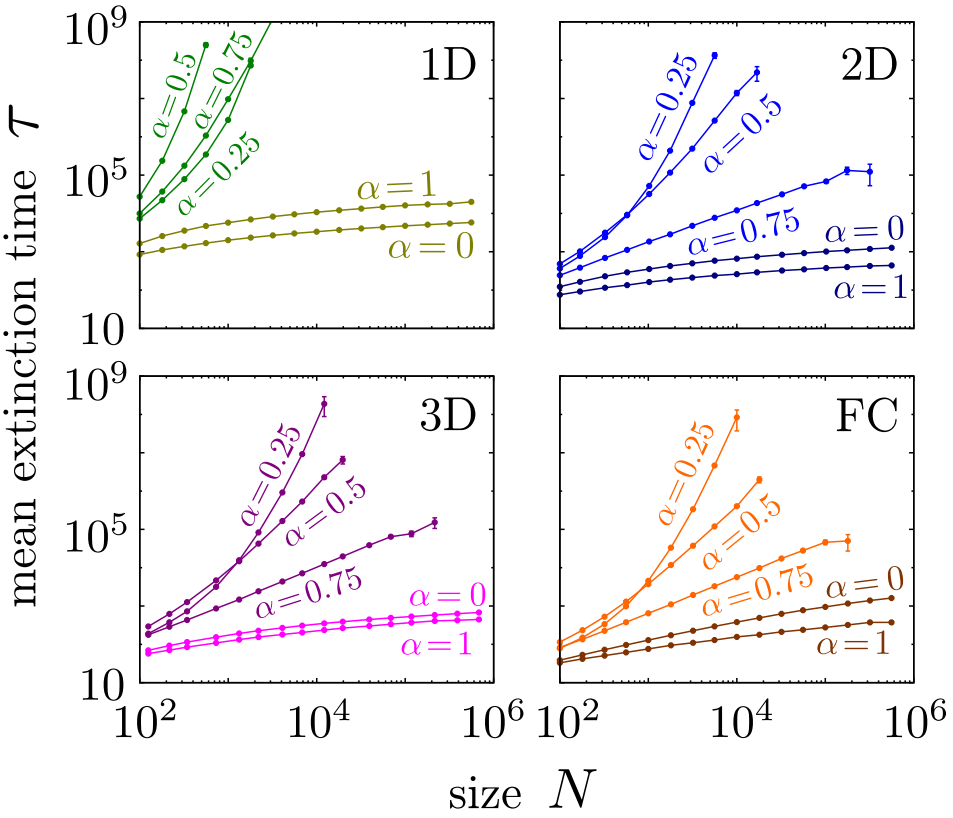


Figure 3.5: **Mean extinction times as a function of system size N for different strategies α and spatial dimensions**: for our parameter choice, the two pure strategies $\alpha = 0, 1$ have a logarithmic dependency (characteristic of subcritical behavior [133]), while a range of hybrid strategies exhibits an exponential or power-law increase typical of active phases [133]. Parameter values $(p_0, \bar{p}, \sigma)$: 1D, $(0.71, 0.80, 0.20)$; 2D, $(0.58, 0.71, 0.29)$; 3D, $(0.54, 0.69, 0.30)$, and FC (fully connected), $(0.499, 0.67, 0.33)$; most error-bars are smaller than symbol sizes.

other mechanisms, but results do not depend on this choice.

Fig. 3.6A clearly illustrates that, for fully connected lattices, bet-hedging strategies permit to achieve much larger values of ρ in more variable environments. The same trend holds for low-dimensional lattices (not shown).

However, an enhancement in the activity does not necessarily mean that the hybrid strategy corresponds to the *best* strategy; for instance, one could argue that extinctions times could be shorter for the hybrid strategy than for the conservative one; contrarily, the risky strategy could be more *invasive* when competing with many other strategies,

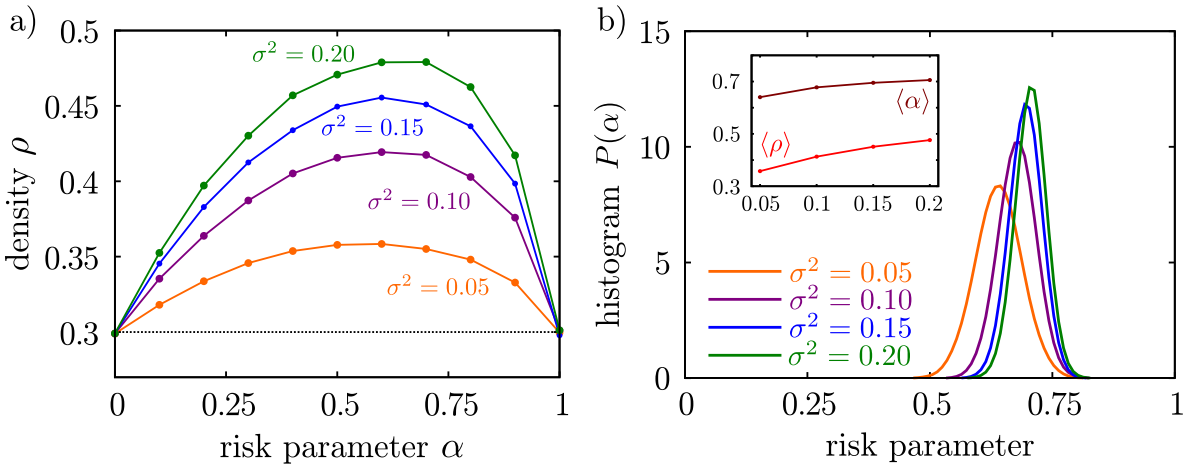


Figure 3.6: **A) Effect of the external-noise variability (σ^2) on the stationary density for different bet-hedging strategies**. Curves are results of Monte Carlo simulations of the fully-connected CP with bet-hedging. As σ^2 is increased, the optimal strategy induces larger stationary densities, even if the pure strategies $\alpha = 0, 1$ lead to the same density, $\langle\rho(\alpha = 0, 1)\rangle = 0.3$. **B) Self-optimization with a genetic algorithm**: each particle inherits the value of α from its ancestor, and mutates with a small Gaussian deviation $\nu = 10^{-3}$. We measure the histogram $P(\alpha)$ at stationarity for each external noise variance σ^2. The resulting distributions are located at the optimal solution; the average density $\langle\rho\rangle$ and risk parameter $\langle\alpha\rangle$ are represented in the inset. Parameter values are$N = 10^4$, $p_0 = 0.567$. $(\bar{p}, \sigma^2)$ are $(0.628, 0.05)$, $(0.699, 0.10)$, $(0.765, 0.15)$ and $(0.825, 0.20)$ in the different curves. Error bars are smaller than symbol sizes in all cases.

as it experiences higher values of the spreading probability $p(t)$ during short periods.

To answer these questions, we implement a genetic algorithm [54, 56] in the dynamics and optimize the strategy in a self-consistent way. In this case, each particle has its own value α and every time it reproduces, the offspring inherits the value α, except for a small Gaussian-distributed random mutation with zero mean and standard deviation ν (with *hard* boundaries conditions at $\alpha = 0, 1$). In this case, we have computed the histogram $P(\alpha)$ in the community once a steady state has been reached, as well as the mean values $\langle\rho\rangle$ and $\langle\alpha\rangle$. As show in fig. 3.6B i) the resulting averaged values coincide with the above-reported optimal ones and ii) the mean α appears a bit biased to $\alpha = 1$ for lower dimension. Interestingly, the distribution appears extremely wide compare to the mutations deviation ($\nu = 10^{-3}$ in the simulations).

3.5.1 Analytics

The previous result can be easily rationalized within the linearized mean-field framework above, eq. 3.4. Using the definition of $G(\alpha)$ and keeping the environmental variance σ^2 as a control parameter, p_0 and $\bar{p}$ can be fixed by imposing identical growth rate of the pure strategies, $G(0) = G(1) \equiv G_{0,1}$. This determines the constraints $p_0 = (G_{0,1} + 1)/2$ and $\bar{p} = (G_{0,1} + 1)/2 + \sigma^2$, that, introduced in eq. 3.5 for the optimal α, lead to the maximum growing rate as a function of σ^2 and $G_{0,1}$:

$$\max(G) = G(\alpha = 1/2) = G_{0,1} + \frac{\sigma^2}{2} \tag{3.8}$$

predicting a linear increase of the optimal G with σ^2. Observe that the optimal value of α in numerical simulations is typically slightly larger than the analytical prediction $\alpha = 1/2$, owing to non-linear saturation effects that have been neglected in the linear approximation.

3.6 Dependence on internal/demographic variability

We now explore the effect of demographic noise. To this aim, the spatial dimension of the systems is varied while keeping a fixed external noise variance σ^2. As above, to ease comparison, we choose the pure strategies for each dimension so that $\langle\rho(0)\rangle = \langle\rho(1)\rangle = 0.3$, and measured computationally $\langle\rho(\alpha)\rangle$ for hybrid strategies.

Fig. 3.6A clearly illustrates that the benefits of bet-hedging are much enhanced as the system dimensionality is decreased, allowing for much larger densities.

For the risky strategy, as the parameter $\bar{p}$ is increased as we decrease the dimension in order to keep the same stationary density(see fig. 3.2), one could argue that it is not fair to use the same environmental variance σ for all the lattices; however, the similar result is obtained (indeed, more pronounced) if we repeat the same experiment but keeping the ratio $\sigma/\bar{p}$ constant.

As for the external noise analysis, we have implemented the genetic algorithm [54, 56] and verified that the emerging distributions of evolutionarily-stable strategies are peaked at the previously obtained optimal strategies. Results are shown in fig. 3.7B,

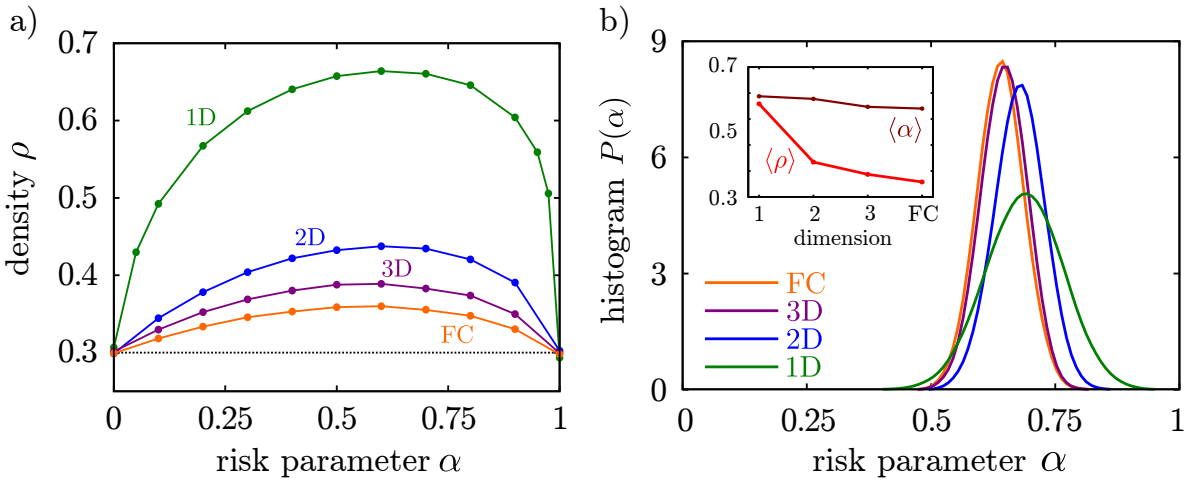


Figure 3.7: **A) Effect of dimensionality at fixed σ^2 on the stationary density for different strategies**: the net benefit of bet-hedging is much enhanced in lower-dimension. Error bars are smaller than symbol sizes in all cases. **B) Self-consistent optimization with the genetic algorithm**: as in fig. 3.6B, the self-tuned strategies are located at the optimal strategies. Parameters: $\langle\rho(\alpha = 0, 1)\rangle = 0.3$, $\sigma^2 = 0.05$, $N = 10^4$ for 1D, 2D and FC, $N = 10648$ for 3D, and $(p_0, \bar{p})$ are 1D:$(0.722, 0.847)$, 2D:$(0.618, 0.704)$, 3D:$(0.594, 0.665)$ and FC:$(0.567, 0.628)$. In the genetic-algorithm, $\nu = 10^{-3}$.

illustrating the robustness of the hybrid strategy.

3.6.1 Analytics

To give a mathematical argument that explains the previous result, we should develop a rigorous analysis including explicitly spatial dependence in eq. 3.9. However, we can avoid this very challenging task, which is beyond the scope of this work, and use a simple and effective argument to shed some light on this finding.

For this purpose, we consider a refined version of eq. 3.4, including the non-linear term ignored above as well as a demographic-noise term with tunable amplitude γ [52, 104]:

$$\dot{\rho}(t) \approx (2p_{\text{av}}(\alpha) - 1)\,\rho - p_{\text{av}}(\alpha)\rho^2 + 2\alpha\sigma\rho\xi(t) + \gamma\sqrt{\rho}\eta(t) \tag{3.9}$$

where $\eta(t)$ is a delta-correlated Gaussian noise and the factor $\sqrt{\rho}$ is a direct consequence of the central limit theorem [52]. Demographic noise is a key ingredient, missing in the

mean-field limit, which phenomenologically accounts for intrinsic fluctuations in low dimensional systems. Demographic noise is known to be much stronger in low dimensions; thus one should expect that the effective noise amplitude γ in our simplified approach should increase as the dimension decreases. Note that, despite of this improvement, eq. 3.9 corresponds to an effective one-variable theory, so it will not capture all the essential features of a low dimensional system, and we can only expect a qualitative description of our results. Moreover, the $\sqrt{\rho}$ term does not permit to compute the exponential growing rate in the same way that we did for the external noise analysis. Therefore, we have to work with stationary densities rather than growing rates; this is why we have included the saturation term.

Associated with eq. 3.9, we write the Fokker-Planck equation[52] for the probability distribution $P(\rho, t)$. To work in the quasi-stationary approximation [52, 43, 104] (i.e. avoiding the absorbing state $\rho = 0$), we include a small and constant drift ε; identifying with $a = (2p_{\rm av}(\alpha)-1)$, $b = p_{\rm av}(\alpha)$ and $\beta = 2\alpha\sigma$, the equation for the density distribution is:

$$\partial_t P(\rho,t) = -\partial_\rho \left[\left(\varepsilon + a\rho - b\rho^2\right) P(\rho,t)\right] + \frac{1}{2}\partial_{\rho^2}\left[\left(\gamma^2\rho + \beta^2\rho^2\right) P(\rho,t)\right]. \tag{3.10}$$

Solving for the stationary solution $P(\rho, t \to \infty)$ (valid when the external flux $\varepsilon \ll 1$), we find:

$$P(\rho, t\to\infty) = \begin{cases} C_1\, \rho^{\frac{2\varepsilon}{\gamma^2}-1} \exp\left(\frac{2a\rho - b\rho^2}{\gamma^2}\right), & \beta = 0 \\ \\ C_2\, \rho^{\frac{2\varepsilon}{\gamma^2}} \left(\gamma^2 + \beta^2\rho\right)^{\frac{2a}{\beta^2} + \frac{2b\gamma^2}{\beta^4} - 1} \exp\left(-\frac{2b\rho}{\beta^2}\right), & \beta > 0, \end{cases} \tag{3.11}$$

where C_1 and C_2 are normalization constants which has to be fixed via numerical integration. At this point we can compute (numerically) the stationary density $\langle\rho\rangle = \int_0^\infty d\rho\, \rho P(\rho, t\to\infty)$. As we have eliminated the absorbing state, we do not have to worry about potential divergences at $\rho = 0$.

To obtain an estimation of the effective values of γ in dimensions $D = 3, 2$ and 1, we keep parameters other than γ fixed as in Monte Carlo simulations in fig. 3.7. We then calculate the quasi-stationary density associated with eq. 3.9 from its associated Fokker-Planck equation [52]. The parameter γ is determined by the condition that the

quasi-stationary density satisfies $\langle\rho(\alpha=0)\rangle = \langle\rho(\alpha=1)\rangle = 0.3$. The resulting values of γ are $\gamma = 0.09, 0.15$ and 0.28, for 3, 2 and 1 dimensions, respectively (each value is an average of two very close results obtained for the two pure strategies).

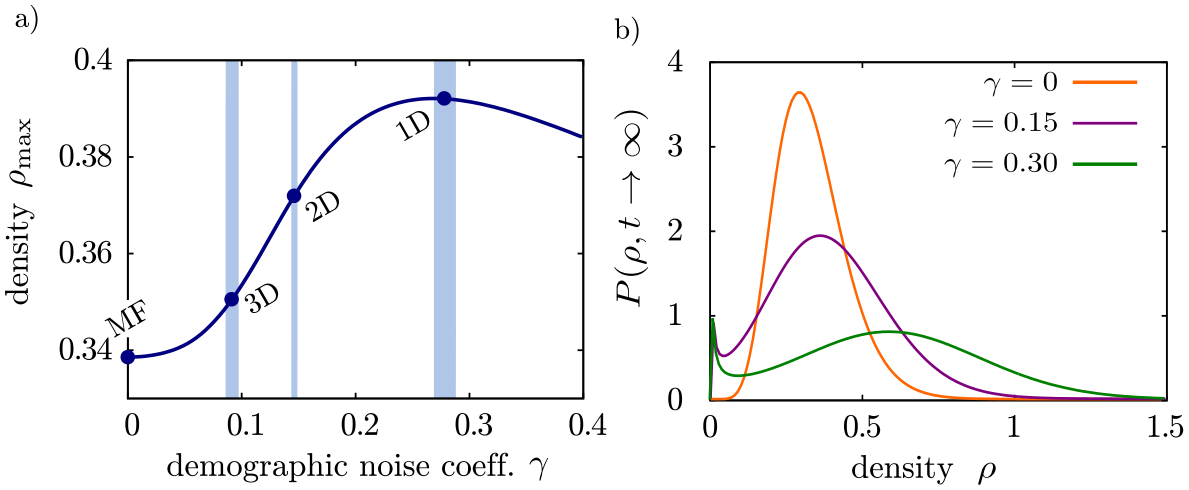


Figure 3.8: **A) Stationary density of the optimal hybrid strategy as a function of the demographic noise amplitude** γ: each point has been computed using the value of γ inferred from eq. 3.11, for a different spatial dimension (1, 2, and 3), and $\gamma = 0$ for the mean-field of infinite dimensional case. Parameters p_0, $\bar{p}$ and σ are the same as in Fig. 3.7 and γ is tuned to produce $\rho(0,1) = 0.3$. The inferred γ changes slightly for $\alpha = 0$ and $\alpha = 1$ as reflected in the errorbars (shaded region). These results confirm that the effective noise amplitude γ increases as the system dimensionality decreases and that the benefits of bet-hedging are enhanced by demographic noise. However, the curve becomes non monotonous for larger values of γ. **B) Density distributions in the optimal strategy** α^* **for different demographic noise amplitudes** γ: We represent the the quasi-stationary solution of 3.11 ($\varepsilon = 10^{-2}$). The calculus fails for higher values of γ, as the probability to decay into the absorbing state (emerging peak at $\rho = 0$) becomes non negligible.

Having determined, for each dimension, the level of demographic fluctuations γ, we compute the maximum density as a function of α in each case. Results are shown in fig. 3.8, which reveals that the benefits of bet-hedging are enhanced for larger demographic noises and thus, for lower spatial dimensions. We remark that this phenomenological theory only provides a qualitative explanation of the phenomenon and does not quantitatively reproduce the actual stationary densities in fig. 3.7.

3.7 Effect of temporal correlations

In the previous sections, our results were derived under the assumption that the external environmental noise has no temporal correlations, as the risky spreading probability is renewed at each time step independently of the previous value. However, a realistic environment may present temporal correlations, thus we wonder how the benefits of bet-hedging are enhanced or reduced when the environment is take in this other way.

A simple way to introduce such correlations is to take a stochastic environment $p(t)$ obeying the Ornstein-Uhlenbeck process, with time-series similar to that of a Brownian particle moving in a parabolic potential. Mathematically, it follows the Langevin equation [52]:

$$\dot{p} = \theta(p - \bar{p}) + \sqrt{2\theta}\sigma\xi(t), \tag{3.12}$$

where $\bar{p}$ and σ represent, as before, the mean and variance of the risky parameter, respectively; with this choice, $p(t)$ is distributed as a Gaussian $N(\bar{p}, \sigma^2)$. The new parameter θ controls the temporal correlations, as $\langle p(0)p(t)\rangle \sim e^{-t/2\theta}$; consequently, $\theta \to 0$ and $\theta \to \infty$ represent the extreme cases of immutable –completely correlated– and constantly changing –non-correlated, i.e. the original implementation– environments, respectively. In our case, $\theta \sim 10$ can be considered as very large.

We run our simulations integrating eq. 3.12 with the exact formula [52]

$$p(t+1) = \bar{p}(1 - e^{-\theta}) + p(t)e^{-\theta} + \sigma\sqrt{1 - e^{-2\theta}}G_t(0, 1), \tag{3.13}$$

where $G_t(0, 1)$ is a Gaussian random number with zero mean and unit variance generated at each time step. Fixing the environmental variance σ^2, we study the effect of temporal correlations on bet-hedging for different values of θ in every dimension. Following the same strategy as above, we tune the parameters p_0 and $\bar{p}$ for each temporal correlation θ to fix the stationary density at $\langle\rho(\alpha = 0, 1)\rangle$, and we measure $\langle\rho(\alpha)\rangle$. Results are summarized in fig. 3.9; some remarks are in order:

- The optimal strategy is always a hybrid strategy between $\alpha = 0$ and $\alpha = 1$. Additionally, curves coincide with those in fig. 3.7 when θ is high ($\theta \sim 10$), as the external environment is poorly correlated.

- When θ increases moderately, the stationary density at the optimal strategy be-

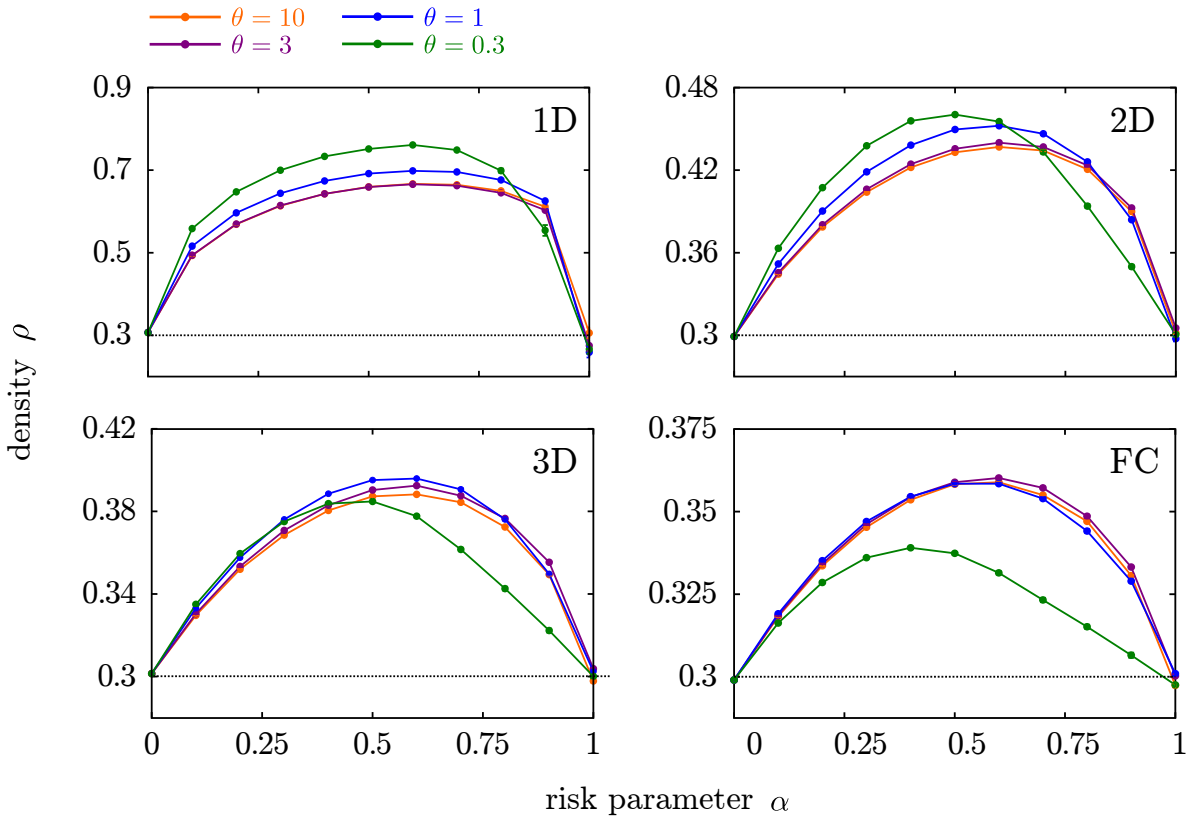


Figure 3.9: **Stationary density as a function of the risk parameter for different lattice topologies in temporal-correlated environments**: the spreading probability of the risky strategy, $p(t)$ obeys now an Ornstein-Uhlenbeck process with mean $\bar{p}$, variance σ^2, and exponential temporal correlations with a characteristic time 2θ. The benefits of the hybrid strategies in the stationary density become enhanced for intermediate values of θ. This result is much intensified in lower dimensions, while it is imperceptible for the 3D and FC networks. Parameters: $\sigma^2 = 0.05$; $(\bar{p}, \theta)$ in 1D: $(10, 0.848)$, $(3, 0.849)$, $(1, 0.876)$, $(0.3, 0.944)$; 2D: $(10, 0.704)$, $(3, 0.706)$, $(1, 0.719)$, $(0.3, 0.745)$; 3D: $(10, 0.665)$, $(3, 0.667)$, $(1, 0.674)$, $(0.3, 0.681)$ and FC: $(10, 0.628)$, $(3, 0.629)$, $(1, 0.631)$, $(0.3, 0.605)$; p_0 and N are taken as in fig. 3.7.

comes larger compared to the non-correlated case. In other words, *bet-hedging strategies are worthier for temporally correlated environments*. This effect becomes stronger for lower dimensions, whereas barely applies to higher dimensional lattices.

- When the correlation is highly increased ($\theta < 0.1$), the situation is reversed

and the benefits of hybrid strategies are reduced compared to the non-correlated case. However, we are not interested in this "frozen" scenario, where the external environment remains almost unchanged during long periods of time. In this highly correlated situation, the dynamics are effectively sampling the stationary densities of processes with different parameters. The inflection point in θ at which this effect appears varies for different dimensions.

- Finally, the optimal strategy becomes more conservative when temporal correlations are added to the environment, with a bias to $\alpha^* \to 0$ when θ increases.

So far we have not developed a proper mathematical analysis to explain this finding. However, it can be easily rationalized within the following argument: temporal correlations affect the external environment in a way that, when it turns to be unfavorable ($p(t) \ll 1$), it continuous to be unfavorable for a period (on the order of $\sim 2\theta$ steps); this can lead to catastrophic consequences in the system, specially for lower dimensions, where internal fluctuations make the system to be less robust; additionally, although favorable periods are also maintained, the effects of a global extinction is not reversible.

