Imperial College of Science, Technology and Medicine
Department of Mathematics

Response Theory and Critical Phenomena for Noisy Interacting Systems

PhD candidate: Niccolò Zagli

PhD supervisors: Grigorios A. Pavliotis and Valerio Lucarini

PhD examiners: Jeroen Lamb and Georg Gottwald

Abstract

In this thesis we investigate critical phenomena for ensembles of identical interacting agents, namely weakly interacting diffusions. These interacting systems undergo two qualitatively different scenarios of criticality, *critical transitions* and *phase transitions*. The former situation conforms to the classical tipping point phenomenology that is observed in finite dimensional systems and originates from a setting where negative feedbacks that stabilise the system progressively loose their efficiency, resulting in amplified fluctuations and correlation properties of the system. On the other hand, *phase transitions* stem from the complex interplay between the agents' own dynamics, the coupling among them and the noise, leading to macroscopic emergent behaviour of the system, and are only observed in the thermodynamic limit. Classically, *phase transitions* are investigated with the use of suitable macroscopic variables, called order parameters, acting as effective reaction coordinates that capture the relevant features of the macroscopic dynamics. However, identifying an order parameter is not always possible. In this thesis we adopt a complementary point of view, based on Linear Response theory, to investigate such critical phenomena. We are able to identify the conditions leading either to a *critical transition* or a *phase transition* in terms of spectral properties of suitable response operators. We associate critical phenomena to settings where the response of the system breaks down. In particular, we are able to characterise these critical scenarios as settings where the complex valued susceptibility of the system develops a non analytical behaviour for real values of frequencies, resulting in a macroscopic resonance of the system. We provide multiple paradigmatic examples of equilibrium and nonequilibrium *phase transitions* where we are able to prove mathematically and numerically the clear signature of a singular behaviour of the susceptibility at the phase transition as the thermodynamic limit is reached. Being associated to spectral properties of suitable operators describing either correlation or response properties, these resonant phenomena do not depend on the specific details of the applied forcing nor on the observable under investigation, allowing one to bypass the problem of the identification of the order parameter for the system.

Contents

Chapter 1

Introduction

1.1 Motivation and Objectives

The investigation of dynamical phenomena in complex networks constructed according to different topologies is an extremely active research area [PG16, KAB$^+$14, YRMK21]. Multiagent systems, i.e. ensembles of interacting dynamical agents, are in this regard a fundamental class of models commonly employed to investigate various phenomena in the natural sciences, social sciences and engineering [NPT10, PT13, KjHGG11]. Applications of multiagent systems are ubiquitous, ranging from the more classical cooperation [Daw83] and synchronisation [ABPV$^+$05, PKRK03] phenomena but also including areas such as management of natural hazards [SGL19] and of climate change impacts [Gei18], consensus formation [GPY17], algorithms for sampling, optimization and the training of neural networks [RVE18, GINR20, BKPP21] and emergent phenomena in neural networks and life sciences [CDPF15, DP19].

It is well known that high dimensional complex systems, possibly featuring interacting degrees of freedom on multiple space and time scales, often exhibit abrupt changes in their behaviour, a phenomenon that is commonly referred to as critical transition [Sch09a]. The concept of critical transition encompasses a variety of fields of science, ranging from climate science, where they are commonly referred to as "tipping points", to ecology ("regime shifts") and neuroscience [FPS18]. Such transitions are often associated to undesirable and catastrophic events [Arn92], such as climate crisis, market crashes, death of ecosystems, etc [Sch09a, Sor06]. Incredible efforts have been devoted to the understanding of the dynamics leading to tipping points, resulting in the following fruitful conceptual classification of scenarios in which a system might lose its stability [AWVC12]:

- Bifurcation (B-)tippings refer to situations where, as the parameters of the system change parametrically, i.e. in an adiabatical way, the reference state becomes unstable and the system is driven towards another, most likely unpredictable, state.

- Noise induced (N-)tippings arise instead when the system features some degree of randomness, leading the multiple stable attractors to become metastable and inducing transitions, on some relevant timescales, between them.

- Shock (S-)tippings originate when the system is subjected to a strong and sudden exogenous impact that drives it outside the basin of attraction of the unperturbed state. In this critical scenario the geometry of the basins becomes a very relevant feature of the dynamics since it is closely related to the concept of "minimal fatal shock" that breaks the stable state of the system [HF20].

- As a last scenario, Rate induced (R-)tippings are not necessarily associated with bifurcation or noise mechanisms but derive from sufficiently rapid changes to an external input, that could potentially lead the system to not being able to track the branch of stable attractors and tip to a different phase space region.

In natural systems, the existence of transient chaos and unstable chaotic saddles has been identified as another source of uncertainty that accompanies multiple scenarios of critical transitions [Feu08, KFT16, HVF21, LB17, LB20]. The importance of developing tools for predicting critical transitions has long been recognized. Such early warning signals include an increase in variance and correlation time of the dynamical variables as the system approaches the transition point [DSvN+08, Kue11, SBB+09].

Multiagent systems can also undergo an alternative scenario of criticality, namely phase transitions, due to the appearance of complex emergent macroscopic behaviours stemming from the coupling among the agents. This qualitative change of the properties of the system derives from a different mathematical mechanism, i.e. exchange of stability in the thermodynamic limit of nonunique stationary distributions as the parameters of the systems vary. Generally, phase transitions are usually investigated by

a. identifying an order parameter, i.e. a suitable observable of the system able to capture some degree of collective behaviour of the possibly high dimensional system,

b. verifying that in the thermodynamic limit, for some value of the parameters of the system, the properties of such an order parameter undergo a sudden change.

An order parameter is essentially a "smart" projection operation from the full, high dimensional dynamics to a low dimensional, effective dynamical system that keeps the relevant macroscopic features of the exact system. As such, the notion of an order parameter is closely linked to the concept of reaction coordinates and dimension reduction techniques for high dimensional systems. It is important to point out that while order parameters can in many cases be easily deduced for equilibrium systems using, e.g. symmetry arguments, the definition of reaction coordinates for nonequilibrium system is far more challenging [MD05, BLP06, Rog21].

The goal of this thesis is to adopt a different and complimentary viewpoint for the investigation of critical phenomena for multiagent systems. Inspired by the success of Linear Response Theory in explaining and predicting tipping points in high dimensional systems, such as the climate [CNK+14, TLD18], we here also associate the nearing of a phase transition with the setting where a very small cause leads to very large effects. More technically, as in the case of critical transitions for finite dimensional systems, we associate phase transitions of the thermodynamic limit of interacting systems to the breakdown of linear response properties and the development of non analytical, singular behaviours of response functions (susceptibilities) of the system.

The methodology we will employ is based, from a mathematical standpoint, on the theory of

Markov semigroups for stochastic differential equations (also known as transfer operator theory in deterministic settings) and in particular to their spectral theory based on a generalisation of the concept of Ruelle Pollicott resonances to stochastic systems. The response of the system to arbitrary (weak) perturbations can be in fact related to spectral properties of the operators describing the unperturbed reference dynamics. As a result, by adopting a response theory perspective, we can identify critical settings in terms of universal properties of the dynamics of the system, stemming from the spectral properties of suitable operators, rather than from the specific details of the dynamics, applied forcings or of the observables under investigation. We also provide extensive numerical simulations for the investigation of critical phenomena in multiagent systems. The numerical perspective we adopt mirrors spectroscopic techniques that are used for investigating the frequency dependence of the optical properties of materials and is aimed at highlighting the development of singularities of the complex valued susceptibility of the system as the thermodynamic limit is reached.

Outline and results of the thesis

We here briefly describe the main structure and achievements of this thesis. The next two chapters provide an extensive summary of the literature regarding, respectively, weakly interacting diffusions and linear response theory for finite dimensional systems. We have tried to provide a coherent review of the topic, with a focus on the investigation on fluctuations and critical phenomena.