In conclusion, bet-hedging strategies represent a reliable possibility when the system has to choose between a conservative, relatively poor solution, and a risky, even correlated in time, one.

3.8 Heterogeneous vs. homogeneous populations

In CP model with hybrid strategies, each particle decides it spreading dynamics at every time step; after many steps, a portion α of times it will choose the risky strategy and $(1-\alpha)$ the conservative one. This scenario corresponds to an individual-based bet-hedging, because each agent can perform, with different probabilities, both dynamics. A different case can be, for instance, a population in which the bet hedging is done at the community level: a portion α of the particles are risky, whereas $(1-\alpha)$ are conservative. In this section, we analyze the difference between exploiting bet-hedging individually or at a community level.

The first case we can study is the simple generalization of the hybrid CP: the system is composed only by particles of conservative type (spreading with constant probability p_0) and risky type (spreading at variable rate $p(t)$). Each time a particle reproduces, the

offspring can be, with probability α, a risky particle, or, complementary with probability $(1-\alpha)$, a conservative one. In this case, α is a control parameter, but, in principle, it does not have to coincide with the relative concentration of risky agents in the network (we use $\langle\alpha\rangle$ to refer this magnitude, understanding that conservative and risky particles have $\alpha = 0, 1$, respectively).

In this situation, we repeated our previous analysis, and measured $\langle\rho(\langle\alpha\rangle)\rangle$ when $\langle\rho(0)\rangle = \langle\rho(1)\rangle = 0.3$. We have not encountered substantial differences respect to our previous findings (not shown), except for a small bias in $\langle\alpha\rangle > \alpha$ due to the high invasivility of the risky dynamics during the periods in which $p(t)$ is large (see below).

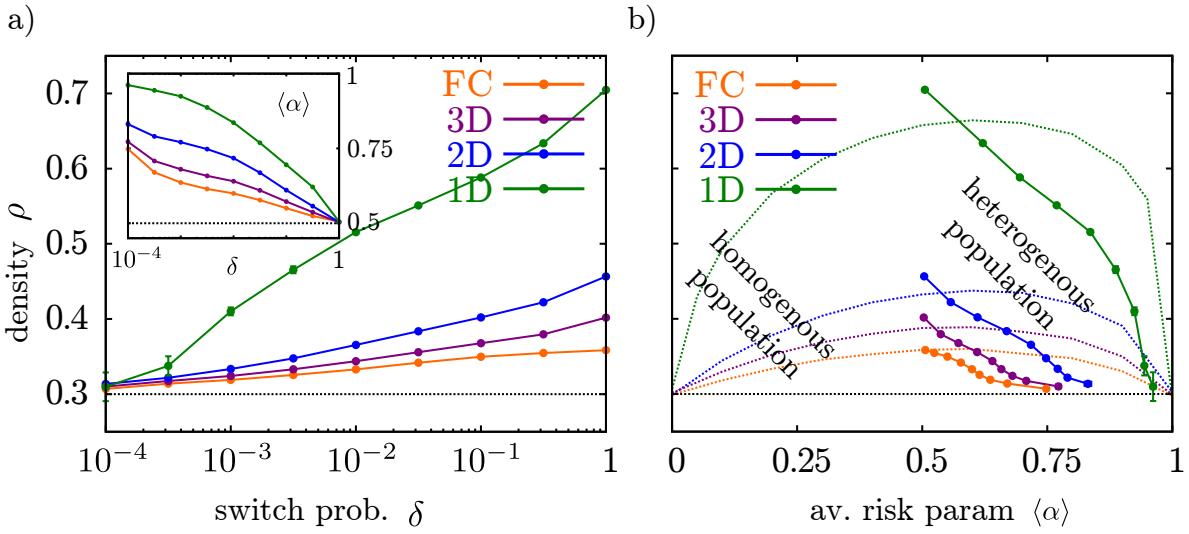


Figure 3.10: **A) Stationary density for heterogeneous populations as a function of the switch parameter** δ: in this implementation of the hybrid CP, there coexist particles which reproduce through either the conservative ($\alpha = 0$) or the risky ($\alpha = 1$) strategy. Offspring inherit the strategy from the progenitor, but it can mutate and switch $\alpha = 0 \leftrightarrow 1$ with a probability δ. Separately ($\delta \to 0$), each strategy leads to $\langle\rho(0,1)\rangle = 0.3$, while a combined system ($\delta > 0$) exhibits a larger density. Inset: average strategy in the community, which is tuned self-consistently for every δ. **B) Stationary density vs. average risky parameter** $\langle\alpha\rangle$: each point corresponds to a pair $(\langle\alpha\rangle, \langle\rho\rangle)$ on the left. To compare with the homogeneous case (all particles with equal strategy α), we include in dashed line the curves from fig. 3.7. Parameters are taken as in fig. 3.7.

More interesting is the case in which particles with $\alpha = 0, 1$ are not well mixed, and spatial correlations are included in the system. A possible implementation of this

idea is: each time a particle reproduces, the offspring inherit the strategy $\alpha = 0$ or $\alpha = 1$ from its ancestor, and with a small probability δ, mutate and switch to the be of the other type. Consequently, if the original particle was a conservative particle, the new one can be, with probability $(1 - \delta)$, still conservative, or, with probability δ, a different, risky one. In this model, $\langle\alpha\rangle$ is set self-consistently and depends on the value of δ; when $\delta \to 1$, it is expected that $\langle\alpha\rangle \to 0.5$.

We run our simulations tuning the parameters to have the same stationary density for each strategy, as before, to be $\langle\rho(\alpha = 0, 1)\rangle = 0.3$. Results are shown in fig. 3.10; summarizing, we find that:

- The stationary density for any hybrid situation is always higher than for the pure strategies, $\langle\rho\rangle > 0.3$.

- Even if both strategies are set to perform the same density separately, the self-tuned average strategy $\langle\alpha\rangle$ is always greater than $1/2$; this constitutes an evidence of the high invasivity of the risky strategy compared to the conservative one.

- Bet-hedging strategies in heterogeneous populations with average risk parameter $\langle\alpha\rangle$ exhibit a lower density than bet-hedging in homogeneous populations with risk parameter α, with the exception of the case $\delta \to 1$. This limit corresponds to a scenario in which individuals always produce 'opposite' offspring, i.e., if the predecessor performs the risky strategy, the new particle has the conservative one, and vice versa.

Such effect for $\delta = 1$ ($\langle\alpha\rangle = 0.5$) could be explained with the following argument: in the homogeneous case, it can happen that, sometimes, a particle and its neighbors choose to perform the same risky strategy, because their decisions are completely uncorrelated. This can be an unwise choice if the environment becomes unfavorable, and consequently, the activity gets highly reduced. A good stratagem to avoid these events is to place a different particle next to the progenitor every time it reproduces. Notice that this argument only applies for networks where spatial correlations are relevant, i.e. not for the mean-field. This is in perfect agreement with fig. 3.10B for the FC network, where the optimal density at $\delta = 1$ is exactly equal than for the homogeneous case (meeting point of orange straight and dashed lines). In conclusion, we could say

the best strategy is the one of having insolent children, doing just the opposite of what their parents do.

Still, heterogeneous populations usually exhibit lower densities than homogeneous communities (rest of points with $\delta < 1$, $\langle\alpha\rangle > 0.5$). As such discrepancy is also found in the mean-field case (FC network), it should be explained apart from the spatial correlations between progenitors and offspring. This effect can be rationalized from a mathematical derivation of the mean-field stochastic equation for the heterogeneous population model (not shown), which leads to a pair of equations for $\dot{\rho}$ and $\langle\dot{\alpha}\rangle$. The first one corresponds to same eq. 3.3 (substituting $\alpha \to \langle\alpha\rangle$), while the dynamics in the second strongly depend on δ. Thus, the behavior of the homogeneous and the heterogeneous population model do not coincide, not even when spatial correlations are neglected, as $\langle\dot{\alpha}\rangle \neq 0$.

3.9 Chapter summary

In this chapter, we have studied a simple model for population dynamics with bet-hedging, in which individuals can choose between a poor but safe reproductive strategy, a better but risky alternative, or a combination of both. In this context, we have sought under which conditions hybrid strategies suppose a substantial benefit respect to the pure dynamics.

In particular:

- The model we have implemented corresponds to the Contact Process, with the particularity that, at each time, particles choose, with probability α, to reproduce with a risky stochastic probability $p(t)$, common to all individuals in the community, or, with complementary probability $(1-\alpha)$, with a fixed conservative probability p_0. The first strategy provides high spreading rates at some times, but, eventually, can turn very unfavorable.

- We have found that, for certain choices of the parameters, a population with a combined dynamics exhibits larger densities, compared to the case with just pure strategies. And, furthermore, we have found situations in which both pure strategies lead to a deterministic extinction, while a combination of them gives a stable population.

- We have analyzed the role of external (environmental) and intrinsic (demographic) fluctuations in enhancing the benefits of bet-hedging strategies. To this end, we have compared different scenarios of environmental variability and spatial dimension (1D, 2D, 3D and fully-connected network). For each of them, we have tuned the pure strategies to give the same stationary density, and we have analyzed how this density increases for different values of the risky parameter α. The main finding is that both environmental and intrinsic fluctuations foster the benefits of performing bet-hedging strategies.

- We have checked the robustness of such hybrid solutions with a genetic algorithm, in which the parameter α is dynamically self-tuned by inheritance and mutations of the community.

- We have studied the system when the environment presents temporal correlations. With this purpose, we have described the risky probability $p(t)$ as an Ornstein-Uhlenbeck process with tunable correlation time scale. We have seen that such correlations increase –when these are moderate– the benefits of bet-hedging strategies, specially in lower dimensional systems.

- Finally, we have analyzed the case in which bet-hedging is performed at the community level, rather than individually. We have studied a mixed population where each individual follows only one of the pure dynamics. We have seen that, in most of the cases, the best option corresponds to bet-hedging at the single individual level. However, a more detailed study should be carried out to compare these two scenarios, for instance, by including the cost of developing and maintaining the hybrid mechanism for each individual.

Summing up, our main finding is that the relative benefits of bet-hedging strategies are strongly enhanced in highly-fluctuating low-dimensional systems. Given that these conditions are often met by biological populations, as for example in bacterial colonies competing at the front of a range expansion in noisy environments [11, 78], our results provide a strong justification for the ubiquitousness of bet-hedging in nature.

In general, these could be condensed as in the popular saying: *"troubled waters, fisherman's gain"*.

Chapter 4

Mixed dispersal strategies

4.1 *'Diversity in times of adversity'*

Plants in the wild, dealing with ever-changing environments, constitute a characteristic example of bet-hedging organisms in nature [26, 135, 29, 35]. When the external conditions are highly unpredictable, individuals cannot accurately predict the environment, and individuals hedge their bets by producing different types of offspring phenotypes, for instance by diversifying their reproduction timing [116, 123].

A particularly interesting example of this risk assessment corresponds to the way in which plants disperse their seeds. In ecology, it is well known that dispersal constitutes a key evolutionary mechanism for the plants, and therefore it has been extensively studied in the literature [28, 112]. Dispersal is generally extended in plants [37, 30] because it provides high competitive advantages such as the colonization of distant areas as well as a reduction of the intra-community reproduction. This selfing, also called *inbreeding*, has been reported to produce a negative effect on the individuals' fitness [113, 112], in contrast with the *outcrossing*, i.e. the mixing with different communities and genotypes. On the other hand, seed dispersal implies an extra cost of resources for the plant, and they are much subjected to external/environmental conditions.

Nature has reached to solve this problem in an efficient way. One can find many examples in which the same plant presents two clearly different reproductive mechanisms [68, 7, 3]: one with open-pollinated external flowers, producing aerial seeds, and another one with subterranean self-pollinated flowers, with internal seeds. Specific examples of

this are the *Lathyrus amphicarpos* and *Vicia amphicarpa* (see fig. 4.1) which can be found, for instance, in the Torcal de Antequera Nature Reserve (southern Spain).

Figure 4.1: **Two examples of species exhibiting a hybrid reproductive mechanism**: (**Left)** Aerial open-pollinated flowers of *Lathyrus amphicarpos* (red) and *Vicia amphicarpa* (violet). **(Right)** Subterranean self-pollinated flower of *Lathyrus amphicarpos*. Photos by Rafael Rubio de Casas.

In the previous section, we provided a general scenario to study the conditions under which bet-hedging strategies can emerge, in a simplistic model of population dynamics, the Contact Process. Our aim now is to study a specific more-realistic model from a computational point of view, and try to understand the empirical fact that certain vegetable species mixed dispersal strategies, consisting in the evolutionary-stable coexistence (even within a single individual) of two different reproduction mechanisms (e.g. two different seed types): a dispersive and a non-dispersive one. To this end, we introduce several biologically feasible mechanisms: i) effect of environmental variability on dispersed seeds, ii) positive correlation between dispersal and outcrossing, and finally iii) negative effect of inbreeding depression.

4.2 Simplistic model with inbreeding depression

We introduce a model in which a community of plants evolves in time through the mechanism of birth, reproduction, competition, and death. These plants are annual and reproduce by means of two different strategies: a conservative one, based on selfing, which is not worthy at a large time scales, and a risky environment-dependent one, based on outcrossing.

Each plant lives in a fixed node of a $L \times L$ square lattice; the maximal occupation per node is restricted to one plant, accounting for local competition for resources (see Fig. 4.2). Plants reproduce by seeds that can be of two different types: *external* seeds, produced through sexual reproduction, and *internal* seeds, produced from auto-fecundation or inbreeding. External seeds travel by air to distant sites –for simplicity, we assume that they can reach any random cell in the lattice–, whereas internal seeds can only be implanted into the maternal or adjacent sites.

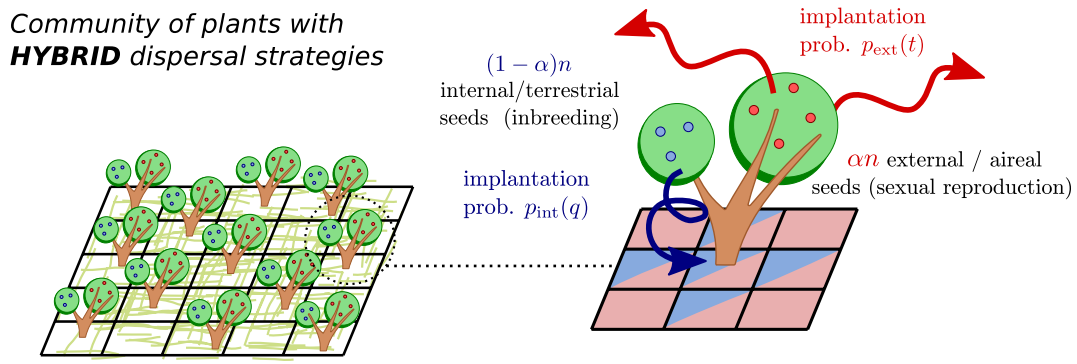


Figure 4.2: **Community of plants with a hybrid dispersal mechanism**: each plant is located at a node of a square $L \times L$ lattice. At each reproduction period, each plant produces αn external/aerial seeds and $(1-\alpha)n$ internal/terrestrial seeds, where α is the bet-hedging parameter. External seeds (produced by sexual reproduction) can arrive far away to any cell in the lattice, but in contrast have a variable environment-dependent implementation prob. $p_{\rm ext}(t)$. On the other hand, internal seeds (produced from inbreeding) are not affect by the environment, but they only can achieve adjacent cells in the lattice. Additionally, its implantation prob. is penalized as they are produced by self-reproduction; this penalization is encoded in a quality q for every plant (and its offspring), which modulates the implantation probability $p_{\rm int}(q) = qp_{\rm int}$; it is reduced after inbreeding $q \to (1-\delta)q$ and reseted to 1 after sexual reproduction.

Individuals have an intrinsic propensity to produce each of the two types of seeds, which is encoded in the bet-hedging parameter α: for $\alpha = 1$ all seeds are of the dispersive type while for $\alpha = 0$ all of them are "internal". For the moment, we keep α as a control parameter in the simulations, common for all plants in the community.

Each type of seed obeys a different implantation dynamics. External seeds do it at a randomly chosen node with a probability $p_{\rm ext}(t)$, which is a fluctuating variable subjected to environmental conditions. In contrast, internal seeds are assumed not to be

influenced by the external environment; on the other hand, their offspring are placed at local neighboring sites and experience an *inbreeding depression* in their "quality", thus their implantation probability $p_{\rm int}$ is reduced after auto fecundation. A much realistic version of this model should include a temporal dependency in the internal implantation probability, but we restrict our analysis to the simplest version of the dynamics.

A dynamical process is implemented as follows. At each discrete time-step t –which represents a reproduction period, i.e. typically one year– all occupied sites are updated according to the following rules:

1. Environmental conditions are set at every t, encoded in the variable $p_{\rm ext}(t)$. We take $p_{\rm ext}(t)$ to be an uncorrelated random variable with uniform distribution in $[\bar{p}_{\rm ext}+\sigma, \bar{p}_{\rm ext}-\sigma]$.

2. Each plant on the lattice produces n seeds. A similar implementation can be done with a variable number of seeds, extracted from a Poisson distribution of mean n, which is more realistic, but our results remain essentially the same. With probability α, each seed comes out to be of the external type, or, complementary with probability $(1-\alpha)$, of the internal type. We take $n = 5$, but note that the relevant parameter corresponds to the ratio number of seeds $\times$ implantation probability.

3. Plants are equipped with an individual *quality* parameter q, different for each one, which is inherited by the seeds and reduced as $q \to (1-\delta)q$ after an inbreeding event, or reset to $q \to 1$ with outcrossing. A more realistic version of this "resetting" mechanism can be implemented, where the quality factor increases progressively after sexual reproduction, but we do not enter into such a detailed level as our results remain qualitatively unaltered.

4. Seeds are dispersed on the lattice: external ones finish into a random site of the whole system with probability $p_{\rm ext}(t)$, and internal seeds into the $(1+4)$ next locations with probability $q \cdot p_{\rm int}$. To reduce the finite-size effects, we choose periodic boundary conditions.

5. All the plants are removed from the ecosystem, i.e. generations are non-overlapping.

6. At each location, only one of the implanted seeds (if there is any) is chosen at randomly and grown at such a place.

This is the simpler version of our hybrid dispersal model and it contains the basic ingredients we are interest in: positive correlation between dispersal and outcrossing, inbreeding depression and environmental variability, as well as intrinsic mechanisms in the lattice such as competition for the resources and saturation.

4.3 Phases of the pure strategies

First we analyze the behavior of the system for the pure strategies $\alpha = 0, 1$. Although this has been already studied in deep in the literature [88], it is a useful reference for our specific model. To proceed, we compute the stationary density $\langle \rho(\alpha = 1) \rangle$ as a function of $(\bar{p}_{\text{ext}}, \sigma)$ and $\langle \rho(\alpha = 0) \rangle$ as a function of (p_{int}, δ). Results, shown in Fig. 4.4, exhibit a continuous phase transition for the density.

Notice that, for any non-vanishing inbreeding depression $\delta > 0$, the system always becomes extinct, regardless the value of p_{int}.

4.4 Benefits of the hybrid strategy

We study now the stationary density $\langle \rho \rangle$ for different values of the bet-hedging parameter α. As we are more interested in the behavior of the system with environmental variability σ and inbreeding depression δ, for simplicity we fix $p_{\text{int}} = p_{\text{ext}}$.

Results are shown in fig. 4.4. We can see that:

- In the absence of any penalization nor variability ($\delta = \sigma = 0$), the dispersive strategy ($\alpha = 1$) leads to a higher density. This is in agreement with the common fact that plants always prefer to disperse as much as possible [37].

- When δ and σ increase (fixing p_{int} and $\bar{p}_{\text{ext}}$), the stationary density of both pure strategies $\alpha = 0, 1$ is strongly decimated. Instead, communities dispersing through a hybrid strategy (intermediate values of α) produce higher densities. This can be rationalized within our previous finding in the Contact Process with hybrid dynamics: the growing rate is a quadratic function of α; therefore, depending

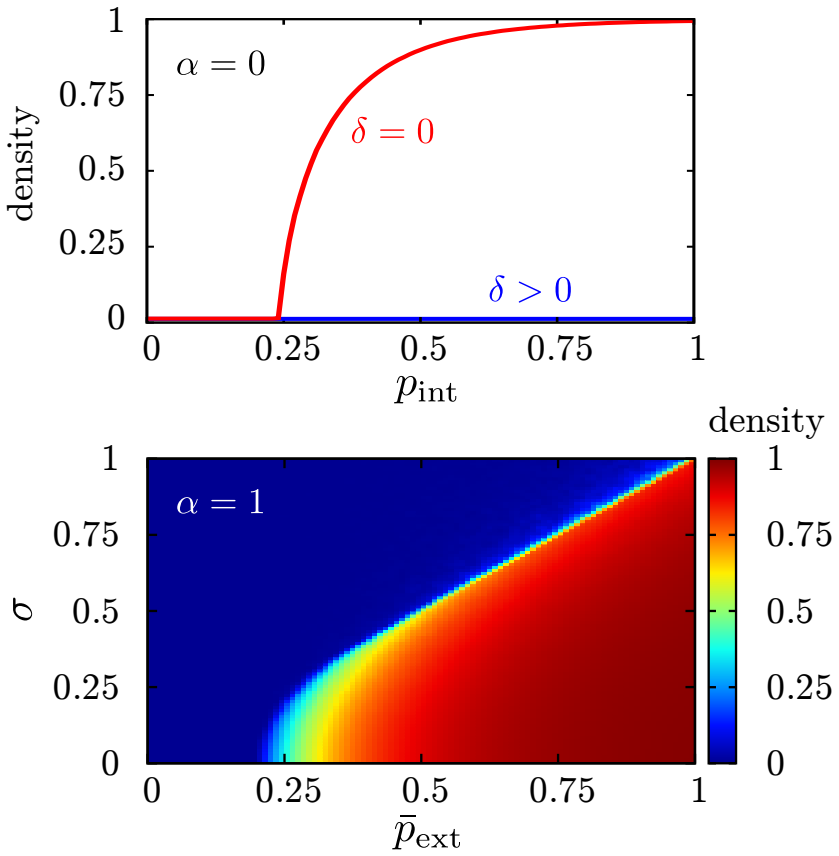


Figure 4.3: **Density phases of the pure strategies**. Both strategies exhibit a phase transition between the *absorbing* state, where the community always become extinct, and an *active* phase, with a maintained stationary density. In the case of the pure internal strategy with some inbreeding forfeit ($\delta > 0$, blue line in the top panel), the plants always disappear after a few generations. In the pure external strategy, greater values of the environmental variability ($\sigma > 0$) are prejudicial for the community, thus expanding the absorbing phase (blue region in the bottom panel).

on the parameters, the system can exhibit larger densities at intermediate values of α than in the extremal points (eq. 3.5). Consequently, *bet-hedging strategies can reduce the catastrophic consequences of inbreeding depression and external variability.*

- There are cases ($\delta > 0, \sigma \gtrsim 0.3$) in which the community becomes extinct if plants disperse through one of the pure strategies, while it survives indefinitely when they do so through a mixture of them. This constitutes another instance of

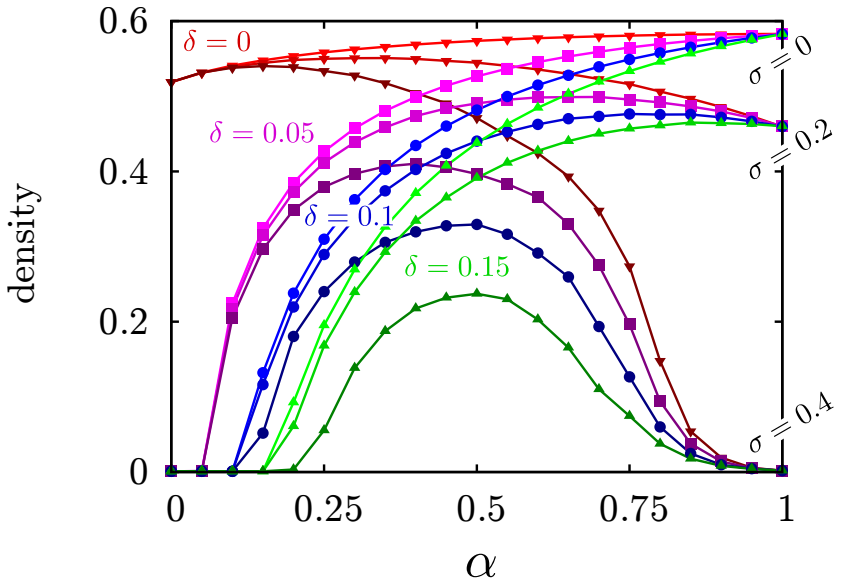


Figure 4.4: **Stationary density as a function of the bet-hedging parameter α in different environments and types of inbreeding**: We compute the density for different values of the environmental variability σ (red, purple, blue and green curves) and inbreeding depression δ (light to dark curves for each color set). Systems with one of the pure strategies $\alpha = 0, 1$ are dramatically affected when increasing δ or σ, respectively (extremal points), while hybrid strategies do not (central region). Parameters are set to $p_{\text{int}} = \bar{p}_{\text{ext}} = 0.3$ and $L = 100$ ($N = 10^4$).

the Parrodo's paradox [57], that we already studied in section 3.4.

- For low environmental variability, the optimal strategy can be located at some intermediate value of α, but, compared to the pure strategy $\alpha = 1$, it essentially leads to the same stationary density (see middle purple curve in fig. 4.4), so one could argue that it is not worthy to develop a hybrid mechanism. On the other hand, for high environmental variability (and inbreeding depression $\delta > 0$), the optimal hybrid strategy produces a much higher density compared to the pure ones (see bottom blue curve in fig. 4.4). This effect is analyzed in detail in the following section.

4.4.1 Relative gain enhancement of hybrid strategies

Fig. 4.4 illustrates that, for certain choices of the parameters, hybrid strategies not only become optimal, but also lead to much higher densities, compared to the pure strategies.

To measure this 'relative gain' in an objective manner, we define two metrics. The first one corresponds to the stationary density of the optimal strategy α^*, compared to the pure ones:

$$\Delta\rho_{0,1} = \rho(\alpha^*) - \max\left[\rho(0), \rho(1)\right]. \tag{4.1}$$

However, as we have seen from the simulations, $\rho(\alpha = 0) = 0$ for any $\delta > 0$, so we are loosing a lot of information about the non-dispersive strategy. One possibility is to compare the optimal strategy α^* (in the sense of the one producing the larger density) with two similar strategies $\alpha^* \pm \Delta\alpha$. Mathematically, this is the curvature of $\rho(\alpha)$ at α^*, measured throw the discretized second derivative, or Laplacian:

$$\Delta\rho_{\Delta\alpha} = \frac{1}{\Delta\alpha^2}\left(\rho(\alpha^*) - \frac{\rho(\alpha^* + \Delta\alpha) + \rho(\alpha^* - \Delta\alpha)}{2}\right). \tag{4.2}$$

Note that, for convenience, we have multiplied the original Laplacian by a factor $-1/2$.

Using data from fig. 4.4 (with more resolution), we have computed the measures above as a function of (δ, σ). Results are shown in Fig. 4.5. In both cases, although surface plots differ qualitatively, the same conclusion holds: *bet-hedging strategies are more likely to be developed in systems with low –but greater than zero- inbreeding depression δ and high environmental variability σ.*

4.4.2 Reduction of the mean extinction time

We study how the mean-extinction time can be reduced when plants perform a hybrid dispersal strategy. It can happen that, in a temporal step, all plants and seeds are removed from the lattice; consequently, no more individuals can be created unless we include other mechanisms such as external migration. This collapse can be found when parameters lie in the absorbing region (i.e. $\rho = 0$, see. fig. 4.3) or in the active phase ($\rho > 0$) for small system sizes, where intrinsic fluctuations can lead to a global extinction [88]. For larger system sizes in the active phase, the mean extinction time

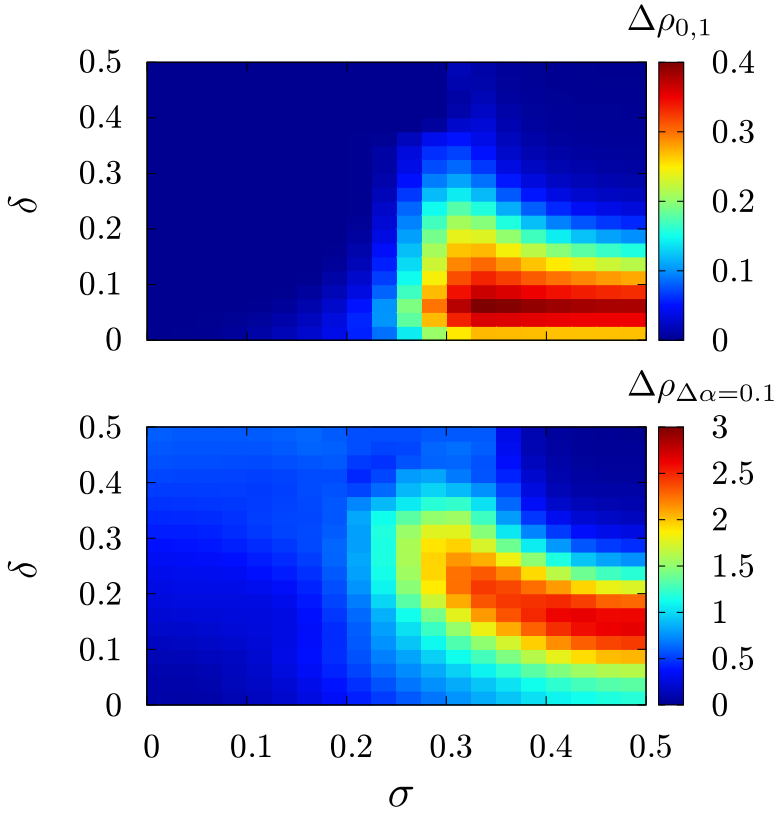


Figure 4.5: **Relative gain enhancement of hybrid strategies**. **(Top)** Density of the optimal strategy α^*, compared to the best pure strategies. **(Bottom)** Density of the optimal strategy, compared to similar strategies $\alpha^* \pm \Delta\alpha$ ($\Delta\alpha = 0.1$ here). Note that lower resolution, in contrast to previous figures, comes from that, for each point (δ, σ), the whole function $\langle \rho(\alpha|\delta, \sigma) \rangle$ has to be computed.

increases exponentially with L (see section 3.4), thus the system can be considered, for all purposes, stable and robust.

To this end, we compute the mean-extinction time in a small community (10×10) for different values of α; we have not analyze the behavior with L in this case. Fig. 4.6 illustrates that i) as expected, the extinction times are reduced when increasing both the inbreeding depression (δ) and environmental variability (σ), and more interesting ii) the *optimal* strategy, in this case the one which survives more generations, always lies at intermediate values of α.

At this point, it does not seem clear what is the best strategy: the one producing higher densities, or the one where the system survives more? As the answer does not

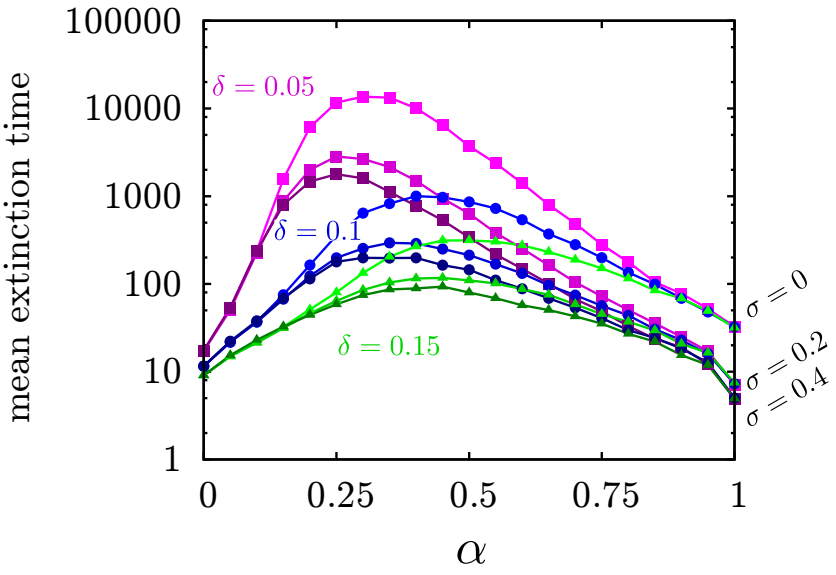


Figure 4.6: **Mean extinction time for hybrid strategies as a function of the bet-hedging parameter** α: Parameters of the pure strategies are set to $p_{\text{int}} = \bar{p}_{\text{ext}} = 0.3$. In this case, as computational times grow exponentially with the system size, we restricted ourselves to a small system size $L = 10$ ($N = 100$) for technical reasons.

seem evident, in the next section we implement a genetic algorithm-based dynamics to study the self-organized stable strategy.

4.5 Evolutionary stable strategy

To identify the *best* strategy in terms of stability of the ecosystem, we simulate a second variant of the model in which α becomes a dynamical variable, tuned self-consistently through inheritance and mutation, i.e. with a genetic algorithm [54, 56].

In this case, every plant is characterized by its own parameter α; when it reproduces, offspring inherit its α with a small random mutation, that we suppose Gaussian distributed with zero mean and standard deviation ν. To ensure that α belongs in the interval $[0, 1]$, we impose *by hand* that, every time the mixture parameter escapes from it, we put it at either 0 or 1, respectively (*hard* bounds). Other inheritance rules could be used for the same purpose in a more natural way (*soft* bounds), for instance by implementing an α-dependent mutating rule. Here, we restrict the analysis to the simpler rule.

Note that this genetic algorithm constitutes a simple manner of measuring the stable, self-organized strategy, rather than representing a realistic evolutionary dynamics. In fact, one could argue that mutations from $\alpha = 0,1$ to $\alpha' \neq 0,1$ would require an extra cost, because plants have to develop and maintain the hybrid mechanism (flower and fruit tissues, architectural traits, etc.). However, the type of dynamics including costs are far from the goal of this work and we will keep it for a future study.

After some generations, and regardless of the initial condition, the distribution of αs in the community, $P(\alpha)$, reaches a stationary shape. In our simulations, we have chosen a mutation deviation $\nu = 10^{-3}$, which, although too high and biologically unrealistic, gives a good trade-off between accuracy and short convergence times.

Results for the stationary density, as well as the mean and standard deviation of α in the community, are plotted in Fig. 4.7. From top to bottom, we find:

- The density is non zero even for higher values of the inbreeding depression δ. However, the population always becomes extinct when both δ, σ are high enough.

- For low environmental variability ($\sigma \ll 1$), plants prefer the spreading strategy (red region in the middle panel), while hybrid strategies are only selected for high values of σ (green region). Note that the internal pure strategy $\alpha = 0$ (blue region) is rarely selected, highlighting its poor dispersal efficiency.

- Additionally, distributions are highly peaked at the optimal α^*, as inferred from the small value of the deviation of α (bottom panel). Therefore, we do not encounter heterogeneous populations from the genetic algorithm implementation, where many different strategies may coexist; instead, populations appear to be very homogeneous.

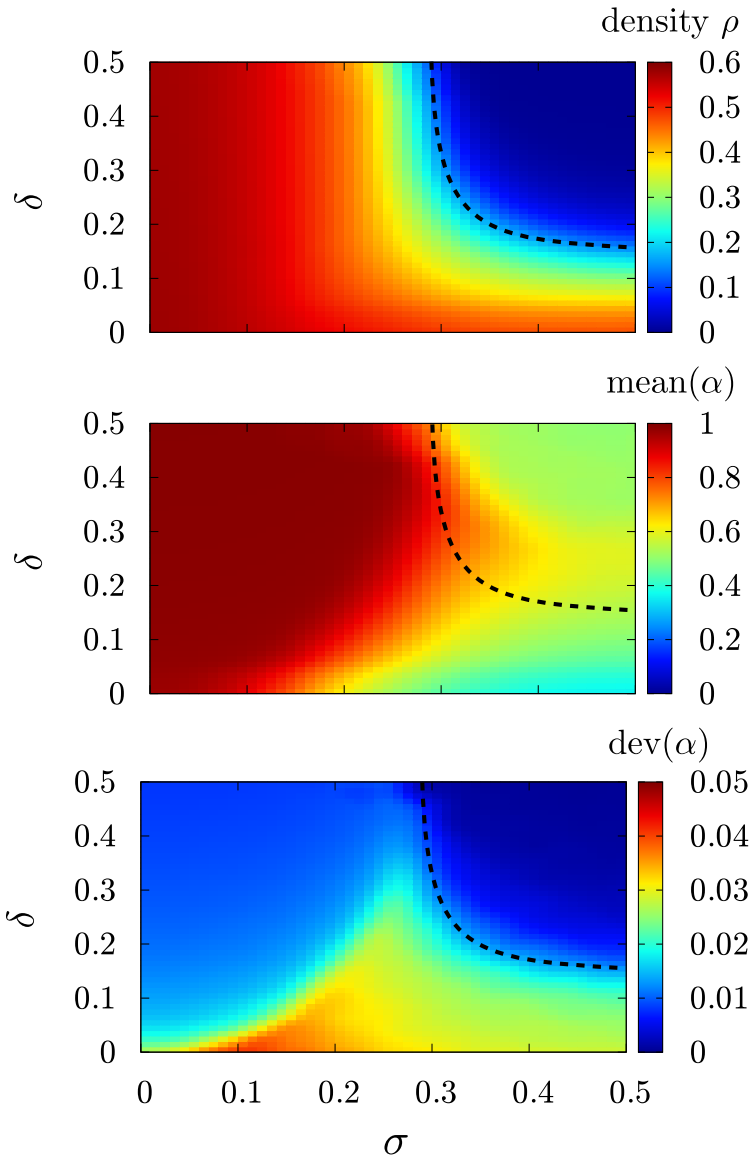


Figure 4.7: **Self-tuned strategy with a genetic algorithm**: the bet-hedging parameter α is optimized dynamically through a genetic algorithm. Individuals inherits its strategy α from its ancestor with a small Gaussian mutation of zero mean and deviation ν, with *hard* bound conditions in $\alpha = 0, 1$. **(Top)** Stationary density for different values of δ and σ. The dashed line separates the absorbing from the active phase. **(Middle)** Mean strategy in the community. **(Bottom)**: Deviation of α; as the strategy is well defined (narrow histograms), so we do not find heterogeneous populations. Parameters are set to $L = 100$, $p_{\text{int}} = \bar{p}_{\text{ext}} = 0.3$, $\nu = 10^{-3}$.

4.6 Chapter summary

In this Chapter, we have analyzed the case of mixed dispersal strategies developed by plants, as a specific example of bet-hedging in nature. Plants with hybrid dispersive mechanisms reproduce simultaneously trough a highly risky but profitable strategy (with dispersive seeds), and a non-variable but poor one (with terrestrial seeds coming from self-reproduction, also called inbreeding). In particular, our work can be summarized as follows:

- We have implement a model in which individuals (plants) reproduce by means of a hybrid dispersal mechanism; the model includes several biologically realistic ingredients, such as competition for the resources, environmental variability and inbreeding depression. The relative amount of energy (seeds) invested on each strategy is controlled by the parameter α.

- In the absence of any environmental variability and inbreeding depression, a pure dispersal strategy leads to a higher density than a non-dispersive one. Environmental variability has a negative effect in dispersal, while inbreeding depression radically leads to extinction in a population performing only the non-dispersive strategy.

- We have computed the population density for different values of the parameter α, with different levels of environmental variability and inbreeding depression. Depending on these values, we find that the optimal strategy can be either dispersal or a combination of both dispersal and non-dispersal; the pure non-dispersive strategy rarely becomes an optimal solution.

- We have seen that, for certain scenarios, a hybrid strategy not only maximizes the density: when compared to any of the pure strategies, its relative benefit is much enhanced. We have measure this relative gain with two different metrics, concluding that larger relative benefits are obtained for highly variable environments and moderate values of the inbreeding depression,

- We have seen how the mean extinction times are reduced when introducing a hybrid mechanism.

- Finally, we have checked the robustness of such optimal strategies with an evolutionary dynamics, where each individual is characterized by its own parameter α inherited from its predecessor with a small mutation. After many generations, the average $\langle\alpha\rangle$ converged to the previous optimal solutions. No stable heterogeneous populations emerged from this implementation, as all individuals exhibited a similar value of α.

In conclusion, our main finding in Chapter 3 still holds, and the benefits of bet-hedging, in particular of mixed dispersal strategies, are increased in highly fluctuating conditions. However, we may ask ourselves why all plants in nature do not exhibit such a mixed dispersal mechanism. The answer might rely in the cost derived of developing and maintain a hybrid reproductive physiology, which has not been included in our analysis. Therefore, plants in nature, through evolution and adaptation, make the best choice depending on the trade-off between costs and benefits. Although we have not taken into account this fact, our study provides a qualitative framework to understand under which conditions individuals are more likely to evolve and develop a hybrid dispersal mechanism.

Chapter 5

Stochastic amplification in neural dynamics

5.1 *'Rythms of the brain'*

During the last decade, empirical evidence has been flourishing that the brain could operate at the vicinity of a critical point [118]. At this point, "the edge of chaos", critical dynamics might be useful for several cognitive tasks, such as optimal computational capabilities[83], optimal transmission and storage of information[10], and maximal sensitivity to sensory stimuli[77].

Special impact has the work of Beggs and Plenz [9]. In these experiments, authors recorded the activity of intermittent neural avalanches on *in vivo* (with slices of mice cortex submerged into cerebrospinal fluid) and *in vitro* experiments (grown neuron cultures). In both cases, neurons exhibited spontaneous activity for hours or days. A multi-electrode grid was introduced to detect the activity as well as its specific localization. The main observations were that i) the activity comes in the form of neural avalanches of variable size, and ii) the distribution of avalanche sizes is a power-law with exponent $-3/2$. This result suggested the presence of self-organized criticality, similarly to sandpiles and earthquakes, in the brain.

We also mention the work of Bialek and coauthors [101], who registered the activity of 40 neurons in the retina and mapped it into a dynamics on a spin-glass lattice. From measures of pairwise correlations between neurons in the real network, they inferred

the parameters for the artificial network in order to exhibit similar activity patterns. After this, they introduced a virtual temperature T in their artificial network, which modulated the strength of spin interactions, and computed the behavior of the system for different values of T (in particular, the specific heat). Interestingly, the spin glass lattice exhibited a critical point at $T = 1$; or, in other words, the original lattice was already poised at criticality.

However, such methodology has been sharply criticized by Mastromatteo and Marsili [89], who suggested that, in these kind of experiments, there is still some external tuning. The hidden parameter corresponds to the time window in which the activity becomes discretized, in order to be mapped in the spin glass dynamics. For instance, if the window is very narrow, most of the temporal "bins" do not catch capture any spike of activity, and hence are empty. On the other hand, if the time window is very broad, all bins contain at least one spike, and therefore all are considered as active. Only for intermediate values, the discretized time series become non-trivial. From this point of view, it is not surprising that, when inferring the Ising parameters, any non-trivial dynamics usually lays close to the critical point.

Today, there is an ongoing discussion about critical behavior and self-organization in spontaneous neural activity[118]. Of particular interest in this context are the so-called Up and Down spontaneous oscillations [84, 85, 98, 39]. For instance, Fujisawa *et al.* [49] showed that there is a maximum susceptibility to external stimuli during Up and Down oscillations; a similar conclusion was obtained in [86], where the authors reached to control the global brain state by the stimulation of one single neuron during Up and Down states.

However, in this Chapter we skip the question of criticality and focus on the spontaneous dynamics itself, specifically for Up and Down states, and we analyze how the mechanism of "stochastic amplification of fluctuations" , previously reported in the context of Ecology[94] and Epidemiology[1], operates on these oscillations. Given its general character, such mechanism could be of broad applicability on unraveling the origin of different rhythms of the brain.

5.1.1 Up and Down states

Cortical neurons are bistable; as a consequence, the neuron membrane potential fluctuates between a quiescent, close to the resting potential, *Down* state, and a highly depolarized, *Up* state [147]. Additionally, the membrane potential alternates between these two metastable states, exhibiting *Up and Down* transitions [143] (see top panel of fig. 5.1). Such bistability has been also observed in a mesoscopic level (at the local field potential, LFP), in the form of a collective oscillatory rhythm [114, 33] (see bottom panel of fig. 5.1). Two different hypothesis for the emergence of this global bistability have been purposed: i) It is a consequence of an intrinsic neuron bistability [114, 142, 8], or instead ii) it is the result of a collective interaction, where balancing mechanisms such as inhibition and synaptic depression play a key role [109, 64, 121, 108, 98]. Even if its nature is not universally agreed upon, most of the existing computational models for cortical Up and Down states feature network rather than cellular mechanisms [143]. We focus on network models in which global bistability emerges as a collective network phenomenon.

Additionally, the importance of Up and Down states relies on is relationship with several cognitive tasks, such as working memory, selective attention and memory consolidation (see [93] and references therein). However, despite a large number of studies on Up and Down oscillations, their function and role at the global network level are not fully understood [93].

Given the apparent dichotomy between slow and high-frequency oscillations, the empirical finding that slow and fast rhythms may coexist might sound surprising, but it has been shown to occur by different experiments. In particular, we would like to clarify two interesting observations:

- High-frequency oscillations have been observed to occur *within* the active, Up intervals of slow oscillations but not in Down states [128, 103, 49, 31]. In particular, the associated Up state power spectra develop a non-trivial peak between 20 and 30 Hz, together with a substantial increase in the spectral power all along the β/γ $(10 - 100$ Hz) range (see fig. 5.2 Left for a reproduction of fig. 1D in [31]).

- While global network measurements reveal robust oscillations in the β/γ range in the Up state, individual membrane potentials or synaptic events detected at the

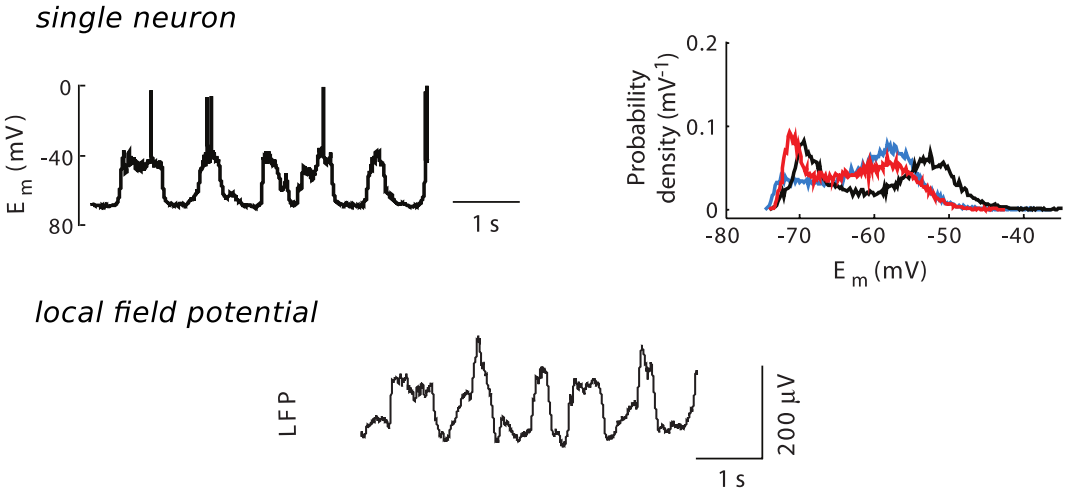


Figure 5.1: **Up and Down states observed in simultaneous whole cell (Top) and local field potential (Bottom) recordings in rat cortex**. In all cases, the characteristic frequency of oscillations lies around $\sim$ 1 Hz. (Top-left panel) Neuron bistability is also evidenced by the distribution of membrane potentials. (colors represent histograms for different neurons) *Source*: taken from [86].

intracellular level do not show any trace of oscillations in this range of frequencies (see fig. 5.2 Right for a reproduction of fig. 4 in [31]).

This suggests that *i)* oscillations within Up states are a collective phenomenon emerging at the network level and *ii)* that there is no global synchronization locking the rhythms of individual neurons to the systemic one.

At the modeling side, several authors have before addressed some of these issues and computed, in particular, the power-spectrum of network oscillations. For instance, Kang *et al.* [74] studied a mean field model in the presence of noise. They performed an analytical calculation of the power spectrum of a Wilson-Cowan-like model with excitatory and inhibitory neurons and showed the emergence of a resonant peak at γ frequency. In a similar model, Wallace *et al.* [138] made the noise variance to scale with the network size and derived analytically the power-spectrum showing that it is possible to have coexistence of high-frequency oscillations for the population without having oscillations for individual neurons. On the other hand, for spiking neural net-

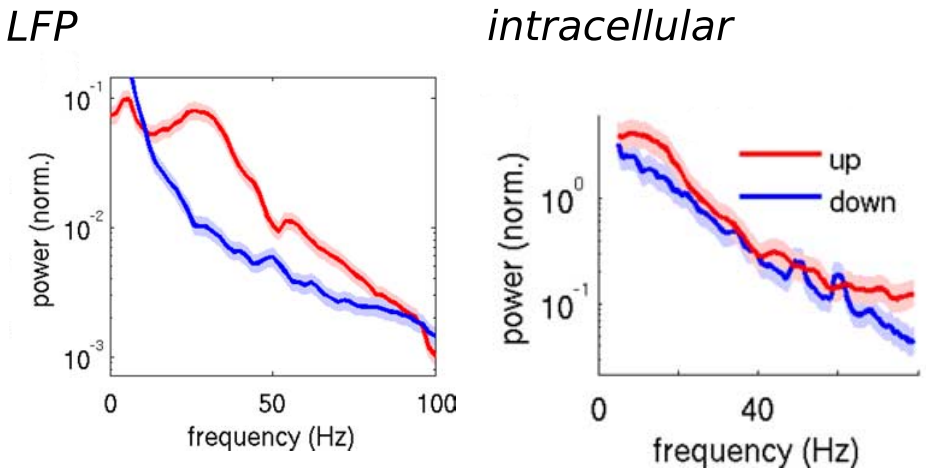


Figure 5.2: **(Left) Power spectra of oscillations within Up (red) and Down (blue) states, measured at the LFP**. A characteristic peak emerges in the range 20−30 Hz for Up states, illustrating the existence of a fast oscillation; none of this occurs for Down states. Such phenomenon has been observed by several experimental groups [128, 103, 49, 31]. **(Bottom) Power spectra of oscillations within Up (red) and Down (blue) states, measured at the intracellular level**. Simultaneous to the LFP measurements, individual records do not exhibit any trace of the previous fast rhythm within Up states, suggesting that such oscillations are a collective phenomenon, emerging at the global network level. *Source*: taken from [31].

works, Spiridon and Gerstner [126] showed that the noise accounting for network-size effects affected the power-spectrum of the population activity. Similarly, and by using a Fokker-Planck formalism, Mattia and Del Giudice [90, 91], described the time evolution of the average network activity in presence of size-effects noise, and analytically derived its power spectrum and their resonant peaks.

Even if much has been written and is known about neural oscillations, the goal of this chapter is to shed some more light on the previously discussed questions by studying general aspects, beyond modeling details, as well as a simple and general theory accounting in general for the above described phenomenology and, in particular, for the asymmetry between Up state and Down state power spectra [61, 60]. Deciphering the cerebral oscillations and their correlates to behavior and function are still major challenges in Neuroscience [23, 73].

5.2 Synaptic dynamics

In this section, we briefly introduce some of the basic concepts in neural physiology. Specifically, we focus on the mechanisms by which neurons communicate to each other, the synapse.

Neurons encode and transmit the information by changing their membrane potential. This potential varies as a consequence of –active– ionic currents flowing across the cellular membrane, thus changing the internal ionic concentration respect to the surrounding media. In this sense, neurons act as a capacitor, storing electric voltage.

During the synapse, the pre-synaptic neuron *sends the information* to the post-synaptic neuron, by means of modulating the membrane potential of the second one. This can be done by releasing neurotransmitters; these chemical substances open or close the ionic flows in the post-synaptic neuron, and consequently, changing its voltage.

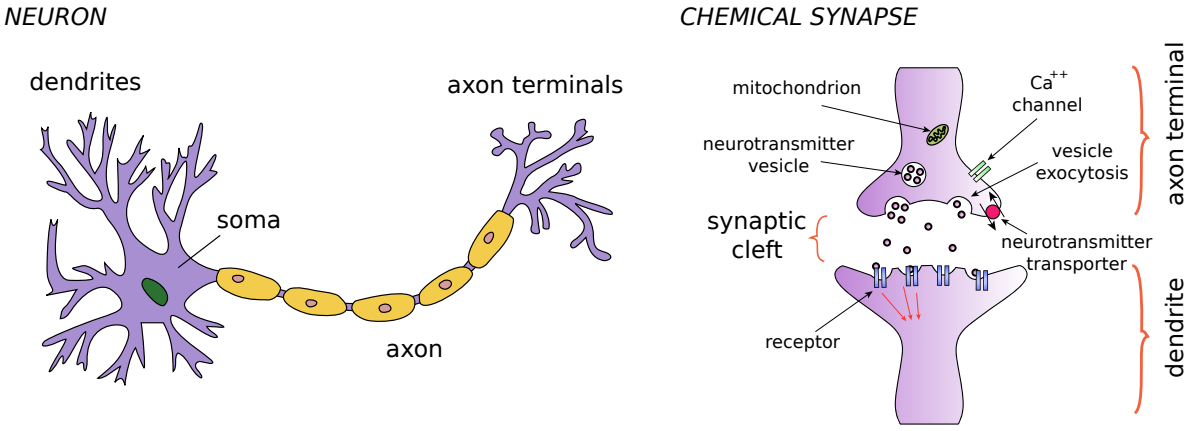


Figure 5.3: **Sketch of a neuron and chemical synapse**. *Source*: Wikipedia (Creative-Commons license).

In a typical neuron (see fig. 5.3) –multiple anatomical shapes could be distinguished–, we find three principal components: i) the *soma*, constituting the main cell body, where the nucleus is located; ii) the *dentrites*, branching many times and collecting the synaptic inputs; and iii) the *axon*, by where the electric impulse is directed to other neurons.