In particular, in chapter 2 we introduce the class of interacting systems, namely weakly interacting diffusions, under investigation. The main feature of weakly interacting diffusions is that, in the thermodynamic limit, they are described by a nonlinear, nonlocal Fokker Planck equation, commonly referred to as McKean Vlasov equation, that supports multistability of invariant measures, as opposed to standard settings for stochastic finite dimensional systems where the physical invariant measure is usually unique and ergodic. The McKean Vlasov equation represents a *Law of Large Numbers* that is obtained as a result of propagation of chaos effects. Firstly, we provide an all rounded introduction on the propagation of chaos phenomenon by approaching it from different perspectives, such as a physics-oriented many-body statistical mechanical approach and a more mathematically rigorous formulation in terms of measure valued processes. These are classical and well known results for weakly interacting diffusions. We then introduce topics that are at the centre of current research from the applied mathematics community. We define the notion of *phase transitions* for the thermodynamic limit of such systems and show how the existence of such *phase transitions*, or, more specifically, the presence of a non convex free energy landscape for equilibrium systems represents the natural obstacle to uniform in time propagation of chaos. In particular we focus on the study of finite size fluctuations by introducing a *Central Limit Theorem* around the McKean Vlasov equation. While non critical fluctuations are expected to be Gaussian, the rigorous study of fluctuations of finite systems at phase transitions (and their universality features) remains an open problem. We conclude the chapter by introducing a *Large Deviation Principle* through a Macroscopic Fluctuation Theory perspective and investigate the statistics of metastability phenomena, such as first passage problems, within this framework. This Chapter should serve as an introduction of the mathematical study of the complex dynamics of finite dimensional weakly interacting systems.

Moreover, chapter 3 is devoted to an introduction of Response Theory, that is, the relationship between forced and unforced fluctuations, for finite dimensional systems. We provide an overview of the topic introducing the notions of Response Formulas and Fluctuation Dissipation Theorems for both deterministic and stochastic systems. While such concepts are classical, well known results in the physics community, they are commonly associated to systems near equilibrium conditions. What might be less known is that such concepts can be defined and studied in very general terms for systems in strong nonequilibrium conditions, as pioneered by David Ruelle's work on deterministic chaotic dynamical systems. In this chapter we review the conditions for which one expects Response Theory and Fluctuation-Dissipation Theorems to hold for general systems. In particular, we stress how smoothness properties of the invariant measure supported by stochastic systems are at the core of the validity of Fluctuation-Dissipation Theorems. The main result we recall in this chapter is the development of a Response Theory by adopting a functional analysis, operator based framework. This allows us to introduce fundamental concepts of ergodic theory such as the spectral decomposition of correlation functions, power spectra and response functions in terms of the stochastic Ruelle Pollicott resonances, that is, the discrete eigenvalues in suitable function spaces of the generator of the stochastic process.

The remaining chapters provide the original work developed in this thesis.
In chapter 4 we investigate from a mathematical perspective the development of critical phenomena of weakly interacting diffusions. In particular we derive linear response formulas for the thermodynamic limit of such systems which, mathematically, is equivalent to developing a response theory for a non linear, non local Partial Differential Equation. The main result of this section consists in obtaining a formula for the *macroscopic* susceptibility $\tilde{\chi}_i(\omega)$ describing the response of the system to general forcings in terms of a decomposition of terms that can be attributed either to external or to endogenous processes. We define critical conditions for the system as settings where response properties of the system break down, that is, settings where the *macroscopic* susceptibility develops a singularity for a real frequency $\omega \in \mathbb{R}$. The decomposition of the *macroscopic* susceptibility allows us to identify conditions leading to either critical transitions, characterised by a divergence of correlation properties of any microscopic degree of freedom, and phase transitions, characterised by a divergence of response properties due to endogenous processes mediated by the coupling among the agents resulting in a critical slowing down of macroscopic observables of the system. Furthermore, we elucidate the properties and differences of critical and phase transitions of interacting systems by applying the spectral decomposition introduced in Chapter 3 to response operators of the system. We also provide dispersion relations for the thermodynamic limit of the ensemble of agents and their modification at the phase transition point, as the singular behaviour of the susceptibility gives an ulterior contribution to the Kramers-Kronig relations.

In Chapter 5 we study, both in a mathematical and numerical way, two paradigmatic examples of equilibrium and nonequilibrium phase transitions. We here apply the theory developed in the previous chapter to provide an analysis, based on the divergence of response operators, of the development of phase transition phenomena for these two examples. However, the main focus of this chapter is to apply a computational probe experiment methodology, mirroring spectroscopic techniques for the investigation of optical materials, to construct response operators for finite dimensional systems. We show how this provides a powerful approach to elucidate the singular behaviour of response functions of the system, characterising the development of a phase transition, as the system reaches the thermodynamic limit. In this chapter we clarify the relationship between smooth microscopic degrees of freedom and their macroscopic singular behaviour. In particular, the linear response experiment methodology that we apply here provides a conceptual foundation for the investigation of phase transitions for finite dimensional systems through the mean field description given by the McKean Vlasov equation.

Lastly, inspired by the close analogy between the notion of order parameters and reaction coordinates of reduced models, in chapter 6 we provide a dimension reduction technique for the thermodynamic limit of weakly interacting diffusions based on a cumulant expansion of the infinite dimensional probability distribution. This well known procedure results in a parametrisation of the dynamics in terms of a low number of cumulants that act as effective reaction coordinates. We show that the low dimensional dynamics returns the correct *diagnostic* properties since it produces a quantitatively accurate representation of the stationary phase diagram of the system that we compare with exact analytical results and numerical simulations. Moreover, we prove that the reduced order dynamics yields the *prognostic*, i.e., time dependent properties too as it provides the correct response of the system to external perturbations. By adopting a linear response perspective for the reduced dynamics and identifying settings where the susceptibilities of the reduced system exhibit a resonant behaviour, we pinpoint and characterise phase transition phenomena and, ultimately, show how the issue of finding an order parameter can be bypassed in this approach.

1.2 Statement of Originality and publications

I declare that this thesis is my own work and that work by others has been properly referenced. The thesis is based on the following three publications [LPZ20, ZLP21, ZPLA23]:

- V. Lucarini, G. A. Pavliotis, and N. Zagli. *Response theory and phase transitions for the thermodynamic limit of interacting identical systems.* Proc. R. Soc. A., 476, 2020.
- Niccolò Zagli, Valerio Lucarini, Grigorios A. Pavliotis. *Spectroscopy of phase transitions for multiagent systems.* Chaos: An Interdisciplinary Journal of Nonlinear Science, 31(6):061103, 2021.
- Niccolò Zagli, Valerio Lucarini, Grigorios A. Pavliotis, Alexander Alecio. *Dimension reduction of noisy interacting systems*, Phys. Rev. Res., 5:013078, Feb 2023.

and has led to writing of the following publication [ZLP23]:

- Niccolò Zagli, Valerio Lucarini, Grigorios A. Pavliotis. *Response Theory Identifies Reaction Coordinates and Explains Critical Phenomena in Noisy Interacting Systems.* arXiv:2303.09047v1 2023.