When the neuron membrane potential is high enough, such electric impulse is sent along the axon –the action potential (AP)–, activating the synaptic connections. The

AP was first modeled by Hodgkin and Huxley [63], who represented the basic ingredients in the AP generation as different components of an electric circuit. A less detailed description is given by the oscillator-based model of FitzHugh and Nagumo [47, 48]. For our purposes, we do not enter into such a detailed level, and simply use the integrate-and-fire (IF) approach [21].

In the synapse, we distinguish two main types: i) the *electric* synapse, consisting in a direct linkage, as the cytoplasms of the pre- and post-synaptic neuron are directly joined, allowing for an ion flux between the two cells;, and ii) the *chemical* synapse, in which neurons are separated by an inter-synaptic gap, and the communication is performed via neurotransmitters. The mathematical models we describe in this chapter are based on this second type of synapse.

In a nutshell, the chemical synapse works as follows: the synaptic terminals are plenty of vesicles fulfilled with neurotransmitters. The action potential induces the open of Ca^{2+} ion channels, which enters into the neuron and releases the vesicle content [13]. Neurotransmitters are received by the post-synaptic neuron, and, depending on its chemical nature, induce the depolarization or hyperpolarization of the second cell (excitatory and inhibitory synapse, respectively).

This process constitutes a dynamical mechanism, as the vesicle release (depletion), the Ca^{2+} concentration, as well as the amount of available neurotransmitters depend on the neuron activity, which changes for each time. For instance, after a depletion, vesicles take a time to recover and be recycled for a latter use; therefore there is an activity-depressor dynamics, called *short-time synaptic depression* (STDP) [131]. On the other hand, if Ca^{2+} ions are already present in the axon terminal, as remains of a previous spike, vesicles are more likely to be released with a second spike; this activity-enhancing mechanism corresponds to the *short-term synaptic facilitation* (STSF)[72].

The synaptic dynamics cited above introduce an activator-depressor mechanism, leading to a plethora of non-linear emerging phenomena in neural dynamics. In particular, STDP and STDF have been modeled to describe Up and Down states [131, 108, 98], which are the central focus of this chapter. In the next section, we describe the models with further detail.

5.3 Models for Up and Down states

Hereafter, we present two different network models reproducing the dynamics of Up and Down states, one based on a mean-field single population model (Model A) and one based on a network of spiking-neurons (Model B). Our strategy is to keep models as simple as possible to uncover the essence of Up and Down oscillations.

5.3.1 Model A: Minimal coarse-grained model

The simplest possible model for Up and Down states including synaptic depression has a deterministic dynamics, and characterizes neural network activity by a global ("mean-field") variable, the population averaged firing rate (which is a proxy for measurements of local field potential). We focus here on the model proposed by Tsodyks *et al.* [131, 132] including short-time synaptic depression as the key regulatory mechanism (in section 5.4.5 we present results for a similar coarse-grained model with inhibition). This model is described by the mean membrane potential v, and the variable u accounting for the strength of synaptic depression (see fig. 5.4). This second variable mimics the amount of available resources (varying between 0 and 1) in the presynaptic terminal to be released after presynaptic stimulation; thus, the larger u, the more synaptic input arriving to the postsynaptic cell.

The mean voltage grows owing to both external and internal inputs, and decreases owing to voltage leakage. On the other hand, synaptic resources are consumed in the process of transmitting information and generating internal activity (providing a self-regulatory mechanism) and spontaneously recover to a target maximum value, fixed here to $u = 1$:

$$\dot{v} = -\frac{v - V_r}{\tau} + p_r u V_{\text{in}} f(v) + I_e \tag{5.1}$$

$$\dot{u} = \frac{1 - u}{\tau_r} - p_r u f(v), \tag{5.2}$$

where $\tau = RC$ (R membrane resistance and C capacitance) and τ_r are the characteristic times of voltage leakage and synaptic recovery, respectively, V_r is the resting potential, p_r is the release fraction indicating the efficiency of synapses, V_{in} is the contribution of internal inputs, f is the network firing rate and I_e represents the external current (in

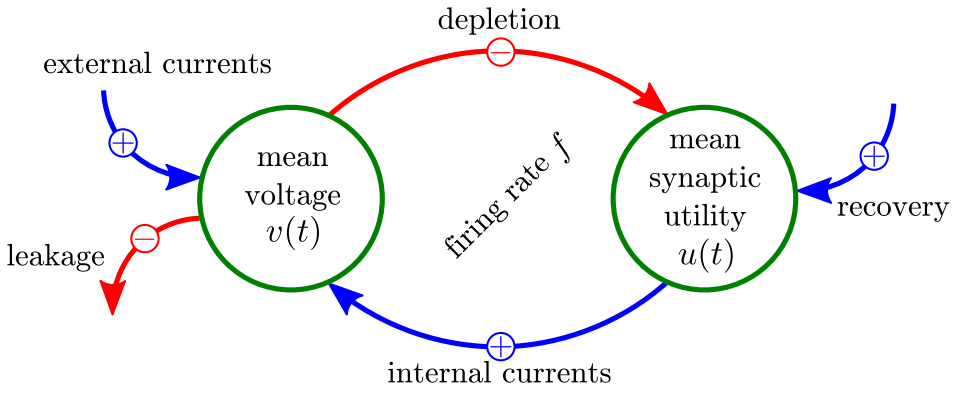


Figure 5.4: **Coarse-grained Model A**[131, 132]: the dynamics is represented by two mean-field variables: the mean voltage, v, and the mean synaptic utility, u. The voltage v is constantly subjected to internal leaks, while the resources u recover to the target value ($u \to 1$). Additionally, the two variables are coupled through the firing rate f, which increases the voltage by consuming the available resources. External currents can be added to increase the activity.

our analysis we simply take $I_e = 0$). The firing rate is assumed to depend on v as a "threhold-linear" gain function:

$$f(v) = \begin{cases} \alpha(v - T) & \text{if} \quad v \geq T \\ 0 & \text{if} \quad v < T, \end{cases} \tag{5.3}$$

with T the threshold value. Although we introduce this function *ad-hoc*, eq. 5.3 is a standard approach in coarse-grained models of neuron populations [141], and actually it can be obtained from a neuron-based model, as we will see in the next section for Model B. The notation and parametrization has been slightly modified with respect to the original model in [64], but the numerical values of parameters are equivalent (see. table 5.1).

In this context, Up and Down states correspond to fixed points of the deterministic dynamics with, respectively, high and low firing-rates. Spontaneous transitions between these two stable states can also be described within this framework by switching-on some stochasticity. Possible sources of noise are network size effects, sparse connectivity, unreliable synaptic connections, background net activity, synapses heterogeneity, or irregular external inputs. An instance of this stochastic approach is the work of Holcman

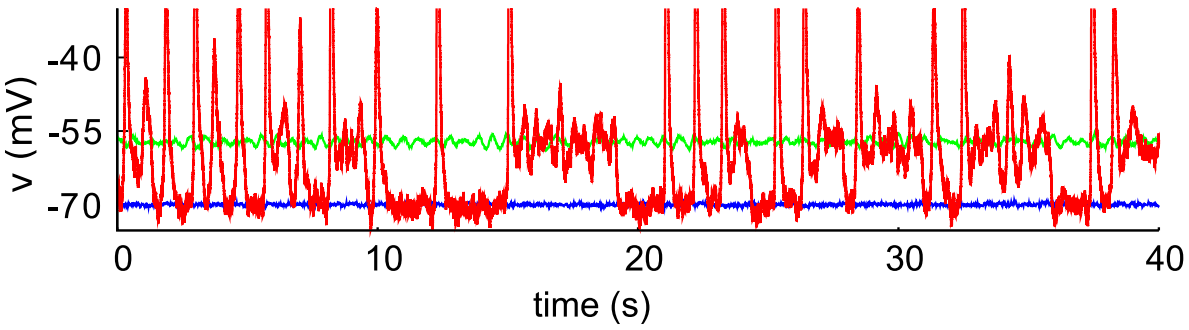


Figure 5.5: **Up and Down states and Up-and-Down transitions in Model A** [64]: time-series for the membrane potential, $v(t)$. Observe the presence of two steady states, a lower one around -70 mV (Down state/blue curve) and a larger one (Up state/green curve) at about -55 mV; these two are obtained for low noise amplitudes ($\sigma_v = 1.5 \cdot 10^{-1}$ mV$\sqrt{\text{Hz}}$, $\sigma_u = 1.5 \cdot 10^{-3}$ mV$\sqrt{\text{Hz}}$) and different initial conditions. Instead, the Up-and-Down state (red curve), corresponds to a high noise amplitude ($\sigma_v = 10$ mV$\sqrt{\text{Hz}}$, $\sigma_u = 0$). Note that, typically, the Up state intervals start with an abrupt spike which parallels empirical observations, as discussed in [64]. Parameters are listed in table 5.1

and Tsodyks [64] (see also [95]) where a noise term was introduced into the above mentioned mean-field model with synaptic depression. Indeed, adding uncorrelated Gaussian white noises, $\eta_v(t)$ and $\eta_u(t)$, of amplitude σ_v and σ_u, respectively, to eqs. 5.1 and 5.2, converts them into a set of stochastic/Langevin equations [64]. While the noiseless version of the model presents bistability, its noisy counterpart exhibits Up-and-Down transitions (see fig. 5.5). The forthcoming results do not depend crucially on the type of noise introduced.

For the chosen parameters (see table 5.1), eqs. 5.1 and 5.2 present two stable fixed points (as well as a saddle-point between them). One of them corresponds to a sustained Up-state with $(v^*, u^*) = (V_r + 12.8 \text{ mV}, 0.188)$, and the other to a Down-state $(v^*, u^*) = (V_r, 1)$. Moreover, if we decrease w_{in}, the system experiments a Hopf-bifurcation, appearing a stable limit cycle with sustained –deterministic– oscillations [64, 108], but we are not interested in this case.

Time-series produced by numerical simulations of Model A are shown in fig. 5.5. Depending on the noise amplitude different outputs are produced. For low noises, either an Up state (with a high firing rate) or a stable Down state (with mean v close to the

resting potential, and therefore with a vanishing firing rate, and mean u close to unity) coexist. For larger noise amplitude, transitions between the fixed points are induced and Up-and-Down oscillations emerge.

Parameter	**Value**
$\tau = RC$	0.05 s
τ_r	0.8 s
V_r	-70 mV
T	-68 mV
V_{in}	252 mV
α	1 Hz/mV
p_r	0.5
I_e	0
σ_v	10 mV$\sqrt{\text{Hz}}$
σ_u	0

Table 5.1: **Parameters for the coarse-grained model (Model A)**. Numerical values are taken from [64].

5.3.2 Model B: Spiking-neuron network model

Millman and coauthors [98] proposed an integrate-and-fire (neuron-level) generalization of the model above, including some additional realistic factors. These refinements allow us to compare the emerging results with empirical ones not only qualitatively but also quantitatively.

The model (Model B, from now on) consists in a population of N leaky integrate-and-fire neurons, each one connected by excitatory synapses with (on average) another K of them, forming a random (Erdos-Renyi) network. Each neuron is described by a dynamical equation for its membrane potential V_i (with $i = 1, ..., N$) in which V_i increases owing to (i) external (stochastic) Poisson-distributed inputs arriving at rate f_e and (ii) internal inputs from connected spiking pre-synaptic neurons, and decreases owing to voltage leakage (see fig. 5.6 for a graphical explanation).

When a neuron membrane potential V_i reaches a threshold value T, the neuron fires: V_i is reset to V_r and its dynamics is switched-off during a refractory period τ_{rp}. When a (pre-synaptic) neuron fires, it may open –with probability p_r– each of the n_r

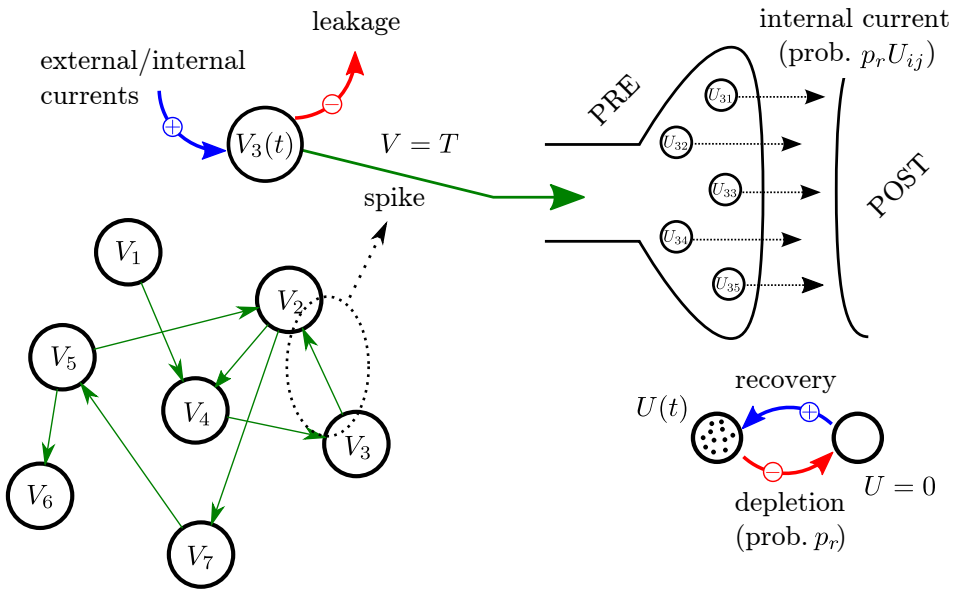


Figure 5.6: **Spiking-neuron network Model B**[98]: integrate-and-fire excitatory neurons interact in a Erdos-Renyi network. Each neuron i increases its membrane potential V_i by internal and external currents, and decreases by leaks. When V_i reaches a threshold value T, it spikes, releasing the vesicles at every synapse j with a probability p_r and resetting its voltage during the refractory period. Each vesicle introduces a current in the post-synaptic neuron, but, to account for the synaptic depression, this happens with a probability U_{ij} (available resources). U_{ij} is reseted to zero at every release and recovers to the target value $U_{ij} \to 1$. In this individual-based model, the network firing rate emerges as a consequence of the internal dynamics.

release sites existing per synapse, inducing a current in the corresponding postsynaptic neuron. External (resp. internal) inputs, $I_{\rm e}$ (resp. $I_{\rm in}$) are modeled by exponentials of amplitude $w_{\rm e}$ (resp. $w_{\rm in}$) and time decay constant τ_s:

$$I_{\rm e/in}(t) = w_{\rm e/in} \exp\left(-\frac{t-t_s}{\tau_s}\right), \qquad t > t_s \tag{5.4}$$

where t_s is the corresponding spiking time.

Similarly to Model A, a variable $U_{ij} \in [0,1]$ (for neuron i and release site j) such that the release probability is modulated by U_{ij}, i.e. $p_r \to p_r U_{ij}$, allows to implement short-time synaptic depression. U_{ij} is set to 0 immediately after a release and recovers exponentially to 1 at constant rate, τ_r.

Mathematically, the mechanisms described above are summarized in the set of equations for the membrane potential V_i of neuron i and the synaptic utility U_{ij} for each release site j:

$$\dot{V}_i = -\frac{V_i - V_r}{RC} + \frac{1}{C}\sum_k I^k_{\mathrm{e}_i}(t) + \frac{1}{C}\sum_{\substack{i'j' \\ \text{linking } i}}\sum_k \Theta(p_r U_{i'j'}(t^k_{s_{i'}}) - \zeta^k_{i'j'}) I^k_{\mathrm{in}_{i'}}(t) \quad (5.5)$$

$$\dot{U}_{ij} = \frac{1-U_{ij}}{\tau_r} - \sum_k U_{ij}\Theta(p_r - \zeta_{ij}(t))\delta(t - t^k_{s_i}), \quad (5.6)$$

where $\zeta_{ij}(t)$ is a uniform random variable in $[0,1]$ and $\Theta(x)$ the Heaviside step function. Indexes i and i' run over different neurons, j and j' over the n_r release sites, and k over spikes occurring at times t^k_s. The first term in the r.h.s. of the first equation describes the leakage, the second is the sum over external inputs, and the third represents the internal currents arriving from every release site j' in (pre-synaptic) neuron i' to (post-synaptic) neuron i. On the second equation, the first term in the r.h.s. corresponds to the recovery of the synaptic resources, while the second comes from the depletion produced by each spike.

We have scrutinized Model B by numerically integrating the corresponding integrate-and-fire stochastic equations on sparse random networks as well as on regular networks. Parameters are fixed –mostly as in [98]– to neuro-biologically realistic values (see table 5.2). We compute numerically membrane-potential and synaptic-resource time-series for each individual neuron as well as for the network as a whole. The release probability, p_r, is kept as a control parameter [131]: for intermediate values as $p_r = 0.3$ the system exhibits Up-Down transitions as illustrated in fig. 5.7; for larger values (e.g. $p_r = 0.5$) it remains steadily in the Up state, while for sufficiently low ones ($p_r = 0.2$) only Down states are observed (see fig. 5.7).

From the complex dynamics described by eqs. 5.5 and 5.6, Millman *et al.* [98] derive a theoretical prediction of the possible attractors of the dynamics. This is only an approximation, but it provides very accurate results. In a nutshell, the approach consists in converting eqs. 5.5 and 5.6 into a unique Langevin equation for the individual membrane potential; from it authors compute the associated stationary probability distribution of potentials, obtaining two possible solutions, idenfitied with the Up and Down state, respectively. In the next point we reproduce their method, but the reader

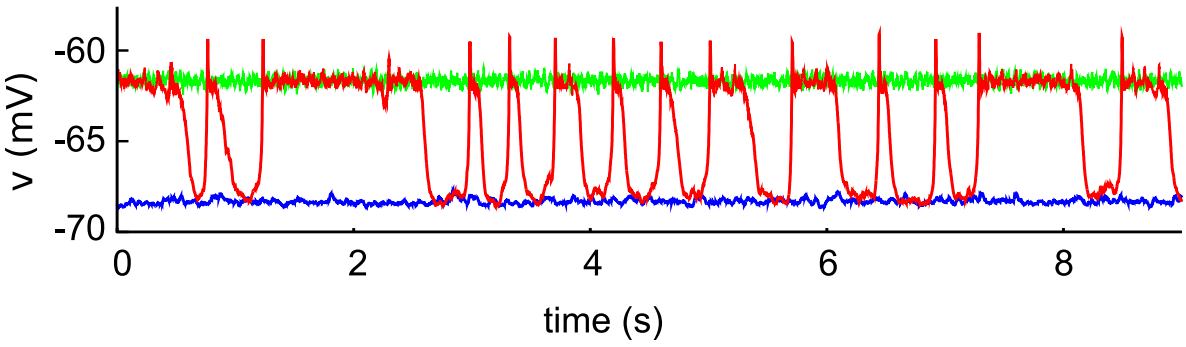


Figure 5.7: **Up and Down states and Up-and-Down transitions in Model B** [98]: Time series of membrane potential. Curves and color code are as for Model A. For $p_r = 0.3$ the system exhibits Up-and-Down transitions, for larger (smaller) values as $p_r = 0.5$ ($p_r = 0.2$), it remains steadily in the Up (Down) state. Parameters are listed in table 5.2

Parameter	**Value**	**Parameter**	**Value**
$\tau = RC$	0.02 s	w_{e}	95 pA
τ_r	0.1 s	w_{in}	50 pA
τ_s	5 ms	f_e	5 Hz
τ_{rp}	1 ms	n_r	6
V_r	-70 mV	K	7.5
T	-50 mV	N	1000
C	30 pF		

Table 5.2: **Parameter values for the spiking-neuron network model (Model B)**, taken from [98].

can skip this part and go directly to Section 5.4

Effective current contribution

To proceed, as the external inputs are Poisson-distributed for each neuron, the number of inputs per unit of time at each terminal can be approximated as a Gaussian noise with identical mean and variance f_e:

$$\frac{1}{C}\sum_k I^k_{e,i}(t) \simeq \overline{V}_e \left(f_e + \sqrt{f_e}\eta_{e,i}(t)\right), \tag{5.7}$$

where $\eta_{e,i}(t)$ is a delta-correlated Gaussian noise, f_e is the spiking frequency of external inputs and $\overline{V}_e$ is the mean contribution of one external spike, defined as $\overline{V}_e = \langle \int_0^\Delta I_e(t)/C\, dt \rangle_\Delta$ (Δ identifies the input duration, see below).

Equivalently, we can do a similar approximation for the internal currents. The simplest way is to suppose that, at the stationary state, the probability for the post-synaptic neuron to receive a spike from the pre-synaptic on is constant in time, and therefore it is Poisson-distributed (however, this is not completely true). Again, we approximate the number of spikes per unit of time as a Gaussian noise with equal mean and variance f:

$$\frac{1}{C} \sum_{i'j'} \sum_k \Theta(p_r U_{i'j'}(t^k_{s_{i'}}) - \zeta^k_{i'j'}) I^k_{\text{in}}(t) \simeq \sum_{i'=1}^K n_r p_r u \overline{V}_{\text{in}} \left(f + \sqrt{f} \eta_{\text{in},i'}(t) \right), \tag{5.8}$$

where $\eta_{\text{in},i'}(t)$ are Gaussian white noises (independent for each pre-synaptic neuron), f is the spiking frequency of the pre-synaptic neurons and $\overline{V}_{\text{in}}$ the mean contribution of one internal spike. As the topology is Erdos-Renyi, all the neurons are equivalent and f identifies the spiking frequency of the whole network. Finally, we have substituted the summation over release sites by the average contribution $n_r p_r u$, where u is the mean synaptic utility (as in Model A), $u = \langle U_{ij}(t) \rangle_{i,j,t}$. The substitution is only valid in the limit of many release vesicles, $n_r \gg 1$, otherwise we should include other stochastic terms. Notice that we have dropped the temporal dynamics in the synaptic utility u, as we are more interested in the V variable.

We can plug the equations above in the original eq. 5.5, obtaining the Langevin equation:

$$\dot{V}_i = \overline{V}_e \left(f_e + \sqrt{f_e} \eta_{e,i}(t) \right) + \sum_{i'=1}^K n_r p_r u \overline{V}_{\text{in}} \left(f + \sqrt{f} \eta_{\text{in},i'}(t) \right), \tag{5.9}$$

with the extra condition that if $V_i(t) = T \Longrightarrow V_i(t+\tau_{rp}) = V_r$. However, we still need to compute some of the quantities introduced above. The frequency of external inputs is a parameter of the system, but not the network frequency f. This is easier to understand from the Fokker-Planck scenario, that we describe in the next section, so for the while we omit this calculation. Finally, the mean synaptic utility can be computed taking

the average in eq. 5.6,

$$\dot{u} = \langle \dot{U}_{ij} \rangle_{i,j} \simeq \frac{1-u}{\tau_r} - p_r \left(uf + \sqrt{f} \langle \langle U_{ij} \rangle_j \cdot \xi_{\text{in},i}(t) \rangle_i \right). \tag{5.10}$$

This replacement is valid when the number of release vesicles n_r is large, while otherwise it is just an approximation. In the limit $N \gg 1$, it becomes

$$\dot{u} = \frac{1-u}{\tau_r} - p_r u f, \tag{5.11}$$

with its correspondent stationary point at

$$\dot{u} = 0 \Longrightarrow u = \frac{1}{1 + p_r \tau_r f}. \tag{5.12}$$

In the original work [98], the authors computed the average contribution per spike by integrating over an infinite time the exponential function in eq. 5.4. However, when a neuron fires, it is kept silent during the refractory period and it "ignores" all the arriving currents. As the tails of exponential (internal or external) currents can be interrupted by this mechanism, if we want to compute the overall contribution of the spike to the membrane potential, it is not a good approximation to take the whole integral.

We have obtained a simple and better estimation as follows. In the mean-field approach, a neuron fires every f^{-1} seconds; meanwhile, the incoming currents contribute to increment its membrane potential. Therefore, one can consider that the neuron integrates on average from its spiking-time $t = 0$ to some effective final time $t = f^{-1}$. The contribution of an incoming spike arriving at $t = t_s$, $V_{\text{e/in}}(t_s) = C^{-1} \int I_{\text{e/in}}(t) dt$, can be computed as

$$V_{\text{e/in}}(t_s) = \int_{t_s}^{f^{-1}} \frac{w_{\text{e/in}}}{C} \exp\left(-\frac{t - t_s}{\tau_s} \right) dt = \frac{w_{\text{e/in}} \tau_s}{C} \left[1 - \exp\left(-\frac{f^{-1} - t_s}{\tau_s} \right) \right], \tag{5.13}$$

where t_s is an stochastic variable. The simplest estimation is to suppose a uniform distribution for the incoming times t_s in the interval $[0, f^{-1}]$. Then, the mean value of

eq. 5.13 is:

$$\overline{V}_{\mathrm{e/in}}(f) = \int_0^{f^{-1}} \frac{1}{f^{-1}} V_{\mathrm{e/in}}(t_s)\, dt_s = \frac{w_{\mathrm{e/in}}\tau_s}{C}\left[1 - f\tau_s\left(1 - \exp\left(-\frac{1}{f\tau_s}\right)\right)\right]. \quad (5.14)$$

Observe that an extra factor, which depends on the firing rate f, appears multiplying the mean value of the exponential, $w_{\mathrm{e/in}}\tau_s/C$. This factor was omitted in the original description of the model [98], and it supposes a better estimation of the contribution of the incoming currents. For instance, it is re-scaled by a more than a 50 % for typical values of $f = 100$ Hz and $\tau_s = 5$ ms.

Self-consistent solution

For the Langevin equation for $\dot{V}_i$, eq. 5.9, the Fokker-Planck equation proposed in [98] to describe the membrane potential is:

$$\begin{aligned}\frac{\partial P(V,t)}{\partial t} &= -\frac{\partial}{\partial V}\left[\left(-\frac{V - V_r}{RC} + \overline{V}_{\mathrm{e}} f_e + K n_r p_r u \overline{V}_{\mathrm{in}} f\right) P(V,t)\right] + \\ &\quad \frac{1}{2}\left(\overline{V}_{\mathrm{e}}^2 f_e + K\left(n_r p_r u \overline{V}_{\mathrm{in}}\right)^2 f\right)\frac{\partial^2 P(V,t)}{\partial^2 V}. \end{aligned} \quad (5.15)$$

We have dropped the sub-index i, because the equation is equivalent for all the neurons. The first drift term in eq. 5.15 includes the potential leakage and the external and internal input contributions. The second diffusive term stems from the Poisson-like nature assumed for both external and internal spikes . Note that this equation describes the membrane potential of one typical neuron, but not the mean-voltage in the network.

Let remark something about the diffusive term: in the fully connected case $K = N$, assuming that internal input amplitudes are rescaled by the average connectivity (i.e. $w_{\mathrm{in}} \to w_{\mathrm{in}}/K$) in order to keep the total signal per spike constant, the internal noise disappears in the infinite size limit. In other words, the internal contribution to the diffusion term, proportional to $\overline{V}_{\mathrm{in}}$, comes from the finite connectivity of each individual neuron in sparse networks. Similarly, in the absence of external stochasticity, the external contribution to the diffusion term, proportional to $\overline{V}_{\mathrm{e}}$ would disappear for a homogeneously distributed excitation. If the two previous conditions hold, the dynamics becomes purely deterministic.

Eqs. 5.15 needs to be complemented with the following constrains:

- A neuron spikes when it achieves the threshold potential, resetting the membrane potential after the refractory period [20, 98]. Mathematically, this corresponds to the conservation of the total flux, defined as

$$\begin{aligned} F(V,t) &= \left(-\frac{V-V_r}{RC} + \overline{V}_{\mathrm{e}} f_e + K n_r p_r u \overline{V}_{\mathrm{in}} f\right) P(V,t) \\ &\quad -\frac{1}{2}\left(\overline{V}_{\mathrm{e}}^2 f_e + K\left(n_r p_r u \overline{V}_{\mathrm{in}}\right)^2 f\right) \frac{\partial P(V,t)}{\partial V}. \end{aligned} \tag{5.16}$$

  The boundary condition is then written as $F(T,t) = F(V_r, t+\tau_{rp})$.

- Finally, the firing rate, f, is computed as the fraction of neurons overcoming T per unit time, i.e. the outgoing probability flux:

$$f(t) = F(T,t). \tag{5.17}$$

As the dynamics depends on the probability flux f, which on its turn is fixed by the overall dynamics, the Fokker-Planck equation needs to be solved self-consistently. This can be done numerically (for instance, with an Euler-implicit method) giving results in agreement with those in [98]: there are two different stable states for the probability distribution (see figure 5.8), which correspond to the Up and Down states. Finally, although the dynamics have been much simplified, the real dynamics is more complex and operates far away from the stationarity. Then, transitions between these two attractors are expected to occur for a broad regime of parameter, exhibiting the Up-and-Down transitions, as it actually happens (see fig. 5.7).

5.4 Stochastic amplification of fluctuations (SAF)

In this section we review the phenomenon of stochastic amplification of fluctuations (SAF). This mechanism has been reported in the context of Ecology [94] (see also [105] for an earlier reference) and Epidemiology [1]. The purpose of the present chapter is to apply it to the Up and Down states [61], but mainly, to claim the general applicability of this phenomenon to the field of Neuroscience.

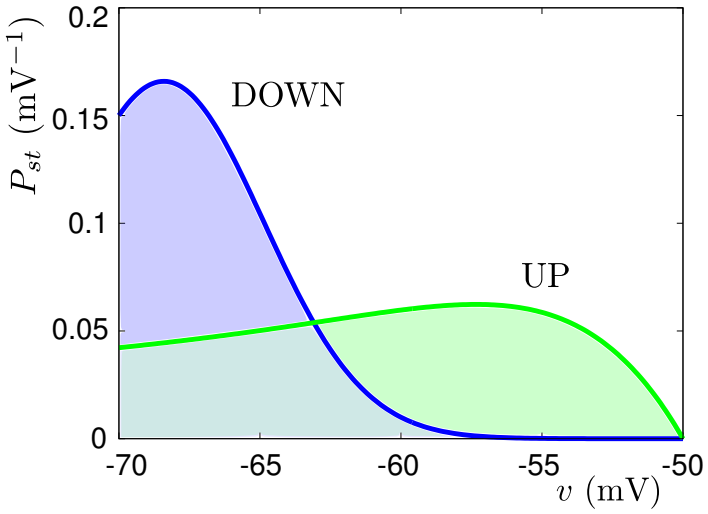


Figure 5.8: **Solutions for the membrane potential distributions described by equation 5.15** with parameters of Table 5.2. In the Down-state, membrane potentials are closer to V_r, and the slope in T gives a low firing rate $f = 0.00022$ Hz, while in the Up-state, the membrane potential biases to the threshold value T, giving $f = 74.9$ Hz. Table 5.3 includes all the relevant values computed from such distributions.

It is well known in Ecology that many populations with prey-predator interactions exhibit oscillations in nature [12]. This fact can be rationalized in a very intuitive manner: for example, imagine a population of rabbits and lynxes, with populations equilibrated at their stationary values. If, for any reason, certain year is particularly good for the rabbits (for instance, the government prohibits the use of dangerous pesticides), the population highly increases; this is a good new for the population of lynxes, so they also proliferate. But, at a certain point, such growth of the lynx population is in detrimental of the rabbits, reducing their abundance; so on, an oscillation comes out from the fluctuations around the stationary density of preys and predators.

It is surprising how these oscillations are hard to be found out from a set of deterministic equations. Detailed mechanisms, such as predator satiation [92], have to be incorporated into the basic prey-predator dynamics to reproduce the oscillations described above. In contrast, McKane and Newman show in [94] that this phenomenon can be achieved from a stochastic model in a very natural way.

Let consider a pair of Langevin equations, representing the coarse-grained dynamics of two relevant variables in the system (for instance the population densities of preys and predators):

$$\dot{v} = g_v(v,u) + \sigma_v \eta_v(t) \tag{5.18}$$

$$\dot{u} = g_u(v,u) + \sigma_u \eta_u(t) \tag{5.19}$$

with additive Gaussian white noises $\eta_{v,u}(t)$. These equations can be understood to be a simplified approach of a microscopic interaction (described in terms of a master equation), but for us, it constitutes the starting point of our analysis.

Let suppose that this dynamics leads to a stationary solution, and the pair (v,u) ends and fluctuates around the fixed point (v^*,u^*). We are interested in the nature of the attractor, so we first perform a standard linear stability analysis. Defining $x = v - v^*$ and $y = u - u^*$, one can linearize the deterministic part of the dynamics, leading to the pair of linear Langevin equations

$$\dot{x} = a_{vv}x + a_{vu}y + \sigma_v \eta_v(t) \tag{5.20}$$

$$\dot{y} = a_{uv}x + a_{uu}y + \sigma_u \eta_u(t), \tag{5.21}$$

where $a_{zz'} = \left.\frac{\partial g_z(v,u)}{\partial z'}\right|_{(v^*,u^*)}$ (z and z' standing for either v or u) are the elements of the Jacobian matrix, A, evaluated at the fixed point. This matrix has two eigenvalues, $\lambda_\pm$, that can be expressed as

$$\lambda_\pm = \Gamma/2 \pm \sqrt{\Gamma^2/4 - \Omega^2}, \tag{5.22}$$

with $\Omega^2 = \det(A) = a_{vv}a_{uu} - a_{vu}a_{uv}$ and $\Gamma = \mathrm{Tr}(A) = a_{vv} + a_{uu}$.

A useful tool to identify oscillations in noisy time-series is the power spectrum $P_x(w) = \langle |\widetilde{x}(w)|^2 \rangle$, where $\widetilde{x}(w)$ is the Fourier transform of $x(t)$, $\mathcal{F}[x] = \widetilde{x}$ (similarly $P_y(w)$ for $y(t)$), and $\langle . \rangle$ stands for independent runs average. The goal is to seek for the possible peaks in the power spectrum, identifying the characteristic frequencies. Fourier transforming eqs. 5.20 and 5.21, and solving for $\tilde{x}(w)$ and $\tilde{y}(w)$ and averaging its squared modulus, we find

$$P_z(\omega) = \frac{\alpha_z + \sigma_z^2 \omega^2}{\left[\Omega^2 - \omega^2\right]^2 + \Gamma^2 \omega^2} \tag{5.23}$$

where z stands for x or y, and $\alpha_x = a_{vu}^2\sigma_y^2 + a_{uu}^2\sigma_x^2$, $\alpha_y = a_{uv}^2\sigma_x^2 + a_{vv}^2\sigma_y^2$.

For low values of the noise amplitude ($\sigma_z \ll 1$), we can neglect the dependence of ω in the numerator, and the maximum of the power spectra can be found where the denominator minimizes, resulting at the frequency:

$$\omega_0 = \sqrt{\Omega^2 - \Gamma^2/2} = \sqrt{-a_{vu}a_{uv} - (a_{vv}^2 + a_{uu}^2)/2}. \tag{5.24}$$

If ω_0 is not a real number, no peak is found, and therefore the power spectrum is monotonous. In particular, to have a real ω_0 requires that both a_{vu} and a_{uv} are non-vanishing and of opposite sign; when this happens, both eigenvalues $\lambda_\pm$ are complex. A detailed analysis of the conditions needed to achieve a peak in the power-spectrum is described in the Appendix D.

Some remarks are in order:

- The presence of a non-trivial peak in the spectrum of fluctuations reflects the existence of *quasi-cycles* of a leading characteristic frequency coexisting with many other frequencies (as the peak is broad), producing a complex oscillatory pattern.

- The location of ω_0 does not depend on the noise amplitude; it is determined by the elements of the Jacobian matrix A, i.e. by the deterministic dynamics. This is interesting because, sometimes, only the deterministic part of the dynamics are known.

- Notice that, even if the peak location ω_0 is noise independent (as long as the noise amplitude does not vanish) the very presence of a peak is a noise induced effect: in the noiseless limit the system reaches a fixed point, without any oscillations.

- Additionally, the characteristic frequency is equal for the two variables $x(t)$ and $y(t)$. Moreover, only one stochastic source is needed to perform such oscillations, thus similar results are found by making $\sigma_v = 0$ or $\sigma_u = 0$ (but not both).

- The peak only appears when the eigenvalues $\lambda_\pm$ are complex, i.e. the relaxation towards the stable fixed point should be in the form of damped oscillations (spiral trajectories). However, ω_0 differs from the damping frequency ω_d: $\omega_d = \sqrt{\Omega^2 - \Gamma^2/4} \neq \omega_0$.

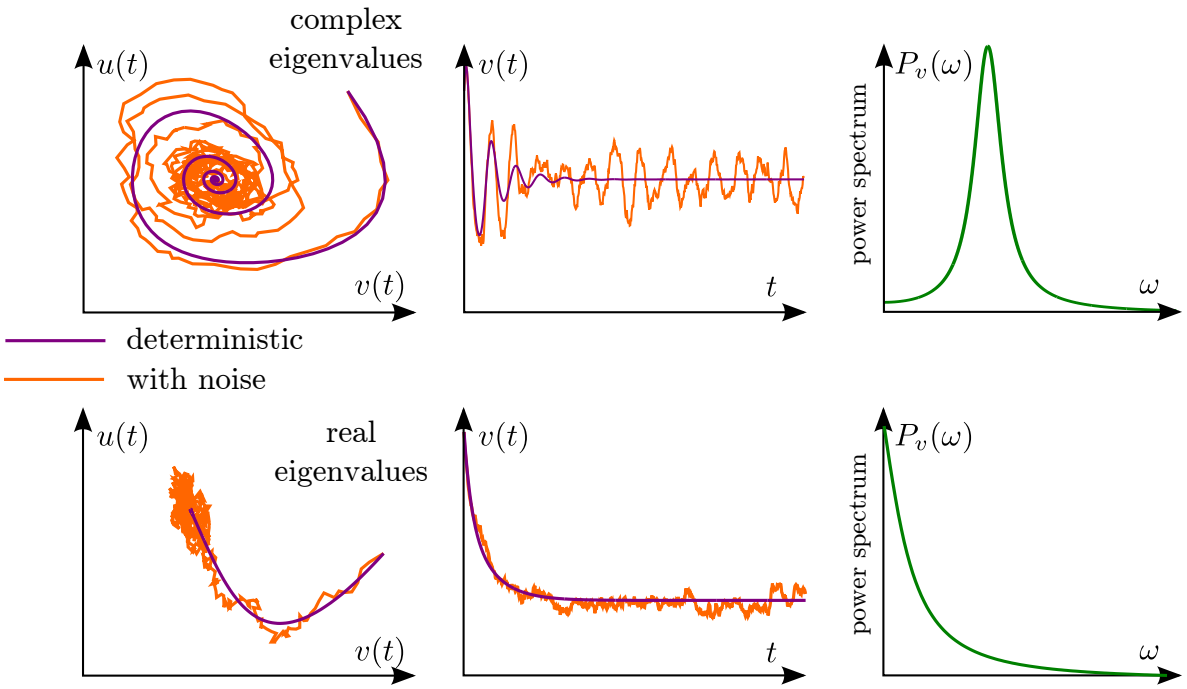


Figure 5.9: **Mechanism of stochastic amplification of fluctuations**: two coupled dynamical variables $v(t)$ and $u(t)$ relax and fluctuate around a stable fixed point. **(Top)** If the Jacobian matrix (evaluated at the fixed point) has complex eigenvalues, trajectories have spiral form in the (v, u) plane; the system tries to relax to the fixed point, but noise *kicks* it away, leading to an effective *quasi*-oscillation with a characteristic frequency similar, but not equal, to the damping frequency, as it is shown by the peak of the power-spectrum of fluctuations (right). **(Bottom)** None of this occurs for the case in which eigevalues are real (straight trajectories), and consequently, no natural frequencies are found (monotnous power spectrum of fluctuations).

- Finally, if the equations become decoupled, the eigenvalues $\lambda_\pm$ become real, therefore the power-spectrum does not exhibit any peak and the possibility of SAF is lost.

In conclusion, SAF requires the presence of some noise source acting on top of the underlying deterministic stable fixed point with complex eigenvalues $\lambda_\pm$, i.e. with spiral damped trajectories (this is, it is a "focus") with a not too small damping frequency (details are explained in the Appendix D). Noise "kicks" the system away from the fixed point, and amplifies predominantly some frequency which turns out to be slightly

different from the characteristic frequency of the deterministic damped oscillations. It is also noteworthy that a set of at least two coupled equations is required to have complex eigenvalues, and hence, too simplistic models in terms of only one effective variable cannot give raise to SAF.

In the next section, we seek for this mechanism in the models of neural dynamics described above.

5.4.1 SAF in Model A

By performing a linear stability analysis of equations for $\dot{v}$ and $\dot{u}$ (eqs. 5.1 and 5.2) around the stable fixed points (v^*, u^*) as described above, we have measured the power-spectrum at either the Up state and the Down state, both analytically and numerically. The Jacobian matrix is

$$A = \begin{pmatrix} a_{vv} = -\frac{1}{\tau} + p_r u V_{\text{in}} f'(v) & a_{vu} = p_r V_{\text{in}} f(v) \\ a_{uv} = -p_r u f'(v) & a_{uu} = -\frac{1}{\tau_r} - p_r f(v) \end{pmatrix}. \tag{5.25}$$

When evaluating at each stable fixed point, $(v^*_{\text{up}}, u^*_{\text{up}})$ or $(v^*_{\text{down}}, u^*_{\text{down}})$, the firing rate –and its derivative– plays a fundamental role in the non-diagonal terms of the Jacobian matrix. In the case of the Up state, as both $f(v^*), f'(v^*) > 0$, the fixed point turns out to be a focus, with complex eigenvalues, satisfying the conditions for the existence of a non-trivial peak in the power spectra for both v and u. On the other hand, the Down state fixed point (owing to the vanishing firing rate and, therefore, to the absence of crossed coupling terms $a_{vu} = a_{uv} = 0$ in eq. 5.25) is a node with real eigenvalues and, consequently, there is no peak in the power-spectrum.

The power-spectrum of the time series of Model A is illustrated in fig. 5.10. Observe i) the perfect agreement between analytical and numerical results in all cases, ii) the presence of a peak (around 1.6 Hz) for the v power spectrum in the Up state (note that this rhythm is much faster than that of the Up-and-Down transitions, see fig. 5.10), as well as iii) the absence of a similar peak for the Down-state.

Summing up, *a mean-field single-population model in presence of short-term synaptic depression as the key regulatory ingredient reproduces Up-and-Down transitions, with a non-trivial peak in the Up state power spectrum emerging as a consequence of the phenomenon of SAF; moreover, no analogous peak is found in Down states.* Numerical

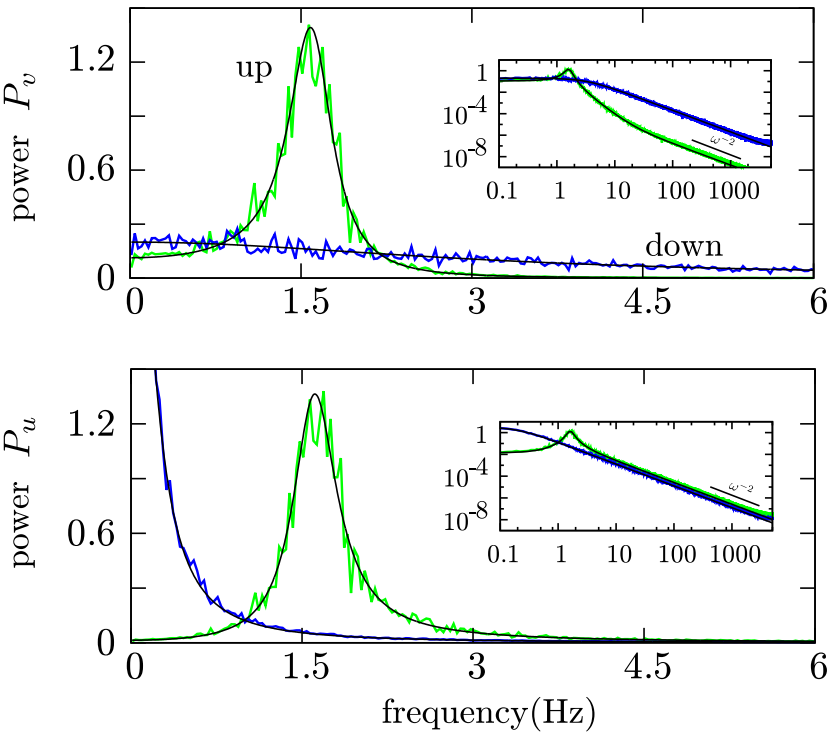


Figure 5.10: **Power spectrum of membrane potential $v(t)$ (Top) and mean synaptic utility $u(t)$ (Bottom) time-series in Up and in Down states computed in Model A**. Curves are normalized to unit area. The main plots show the power-spectra in linear scale: a pronounced peak appears for the Up state (green curve) around ≈ 1.6 Hz. Instead, there is no track of similar peaks for Down states (blue curve). Observe the excellent agreement between simulation results (noisy curves) and analytical results, eq. (5.23) (black lines). Insets represent analogous double logarithmic plots, illustrating in all cases the presence of w^{-2} tails.

results are in full agreement with this theory. To test the robustness of this hypothesis, in section 5.4.4 we have checked how these results hold when the mechanism of synaptic facilitation is turned on in the dynamics. Additionally, in section 5.4.5 we have also considered the mean-field dynamics of a simple model in presence of synaptic inhibition rather than synaptic depression.

Despite this success, the strategy of resorting to simplistic mean-field models presents some undeniable drawbacks: i) given the lack of a detailed correspondence with neurophysiological realistic parameters it is not possible to *quantitatively* compare the results

with experimental ones; ii) noise is implemented in a poorly understood way; and iii) last but not least, mean-field models do not allow for comparison of individual-neuron activity with collective rhythms, which is important to figure out whether single cells frequency lock to emergent oscillations or not. Aimed at overcoming these difficulties, in the next section we present results for a network of spiking-neurons, Model B.

5.4.2 SAF in Model B

The power-spectra of the time-series produced by Model B are illustrated in fig. 5.11 (green for the Up state, blue for the Down one, both in linear and in double-logarithmic scale). Very similar plots can be obtained –in analogy with measurements in [31]– in the Up-intervals within Up-and-Down states. In the Up state, the spectrum exhibits a sharp peak at a frequency around $\sim$ 20 Hz, together with the expected power-law decay. On the other hand, the power spectrum for Down states lacks a similar peak.

In analogy with the mean-field model in the previous section, there is a significant enhancement of the power-spectrum for Up vs Down states in the $\beta-\gamma$ range. Contrary to the model above, the more detailed neuron-level modeling and the use of realistic parameter values allows us to contrast *quantitatively* this result with the empirical findings. Observe that, in this case, there is a remarkable accordance with the observation in [31] (reproduced in fig. 5.2): the peak in the Up state spectrum lies in the range of 20-30 Hz. Let us remark that no parameter fine-tuning has been required to achieve this result.

Furthermore, Millman *et al.* showed in [98] that Up-and-Down states in their model (Model B) are robust against addition of other realistic features such as inhibitory currents, more structured (small-world) network topologies, voltage-dependent membrane resistance and so on. Also, the non-trivial peak of the power-spectra and the associated spectral power enhancement in the β/γ range for Up states, together with the absence of similar traits for Down states, are robust features against the extensions of the model we have scrutinized.

To establish the correspondence between the phenomenology described here and SAF, we need to write down a set of Langevin equations for the mean-field potential and mean synaptic utility. This turns out to be a non-trivial task, which is analyzed step by step in what follows. However, even if the analysis is a bit cumbersome, we

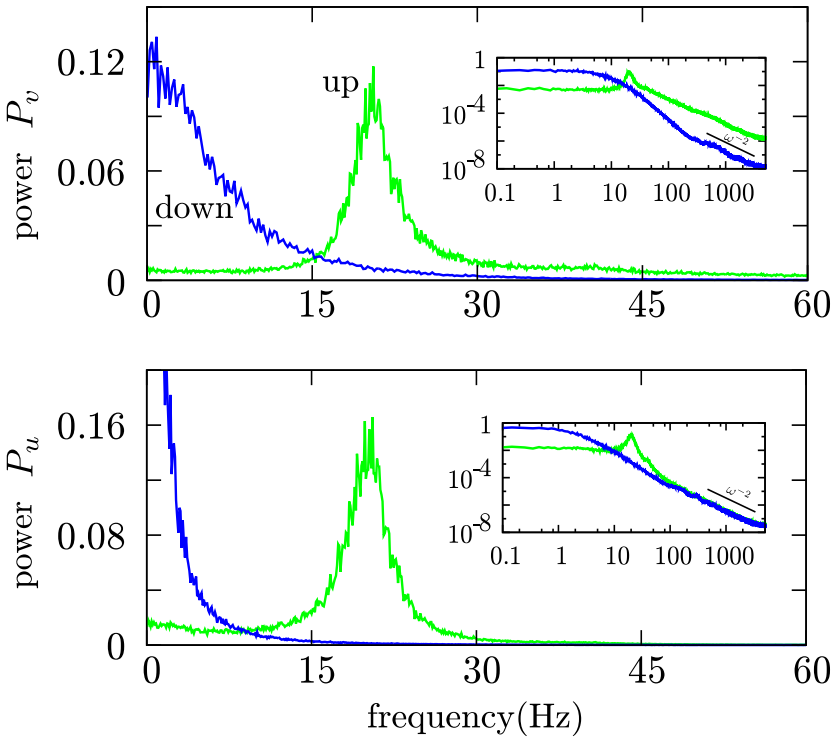


Figure 5.11: **(Top) Power spectrum of mean potential $v(t)$ and (Bottom) mean synaptic utility $u(t)$ time-series in Up and in Down states in Model B.** Similarly to fig. 5.10, a peak appears for the Up state (green curve) around ≈ 20 Hz, while there is no track of similar peaks for Down states (blue curve). In this case, a precise analytical prediction cannot be obtained. Insets represent analogous double logarithmic plots.

anticipate that the mechanism works similarly than for Model A: the two mean variables are coupled through the firing rate; in the Up state, such firing rate is high, leading to complex eigenvalues of the Jacobian matrix, and hence, a non-trivial peak in the power spectrum; in the Down state, the firing rate essentially vanishes, eigenvalues are real, and no characteristic frequencies emerge. The most interesting feature of Model B, apart of the quantitative agreement with experiments, is found at the microscopic level, described in the next Section 5.4.3.

To proceed, the starting point is to obtain the equation for the mean voltage, $v = \sum_{i=1}^{N} V_i/N$. Notice that eq. 5.9 corresponds to a Langevin equation for the membrane potential dynamics of *one* neuron instead of being the coarse-grained variable. For this

purpose, we compute the mean potential variable as

$$v(t) = \int_{V_r}^{T} V P(V,t), \tag{5.26}$$

with $P(V,t)$ the membrane potential distribution. Multiplying the Fokker-Planck eq. 5.15 by V and integrating over all possible values, after integrating by parts, one obtains

$$\begin{aligned} \dot{v} &= -(T-V_r)f(t) - \frac{v(t)-V_r}{RC} + \overline{V}_e f_e + K n_r p_r u(t) \overline{V}_{\text{in}} f(t) \\ &\quad + \left(\overline{V}_e^2 f_e + K \left(n_r p_r u(t) \overline{V}_{\text{in}} \right)^2 f(t) \right) P(V_r,t), \end{aligned} \tag{5.27}$$

where boundary conditions have been imposed and τ_{rp} has been, for simplicity, neglected. Together with eq. 5.11, the pair constitutes the set of equations for the coarse-grained variables (v,u).

Additionally, as we are interested in the limit of many-but-not-infinite neurons (i.e. the local field potential), eq. 5.27 has to be equipped with noise, coming from the finite-size population. The noise amplitude, as shown in fig. 5.12, decreases with the network-size as expected from the central limit theorem [52]. However, we have not written the noise term because i) we do not know its explicit form and ii) we do not actually need it to compute the characteristic frequency of SAF via eq. 5.24, even though it is a noise-induced phenomenon.

When studying the phenomenon of SAF, we need to compute the dynamical response when a fluctuation puts the system away from the stationarity. This is why we have explicitly written all the potential time dependences in eq. 5.27, i particular for $v(t)$, $u(t)$, $f(t)$ and $P(V_r,t)$; also $\overline{V}_{\text{in/e}}$ depend on on time through $f(t)$ (eq. 5.14).

At this point, we have wirtten the dynamical equations for the coarse-grained variables $v(t)$, $u(t)$. Moreover, the self-consistent method used to solve the Fokker-Planck, eq. 5.15, provides the numerical values of the stationary points v^*, u^*, f^* and $P(V_r)^*$ (see fig. 5.8 and table 5.3). Still, we need additional equations for the spiking frequency $f(t)$ and the density of resting neurons $P(V_r,t)$. Instead of adding more equations, we suppose that fluctuations around the stationary point do not dramatically alter the number of neurons in the resting state, as it is view from the simulations, so we keep it constant, $P(V_r,t) = P(V_r)^*$. This approximation cannot be done for the firing rate f,

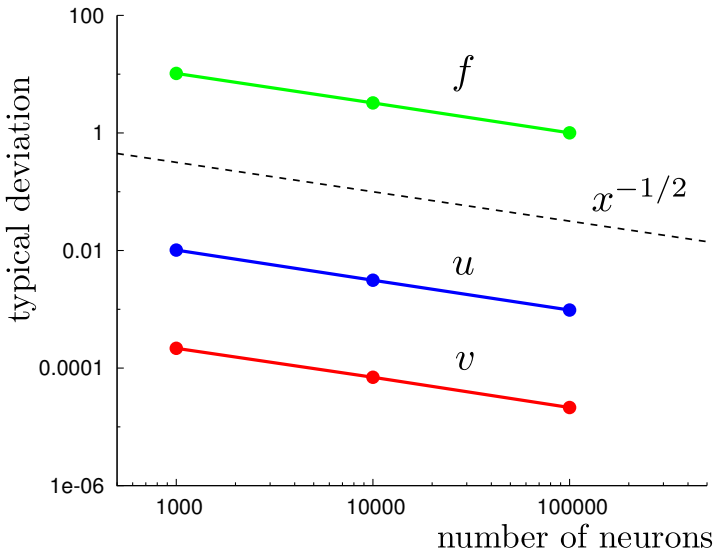


Figure 5.12: **Typical deviation of fluctuations for different variables as a function of the system size**: Simulations -computed for a persistent Up state– show a decay of $1/\sqrt{N}$, as expected on the basis of the central limit theorem.

and inspired by Model A, we suppose a dependency of the firing rate f in the shape of a threshold-lineal function, similar to eq. 5.3.