1.2.1 Copyright statement

Chapter 2

Weakly Interacting Diffusions

2.1 The class of models

We consider an ensemble of N identical interacting M-dimensional systems $\{\mathbf{x}^k(t)\}_{k=1}^N \subset \mathbb{R}^M$. The dynamics of the N particle system (see Figure 2.1) is described by the following stochastic differential equations:

$$\mathrm{d}x_i^k = F_{i,\alpha}(\mathbf{x}^k)\mathrm{d}t - \frac{\theta}{N}\sum_{l=1}^{N} \partial_{x_i^k}\mathcal{U}\left(\mathbf{x}^k - \mathbf{x}^l\right)\mathrm{d}t + \sigma \hat{s}_{ij}(\{\mathbf{x}^l\}_{l=1}^N)\mathrm{d}W_j, \tag{2.1}$$

where $k = 1,\ldots,N$, $i = 1,\ldots,M$ and the Einstein summation convention is used. In the following, we consider a chaotic initial condition, such that all the particles at time $t = 0$ are statistically independent and identically distributed according to a distribution ρ_{in}, that is $\mathbf{x}^k(t = 0) \sim \rho_{in}(\mathbf{x}^k)$. The smooth vector field $\mathbf{F}_\alpha : \mathbb{R}^M \to \mathbb{R}^M$, possibly depending on a parameter α, determines the individual dynamics of each of the interacting systems. Moreover, the N systems undergo an all-to-all coupling given by the pair-wise symmetric interaction potential $\mathcal{U} : \mathbb{R}^M \to \mathbb{R}$, with $\mathcal{U}(\mathbf{x}) = \mathcal{U}(-\mathbf{x})$. The coefficient θ modulates the intensity of such a coupling, which attempts at synchronising all systems. The interaction potential $\mathcal{U}$ can be either long or short range, see the list of examples below. Additionally, $\{W_i\}_{i=1}^N$, are independent Brownian motions. We here use the Ito's conventions. It is well known that different noise conventions result in different stability properties of stochastic differential equations with multiplicative noise. In many systems with a separation of timescale, the Stratonovich convention arises naturally as the timescale separation grows to infinity. However, there are situations where neither the Ito nor the Stratonovic interpretation are correct [Pav14, Ch. 5]. We will not delve here into these interesting modelling issues and we will just mention that there exists a systematic way to pass from different interpretation of the noise by suitably modifying the drift term [Pav14]. The results of this thesis can be applied seamlessly to any interpretation of the noise. We will investigate in chapter 6 a model with multiplicative noise with Stratonovich interpretation. $\hat{s}_{ij}(\cdot)$ is the N-particle volatility matrix, and the parameter $\sigma > 0$ controls the intensity of the stochastic forcing. The N-particle volatility matrix is a block diagonal matrix composed of identical blocks $\mathbf{s} : \mathbb{R}^M \to \mathbb{R}^{M\times M}$

$$\hat{s}\left(\{\mathbf{x^k}\}\right) = \begin{pmatrix} \mathbf{s}(\mathbf{x}^1) & 0 & \dots & 0 \\ 0 & \mathbf{s}(\mathbf{x}^2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{s}(\mathbf{x}^N) \end{pmatrix}, \tag{2.2}$$

where $\mathbf{s}$ represents the one-particle volatility matrix. This preserves the exchangeability properties of the N systems and, in particular, does not introduce a coupling among them through the noise. If $\mathbf{F}_\alpha(\mathbf{y}) = -\nabla V_\alpha(\mathbf{y})$, we interpret V_α as the confining potential [Pav14]. In some cases, equation (2.1) describes an equilibrium statistical mechanical system, in particular if $\mathbf{F}_\alpha = -\nabla V_\alpha(\mathbf{y})$ and $\hat{s}_{ij}$ is proportional to the identity. More generally, equilibrium conditions are realised when the drift term - the deterministic component on the right hand side of equation (2.1) - is proportional to the gradient of a function defined according to the Riemannian metric given by the N-particle diffusion matrix $\hat{\Sigma}_{ij}(\cdot) = \hat{s}_{ik}(\cdot)\hat{s}_{jk}(\cdot)$ [Gra77].

As previously mentioned, interacting systems are ubiquitously used in multiple areas of science. Below we provide a list of common examples of weakly interacting diffusions models. Undoubtedly, the list is not exhaustive but could serve the reader to appreciate the vast range of applications for which such models have been employed.

- Order-Disorder continuous phase transitions in ferromagnetic-like models. The paradigmatic example for such models is the Desai-Zwanzig model [DZ78, Shi87, Daw83] defined by $M = 1$, a local double well dynamics given by a confining potential $V_\alpha(x) = \frac{x^4}{4} - \alpha\frac{x^2}{2}$, a Curie-Weiss quadratic interaction potential $\mathcal{U}(x) = \frac{x^2}{2}$ and thermal noise $\hat{s}_{ij} = \delta_{ij}$, where δ_{ij} is the Kronecker delta. See also section 2.6 for a more in-depth analysis of the model. Moreover, in chapter 6 we will investigate a variant of this model [VdBPAHM94] that features multiplicative noise acting on the system. More complicated settings have also been investigated, such as multiwell and random energy landscapes $V_\alpha(x)$ [GKPY19] or two-scale periodic potentials [GP18].

- Opinion Dynamics models, such as the noisy Hegselmann-Krause model [WLEC17, GPY17, RH02], defined on the torus $\mathbb{T}^M$ rather than $\mathbb{R}^M$ ($M = 1$), and characterised by $F_\alpha(x) = 0$, thermal noise and a short range interaction potential $\mathcal{U}(x) \propto \left(\min\{0, |x| - \frac{R}{2}\}\right)^2$, where R represents the radius of attraction of each agent. Due to the interactions, two agents whose *opinions* are close enough will merge, creating mesoscopic clusters of *opinions*. If the thermal noise is low enough, such clusters perform a motion similar to random walk and will merge as soon they get close to other clusters, eventually reaching a consensus. As the noise increases, the ordering effect given by the coupling becomes progressively less inefficient and the system transitions to an asymptotic state characterised by a uniform distribution of *opinions*. In fact, this model exhibits a discontinuous phase transition for low enough R [CGPS20]. Similar cluster formation phenomena, and relative discontinuous phase transitions, have been studied in [MA01, GRVE22], for 2 dimensional particle systems with gaussian interaction potential.

- Phase synchronisation of globally coupled oscillators such as the (noisy) Kuramoto model [ABPV$^+$05, CGPS20], where the interaction potential is periodic $\mathcal{U}(x) \propto \cos(x)$. Interesting variants of the model have also been studied where the introduction of a new harmonic in the potential [VKP15] gives rise to a non trivial dynamics, featuring self-consistent partial synchrony or switching properties [CP17].

- Synchronisation of periodic or chaotic oscillators [Sak00, PKRK03, Pec98, PC15, ELP17, TKP01]. Given a quadratic interaction potential $\mathcal{U}(x) = \frac{x^2}{2}$ and assuming that the

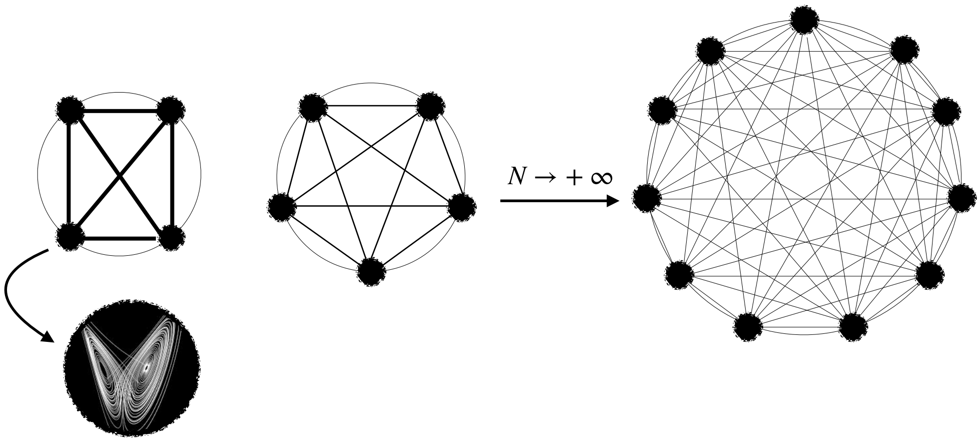


Figure 2.1: Weakly interacting diffusions. The local dynamics of each subsystem, given by the smooth vector field $\mathbf{F}_\alpha(\mathbf{x})$, is in general dissipative and can support a wide range of complex behaviours, including deterministic chaos, as depicted in this figure.