Fig. 5.13 confirms this hypothesis illustrating that $f(t)$ is strongly correlated with $v(t)$, approximately in the shape of a "split" function. The inset of fig. 5.13, where f is plotted as a function of v and u, shows the existence of two well-defined branches, one for up-to-down transitions and another for down-to-up. For simplicity, we keep our analytical study to the one-variable correlated $f(t) = f(v(t))$ and we take the numerical estimation from the simulation in fig. 5.13.

Having an analytical approximation for $f(v)$, it is now possible to perform a lineal

	Up ($p_r = 0.5$)	**Down** ($p_r = 0.2$)
v^* (mV)	-61.16	-66.54
u^*	0.2108	0.999996
f^* (Hz)	74.88	0.000216
$P(V_r)^*$ (V^{-1})	40.75	150.04

Table 5.3: **Results obtained from the Up and Down steady state distributions shown in fig. 5.8**. As the Fokker-Planck constitutes only an approximation, we find discrepancies with the simulations: $\langle v\rangle_{\rm up} = -61.67\,\mathrm{mV}$, $\langle u\rangle_{\rm up} = 0.2352$; $\langle v\rangle_{\rm down} = -68.3\,\mathrm{mV}$, $\langle u\rangle_{\rm down} = 0.997$.

stability analysis. The corresponding elements of the Jacobian matrix A are:

$$\begin{aligned} a_{vv} &= -\frac{1}{RC} + f'\left(-(T-V_r) + \left(1 + n_r p_r u \overline{V}_{\rm in} P(V_r)\right) K n_r p_r u \overline{V}_{\rm in}\right) \\ &\quad + G\left(\overline{V}_e f_e + K n_r p_r u \overline{V}_{\rm in} f + 2\left(\overline{V}_{\rm e}^2 f_e + K (n_r p_r u \overline{V}_{\rm in})^2 f\right)^2 P(V_r)\right) \\ a_{vu} &= \left(1 + n_r p_r \overline{V}_{\rm in} P(V_r)\right) K n_r p_r \overline{V}_{\rm in} f \\ a_{uv} &= -u p_r f' \\ a_{uu} &= -\frac{1}{\tau_r} - p_r f \end{aligned} \tag{5.28}$$

where G is defined as the logarithmic derivative of the re-scaling factor in the incoming currents $\overline{V}_{\rm e/in}$ (eq. 5.14), which depends on f:

$$G(f) = \frac{\overline{V}'_{\rm e/in}}{\overline{V}_{\rm e/in}} = \frac{\tau_s f' \left(\exp\left(-\frac{1}{f\tau_s}\right)\left(1 + \frac{1}{\tau_s f}\right) - 1\right)}{1 - f\tau_s\left(1 - \exp\left(-\frac{1}{f\tau_s}\right)\right)}, \tag{5.29}$$

giving a non-trivial correction.

Finally we evaluate the Jacobian matrix at the fixed points of table 5.3. The argument that follows is similar to that for the Model A, where the firing rate played the relevant role in the non-diagonal terms: at the Up-state fixed point this leads to $a_{vv} = -120.12$ Hz, $a_{vu} = 10.4272$ V·Hz, $a_{uv} = -1355.44$ Hz/V, $a_{uu} = -47.4422$ Hz for the coefficients of the stability matrix, and hence a minimum at the denominator of $P(w)$ at $w_0 = 76.1$ rad/s $\Longrightarrow f_0 = 12.11$ Hz; instead, in the Down-state, the equation for u becomes essentially decoupled from that for v, resulting in a complex ω_0 (even

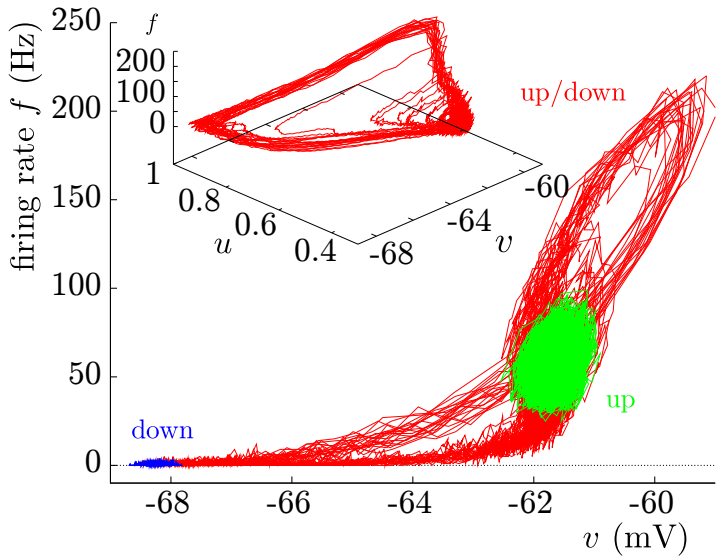


Figure 5.13: **Firing rate correlated to the the mean membrane potential** v: **(Main)** We plot f as a function of v in Model B. Red curve corresponds to the Up-and-Down state ($p_r = 0.3$, $N = 10^3$), blue to the Down-state ($p_r = 0.2$, $N = 10^3$), and green to the Up-state $p_r = 0.5$. We approximate this shape to a "split" function: $f_{\text{up}}(v) = (12.86 \pm 0.05\,\text{Hz/mV})v + (850 \pm 3)\text{Hz}$ for the Up state and $f(v) = 0$ for the Down state. **(Inset)**: f as a function of both v and u for the Up-and-Down state ($p_r = 0.3$) illustrating the origin of the two branches in the main plot.

for small but non zero firing rate, see Appendix D), i.e. in the absence of a non-trivial peak in the spectrum.

In conclusion, we have shown that also for this more complex network model, an analytical approach permits us to elucidate that the phenomenon of SAF is responsible for the non-trivial enhancement of fluctuations in the whole β/γ range for the Up states. The result is still an approximation, because the computed characteristic frequency is located at $f_0 \simeq 12$ Hz, in contrast with the 20 Hz from the simulation. Additionally, the calculation explains why the peak does not appear for the Down state. Similarly to Model A, the firing rate represents the major role in the mechanism, which strongly couples the dynamical variables for the Up states, and vanishes for the Down state, decoupling the system.

5.4.3 SAF as a collective phenomenon (Model B)

Until now, Model B has only been used to illustrate the mechanism of SAF in a more detailed and realistic model. However, its real advantage respect to Model A is that we can look at the single neuron level, and relate the individual rhythms to the collective dynamics. For this purpose, we have analyzed time-series of individual neurons and compared them to that of the global, mean-field variable, $v(t)$.

Fig. 5.14 (left) shows that individual neurons do follow the global trend in Up-and-Down states: global high average membrane potentials correspond to high firing rates at the individual neuron level; the opposite occurs for the Down-states, where the neurons remain essentially silent. On the other hand, and contrary to naive expectations, *within Up states* (as well as within Up periods of up-and-down states), *individual neurons do not lock themselves to the collective quasi-oscillatory rhythm.* As shown in fig. 5.14 (right) individual neurons fire at a much faster pace than that of the global rhythm. These results are in perfect agreement with the experiment, where individual neurons are unlocked to the global wave [31].

Actually, a histogram of the inter-spike intervals for all neurons in the network (fig. 5.15) has an averaged value ≈ 17 ms, corresponding to a frequency $f \approx 60$ Hz. Therefore, given that the peak-frequency of the collective quasi-oscillations is located around 20 Hz each neuron fires on average 3 times before a cycle of the collective rhythm is completed. The same result is obtained by analyzing the power-spectrum for individual neurons (fig. 5.16), exhibiting a peak around $f \approx 60$ Hz and no signs of power enhancement in the 20-30 Hz band.

The fact that individual neurons do not lock to the global quasi-oscillation reminds to what has been called asynchronous-states or sparse-synchronization. Here, a collective rhythm –to which individual neurons do *not* lock– emerges (see [19, 36] for related, though different, phenomena). Observe that in the so-called "fast-oscillations", as described for instance in [19], the emerging global rhythm is much faster than individual neurons, while here, it is the other way around.

Remarkably, the collective oscillations found by Wallace et al. [138], where the underlying mechanism is still SAF, are much faster than individual neuron oscillations, which is just the opposite of what we (and the experiments in [31]) find for Up states.

Certainly, although there is no trace of the global rhythm at the single neuron level,

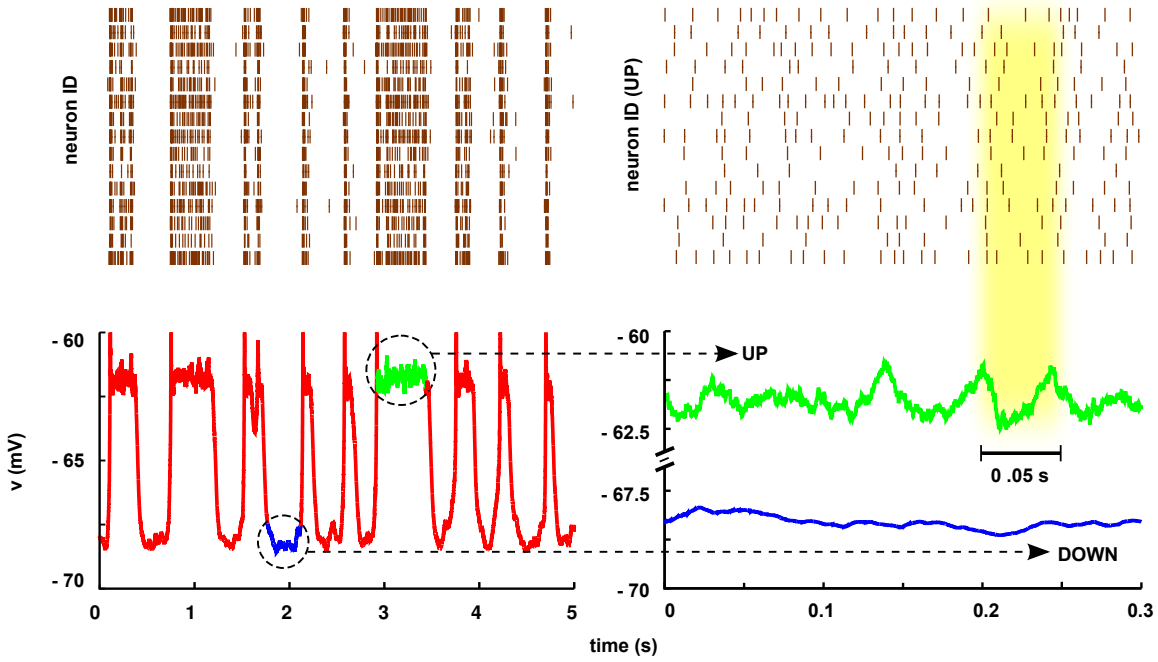


Figure 5.14: **Raster plots and average membrane potential in the spiking-neuron network model (Model B).** Left: (Top) Raster plot of 15 randomly chosen neurons (out of a total of $N = 1000$ neurons in the simulation). Sticks are plotted whenever a neuron spikes. (Bottom) Time-series of the network-averaged membrane potential in the same simulation. Comparison of the two left panels (both of them sharing the same time axis) reveals that individual neurons fire often during Up states, while they are essentially quiescent in Down-state intervals. Right: (Bottom) zoom of an Up interval (green curve) and of a Down interval (blue curve); while the Up state exhibits quasi-oscillations, the Down-state does not. (Top) Raster plot of 15 randomly chosen neurons during the Up state. Remarkably, their spiking frequency is not locked to the collective rhythm: it is about three times faster.

it is clear that the collective quasi-oscillation is produced by the effective activity of N single oscillations in the network. The key point is that, as the Fourier transform is not commutative, $\mathcal{F}\left[\sum_i^N V_i/N\right] \neq \sum_i^N \mathcal{F}\left[V_i\right]/N$, the only way to understand the global oscillation is to perform a theory for the coarse-grained level, where SAF applies.

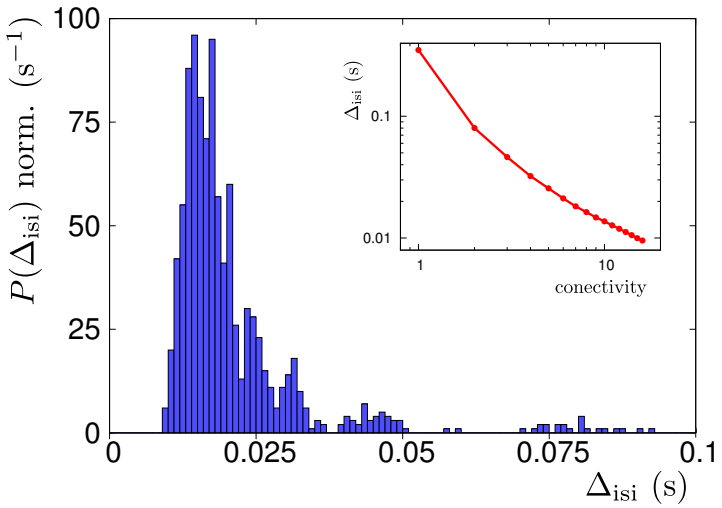


Figure 5.15: **Probability distribution of the inter-spike-intervals in the random network**: its average gives $\langle \Delta_{\rm isi} \rangle \approx 17$ ms, corresponding to a mean firing rate of $f \approx 60$ Hz. This result perfectly agrees with the mean firing rate of fig. 5.16. Heterogeneity in the average inter-spike-intervals stems from the different connectivity degrees, as illustrated in the inset. In the latter, we show the average value of $\Delta_{\rm isi}$ for different (pre-synaptic) connectivity levels in the Erdos-Renyi network of mean degree K.

5.4.4 Robustness of the mechanism: effect of the synaptic facilitation

We have checked the robustness of the mechanism of SAF in the model with synaptic plasticity when we introduce short-term synaptic facilitation [60]. We restrict this extension to the coarse-grained Model A.

Following Tsodyks and Markram[132, 96], we equip the set of eqs. 5.1 and 5.2 with a new equation for the release probability, $p_r = p_r(t)$, which was taken to be a constant in the first implementation. This new equation describe the dynamics of Ca^{++} ions in the synaptic terminal, modulating the probability with which neurotransmitters are released. During silent periods of activity, it decays to its baseline P_0 with time constant τ_f, while in the presence of activity it increases proportionally to $(1 - p_r)$:

$$\dot{p}_r = \frac{P_0 - p_r}{\tau_f} + P_0(1 - p_r) f(v). \tag{5.30}$$

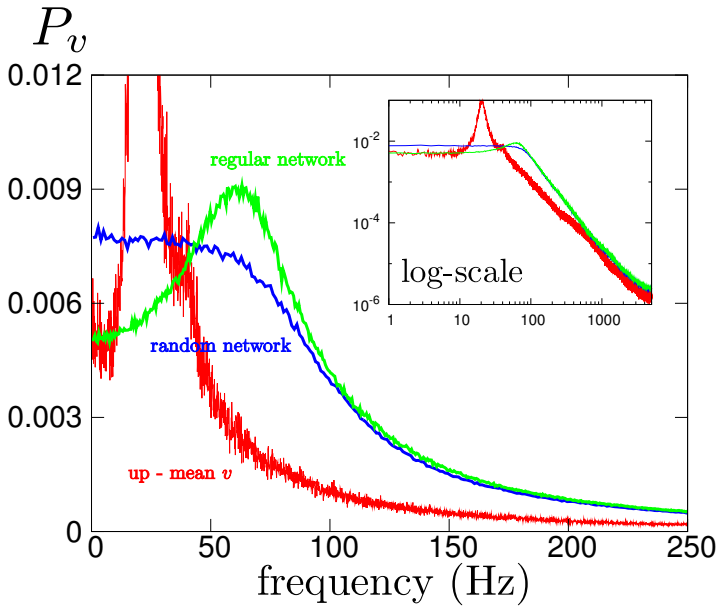


Figure 5.16: **Average power spectra for individual neuron potentials** for both a random network with average connectivity $K = 7.5$ (blue) and a regular network with connectivity $K = 7$ (green) in the Up state. A sharp peak (around 60 Hz) is seen for regular networks; instead in random networks the peak is blurred owing to node-to-node heterogeneity. In any case, there is no peak at the characteristic frequency of the global, network-averaged membrane potential $\simeq 20$ Hz (peak of the red curve): individual neurons do not lock to the collective rhythm within Up states. The inset shows a similar plot in logarithmic scale, putting forward the presence of a distinct peak for regular networks together with a w^{-2} tail for all spectra. All curves are normalized to unit area.

Fixing $P_0 = 0.01$ and $\tau_f = 1.5$ s as in [96], and other parameters as before (see table 5.1), we find that the stable fixed point corresponding to the Up state shifts to $v^* = V_r + 12.5921$ mV, $u^* = 0.2005$ and $p_r^* = 0.4708$. On the other hand, the Down state remains at $v^* = V_r$, $u^* = 1$ and $p_r^* = P_0$. Computing the power spectrum for each variable, we can generalize eq. 5.23 to the case with an arbitrary number of coupled equations, obtaining

$$P_z(\omega) = \frac{[\mathrm{Adj}(A - i\omega\mathbb{1})\langle\vec{\eta}\,\vec{\eta}^t\rangle\mathrm{Adj}(A^t + i\omega\mathbb{1})]_{zz}}{\det\left(A^2 + \omega^2\mathbb{1}\right)}, \tag{5.31}$$

where A is the Jacobian matrix evaluated at the fixed point, Adj stands for the adjoint

matrix (transpose of the cofactors) and $\mathbb{1}$ for the identity matrix. In the limit of small noise amplitude, we find a peak at the frequency that minimizes the denominator $\det(A^2 + \omega^2 \mathbb{1})$. In the system, again, a non-trivial peak appears in the spectra only for the Up state (at 1.4 Hz for the parameters above), while the distribution becomes sharper even if its structure remains essentially unchanged (see red curve in fig. 5.17). Therefore, the mechanism of stochastic amplification of fluctuations described above is robust to the inclusion of synaptic facilitation.

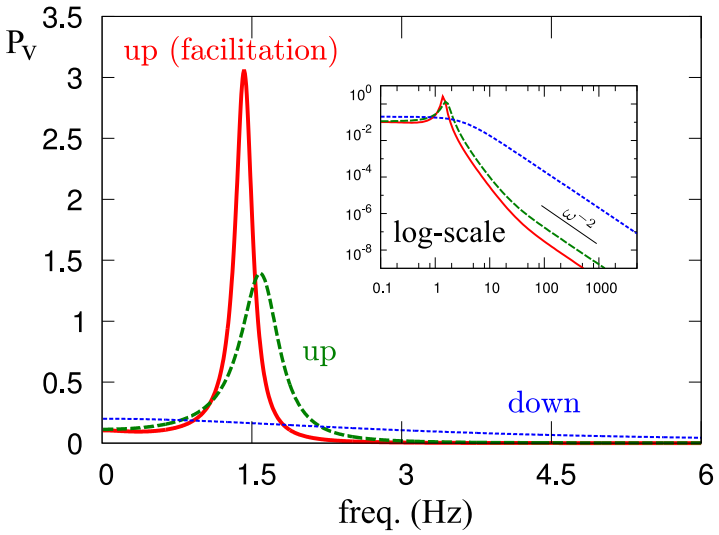


Figure 5.17: **Power spectrum of fluctuations in Model A in the presence of synaptic facilitation**: Red curve represents the power spectrum for the average potential in the Up state when synaptic facilitation is incorporated, while green curve corresponds to the model without it, i.e. with the release probability p_r constant. The peak moves slightly to a lower frequency, and the spectrum becomes sharper. The power spectrum for the Down state remains unaltered (blue curve). The inset represents double-logarithmic plots of the same quantities as in the main plots. All spectra have been generated with $P_0 = 0.01$, $\tau_f = 1.5$ s and $\sigma_z^2 = 0.01 z^*/\tau$ (where z stands for either v, u, p_r) and parameters from table 5.1, and normalized to unit area.

5.4.5 Robustness of the mechanism: SAF in a network of excitatory-inhibitory neurons

Here we show that SAF can be found also in models for Up and Down states relying on populations of both excitatory and inhibitory neurons. We consider the Wilson-Cowan-

like model as described, for example, in [96] (see also [18] and [90, 91, 138] where a similar model has been recently studied). This is a mean-field like model, analogous in this sense to Model A, but with inhibition rather than depression as leading regulatory mechanism.

The model consists of two equations for the mean excitatory and inhibitory firing rates in the network, $E(t)$ and $I(t)$ respectively:

$$\tau_e \dot{E} = -E + g(J_{ee}E - J_{ei}I + E_0) \tag{5.32}$$

$$\tau_i \dot{I} = -I + g(J_{ie}E - J_{ii}I + I_0) \tag{5.33}$$

with a threshold linear response function

$$g(x) = \begin{cases} 0 & x < T \\ \beta(x - T) & x \geq T \end{cases} \tag{5.34}$$

where T is the threshold parameter and the couplings between excitatory/inhibitory are represented by the J constants. For a wide range of parameter values these equations exhibit bistability: there is a stable fixed point with with low-activity regime (Down state) and a second one with a non-vanishing firing rate and significant activity (Up state). Adding Gaussian white noises to eqs. 5.32 and 5.33, the system fluctuates and eventually may jump between the two fixed points. We have verified by means of computer simulations that indeed this model exhibits Up-and-Down transitions, that a non-trivial peak appears for fluctuations within Up states and not for Down states. The chosen parameters are shown in table 5.4.

Trajectories of the deterministic dynamics reveal spiral trajectories (i.e. damped oscillations) near the Up-state fixed point but not in the Down state (straight trajectories corresponding to real eigenvalues). Therefore, one can expect a non-trivial peak to appear in the Up-state power-spectrum but not in the Down one. This can be explicitly seen from a linear stability analysis, so we compute the Jacobian matrix:

$$A = \begin{pmatrix} a_{EE} = -\frac{1}{\tau_e} + J_{ee}\, g' & a_{EI} = -J_{ei}\, g' \\ a_{IE} = J_{ie}\, g' & a_{II} = -\frac{1}{\tau_i} - J_{ii}\, g' \end{pmatrix}. \tag{5.35}$$

When evaluating at the fixed points, $g'_{\text{up}} = \beta$ while $g'_{\text{down}} = 0$. Therefore, A is already

Parameter	Value
τ_e, τ_i	0.01 s
J_{ee}, J_{ie}, J_{ii}	5 mV/Hz
J_{ei}	9 mV/Hz
β	0.5 Hz/mV
T	15 mV
E_0	10 mV
I_0	0 mV

Table 5.4: **Parameter values for the excitation-inhibition model**, taken from [96].

diagonal in the Down state, with both real eigenvalues. From equation 5.24, $\omega_{0,\text{down}}$ turns out to be a complex number (i.e. there is no solution in the maximization problem) and hence, no characteristic frequencies are found in the power spectrum of fluctuations. On the other hand, eigenvalues are complex for the Up state, giving a peak near $\omega_{0,\text{up}} = 200$ rad/s $= 31.8$ Hz. These analytical predictions are in excellent agreement with results of computer simulations for this model (not shown).

5.5 Chapter summary

In this Chapter, we have focused on the study of spontaneous brain activity, in particular in the so-called Up and Down oscillations, in which a population of neurons collectively alternate between intervals of high activity (Up states) and completely silent periods (Down states). Specifically, we have been attracted by the experimental evidence that rapid collective oscillations (20-30 Hz) appear within Up intervals, but do not for Downs [31]. Additionally, such rapid oscillations are only found at the meso-scale (local field potential), but not at the micro-scale (intracellular records) [31]. Our main finding is that this phenomenon can be rationalized under the mechanism of stochastic amplification of fluctuations [61, 60].

In particular,

- We have introduced two computational models with different levels of complexity, both of them including short-term synaptic depression (STDP) as the main regulatory mechanism. The first model [131, 64] is a mean-field like model, with

only two coupled variable, whereas in the second one [98] includes biologically realistic features, and the dynamics is described for each neuron forming part of the network. In both models, Up and Down states correspond with two different attractors of the dynamics, and Up-and-Down transitions emerge in the presence of stochastic sources.

- We have described the mechanism of stochastic amplification of fluctuations (SAF), previously reported in the contexts of Ecology [94] and Epidemiology [1]. For systems in which the dynamics are given by two (or more) coupled stochastic variables, fluctuations around a stable attractor can lead to different patters, and noise-induced oscillations can emerge depending on the deterministic dynamics.

- We have computed the power spectrum of fluctuations for either Up and Down states, in the two models described above. A non-trivial peak appears for Up states, but not for Downs (as in [31]); moreover, the peak location coincides with the experimental findings (20-30 Hz) for the most realistic model. Such peak can be easily explained with the theory of SAF: in the Up state, the stability matrix of the attractor has complex eigenvalues (spiral deterministic trajectories), and fluctuations around such attractor present a characteristic frequency of oscillation; for the Down state, eigenvalues are real, and none of the previous applies. The nature of such eigenvalues is controlled by the spiking firing rate.

- For the single neuron-based model, we have seen that individual rhythms do not lock to the collective characteristic oscillations during Up intervals (as found in [31]), in agreement with the fact that SAF only applies for the coarse-grained variables.

- Finally, we have checked the robustness of the emergence of quasi-oscillations described by SAF against specificities of the model, in particular under the addition of synaptic facilitation and when the self-regulatory mechanism is based on inhibition rather than on synaptic depression, obtaining the same results.

Given the general character of SAF, we think that this mechanism could be of relevant interest for the field of Neuroscience.

Chapter 6

Conclusions

In the study of collective phenomena, phase transitions and self-organized dynamics have attracted particular attention. From their study, we have learned that simple rules for the microscopic interactions can lead to a plethora of diverse and complex patterns at the global scale. Here, we highlight the theory of self-organized criticality [5], which has set up a general framework to understand why many systems in nature, such as earthquakes or the distributions of rainfalls exhibit scale invariance, large scale correlations and heterogeneous responses in a wide spectrum. Furthermore, during the last decade, it has been conjectured, based on several empirical evidences [118, 107, 15, 24, 50, 79], that living systems might benefit from having attributes akin to criticality [101], such as a large repertoire of dynamical responses, a high sensitivity to environmental changes and an efficient management of the information.

In this thesis, first, we have addressed, from a general point of view, the hypothesis that many aspects of living systems seem to operate at the vicinity of critical points [101]. Under the assumption that biological systems need to construct good representations of the outer complex world, and that such representations are encoded in terms of probability distributions, we have shown that the encoding probability distributions do necessarily lie in the vicinity of critical points [62].

In the presence of broadly different ever-changing heterogeneous environments, computational evolutionary and adaptive models have vividly illustrated how a collection of living systems eventually cluster near the critical state, while they do not for homogeneous and predictable environments. This result could be rationalized, in a sense, from

the known fact that most different dynamics are concentrated at the critical point [89]. A more accurate convergence to criticality is found in a co-evolutionary/co-adaptive set-up in which individuals evolve/adapt to represent with fidelity other agents in the community, thereby creating a collective "language", which turns out to be critical; this phenomenon cannot be understood in terms of [89], and therefore constitutes a novel mechanism for self-organization, which living systems in nature may have learned to exploit.

These ideas apply straightforwardly to genetic and neural networks –where they could contribute to a better understanding of why neural activity seems to be tuned to criticality– but have a broader range of implications for general complex adaptive systems [56]. For example, our framework could be applicable to some bacterial communities [81] and viral populations [129] for which a huge phenotypic variability has been empirically observed. Such a large phenotypic diversification can be seen as a form of "bet hedging", an adaptive survival strategy analogous to stock-market portfolio management [144], in which individuals diversify their "assets" to maximize their long-term growth rate in the presence of environmental uncertainty. In our context, such variability turns out to be a direct consequence of individuals in the community being critical.

Bet-hedging strategies have attracted our attention in the general context of population dynamics. To reach a better understanding of them, we have proposed a simple birth-death stochastic model in which individuals can choose between a poor but safe strategy, a better but risky alternative (environmental dependent), or a combination of both. For instance, this general framework could be applied to micro-organisms able to metabolize two different resources, with different levels of risks.

Our main finding is that the benefits of developing bet-hedging strategies are strongly enhanced in highly fluctuating environments, as well as for low-dimensional systems, where intrinsic fluctuations play a key role. Such conditions are usually faced by biological populations, for instance, in bacterial colonies competing at the front of a range expansion in unpredictable environments [11, 78, 139]; therefore, our results provide a justification for the ubiquitousness of bet-hedging strategies in nature.

Additionally, our work may offer a physical framework to answer different questions, for instance, what is the difference between exploiting bet-hedging individually or at a

community level. We have briefly addressed this question, leading to the conclusion that individual-based bet-hedging provides the optimal strategy; however, further realistic features should be properly taken into account to extract a more solid conclusion. For instance, by including the cost derived of maintaining a hybrid mechanism when bet-hedging is performed individually.

We have developed a specific model for bet-hedging strategies in plants, in particular for mixed dispersal strategies. Here, organisms diversify their reproduction possibilities by mean of two different kind of seeds, dispersive and non-dispersive. Including some realistic features of these ecosystems, such as the negative effect of self-reproduction (inbreeding depression), we have studied under which conditions bet-hedging strategies lead to a substantial increase of the population growth rate.

In general, the previous result still holds, and the benefits of mixed dispersal strategies are enhanced by environmental fluctuations. Moreover, depending on the environmental variability and the strenth of inbreeding depression, the optimal strategy can either ensure a high dispersal propensity, or a mixed situation in which both of the previous dispersal mechanisms are employed; populations presenting only the pure non-dispersing strategy are rarely found. Additionally, we have shown that, for highly unpredictable environments and moderate inbreeding depression, hybrid dispersal mechanisms are not only optimal, but also provide a significant enhancement of the density compared to the standard only-dispersive strategy. Our findings suggest that plants living in such circumstances are more likely to evolve and develop a hybrid reproductive mechanism in order to survive.

Finally, returning to the starting topic about critical behavior exhibited by many living systems, we have focused on spontaneous brain activity, as another instance of critical dynamics [118]. In particular, we have centered our study on the spontaneous rhythms in the cortex called Up and Down states [147, 143]. However, rather than focusing on the *critical* question, we have studied the dynamics of Up and Down states itself, trying to explain the intriguing patterns emerging within such Up-and-Down oscillations.

By implementing diverse computational models for Up and Down oscillations (with different levels of complexity), we have measured the fluctuations of the activity for either Up (active) and Down (silent) periods. Our computational experiments show

a non-trivial peak in the power-spectrum of Up states, indicating the existence of a characteristic frequency, while no similar peak is found for Down states. These results are in excellent accordance with the experimental findings of several research groups [128, 103, 49, 31].

Our main contribution has been to show that the reported phenomenology can be perfectly explained by the mechanism of "stochastic amplification of fluctuations", previously described in the context of Ecology [94] and Epidemiology [1]. This mechanism consists in the resonant amplification of some frequencies in the spectra of stochastic systems when the corresponding fixed-point of its deterministic dynamics is a focus (i.e. with complex associated eigenvalues). The presence of any source of noise kicks the system away from the deterministic fixed point leading to a non-trivial power-spectrum. In the system we are studying, such circumstances only occur during Up periods, but never for Downs. In addition, this explanation is robust beyond modeling specificities.

Furthermore, we have shown that the mechanism of stochastic amplification of fluctuations operates for global variables but not for individual neurons, which perfectly accounts for empirical findings in [31]. Such collective behavior is related to what has been called asynchronous-states or sparse-synchronization in which a global rhythm –to which individual neurons do *not* lock– emerges [19, 36].

Summing up, by using different tools from statistical mechanics, information theory, game theory and stochastic processes, this thesis has tried to identify several underlying mechanisms which allow biological systems to successfully operate in their everyday life. Given their general character, it might be expected that living systems in nature have learned, throughout the course of adaptation and evolution, to take advantage of these mechanisms in a wide range of different contexts. As if, in some sense, the intricate world of living matter interactions would act as a smart collective *complexsystemist*.

Conclusiones

En el estudio de los fenómenos colectivos, las transiciones de fase y las dinámicas de auto-organización han atraído gran parte de la atención. De su estudio hemos aprendido que reglas simples para las interacciones microscópicas pueden lar lugar a complejos patrones a nivel global. Aquí, destacamos la teoría de la Criticidad Auto-Organizada [5], la cual ha conseguido establecer un marco general sobre el que entender por qué diferentes sistemas en la naturaleza, como los terremotos o las distribuciones de precipitaciones, muestran invariancia de escala, correlaciones a larga escala y respuestas heterogéneas en un amplio espectro. Además, durante la última década, se ha especulado, basándose en evidencias empíricas [118, 107, 15, 24, 50, 79], que los sistemas biológicos podrían beneficiarse de tener atributos de criticidad [101], como un gran repertorio de respuestas dinámicas, alta sensibilidad a cambios ambientales y un manejo eficiente de la información.

En esta tesis, primeramente, hemos abordado, desde un punto de vista general, la hipótesis de que muchos aspectos de los sistemas biológicos parecen operar en la cercanía de puntos críticos [101]. Bajo la suposición de que los sistemas vivos necesitan construir buenas representaciones del complejo mundo exterior, y que tales representaciones están codificadas en términos de distribuciones de probabilidad, hemos mostrado que dicha codificación necesariamente cae cerca de un punto crítico [62].

En presencia de ambientes externos suficientemente heterogéneos y cambiantes, diferentes modelos computacionales evolutivos y adaptativos han ilustrado cómo una colección de sistemas vivos convergen hacia el estado crítico, mientras que no lo hacen para ambientes homogéneos y predecibles. Este resultado podría entenderse, en cierto modo, a partir del hecho conocido de que la mayoría de "dinámicas diferentes" se localizan en torno al punto crítico [89].

Una convergencia más precisa hacia la criticidad se encuentra en un escenario co-evolutivo/co-adaptativo, en el cual los individuos evolucionan/adaptan para representar con fidelidad al resto de agentes en la comunidad, creando un "lenguaje" colectivo, que resulta ser crítico [62]. Este fenómeno no puede entenderse en los términos anteriormente descritos [89], y por tanto constituye un nuevo mecanismo para auto-organización que los sistemas biológicos podrían haber aprendido a explotar.

Estas ideas encuentran aplicación directa en redes genéticas y neuronales –donde podrían contribuir a un mejor entendimiento de por qué la actividad cerebral parece estar tuneada a la criticidad– pero tienen un rango más amplio de aplicaciones en sistemas complejos adaptativos. Por ejemplo, nuestro marco teórico podría ser aplicado a ciertas comunidades de bacterias [81] y poblaciones de virus [129], para las cuales se ha observado empíricamente una elevada variabilidad fenotípica. Dicha variabilidad puede verse como una forma de "cobertura de riesgos" ("bet-hedging" en inglés), una estrategia de supervivencia similar a la gestión del mercado de acciones en bolsa [144], en la cual los individuos diversifican sus "activos" para maximizar el crecimiento a largo plazo en presencia de ambientes impredecibles. En nuestro contexto, dicha variabilidad resulta ser una consecuencia directa de individuos en una comunidad en el punto crítico.

Las estrategias de "bet-hedging" han atraído nuestra atención a su análisis general en el contexto de dinámica de poblaciones. Para obtener un mejor entendimiento de ellas, hemos propuesto un modelo estocástico sencillo de reproducción y muerte en el que los individuos pueden elegir entre una estrategia pobre pero segura, una mejor pero arriesgada (dependiente del medio ambiente), o una combinación de ambas. Por ejemplo, este escenario podría ser aplicado a micro-organismos que son capaces de metabolizar dos fuentes de alimento diferentes, con distintos niveles de riesgo.

Nuestro principal descubrimiento es que los beneficios de desarrollar estrategias de bet-hedging están fuertemente potenciadas en medios ambientes altamente fluctuantes y en bajas dimensiones espaciales (donde las fluctuaciones intrínsecas juegan un papel clave). Tales condiciones son justamente con las que normalmente se encuentran las comunidades de individuos en la naturaleza, por ejemplo, en colonias de bacterias compitiendo en el frente de expansión bajo entornos impredecibles [11, 78, 139]. Así pues, nuestro resultado proporciona una justificación para la ubiquidad de estrategias de bet-hedging en la naturaleza.

Además, nuestro trabajo aporta un marco general con el que responder a diferentes preguntas, por ejemplo, cuál es la diferencia entre explotar una estrategia de bet-hedging a nivel individual o a nivel de comunidad. Nosotros hemos abordado brevemente la cuestión, llegando a la conclusión de que el bet-hedging llevado a cabo individualmente constituye la estrategia óptima; sin embargo, deberíamos incluir más ingredientes realistas al modelo para extraer una conclusión sólida. Por ejemplo, incluyendo el coste derivado de mantener un mecanismo híbrido para cada agente.

También hemos desarrollado un modelo específico para estrategias de bet-hedging en plantas, en particular para estrategias dispersivas mixtas. Aquí, los organismos diversifican sus posibilidades de reproducción a través de dos tipos de semillas, dispersivas y no-dispersivas. Incluyendo algunos elementos realistas de estos ecosistemas, como el efecto negativo de la auto-reproducción (depresión endogámica), hemos estudiado bajo qué circunstancias las estrategias de bet-hedging llevan a un incremento substancial del ritmo de crecimiento de la población.

En general, el resultado anterior se mantiene, y los beneficios de las estrategias mixtas dispersivas se potencian por las fluctuaciones ambientales. Además, dependiendo de la intensidad relativa de la variabilidad ambiental y la depresión endogámica, la estrategia óptima puede asegurar, o bien una alta propensión dispersiva, o bien una situación mixta en la que los dos mecanismos de dispersión previamente descritos son utilizados. Las poblaciones que solamente presentan la estrategia pura no-dispersiva son raramente viables. Adicionalmente, hemos mostrado que, para entornos altamente impredecibles, los mecanismos de dispersión híbridos no solo son óptimos, sino que además producen un beneficio relativo muy significativo comparado con la estrategia solamente dispersiva estándar. Nuestro resultado sugiere que las plantas que viven en dichas circunstancias tienen más probabilidad de evolucionar y desarrollar un mecanismo reproductivo híbrido para sobrevivir.

Por último, volviendo al tema del inicio sobre el comportamiento crítico que presentan muchos sistemas biológicos, nos hemos concentrado en el estudio de la actividad espontánea del cerebro, como otro ejemplo más de dinámica crítica [118]. En particular, hemos llevado nuestro estudio a los ritmos espontáneos del córtex llamados estados Up-Down [147, 143]. Sin embargo, más que continuar investigando en la pregunta *crítica*, nos hemos focalizado en la dinámica misma de los estados Up-Down, con el objetivo de

entender los interesantes patrones que emergen en dichas oscilaciones.

Implementando diferentes modelos computaciones que describen estados Up-Down (con diferentes niveles de complejidad), hemos medido las fluctuaciones de la actividad durante cada período Up (activo) o Down (inactivo). Nuestros experimentos computaciones muestran un pico no trivial en el espectro de potencias de los estados Up, indicando la existencia de una frecuencia característica, mientras que dicho pico no aparece para los estados Down. Estos resultados están en perfecta concordancia con los medidos experimentalmente por diferentes grupos [128, 103, 49, 31].

Nuestra contribución principal ha sido mostrar que tal fenomenología puede ser perfectamente explicada por el mecanismo de "amplificación estocástica de las fluctuaciones", anteriormente descrito en el contexto de la ecología [94] y la epidemiología [1]. Este mecanismo consiste en la amplificación resonante de algunas frecuencias en el espectro de sistemas estocásticos cuando el correspondiente punto fijo de su dinámica determinista es un foco (es decir, un punto fijo con autovalores complejos). La presencia de cualquier fuente de ruido saca al sistema del punto fijo, dando lugar a un espectro de potencias no trivial. En nuestro sistema en cuestión, dichas circunstancias solo se dan durante los períodos Up, pero nunca para los Down. Además, esta explicación es robusta más allá de detalles del modelo.

Por último, hemos mostrado que el mecanismo de amplificación estocástica de las fluctuaciones opera para las variables globales pero no para las neuronas individuales, lo cual encaja perfectamente con los resultados experimentales [31], en los que no se observa dicha resonancia a nivel celular. Este fenómeno colectivo está relacionado con lo que han sido llamados "estados asíncronos" [19, 36], en los cuales un ritmo global, no seguido por las neuronas individuales, emerge a gran escala.

En resumen, utilizando diferentes herramientas de mecánica estadística, teoría de la información, teoría de juegos y procesos estocásticos, esta tesis ha tratado de identificar varios mecanismos que permiten a los sistemas biológicos operar de forma satisfactoria en su día a día. Dado su carácter general, cabe esperar que los sistemas en la naturaleza hayan aprendido, a través de la adaptación y la evolución, a sacar provecho de estos mecanismos en un amplio rango de contextos. Como si el intrincado mundo de las interacciones de la materia viva se comportara, de alguna forma, como un *complexsystemista* colectivo e inteligente.

Appendix A

Approach to Sanov's theorem

In Chapter 2 we introduced the Kullback-Leibler divergence (KLD, eq. 2.8) as a measure of the *closeness* between two probability distributions P and Q, from one to the other. As this statement is well understood from the Sanov's theorem, in this appendix we briefly sketch the *derivation for physicists* suggested in [97].

Le consider a discrete stochastic variable x distributed as $P(x)$, that we call *original* probability distribution. Then, we extract a sequence of T independent samples $\{x_1, x_2, x_3, ..., x_T\}$, and we compute the *experimental* probability distribution $Q(x)$, i.e. the relative frequency histogram. We expect that $Q \to P$ for large datasets, but the result differs for any finite sample. Therefore, the experimental distribution Q is another stochastic variable, distributed as $\mathcal{L}(Q|P)$. We can compute this function as the probability of generate a sequence of P compatible with Q:

$$\mathcal{L}(Q|P) = \sum_{x_1,...,x_T} P(x_1, x_2, ..., x_T) \prod_x \delta_{TQ(x),\sum_{i=1}^T \delta_{x,x_i}}. \tag{A.1}$$

As realizations are independent, $P(x_1, ..., x_T) = \prod_{k=1}^T P(x_k)$. To perform this calculation, we use the integral representation of the Kronecker delta:

$$\int_0^{2\pi} d\lambda \frac{\exp(ia\lambda)}{2\pi} = \delta_{a,0}\,, \qquad a \in \mathbb{Z}. \tag{A.2}$$

As there is one constraint for each value of x, we introduce the integral representation

in eq. A.1 with a set of new variables $\boldsymbol{\lambda} = \{\lambda(x)\}_x$:

$$\mathcal{L}(Q|P) = \sum_{x_1,\dots,x_T} \prod_{k=1}^{T} P(x_k) \prod_x \int_0^{2\pi} \frac{d\lambda(x)}{2\pi} \exp\left(i\lambda(x) \left[TQ(x) - \sum_{i=1}^{T} \delta_{x,x_i} \right] \right). \quad \text{(A.3)}$$

Note that the term in square brackets is always an integer number. This equation can be rearranged as

$$\mathcal{L}(Q|P) = \int_0^{2\pi} \dots \int_0^{2\pi} \left(\prod_x \frac{d\lambda(x)}{2\pi} \right) e^{T\left(i\sum_x \lambda(x)Q(x) + \log\left[\sum_{x'} P(x') \exp\left(-i\lambda(x')\right)\right]\right)}, \quad \text{(A.4)}$$

and solved by means of the saddle point approximation for $T \gg 1$[22],

$$\int_I d\boldsymbol{\lambda} \exp\left(TF(\boldsymbol{\lambda})\right) \simeq \sqrt{\left(\prod_x \frac{2\pi}{T} \right) \det \left(\left. \frac{\partial^2 F}{\partial\lambda(x)\lambda(x')} \right|_{\boldsymbol{\lambda}^*} \right)^{-1}} \exp\left(TF(\boldsymbol{\lambda}^*)\right) \quad \text{(A.5)}$$

with I the interval of integration and $\boldsymbol{\lambda}^*$ the value which maximizes $F(\boldsymbol{\lambda})$. Identifying for our case $F(\boldsymbol{\lambda}) = i\sum_x \lambda(x)Q(x) + \log\left[\sum_{x'} P(x') \exp\left(-i\lambda(x')\right)\right]$, the condition of extrema leads to

$$\left. \frac{\partial F}{\partial\lambda(x)} \right|_{\boldsymbol{\lambda}^*} = 0 \Longrightarrow -i\lambda^*(x) = \log \frac{Q(x)}{P(x)} + \log \left[\sum_{x'} P(x') e^{-i\lambda^*(x')} \right] \quad \text{(A.6)}$$

and it can be checked that $\boldsymbol{\lambda}^*$ represents a maximum. Note that, although eq. A.6 pretends to be a non-explicit result for $\lambda^*(x)$, the last term is common to all of them, and therefore it cannot depend on x, only on the specific choice of P and Q.

Finally, evaluating at the maximum we find that $F(\boldsymbol{\lambda}^*) = -\sum_x Q(x) \log \frac{Q(x)}{P(x)} + F_0(P)$, where the constant F_0 do not depend on Q. Consequently, the leading exponential term in eq. A.5 is

$$\mathcal{L}(Q|P) \sim \exp\left(-TD(Q|P)\right) \quad \text{(A.7)}$$

where $D(Q|P) = \sum_x Q(x) \log \frac{Q(x)}{P(x)}$ is the Kullback-Leibler divergence.

Let us remark that problem we have described is intrinsically non-symmetric, i.e. the probability that Q generates characteristic states of P differs from the probability that

P generates such states of Q. The Kullback-Leibler divergence, as it is non-symmetric and therefore, not a proper distance, takes account of this dissimilarity.

In conclusion, Sanov's theorem states that, given a set of realizations generated from a distribution $P(x)$, the probability of finding a dataset with relative frequency histogram $Q(x) \neq P(x)$ decays exponentially with the number of samples generated. However, some histograms are more likely to be found than others, in particular those with lower KLD respect to the original $P(x)$. Thus, the KLD gives a measure for the closeness between two probability distributions.

Appendix B

Fisher information and Cramérs-Rao inequality

The Fisher Information (FI) was defined in Chapter 2 in eq. 2.18 for a system of states $\mathbf{s}$, randomly distributed as $P(\mathbf{s}|\boldsymbol{\beta})$; $\boldsymbol{\beta}$ is a multivariate parameter encoding some tunable behaviour in the dynamics. In this context, we stated that the FI measures the *amount of information* encoded in $\mathbf{s}$ about the internal parameters $\boldsymbol{\beta}$. This can be rationalized within the Cramér-Rao inequality [34], that we try to illustrate –not rigorously– in this appendix.

In a nutshell, the inequality states that the error made when we infer the parameter $\boldsymbol{\beta}$ from one realization $\boldsymbol{s}$, on average, is always greater or equal than the inverse of the FI.

Let consider a estimator of $\boldsymbol{\beta}^{(1)}$ for the set of realizations $(\mathbf{s}_1, ..., \mathbf{s}_n)$, which is a function that converts a set of samples into a single value $\boldsymbol{\beta}$. This estimator is expected to give an approximate value of the "real"parameter $\boldsymbol{\beta}$, hence the name *estimator*, but it may differ for different realizations of the dynamics. Consequently, the error $\left(\boldsymbol{\beta}^{(1)} - \boldsymbol{\beta}\right)$ is, in general, a stochastic variable.

We first want the estimator to be *unbiased*, i.e. with a zero error on average:

$$\langle \boldsymbol{\beta}^{(1)} - \boldsymbol{\beta} \rangle = \sum_{\mathbf{s}_1,...,\mathbf{s}_n} P(\mathbf{s}_1, ..., \mathbf{s}_2|\boldsymbol{\beta}) \left(\boldsymbol{\beta}^{(1)} - \boldsymbol{\beta}\right) = \sum_{\mathbf{s}_1,...,\mathbf{s}_n} \prod_{i=1}^{n} P(\mathbf{s}_i|\boldsymbol{\beta}) \left(\boldsymbol{\beta}^{(1)} - \boldsymbol{\beta}\right) = 0, \tag{B.1}$$

where we have introduced the independence between realizations $P(\mathbf{s}_1, ..., \mathbf{s}_2|\boldsymbol{\beta}) =$

$\prod_{i=1}^{n} P(\mathbf{s}_i|\boldsymbol{\beta})$.

Under this condition, it can be proved [34] that the covariance matrix for the error variable obeys the following (Cramers-Rao) inequality:

$$\langle \beta_\mu^{(1)} \beta_\nu^{(1)} \rangle \geq \frac{1}{n} \chi_{\mu\nu}(\boldsymbol{\beta})^{-1}, \tag{B.2}$$

where $\chi_{\mu\nu}$ is the FI (defined in eq. 2.18) and μ and ν stand for different components of the multivariate parameter $\boldsymbol{\beta}$. If the inequality becomes an equality, the estimator is called *efficient*.

This result tells us that, for high values of FI –as for instance in a critical point–, the error made when we infer the parameter $\boldsymbol{\beta}$ with an efficient estimator is low. On the other hand, if the FI is low, we expect a non accurate result in the inferred parameter. In other words, the FI measures the information contained in $\mathbf{s}$ about the parameter $\boldsymbol{\beta}$.

Example

Let consider the case of an exponential distribution $P(s|\beta) = \frac{1}{\beta}\exp(-\frac{s}{\beta})$, defined in the continuous range $[0, \infty)$ for $\beta > 0$. We take n independent realizations $\{s_1, ..., s_n\}$ of $P(s|\beta)$, and we try to infer, from these samples, the parameter β with two different estimators:

$$\beta^{(1)}(s_1, ..., s_n) = s_1 \tag{B.3}$$

$$\beta^{(2)}(s_1, ..., s_n) = \frac{1}{n}\sum_{i=1}^{n} s_i. \tag{B.4}$$

First, we check that estimators are unbiased:

$$\langle \beta^{(1)} \rangle = \int_0^\infty ds_1 ... ds_n \prod_{i=1}^{n} \left[\frac{1}{\beta}\exp\left(-\frac{s_i}{\beta}\right)\right](s_1 - \beta) = 0, \tag{B.5}$$

$$\langle \beta^{(2)} \rangle = \int_0^\infty ds_1 ... ds_n \prod_{i=1}^{n} \left[\frac{1}{\beta}\exp\left(-\frac{s_i}{\beta}\right)\right]\left(\frac{1}{n}\sum_{i=1}^{n} s_i - \beta\right) = 0. \tag{B.6}$$

Secondly, we compute the error variance for each estimator:

$$\langle\left(\beta^{(1)}\right)^2\rangle = \int_0^\infty ds_1...ds_n \prod_{i=1}^n \left[\frac{1}{\beta}\exp\left(-\frac{s_i}{\beta}\right)\right](s_1-\beta)^2 = \beta^2, \tag{B.7}$$

$$\langle\left(\beta^{(2)}\right)^2\rangle = \int_0^\infty ds_1...ds_n \prod_{i=1}^n \left[\frac{1}{\beta}\exp\left(-\frac{s_i}{\beta}\right)\right]\left(\frac{1}{n}\sum_{i=1}^n s_i-\beta\right)^2 = \frac{\beta^2}{n}. \tag{B.8}$$

In consequence, as the error for the second case has a lower variance, $\beta^{(2)}$ constitutes a better estimator than $\beta^{(1)}$.

Finally, we compute the FI of the exponential distribution:

$$\chi(\beta) = \int_0^\infty dsP(s|\beta)\left(\frac{\partial}{\partial\beta}\log P(s|\beta)\right)^2 = \frac{1}{\beta^2}, \tag{B.9}$$

which is a decreasing function of β; then, realizations are maximum informative when $\beta \to 0^+$.

We see that the Cramér-Rao inequality is always satisfied, as $\text{var}\left(\beta^{(1,2)}\right) \geq \chi^{-1}(\beta)/n = \beta^2/n$. The equality is only reached by $\beta^{(2)}$; then, among these estimators, only $\beta^{(2)}$ is efficient.

Appendix C

Discrete time calculation for the Hybrid Contact Process

In chapter 3, we based our analysis of the contact process with two dynamics in the stochastic, continuous in time, Langevin picture. In the literature, however, these processes are usually explained using discrete-time multiplicative descriptions. In particular, Kelly showed that, in the context of a gambler investing his/her capital on two different games, the exponential growth rate of the capital corresponds to the *geometric mean* of the respective game rates [76].

We have not seen any similarity with this claim as we are dealing with a continuous rather than a discrete-in-time equation, but we can re-formulate our results to this other framework. To this end, we introduce the discrete ($\Delta t = 1$) analogous of eq. 3.4:

$$N_{t+1} = 2p_{\text{av}}(\alpha)N_t + 2\alpha\sigma N_t r_t, \tag{C.1}$$

with r_t a Gaussian distributed variable with 0 mean and unit variance (i.e. not a white noise) and where we have switched the notation to the number of particles $N_t = \rho(t)N$.

We also discretize the range of values of r, $r = (r_1, ..., r_R)$. With this, starting from some initial condition N_0, the previous equation becomes

$$N_t = \prod_{i=1}^{R} \left[2p_{\text{av}}(\alpha) + 2\alpha\sigma r_i\right]^{n_i} N_0, \tag{C.2}$$

where n_i is the number of times we got r_i, and therefore $\sum_i n_i = t$. The exponential growth rate is derived from its definition $G = \lim_{t\to\infty} \frac{1}{t} \log \frac{N_t}{N_0}$:

$$G(\alpha) = \sum_{i=1}^{R} \frac{n_i}{t} \log \left[2p_{\text{av}}(\alpha) + 2\alpha\sigma r_i\right], \tag{C.3}$$

and, going back to the continuous random Gaussian variable, it reads:

$$G(\alpha) = \int_{-\infty}^{\infty} dr \frac{e^{-\frac{r^2}{2}}}{\sqrt{2\pi}} \log \left[2p_{\text{av}}(\alpha) + 2\alpha\sigma r_i\right]. \tag{C.4}$$

This result is exactly the geometric mean of the different game rates that we can find in the system, i.e. all possible combinations of the conservative strategy p_0 and a risky one with $p = \bar{p} + \sigma r$, with r Gaussian distributed. It has to be solved numerically, but we can compute an approximation in the limit of $\sigma \ll 1$. Expanding the logarithm up to order r^2, we obtain:

$$G(\alpha) = \log\left(2p_{\text{av}}(\alpha)\right) - \frac{1}{2}\left(\frac{\alpha\sigma}{p_{\text{av}}(\alpha)}\right)^2. \tag{C.5}$$

Fig C.1 shows the growth rate for different values of α obtained via eq. C.5 and its analogous eq. 3.5 for a certain choice of the parameters p_0, $\bar{p}$ and σ, illustrating that functions almost coincide.