local dynamics $\mathrm{d}\mathbf{x}/\mathrm{d}t = \mathbf{F}_\alpha(\mathbf{x})$ supports a unique attractor that is chaotic, the theory of synchronisation provides a strong link between synchronisation regimes and Lyapunov exponents of $\mathbf{F}_\alpha(\mathbf{x})$. In particular, the N systems undergo complete synchronization for any $N \geq 2$ in the absence of noise ($\sigma = 0$) if $\theta > \Lambda_1$, where $\Lambda_1 > 0$ is the first Lyapunov exponent of the dynamics given by $\mathbf{F}_\alpha(\mathbf{x})$. We observe that the onset of synchronisation has also been investigated for more complicated global coupling in [OSBA02].

- Other applications of identical interacting agents include life sciences [DP19], formation of swarms [CCH14], collective periodic behaviours for non oscillatory agents [CDPF15], bacterial chemotaxis, dynamical network [CGPS20], self gravitating systems [Cha14, TBDR05], optimisation and sampling [RVE18, KPP19, GINR20]. Furthermore, such particle systems are used very commonly for general diffusion-aggregation problems, see [CCY19] and references therein and [EK16].

2.2 The thermodynamic (mean field) limit

An arbitrary state, at any arbitrary time t, of the N particle system is uniquely determined by specifying the whole collection $\{\mathbf{x}^k(t)\}_{k=1}^N$, solution of equations (2.1). Such an approach is clearly unfeasible as soon as the number of particles exceeds a certain (low) threshold, since one has to track the time evolution of $N \times M$ degrees of freedom. An alternative perspective is to consider the statistical properties of the system by introducing a measure μ_N for the system that satisfy the N particle Fokker-Planck equation [Ris89, Pav14]

$$\frac{\partial \mu_N}{\partial t} = \mathcal{L}_N \mu_N, \tag{2.3}$$

where the N particle Fokker-Planck operator is given by

$$\mathcal{L}_N(\cdot) = -\sum_{k=1}^{N}\sum_{i=1}^{M} \frac{\partial}{\partial x_i^k}\left(\left(F_{i,\alpha}(\mathbf{x}^k) - \frac{\theta}{N}\sum_{l=1}^{N} \partial_{x_i^k}\mathcal{U}\left(\mathbf{x}^k - \mathbf{x}^l\right)\right)\cdot\right) + \frac{\sigma^2}{2}\sum_{k}^{N}\sum_{i,j=1}^{M} \frac{\partial^2}{\partial x_i^k x_j^k}\left(s_{ij}(\mathbf{x}^k)\cdot\right). \tag{2.4}$$

We assume that the above N particle Fokker Planck equation describes a hypoelliptic diffusion process[1], giving rise to a smooth probability distribution [Pav14, Ch. 6]. In this case, the measure μ_N is absolutely continuous to the Lebesgue measure and it is possible to define the N particle probability distribution $\rho_N(\{\mathbf{x}^k)\}, t)$ of the system satisfying the same Fokker-Planck equation

$$\begin{aligned} \frac{\partial \rho_N}{\partial t} &= \mathcal{L}_N \rho_N, \\ \rho_N\left(\{\mathbf{x}^k\}, 0\right) &= \prod_{k=1}^{N} \rho_{in}(\mathbf{x}^k). \end{aligned} \tag{2.5}$$

The above Fokker Planck equation describes a system of N *exchangeable* agents. More specifically, a system $\{\mathbf{x}^k(\cdot)\}_{k=1}^N$ is said to be exchangeable if the probability law of $\{\mathbf{x}^k(\cdot)\}_{k=1}^N$ is identical to that of $\{\mathbf{x}^{\pi(k)}(\cdot)\}_{k=1}^N$ for every permutation π of $1, 2, \ldots, N$, that is $\rho_N\left(\{\mathbf{x}^k\}, t\right) = \rho_N\left(\{\mathbf{x}^{\pi(k)}, t\right)$. As a matter of fact, equations (2.1) with a chaotic initial condition, or equiv-

[1]The precise mathematical definition of a hypoelliptic diffusion process and the conditions on the drift and diffusion coefficients guaranteeing hypoellipticity, namely Hörmander's condition, can be found in chapter 3.

alently equations (2.3), define an exchangeable system $\{\mathbf{x}^k(t)\}_{k=1}^N$. We remark that the mean field techniques described in this section to obtain the thermodynamic limit of the ensemble of agents can be extended to more general settings where the *exchangeability* assumption is no longer strictly valid. One could potentially consider settings, see for example [DP19, PH96], where a source of quenched disorder is introduced in the dynamics (2.1) as each agent is set to depend on a (set of) parameter(s) $\mathbf{h}^k$ drawn from a known probability distribution $g(\mathbf{h})$, that is $\mathbf{h}^k \sim g(\mathbf{h})\ \forall k = 1, \dots, N$. Such microscopic architecture of the system given by $g(\mathbf{h})$ can be interpreted either as an intrinsic property of the system, such as the natural frequencies of an ensemble of oscillators, or as a model error feature arising from partial knowledge of the microscopic properties of the agents. A generalisation of the result of the thesis to such local quenched disorder has been obtained by the author (et al.) in [ZLP23].

Moreover, an interesting topic that has recently become an active research area is represented by *weakly interacting diffusion on graphs*, that is, ensemble of weakly diffusions whose interactions are mediated by an underlying microscopic network [JPS22, CDG20, Cop22, BW22] with a graphon structure [Lov12]. Furthermore we assume that, fixed N, the Fokker-Planck equation (2.5) describes an ergodic process and, in particular, admits an unique stationary solution $\bar{\rho}_N$ such that

$$\lim_{t \to +\infty} \rho_N(\{\mathbf{x}^k)\}, t) = \bar{\rho}_N. \tag{2.6}$$

Ergodicity and uniqueness of the stationary solution can be related to spectral properties of the N particle Fokker-Planck operator, see chapter 3 for an in-depth analysis of spectral properties of transfer operators. It is clear that the stationary distribution $\bar{\rho}_N$, even if it's known, contains an incredible amount of information and its computation can be demanding. Indeed, we are usually not interested in the distribution of the states of all the particles but in the distribution of suitable averages of these states. As is customary in kinetic theory, we consider a subset of n particles and we denote it as $\mathbf{X}^{(n)} = (\mathbf{x}^1, \dots, \mathbf{x}^n)$. We then introduce the reduced probability distributions

$$\rho_n(\mathbf{X}^{(n)}, t) := \frac{N!}{n!\,(N-n)!} \int \rho_N\left(\{\mathbf{x}^k\}\right) \mathrm{d}\mathbf{X}^{(n)}, \tag{2.7}$$