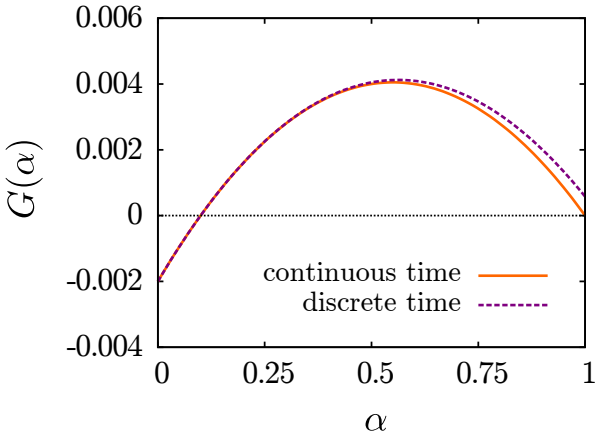


Figure C.1: **Growth rate for hybrid strategies obtained from the discrete and continuous time descriptions**. Parameters: $p_0 = 0.499, \bar{p} = 0.51, \sigma = 0.1$.

Appendix D

Conditions for stochastic amplification

Let us consider the stability (Jacobian) matrix, A, of a two-variable dynamical system (as in eqs. 5.18, 5.19) evaluated at a fixed point, and let λ_1 and λ_2 be its associated eigenvalues. In general, they can be written as complex numbers

$$\lambda_{1,2} = \lambda^R_{1,2} + i\lambda^I_{1,2}. \tag{D.1}$$

As A is a real matrix, its determinant and its trace are both real. This imposes some constraints on the eigenvalues: $\mathrm{Tr}(A) = \lambda^R_1 + \lambda^R_2 + i(\lambda^I_1 + \lambda^I_2) \in \mathbb{R}$ and hence

$$\lambda^I_1 = -\lambda^I_2 \equiv \lambda^I. \tag{D.2}$$

Similarly, $\det(A) = \lambda^R_1\lambda^R_2 - \lambda^I_1\lambda^I_2 + i(\lambda^R_1\lambda^I_2 + \lambda^R_2\lambda^I_1) \in \mathbb{R}$, and therefore

$$\lambda^R_1 = \lambda^R_2 \equiv \lambda^R \quad \text{if} \;\; \lambda^I \neq 0. \tag{D.3}$$

As shown in Chapter 5, the power-spectrum of fluctuations has a maximum around $\omega_0 = \sqrt{\det(A) - (\mathrm{Tr}A)^2/2}$ (eq. 5.24); taking into account eqs. D.2 and D.3, the location of the peak can be rewritten as

$$\omega_0 = \sqrt{(\lambda^I)^2 - (\lambda^R)^2}, \tag{D.4}$$

which provides a direct way to compute ω_0. In particular, observe that $|\lambda^I| > |\lambda^R|$ is a necessary and sufficient condition for a non-trivial maximum to exist, and hence, the system does not exhibit stochastic amplification if λ^I is zero or not sufficiently large. Note that, if A is diagonal (i.e. the two equations become decoupled), $\lambda^I = 0$ and no stochastic amplification can occur. Indeed, it suffices that only one of the non-diagonal terms of A is zero to rule out stochastic amplification.

Stochastic amplification of fluctuations occurs when the deterministic system falls with damped oscillations (spiral decay towards the focus, as corresponds to complex eigenvalues). Noise perturbs trajectories, kicking them away from the focus and sustaining oscillations. It is noteworthy that the selected oscillation frequency does *not* coincide with that of the transitory deterministic dynamics, $\omega_d = \sqrt{\det(A) - \mathrm{Tr}(A)^2/4} = |\lambda^I|$, and consequently, $\omega_0 < \omega_d$.

Bibliography

[1] D. Alonso, A. J. McKane, and M. Pascual. Stochastic amplification in epidemics. *J. R. Soc. Interface*, 4:575–58, 2007.

[2] D. F. Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *The Journal of chemical physics*, 127(21):214107, 2007.

[3] J. R. Auld and R. Rubio de Casas. The correlated evolution of dispersal and mating-system traits. *Evolutionary Biology*, 40(2):185–193, 2013.

[4] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*, 59(4):381–384, 1987.

[5] P. Bak. *How nature works: the science of self-organized criticality*. Copernicus New York, 1996.

[6] E. Balleza, E. R. Alvarez-Buylla, A. Chaos, S. Kauffman, I. Shmulevich, and M. Aldana. Critical dynamics in genetic regulatory networks: examples from four kingdoms. *PLoS One*, 3(6):e2456, 2008.

[7] N. P. Barker. A review and survey of basicarpy, geocarpy, and amphicarpy in the african and madagascan flora. *Annals of the Missouri Botanical Garden*, pages 445–462, 2005.

[8] M. Bazhenov, I. Timofeev, M. Steriade, and T. Sejnowski. Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. *J. Neurosci.*, 22(19):8691–8704, 2002.

[9] J. Beggs and D. Plenz. Neuronal avalanches in neocortical circuits. *The Journal of neuroscience*, 23(35):11167–11177, 2003.

[10] J. M. Beggs. The criticality hypothesis: How local cortical networks might optimize information processing. *Phil. Trans. R. Soc. A*, 366(1864):329–343, 2008.

[11] E. Ben-Jacob, O. Schochet, A. Tenenbaum, I. Cohen, A. Czirok, and T. Vicsek. Generic modelling of cooperative growth patterns in bacterial colonies. *Nature*, 368(6466):46–49, 1994.

[12] A. A. Berryman. *Population cycles: the case for trophic interactions*. Oxford University Press New York, NY, USA, 2002.

[13] R. Bertram, A. Sherman, and E. F. Stanley. Single-domain/bound calcium hypothesis of transmitter release and facilitation. *Journal of Neurophysiology*, 75(5):1919–1931, 1996.

[14] N. Bertschinger and T. Natschlager. Real-time computation at the edge of chaos in recurrent neural networks. *Neural Computation*, 16(7):1413–1436, 2004.

[15] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. Walczak. Statistical mechanics for natural flocks of birds. *Proc Nat Acad of Sci*, 109(13):4786–4791, 2012.

[16] J. Binney, N. Dowrick, A. Fisher, and M. Newman. *The Theory of Critical Phenomena*. Oxford University Press, Oxford, 1993.

[17] J. A. Bonachela, S. de Franciscis, J. J. Torres, and M. A. Munoz. Self-organization without conservation: are neuronal avalanches generically critical? *Journal of Statistical Mechanics: Theory and Experiment*, 2010(02):P02015, 2010.

[18] P. C. Bressloff. Metastable states and quasicycles in a stochastic Wilson-Cowan model of neuronal population dynamics. *Phys. Rev. E*, 82:051903, Nov 2010.

[19] N. Brunel and V. Hakim. Sparsely synchronized neuronal oscillations. *Chaos*, 18(1):015113, 2008.

[20] N. Brunel and D. Hansel. How noise affects the synchronization properties of recurrent networks of inhibitory neurons. *Neural Comput.*, 18(5):1066–1110, May 2006.

[21] N. Brunel and M. C. Van Rossum. Lapicque's 1907 paper: from frogs to integrate-and-fire. *Biological cybernetics*, 97(5-6):337–339, 2007.

[22] R. W. Butler. *Saddlepoint approximations with applications*. Cambridge University Press Cambridge, 2007.

[23] G. Buzsáki. *Rhythms of the Brain*. Oxford University Press, USA, aug 2006.

[24] X. Chen, X. Dong, A. Be'er, H. Swinney, and H. Zhang. Scale-invariant correlations in dynamic bacterial clusters. *Phys Rev Lett*, 108(14):148101, 2012.

[25] D. R. Chialvo. Emergent complex neural dynamics. *Nature Physics*, 6:744–750, 2010.

[26] D. Z. Childs, C. Metcalf, and M. Rees. Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. *Proceedings of the Royal Society B: Biological Sciences*, page rspb20100707, 2010.

[27] J. Claycomb, K. Bassler, J. Miller Jr, M. Nersesyan, and D. Luss. Avalanche behavior in the dynamics of chemical reactions. *Physical review letters*, 87(17):178303, 2001.

[28] J. Clobert, M. Baguette, T. G. Benton, J. M. Bullock, and S. Ducatez. *Dispersal ecology and evolution*. Oxford University Press, 2012.

[29] D. Cohen and S. A. Levin. Dispersal in patchy environments: the effects of temporal and spatial structure. *Theoretical Population Biology*, 39(1):63–99, 1991.

[30] H. N. Comins, W. D. Hamilton, and R. M. May. Evolutionarily stable dispersal strategies. *Journal of Theoretical Biology*, 82(2):205–230, 1980.

[31] A. Compte, R. Reig, V. F., Descalzo, M. A. Harvey, G. D. Puccini, and M. V. Sanchez-Vives. Spontaneous high-frequency (10-80 hz) oscillations during up states in the cerebral cortex in vitro. *J. Neurosci.*, 17:13828–13844, 2008.

[32] A. Corral. Comment on "do earthquakes exhibit self-organized criticality?". *Physical review letters*, 95(15):159801, 2005.

[33] R. Cossart, D. Aronov, and R. Yuste. Attractor dynamics of network up states in the neocortex. *Nature*, 423:283–288, 2003.

[34] T. M. Cover and J. Thomas. *Elements of Information Theory*. Wiley, 1991.

[35] A. J. Crean and D. J. Marshall. Coping with environmental uncertainty: dynamic bet hedging as a maternal effect. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1520):1087–1096, 2009.

[36] J. Csicsvari, H. Hirase, A. Czurko, A. Mamiya, and G. Buzsáki. Fast network oscillations in the hippocampal CA1 region of the behaving rat. *J. Neurosci.*, 19(RC20):1–4, 1999.

[37] H. W. D and R. M. May. Dispersal in stable habitats. *Nature*, 269(5629):578–581, 1977.

[38] C. R. Darwin. *The Origin of Species*, volume Vol. 11. The Harvard Classics. New York, 1909.

[39] L. de Arcangelis, C. Perrone-Capano, and H. J. Herrmann. Self-organized criticality model for brain plasticity. *Physical review letters*, 96(2):028107, 2006.

[40] I. G. de Jong, P. Haccou, and O. P. Kuipers. Bet hedging or not? a guide to proper classification of microbial survival strategies. *Bioessays*, 33(3):215–223, 2011.

[41] H. De Jong. Modeling and simulation of genetic regulatory systems: a literature review. *Journal of computational biology*, 9(1):67–103, 2002.

[42] M. M. de Oliveira and R. Dickman. How to simulate the quasistationary state. *Physical Review E*, 71(1):016129, 2005.

[43] R. Dickman and R. Vidigal. Quasi-stationary distributions for stochastic processes with an absorbing state. *Journal of Physics A: Mathematical and General*, 35(5):1147, 2002.

[44] R. Dickman, M. A. Muñoz, A. Vespignani, and S. Zapperi. Paths to self-organized criticality. *Brazilian Journal of Physics*, 30(1):27–41, 2000.

[45] J. A. Edlund, N. Chaumont, A. Hintze, C. Koch, G. Tononi, and C. Adami. Integrated information increases with fitness in the evolution of animats. *PLoS Comput Biol*, 7(10):e1002236, 10 2011.

[46] R. Fernholz and B. Shay. Stochastic portfolio theory and stock market equilibrium. *The Journal of Finance*, 37(2):615–624, 1982.

[47] R. FitzHugh. Mathematical models of threshold phenomena in the nerve membrane. *The bulletin of mathematical biophysics*, 17(4):257–278, 1955.

[48] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal*, 1(6):445–466, 1961.

[49] S. Fujisawa, N. Matsuki, and Y. Ikegaya. Single neurons can induce phase transitions of cortical recurrent networks with multiple internal states. *Cerebral Cortex*, 16(5):639–654, 2006.

[50] C. Furusawa and K. Kaneko. Adaptation to optimal cell growth through self-organized criticality. *Phys Rev Lett*, 108(20):208103, 2012.

[51] J. Garcia-Ojalvo, M. B. Elowitz, and S. H. Strogatz. Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America*, 101(30):10955–10960, 2004.

[52] C. Gardiner. *Stochastic methods*. Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1985.

[53] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

[54] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional, 1989.

[55] A. Goudarzi, C. Teuscher, N. Gulbahce, and T. Rohlf. Emergent criticality through adaptive information processing in boolean networks. *Phys Rev Lett*, 108(12):12802, 2012.

[56] C. Gros. *Complex and adaptive dynamical systems: A primer*. Springer, 2013.

[57] G. P. Harmer and D. Abbott. Game theory: Losing strategies can win by parrondo's paradox. *Nature*, 402(6764):864–864, 1999.

[58] S. Heilmann, K. Sneppen, and S. Krishna. Sustainability of virulence in a phage-bacterial ecosystem. *Journal of virology*, 84(6):3016–3022, 2010.

[59] M. Henkel, H. Hinrichsen, S. Lübeck, and M. Pleimling. *Non-equilibrium phase transitions*, volume 1. Springer, 2008.

[60] J. Hidalgo, L. Seoane, J. Cortés, and M. Muñoz. Does the phenomenon of stochastic amplification of fluctuations play a relevant role in cortical dynamics? In *AIP Conf. Proc*, volume 1510, page 94, 2012.

[61] J. Hidalgo, L. F. Seoane, J. M. Cortés, and M. A. Muñoz. Stochastic amplification of fluctuations in cortical up-states. *PloS one*, 7(8):e40710, 2012.

[62] J. Hidalgo, J. Grilli, S. Suweis, M. A. Muñoz, J. R. Banavar, and A. Maritan. Information-based fitness and the emergence of criticality in living systems. *Proceedings of the National Academy of Sciences*, 111(28):10095–10100, 2014.

[63] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.

[64] D. Holcman and M. Tsodyks. The emergence of up and down states in cortical networks. *PLoS Comput. Biol.*, 2(3):e23, 03 2006.

[65] K. R. Hopper. Risk-spreading and bet-hedging in insect population biology. *Annual review of entomology*, 44(1):535–560, 1999.

[66] S. Huang. Cell lineage determination in state space: a systems view brings flexibility to dogmatic canonical rules. *PLoS biology*, 8(5):e1000380, 2010.

[67] D. Hughes, M. Paczuski, R. Dendy, P. Helander, and K. McClements. Solar flares as cascades of reconnecting magnetic loops. *Physical review letters*, 90(13):131101, 2003.

[68] E. Imbert. Ecological consequences and ontogeny of seed heteromorphism. *Perspectives in Plant Ecology, Evolution and Systematics*, 5(1):13–36, 2002.

[69] V. A. Jansen and J. Yoshimura. Populations can persist in an environment consisting of sink habitats only. *Proceedings of the National Academy of Sciences*, 95(7):3696–3698, 1998.

[70] H. J. Jensen. *Self-organized criticality: emergent complex behavior in physical and biological systems*. Cambridge university press, 1998.

[71] F. Jinghua, M. Ta-Chung, R. Rittel, and K. Tabelow. Criticality in quark-gluon systems far beyond thermal and chemical equilibrium. *Physical review letters*, 86(10):1961, 2001.

[72] H. Kamiya and R. S. Zucker. Residual ca2+ and short-term synaptic plasticity. *Nature*, 371(6498):603–605, 1994.

[73] E. Kandel, J. Schwartz, and T. Jessel. *Principles of Neural Science*. McGraw-Hill, New York, 2000.

[74] K. Kang, M. Shelley, J. Henrie, and R. Shapley. LFP spectral peaks in V1 cortex: network resonance and cortico-cortical feedback. *J Comput Neurosci*, 29:495–507, 2009.

[75] S. Kauffman. *The origins of order: Self-organization and selection in evolution*. Oxford University Press, USA, 1993.

[76] J. L. Kelly. A new interpretation of information rate. *The Bell System Technical Journal*, 35(4):917–926, 1956.

[77] O. Kinouchi and M. Copelli. Optimal dynamical range of excitable networks at criticality. *Nature Physics*, 2(5):348–351, 2006.

[78] K. S. Korolev, M. Avlund, O. Hallatschek, and D. R. Nelson. Genetic demixing and evolution in linear stepping stone models. *Rev. Mod. Phys.*, 82:1691–1718, May 2010.

[79] D. Krotov, J. O. Dubuis, T. Gregor, and W. Bialek. Morphogenesis at criticality. *Proceedings of the National Academy of Sciences*, 111(10):3683–3688, 2014.

[80] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[81] E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–2078, 2005.

[82] C. Langton. Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D: Nonlinear Phenomena*, 42(1):12–37, 1990.

[83] R. Legenstein and W. Maass. Edge of chaos and prediction of computational performance for neural circuit models. *Neural Networks*, 20(3):323–334, 2007.

[84] A. Levina, J. M. Herrmann, and T. Geisel. Dynamical synapses causing self-organized criticality in neural networks. *Nat. Phys.*, 3(11):857–860, 2007.

[85] A. Levina, J. M. Herrmann, and T. Geisel. Phase transitions towards criticality in a neural system with adaptive interactions. *Phys. Rev. Lett.*, 102(11):118110, Mar 2009.

[86] C.-y. T. Li, M.-m. Poo, and Y. Dan. Burst spiking of a single cortical neuron modifies global brain state. *Science*, 324(5927):643–646, 2009.

[87] B. B. Mandelbrot. *The fractal geometry of nature*, volume 173. Macmillan, 1983.

[88] J. Marro and R. Dickman. *Nonequilibrium phase transitions in lattice models*. Cambridge University Press, 2005.

[89] I. Mastromatteo and M. Marsili. On the criticality of inferred models. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(10):P10012, 2011.

[90] M. Mattia and P. Del Giudice. Population dynamics of interacting spiking neurons. *Physical Review E*, 66(5):051917, 2002.

[91] M. Mattia and P. D. Giudice. Finite-size dynamics of inhibitory and excitatory interacting spiking neurons. *Phys Rev E*, 70:052903, 2004.

[92] J. Maynard Smith. *Models in ecology*. Cambridge University Press, 1974.

[93] D. A. McCormick. Neuronal networks: flip-flops in the brain. *Curr. Biol.*, 15(8):R294–R296, 2005.

[94] A. J. McKane and T. J. Newman. Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett.*, 94(21):218102, Jun 2005.

[95] J. F. Mejias, H. J. Kappen, and J. J. Torres. Irregular dynamics in up and down cortical states. *PLoS ONE*, 5(11):e13651, 2010.

[96] O. Melamed, O. Barak, G. Silberberg, H. Markram, and M. Tsodyks. Slow oscillations in neural networks with facilitating synapses. *J. Comput. Neurosci.*, 25:308–316, 2008. 10.1007/s10827-008-0080-z.

[97] M. Mezard and A. Montanari. *Information, physics, and computation*. OUP Oxford, 2009.

[98] D. Millman, S. Mihalas, A. Kirkwood, and E. Niebur. Self-organized criticality occurs in non-conservative neuronal networks during 'up' states. *Nat. Phys.*, 6(10):801–805, 2010.

[99] M. Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet mathematics*, 1(2):226–251, 2004.

[100] J. Monod. The growth of bacterial cultures. *Annual Reviews in Microbiology*, 3(1):371–394, 1949.

[101] T. Mora and W. Bialek. Are biological systems poised at criticality? *Journal of Statistical Physics*, 144(2):268–302, 2011.

[102] T. Mora, A. Walczak, W. Bialek, and C. Callan Jr. Maximum entropy models for antibody diversity. *Proceedings of the National Academy of Sciences*, 107(12):5405–5410, 2010.

[103] M. Mukovski, S. Chauvette, I. Timofeev, and M. Volgushev. Detection of active and silent states in neocortical neurons from the field potential signal during slow-wave sleep. *Cereb. Cortex*, 17(2):400–414, 2007.

[104] M. A. Muñoz. Nature of different types of absorbing states. *Physical Review E*, 57(2):1377, 1998.

[105] R. Nisbet and W. Gurney. A simple mechanism for population cycles. *Nature*, 263:319–320, 1976.

[106] M. A. Nowak. *Evolutionary dynamics*. Harvard University Press, 2006.

[107] M. Nykter, N. Price, M. Aldana, S. Ramsey, S. A. Kauffman, L. E. Hood, O. Yli-Harja, and I. Shmulevich. Gene expression dynamics in the macrophage exhibit criticality. *Proc Nat Acad of Sci*, 105(6):1897–1900, 2008.

[108] L. Pantic, J. Torres, H. Kappen, and S. Gielen. Associative memory with dynamic synapses. *Neural Comp.*, 14:2903–2923, 2002.

[109] N. Parga and L. F. Abbott. Network model of spontaneous activity exhibiting synchronous transitions between up and down states. *Front. Neurosci.*, 1(1):57–66, 2007.

[110] J. M. Parrondo, G. P. Harmer, and D. Abbott. New paradoxical games based on brownian ratchets. *Physical Review Letters*, 85(24):5226, 2000.

[111] O. Peters, C. Hertlein, and K. Christensen. A complexity view of rainfall. *Physical review letters*, 88(1):018701, 2001.

[112] O. Ronce. How does it feel to be like a rolling stone? ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*, pages 231–253, 2007.

[113] D. Roze and F. Rousset. Inbreeding depression and the evolution of dispersal rates: a multilocus model. *The American Naturalist*, 166(6):708–721, 2005.

[114] M. V. Sanchez-Vives and D. A. Mccormick. Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat. Neurosci.*, 3(10):1027–1034, 2000.

[115] I. N. Sanov. *On the probability of large deviations of random variables*, volume 42. 1957.

[116] A. Satake, A. Sasaki, and Y. Iwasa. Variable timing of reproduction in unpredictable environments: Adaption of flood plain plants. *Theoretical Population Biology*, 60(1):1–15, 2001.

[117] F. Sattin and M. Baiesi. Self-organized-criticality model consistent with statistical properties of edge turbulence in a fusion plasma. *Physical review letters*, 96(10):105005, 2006.

[118] H. G. Schuster, D. Plenz, and E. Niebur. *Criticality in Neural Systems*. John Wiley & Sons, 2014.

[119] D. J. Schwab, I. Nemenman, and P. Mehta. Zipf's law and criticality in multivariate data without fine-tuning. *arXiv preprint arXiv:1310.0448*, 2013.

[120] J. Seger. What is bet-hedging? *Oxford surveys in evolutionary biology*, 4:182–211, 1987.

[121] T. Sejnowski and O. Paulsen. Network oscillations: emerging computational principles. *J. Neurosci.*, 26(6):1673–1676, 2006.

[122] W. L. Shew and D. Plenz. The functional benefits of criticality in the cortex. *The Neuroscientist*, 19(1):88–100, 2013.

[123] A. M. Simons and M. O. Johnston. Environmental and genetic sources of diversification in the timing of seed germination: implications for the evolution of bet hedging. *Evolution*, 60(11):2280–2292, 2006.

[124] J. M. Smith. *Evolution and the Theory of Games*. Cambridge university press, 1982.

[125] A. Solopova, J. van Gestel, F. J. Weissing, H. Bachmann, B. Teusink, J. Kok, and O. P. Kuipers. Bet-hedging during bacterial diauxic shift. *Proceedings of the National Academy of Sciences*, 111(20):7427–7432, 2014.

[126] M. Spiridon and W. Gerstner. Noise spectrum and signal transmission through a population of spiking neurons. *Network*, 263:257–272, 1999.

[127] H. E. Stanley. *Introduction to phase transitions and critical phenomena*. Oxford University Press, 1987.

[128] M. Steriade, F. Amzica, and D. Contreras. Synchronization of fast (30-40 Hz) spontaneous cortical rhythms during brain activation. *J. Neurosci.*, 16(1):392–417, 1996.

[129] M. P. Stumpf, Z. Laidlaw, and V. A. Jansen. Herpes viruses hedge their bets. *Proceedings of the National Academy of Sciences*, 99(23):15234–15237, 2002.

[130] V. Tangri, A. Das, P. Kaw, and R. Singh. Continuum self-organized-criticality model of turbulent heat transport in tokamaks. *Physical Review Letters*, 91(2):025001, 2003.

[131] M. V. Tsodyks and H. Markram. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA*, 94(2):719–723, 1997.

[132] M. V. Tsodyks, K. Pawelzik, and H. Markram. Neural networks with dynamic synapses. *Neural Comput.*, 10(4):821–835, 1998.

[133] F. Vazquez, J. A. Bonachela, C. López, and M. A. Muñoz. Temporal griffiths phases. *Physical review letters*, 106(23):235702, 2011.

[134] J.-W. Veening, W. K. Smits, and O. P. Kuipers. Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.*, 62:193–210, 2008.

[135] D. Venable and J. Brown. The population-dynamic functions of seed dispersal. *Vegetatio*, 107(1):31–55, 1993.

[136] T. Vicsek and A. Zafeiris. Collective motion. *Physics Reports*, 517(3):71–140, 2012.

[137] A. Wagner. *Robustness and evolvability in living systems*. Princeton University Press Princeton, 2005.

[138] E. Wallace, M. Benayoun, W. van Drongelen, and J. Cowan. Emergent Oscillations in Networks of Stochastic Spiking Neurons. *PLoS ONE*, 6:e14804, 2011.

[139] M. F. Weber, G. Poxleitner, E. Hebisch, E. Frey, and M. Opitz. Chemical warfare and survival strategies in bacterial range expansions. *Journal of The Royal Society Interface*, 11(96):20140172, 2014.

[140] P. D. Williams and A. Hastings. Paradoxical persistence through mixed-system dynamics: towards a unified perspective of reversal behaviours in evolutionary ecology. *Proceedings of the Royal Society B: Biological Sciences*, page rspb20102074, 2011.

[141] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12(1):1–24, 1972.

[142] C. Wilson and Y. Kawaguchi. The origins of two-state spontaneous membrane potential flucuations of neostriatal spiny neurons. *J. Neurosci.*, 16(7):2397–2410, 1997.

[143] C. Wilson. Up and down states. *Scholarpedia Journal*, 3(6):1410, 2008.

[144] D. M. Wolf, V. V. Vazirani, and A. P. Arkin. Diversity in times of adversity: probabilistic strategies in microbial survival games. *Journal of theoretical biology*, 234(2):227–253, 2005.

[145] D. M. Wolf, V. V. Vazirani, and A. P. Arkin. A microbial modified prisoner's dilemma game: how frequency-dependent selection can lead to random phase variation. *Journal of theoretical biology*, 234(2):255–262, 2005.

[146] X. Yang, S. Du, and J. Ma. Do earthquakes exhibit self-organized criticality? *Physical review letters*, 92(22):228501, 2004.

[147] R. Yuste, J. N. MacLean, J. Smith, and A. Lansner. The cortex as a central pattern generator. *Nat. Rev. Neurosci.*, 6:477–483, 2005.