where the integration is performed over all the possible values of $(\mathbf{x}^1, \dots, \mathbf{x}^n)$. The hierarchy of equations one can obtain for the ρ_n with $n = 1, \dots, N$ starting from (2.3) is commonly referred to in statistical physics and kinetic theory as BBGKY (Bogoliubov–Born–Green–Kirkwood–Yvon) hierarchy. In particular, it is known that the full N particle probability distribution ρ_N (and thus all the properties of the reduced distributions ρ_n) can be written as a functional of the one particle reduced distribution ρ_1 both in equilibrium and non equilibrium settings [CF05, HK64, Mer65]. However, such functional is usually unknown. The goal of Dynamic Density Functional Theory, see e.g. [GNS$^+$12, GPK12], is to specify this functional by looking for suitable approximations of higher order reduced probabilities ρ_n in terms of the one particle distribution ρ_1. It is then clear that a fundamental issue is to obtain an evolution equation for ρ_1. However, such an equation depends in general on the full distribution ρ_N. In the particular case of pairwise interaction we are considering, it suffices to know the two body distribution ρ_2. Indeed, by integrating the Fokker-Planck equation (2.3) over all the possible values of $\mathbf{x}^k$ with $k \neq 1$ and considering exchangeability properties of the system, we obtain, see also e.g. [Cha08, GPK12], an equation for the one particle distribution ρ_1

$$\frac{\partial \rho_1(\mathbf{x},t)}{\partial t} = -\nabla \cdot \left[\rho_1(\mathbf{x},t)\mathbf{F}_\alpha(\mathbf{x}) - \theta \frac{N-1}{N} \int \nabla \mathcal{U}\left(\mathbf{x}-\mathbf{y}\right) \rho_2(\mathbf{x},\mathbf{y}) \mathrm{d}\mathbf{y} \right] + \frac{\sigma^2}{2} \tilde{\Delta} \rho_1(\mathbf{x},t), \quad (2.8)$$

where $\tilde{\Delta}$ is a linear diffusion operator so that $\tilde{\Delta}\rho(\mathbf{x},t) = \sum_{i=1}^{M} \sum_{j=1}^{M} \partial_{x_i} \partial_{x_j} \left(s_{ij}(\mathbf{x}) \rho(\mathbf{x},t) \right)$, which coincides with the standard M-dimensional Laplacian ($\tilde{\Delta} = \Delta$) if the matrix s_{ij} is the identity matrix. As previously mentioned, all the reduced distributions ρ_n can be written as a functional of the one particle distribution ρ_1. In particular we can write the two body probability distribution as

$$\rho_2(\mathbf{x},\mathbf{y},t) = \rho_1(\mathbf{x},t)\rho_1(\mathbf{y},t) g\left(\mathbf{x},\mathbf{y};[\rho_1]\right), \quad (2.9)$$

where g is a functional of ρ_1. The functional $g-1$ is usually referred to as pair correlation function and, in this setting, gives a measure of the phase space correlation properties among the particles. The simplest approach in dynamical density functional theory is to consider a mean field approximation where correlation among particles are neglected by setting $g \equiv 1$, such that $\rho_2(\mathbf{x},\mathbf{y},t) \approx \rho_1(\mathbf{x},t)\rho_1(\mathbf{y},t)$. In this case, from (2.9) and (2.8), it is easy to see that one obtains, see also [MA01],

$$\frac{\partial \rho_1(\mathbf{x},t)}{\partial t} = -\nabla \cdot \left[\rho_1(\mathbf{x},t)\left(\mathbf{F}_\alpha(\mathbf{x}) - \theta \nabla \mathcal{U} \star \rho_1\right)\right] + \frac{\sigma^2}{2}\tilde{\Delta}\rho_1(\mathbf{x},t) \tag{2.10}$$

where we have assumed that for big enough N we can approximate $\frac{N-1}{N} \approx 1$ and the symbol $\star$ denotes a convolution operation, such that $\nabla \mathcal{U} \star \rho_1 = \int \nabla \mathcal{U}(\mathbf{x}-\mathbf{y})\rho_1(\mathbf{y},t)\mathrm{d}\mathbf{y}$.

The same result can be obtained with an alternative approach that we illustrate below. As a matter of fact, such mean field approximation is most often (see caveats in section 2.4) correct for the thermodynamic limit $N \to +\infty$ of weakly interacting diffusions. In this regard, it is more convenient to adopt a different statistical approach to the description of the N particle system. Given that the particles $\{\mathbf{x}^k\}$ are exchangeable, the system can be described in terms of its empirical measure

$$X_N(t;A) = \frac{1}{N}\sum_{k=1}^{N} \mathbf{1}_A(\mathbf{x}^k(t)), \tag{2.11}$$

where $\mathbf{1}_A(\cdot)$ represents the indicator function of a set $A \subset \mathbb{R}^M$. Physically, the empirical measure X_N represents the proportion of particles in a region A of the "one-particle" phase space. A fundamental difference with respect to the previous approach is that X_N is a random object whereas the N particle distribution ρ_N and the reduced distributions ρ_n are deterministic. As a fundamental consequence of exchangeability features of the N particle system, the properties of the empirical measure $X_N(t,\cdot)$ can be investigated through the mathematical framework of probability-measured-valued processes theory [Daw83]. The empirical measure contains all the statistical information of the system, in particular it can be shown that the N particle distribution ρ_N can be obtained from the empirical measure [CD22]. Since the seminal papers of McKean [McK66, MJ67], it is known that, under general conditions for the drift and diffusion coefficients, it is possible to prove a *Law of Large Numbers* for the empirical measure X_N. Fixed a time interval $[0,T]$, the empirical measure X_N converges weakly for $N \to +\infty$, see e.g. [Oel84, Szn89, Kol10, DP19], to the one particle measure $\mu(\mathrm{d}\mathbf{x},t) = \rho(\mathbf{x},t)\mathrm{d}\mathbf{x}$ satisfying the following McKean-Vlasov equation, which is a nonlinear and nonlocal Fokker-Planck equation

[Fra05]

$$\begin{aligned}\frac{\partial \rho(\mathbf{x},t)}{\partial t} &= -\nabla\cdot\left[\rho(\mathbf{x},t)\left(\mathbf{F}_\alpha(\mathbf{x}) - \theta\nabla\mathcal{U}\star\rho\right)\right] + \frac{\sigma^2}{2}\tilde{\Delta}\rho(\mathbf{x},t),\\ \rho(\mathbf{x},0) &= \rho_{in}(\mathbf{x}), \quad t\in[0,T].\end{aligned} \tag{2.12}$$

Propagation of chaos in terms of empirical measure implies that, for any $k \in \mathbb{N}$

$$\mathbb{E}\left(X_N\left(t,\mathrm{d}\mathbf{x}_1\right)\ldots X_N\left(t,\mathrm{d}\mathbf{x}_k\right)\right) \xrightarrow[w.c.]{N\to+\infty} \prod_{j=1}^{k}\rho\left(\mathbf{x}_j,t\right)\mathrm{d}\mathbf{x}_j \quad \text{with} \quad t\in[0,T], \tag{2.13}$$

where $\rho(\cdot,t)$ is given by the McKean Vlasov equation (2.12) and the arrow denotes convergence as in weak convergence of probability measures. Furthermore, the expectation value is taken with respect to the N particle probability distribution ρ_N. In order to make a stronger comparison with the kinetic theory approach we have adopted above, we also state the propagation of chaos property in terms of the reduced probability distributions ρ_n. Given any $n \in \mathbb{N}$, propagation of chaos is verified when the reduced probability distribution satisfies a factorisation property

$$\rho_n(\mathbf{X}^{(n)},t) \xrightarrow[w.c.]{N\to+\infty} \prod_{j=1}^{n}\rho(\mathbf{x}_j,t), \quad \text{with} \quad t\in[0,T], \tag{2.14}$$

The *Law of Large Numbers* (2.12) and the propagation of chaos properties (2.13) and (2.14) are all equivalent [CD22, Daw83]. Let us observe that propagation of chaos in weakly interacting diffusions can be interpreted as a natural closure scheme for the BBGKY hierarchy. Further details about propagation of chaos and the link between the microscopic dynamics (2.1) and the McKean-Vlasov equation can be found in section 2.4.

2.2.1 Mean Field Dynamics

We here investigate in greater details the McKean-Vlasov dynamics (2.12). If $\sigma = 0$, we are considering a nonlinear Liouville equation - sometimes called Vlasov equation. In what follows, we refer to the case $\sigma > 0$. As a matter of fact, if $\sigma = 0$ it is not always possible to assume a smooth probability distribution. In the following we will always assume that the limiting one-particle measure μ is absolutely continuous with respect to the Lebesgue measure with a smooth probability distribution $\rho(\mathbf{x}, t)$. Conditions detailing the well-posedness of the $N \to +\infty$ problem can be found in e.g. [Kol10, DP19, DGPS23, DGP21].

The non linear, non local feature of equation (2.12) originates from the convolution product of the interaction potential $\mathcal{U}(\mathbf{x})$ and the probability distribution $\rho(\mathbf{x}, t)$. In absence of coupling, $\theta = 0$, the equation (2.12) corresponds to the linear Fokker Planck equation describing one of the N identical decoupled particles, see equation (2.1).

We observe that the standard theory for Fokker Planck equations [Ris89] does not hold for the McKean-Vlasov equation (2.12). In particular, in the $N \to +\infty$ limit, the system can support multiple coexisting stationary solutions (invariant measures) characterised by their corresponding basins of attraction. As opposed to the standard finite dimensional case, where there is convergence to the unique stationary measure for virtually any initial conditions [Ris89], the choice of the initial condition $\rho_{in}(\mathbf{x})$ for the dynamics is here very relevant. The change in stability or the (dis)appearance of such invariant measures, as the parameters of the system are varied, can be interpreted as phase transitions, see e.g. [Daw83, Shi87, DGPS23]. More details about the difference between the finite dynamics given by (2.1) and the infinite dimensional dynamics described by the McKean Vlasov equation can be found in section 2.3. To further elucidate on the properties of the McKean Vlasov equation it is important to notice that, by itself, equation (2.12) does not specify a stochastic process [Fra01].

Firstly, we write the McKean-Vlasov equation as

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} = \mathcal{L}_{\rho(\mathbf{x},t)} \rho(\mathbf{x}, t), \tag{2.15}$$

where $\mathcal{L}_{\rho(\mathbf{x},t)}$ is an integro-differential operator defined by

$$\mathcal{L}_{\rho(\mathbf{x},t)} \psi = -\nabla \cdot \left[\psi(\mathbf{x}, t) \left(\mathbf{F}_\alpha(\mathbf{x}) - \theta \nabla \mathcal{U} \star \rho \right) \right] + \frac{\sigma^2}{2} \tilde{\Delta} \psi, \tag{2.16}$$

and $\psi(\mathbf{x}, t)$ is a smooth function. In order to specify the full Markov process associated to (2.12), it is necessary to introduce the evolution equation for the conditional probability distribution

$\rho(\mathbf{x},t|\mathbf{x}',t')$ that reads [Fra05], for $t \geq t'$,

$$\begin{aligned} \frac{\partial \rho(\mathbf{x},t|\mathbf{x}',t')}{\partial t} &= \mathcal{L}_{\rho(\mathbf{x},t)}\rho(\mathbf{x},t|\mathbf{x}',t'), \\ \rho(\mathbf{x},t'|\mathbf{x}',t') &= \delta(\mathbf{x}-\mathbf{x}'), \end{aligned} \tag{2.17}$$

where $\delta(\mathbf{x})$ is the Dirac-delta distribution. Once the evolution of the one point distribution function $\rho(\mathbf{x},t)$ and the conditional distribution function $\rho(\mathbf{x},t|\mathbf{x}',t')$ are specified, the hierarchy of equations for higher order distribution functions $\rho(\mathbf{x},t;\mathbf{x}_{n-1},t_{n-1};\dots;\mathbf{x}_1,t_1)$ follows from the Markovian feature of the dynamics and can be written in the same form as (2.15) [Fra05]

$$\frac{\partial}{\partial t}\rho(\mathbf{x},t;\mathbf{x}_{n-1},t_{n-1};\dots;\mathbf{x}_1,t_1) = \mathcal{L}_{\rho(\mathbf{x},t)}\rho(\mathbf{x},t;\mathbf{x}_{n-1},t_{n-1};\dots;\mathbf{x}_1,t_1). \tag{2.18}$$

We remark that the above equations are nonlinear with respect to $\rho(\mathbf{x},t)$ but are instead linear with respect to $\rho(\mathbf{x},t|\mathbf{x}',t')$ and the higher order distribution functions.

It is possible to investigate the dynamics given by the McKean-Vlasov equation in the mathematically rigorous framework of nonlinear Markov processes, see e.g. [Kol10, Daw83, MJ67] for a discussion about nonlinear Markov semigroup theory.

A stationary solution $\rho_0(\mathbf{x})$ of the McKean-Vlasov equation satisfies the eigenvalue problem $\mathcal{L}_{\rho_0(\mathbf{x})}\rho_0(\mathbf{x}) = 0$, with $\mathcal{L}_{\rho_0(\mathbf{x})}$ being an usual Fokker-Planck operator obtained by evaluating the operator $\mathcal{L}_{\rho(\mathbf{x},t)}$ at stationarity. Consequently, in stationary conditions, the conditional probability can be written as $\rho_0(\mathbf{x},t|\mathbf{x}',t') = e^{(t-t')\mathcal{L}_{\rho_0(\mathbf{x})}}\delta(\mathbf{x}-\mathbf{x}')$, see equation (2.17). We define the *mean field* correlation function between two observables[2] as

$$C_{AB}(t) = \langle A(\mathbf{x}(t))B(\mathbf{x}(0))\rangle_0 = \int\int A(\mathbf{x})B(\mathbf{x}')\rho_0(\mathbf{x},t;\mathbf{x}',0)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'. \tag{2.19}$$

As previously noticed, the joint probability $\rho_0(\mathbf{x},t;\mathbf{x}',0)$ of a Markov process can be written as

$$\rho_0(\mathbf{x},t;\mathbf{x}',0) = \rho_0(\mathbf{x},t|\mathbf{x}',0)\rho_0(\mathbf{x}') = e^{t\mathcal{L}_{\rho_0(\mathbf{x})}}\delta(\mathbf{x}-\mathbf{x}')\rho_0(\mathbf{x}'). \tag{2.20}$$

[2] As customary, we have assumed that the expectation value of $\langle A\rangle_0 = \int A(\mathbf{x})\rho_0(\mathbf{x})\mathrm{d}\mathbf{x}$ vanishes. An analogous assumption is made on the observable $B(\mathbf{x})$. Of course, one can always redefine the observables such that the above conditions are met. Furthermore, we have used the fact that at stationarity correlation functions only depend on the time differences, e.g., $\langle A(\mathbf{x}(t))B(\mathbf{x}(t'))\rangle_0 = \langle A(\mathbf{x}(t-t'))B(\mathbf{x}(0))\rangle_0$ given any $t > t'$.

The correlation function can then be written as

$$C_{AB}(t) = \int\int A(\mathbf{x})B(\mathbf{x}')e^{t\mathcal{L}_{\rho_0(\mathbf{x})}}\delta(\mathbf{x}-\mathbf{x}')\rho_0(\mathbf{x}')\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{x}'. \tag{2.21}$$

Performing the integration on $\mathbf{x}'$ in the previous expression, we can obtain a formula for the *mean field* stationary correlation function in terms of the operator $\mathcal{L}_{\rho_0(\mathbf{x})}$

$$C_{AB}(t) = \int A(\mathbf{x})e^{t\mathcal{L}_{\rho_0(\mathbf{x})}}B(\mathbf{x})\rho_0(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{2.22}$$

2.3 The Origin of Phase Transitions

In this section we will study the mechanism underlying the development of phase transitions for the $N \to +\infty$ limit of equations (2.1). We will assume that the dynamics of the agents is given by a confining potential $V_\alpha(\mathbf{x})$ such that $\mathbf{F}_\alpha(\mathbf{x}) = -\nabla V_\alpha(\mathbf{x})$ and that the system is subject to thermal noise with diagonal diffusion matrix $s_{ij}(x) = \delta_{ij}$. We have made these assumptions because, in these settings, it is possible to find a characterisation of phase transitions in terms of convexity properties of the confining and interaction potentials. However, most of the qualitative features regarding phase transitions apply to a general non-equilibrium setting, see Chapter 4. With the above assumptions, the equations of motions (2.1) describe an equilibrium statistical mechanics system. Indeed, it is possible to define an Hamiltonian function

$$\mathcal{H}_N(\{\mathbf{x}^k\}) = \sum_{k=1}^{N} V_\alpha(\mathbf{x}^k) + \frac{\theta}{2N}\sum_{k,l=1}^{N}\mathcal{U}(\mathbf{x}^k - \mathbf{x}^l), \tag{2.23}$$

such that equations (2.1) can be written as

$$\mathrm{d}x_i^k = -\frac{\partial \mathcal{H}_N}{\partial x_i^k}\mathrm{d}t + \sigma \mathrm{d}W_i. \tag{2.24}$$

The linear N-particle Fokker-Planck equation associated to the previous stochastic differential equations is

$$\frac{\partial \rho_N}{\partial t} = \nabla \cdot (\rho_N \nabla \mathcal{H}_N) + \frac{\sigma^2}{2}\Delta\rho_N, \tag{2.25}$$

where $\rho_N = \rho_N(\{\mathbf{x}^k\}, t)$ is the probability distribution of the N particle system. If the potentials have nice confining properties such that mass does not escape to infinity [DGPS23], the unique stationary solution of (2.25) is given by the Gibbs measure

$$M_N(\{\mathbf{x}^k\}) = \frac{e^{-\beta \mathcal{H}_N}}{Z_N} \qquad \text{with} \qquad Z_N = \int_{\mathbb{R}^M} \cdots \int_{\mathbb{R}^M} e^{-\beta \mathcal{H}_N} \prod_{k=1}^{N} \mathrm{d}\mathbf{x}^k, \tag{2.26}$$

where Z_N is the partition function of the N-particle system and $\beta = \frac{2}{\sigma^2}$ is the inverse temperature. We remark that, for any finite N, (2.26) is the only solution of the Fokker-Planck equation and that for any initial condition

$$\lim_{t \to +\infty} \rho_N(\{\mathbf{x}^k\}) = M_N\left(\{\mathbf{x}^k\}\right). \tag{2.27}$$

Consequently no phase transition can be observed in finite systems. Transition points for finite dimensional stochastic systems correspond to points where the topological structure of the *unique* invariant measure changes [HL84], [Pav14, Sec. 5.4]. In the classical literature of equilibrium statistical mechanics of lattice systems, phase transitions are identified by the non-uniqueness of the *infinite-volume* grand canonical Gibbs measure [Geo11].
The existence of phase transitions for weakly interacting diffusions can be investigated through the analysis of the stationary solutions of the McKean Vlasov equation (2.12), that, with the assumptions we made at the beginning of this section, can be written as [CCY19, CGPS20, DGPS23]

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta F}{\delta \rho}\right), \tag{2.28}$$

where we have introduced the free energy functional $F[\rho]$ defined as

$$F[\rho] = \int V_\alpha(\mathbf{x})\rho(\mathbf{x})\mathrm{d}\mathbf{x} + \frac{\theta}{2}\int\int \rho(\mathbf{x})\mathcal{U}\left(\mathbf{x} - \mathbf{y}\right)\rho(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} + \beta^{-1}\int \rho(\mathbf{x}) \ln \rho(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{2.29}$$

The above equation provides a meaningful interpretation of the energy budget in the system. The first term represents the internal energy associated to the local potential $V_\alpha(x)$. The second term is the energy originating by the interaction among the agents and, lastly, the third term is the (negative of the) entropy associated to the probability distribution ρ. Remarkably, equation (2.28) belongs to a rich class of dissipative Partial Differential Equations, including the heat equation, the porous medium equation and the diffusion-aggregation equation, that are gradient flows with respect to the Wasserstein metric on the space of probability measure with finite second moment. The Wasserstein metric is a suitable distance function that can be defined between two measures, for more details about Wasserstein gradient flows in the context of interacting particle systems we refer the reader to [CCY19] and references therein. Convexity properties of the free energy $F[\rho]$ are closely related to the existence of phase transitions for the infinite dimensional system. Indeed, we can evaluate the time derivative of $F[\rho]$ along solutions of (2.28) [CCY19, CGPS20]

$$\frac{\mathrm{d}F[\rho]}{\mathrm{d}t} = -\int_{\mathbb{R}^M} \rho(\mathbf{x}) \left(\nabla \frac{\delta F}{\delta \rho}\right)^2 \mathrm{d}\mathbf{x} \leq 0. \tag{2.30}$$

Convexity properties of the free energy functional provide a one-to-one characterisation of the stability properties of the stationary solutions. If there exists only an unique minimiser of the free energy, the dynamics converge exponentially fast, in relative entropy, to the unique stationary state and the rate of convergence to equilibrium can be established [Mal01], see also section 2.4 . Such a situation arises when the confining and interaction potential satisfy suitable (strong) convexity properties [Mal01, Tam84]. In this case the asymptotic limit $t \to +\infty$ and the thermodynamic limit $N \to +\infty$ commute, meaning that the asymptotic dynamics given by (2.25) and (2.28) agree for any initial condition.

However, the minimiser is not necessarily unique and multiple stationary solutions can coexist. In particular, when the interaction potential is a convex polynomial, equation (2.28) has as many stationary solutions, for low enough temperatures $\beta \to +\infty$, as extremal points of the (polynomial) confining potential $V_\alpha(\mathbf{x})$ [Tug14]. For a non convex free energy $F[\rho]$ landscape, the asymptotic limit $t \to +\infty$ of the McKean Vlasov equation depends on the initial condition. The infinite dimensional dynamics (2.28) does not approximate the particle dynamics (2.25) for arbitrary long times, see sections 2.4 and 2.5. Here, the asymptotic limit and the thermodynamics limit do not commute. The non convexity of $F[\rho]$ represents the natural obstacle for the commutativity of such limits. Different results are to be expected depending on the order of the limits, especially when other limiting procedures, such as the homogenisation limit of the

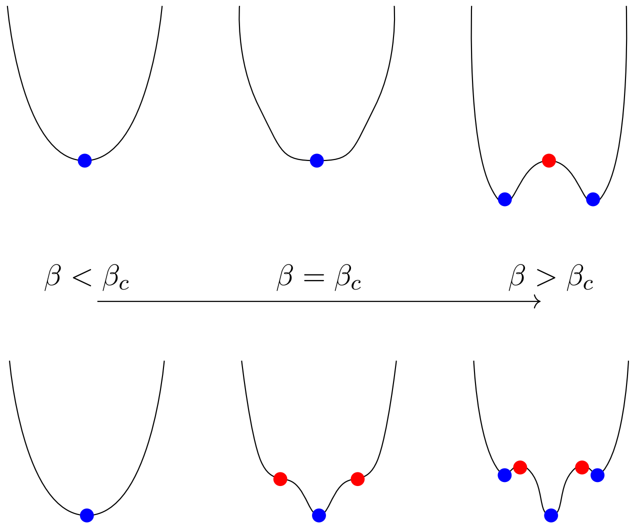


Figure 2.2: A rough schematic showing two possible kinds of phase transition: The upper diagram shows a typical continuous phase transition. In this setting, the unique critical point (shown in blue) loses its local stability through a local (pitchfork) bifurcation which gives rise to new locally stable critical points. The lower diagram shows a typical discontinuous phase transition. In this setting, the unique critical point retains its local stability but new critical points arise in the free energy landscape through a saddle node bifurcation. Figure and caption are (with permission) from [DGPS23].

confining potential [GKPY19] or the approaching of the phase transition point limit [Cha14], are considered.

It is possible to obtain another characterisation of stationary solutions that will be useful for our later analysis. If we write (2.28) explicitly we obtain

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\left(\nabla V_\alpha + \theta \nabla \mathcal{U} \star \rho \right) \rho + \beta^{-1} \Delta \rho \right). \tag{2.31}$$

It is clear from the above equation that any stationary solution $\rho_0(\mathbf{x})$ of the McKean-Vlasov equation satisfies the Kirkwood-Monroe equation [KM41], that is, the self consistency equation

$$\rho_0(\mathbf{x}) = \frac{1}{Z} e^{-\beta(V_\alpha(\mathbf{x}) + \mathcal{U} \star \rho_0)}, \quad Z = \int_{\mathbb{R}^M} e^{-\beta(V_\alpha(\mathbf{x}) + \mathcal{U} \star \rho_0)}. \tag{2.32}$$

The two characterisations of the stationary solutions of the McKean Vlasov equation as extremal points of $F[\rho]$ or solution of the self consistency equation and are in fact equivalent. We refer the reader to [CGPS20] for a thorough investigation of these two approaches for phase transitions on the torus $\mathbb{T}^M$.

We want to conclude this section by providing a brief physical intuition of the origin of phase transitions for interacting systems with non-convex potentials $V_\alpha(\mathbf{x})$ or $\mathcal{U}(\mathbf{x})$. For sufficiently high temperatures, the diffusion is strong enough that the expected escape time from the minima of the potentials is bounded uniformly in the number of particles N. In this case, the mean field free energy $F[\rho]$ is convex and the stationary solution of the McKean Vlasov equation is unique. Indeed, it is possible to show that in such case the self consistency equation (2.32) has a unique solution, see [DGPS23].

On the other hand, for low temperatures the particles can get trapped for arbitrarily long times in the minima of the potentials and *condense* [BM21] . Consequently, the free energy functional $F[\rho]$ has more than one minimisers, corresponding to different stationary solutions of the McKean Vlasov equation.

2.4 Propagation of chaos, fluctuations and critical slowing down

We here investigate in greater detail the convergence properties of the N particle system (2.1) as $N \to +\infty$. The propagation of chaos features presented in section 2.2 indicate that, if we fix a time interval $[0, T]$, then there exists a N_0 such that, when $N > N_0$, the microscopic system (2.1) and the McKean Vlasov equation (2.12) are close for any $t \in [0, T]$. It turns out that the choice of the observation time T is a fundamental issue in the investigation of the dynamics of weakly interacting diffusions. The threshold value N_0 might depend on the timescale T and one might expect to observe a monotonic increasing dependence $N_0 = N_0(T)$. On the other hand, given a fixed and large number of particles N, the dynamics of the finite particle system might start deviating from the mean field limit as the time T increases: the long time behaviour of (2.1) is not reflected by (2.12).

When no deviation from the mean field limit is observed, we say that uniform in time propagation of chaos ($T = +\infty$) is verified. Uniform in time propagation of chaos implies that the stationary measure of the N particle system is close to products of stationary measures of the McKean-Vlasov equation at any time $t \geq 0$. When this fails, one is usually interested in identifying the time scale in which propagation of chaos is approximately valid and then determine the behaviour of the N particle system beyond this time scale.

Fluctuations of the N particle system around the mean field limit can be studied by taking a step further with respect to the *Law of Large Numbers* (the McKean Vlasov equation) and formulate a *Central Limit Theorem* (CLT) or a *Large Deviation Principle* (LDP) [Daw83, DP19]. We here investigate the CLT whereas in section 2.5 we will introduce the LDP. Fluctuations are usually studied by introducing an empirical fluctuation process Y_N, that, in the context of measure-valued processes, can be defined as

$$Y_N = \sqrt{N}\left(X_N(t, \mathrm{d}\mathbf{x}) - \rho\left(\mathbf{x}, t\right) \mathrm{d}\mathbf{x}\right), \tag{2.33}$$

where $\rho(\mathbf{x}, t)$ is the solution of the McKean Vlasov equation. One is usually interested in the asymptotic behaviour $N \to +\infty$ of the fluctuations Y_N. It is possible to prove under remarkably general assumptions that, fixed a time interval $[0, T]$, a *Central Limit Theorem* applies, meaning that the fluctuation process converges weakly to a mean zero Gaussian process, see e.g. [CE88, JM98, PH96, Daw83]. This means that, in the interval $[0, T]$, the empirical measure exhibits gaussian fluctuations around the mean field limit (2.12). As mentioned before, the observation time T plays an important role. Typically, when the mean field limit admits two or more invariant measures, the empirical measure X_N will fluctuate close to one of these invariant states but will then perform, after an N-dependent timescale, a transition to another invariant solution of the McKean Vlasov equation. The scaling (2.33) of the fluctuation process turns out to be valid only in settings far from a phase transition point. At the transition point, the fluctuations get amplified by the collective interaction among the subsystems and become relevant at all macroscopic scales. Due to the strong correlation among the subsystems, such critical fluctuations become of bigger amplitude, persistent, non-Gaussian in time, and they feature *critical slowing down*, namely they are associated with a longer N-dependent timescale [Daw83, CE88].

Phase transitions or, more generally, multistability of the McKean-Vlasov equation (2.12), are intimately related to propagation of chaos properties, fluctuations and asymptotic properties of weakly interacting diffusions. Recently, the authors in [DGPS23] have shown that for weakly interacting diffusions describing equilibrium statistical mechanics systems, the existence of multiple invariant solutions of the McKean-Vlasov equation constitute the natural obstruction to uniform in time propagation of chaos and Gaussianity of fluctuations. Below we briefly review their work as it will allow us to introduce important concepts, such as Logarithmic Sobolev Inequalities, for the quantitative study of the long time limit of the dynamics.

In order to do so, we will assume, as in the previous section, that the local dynamics is given by a confining potential, $\mathbf{F}_\alpha(\mathbf{x}) = -\nabla V_\alpha(\mathbf{x})$, and that the diffusion is given by thermal noise, that is $s_{ij} = \delta_{ij}$. The potentials are required to satisfy a few conditions in order to guarantee the well-posedness of the dynamics. These are quite generic assumptions[3] and the interested reader is referred to [DGPS23].

The goal of this section is to investigate the long time behaviour of the N particle system (2.24), in particular its convergence properties towards asymptotic states of the mean field dynamics (2.31). In section 2.3 we have already observed that when the McKean-Vlasov equation supports multiple stationary solutions the $t \to +\infty$ and $N \to +\infty$ do not commute. We here provide a quantitative analysis of such behaviour.

As previously mentioned, the long time statistical behaviour of the N particle system (2.24) is given by the Gibbs measure M_N defined in (2.26). We apply standard relative entropy techniques [Var91] and define the scaled relative entropy of the N particle distribution ρ_N with respect to the equilibrium Gibbs measure M_N

$$\mathcal{E}\left(\rho_N|M_N\right) \coloneqq \frac{1}{N}\int \ln\left(\frac{\rho_N(\mathbf{x},t)}{M_N(\mathbf{x})}\right)\rho_N(\mathbf{x},t)\mathrm{d}\mathbf{x}, \tag{2.34}$$

where the scaling $\frac{1}{N}$ is relevant when one considers the thermodynamic limit $N \to +\infty$. We now consider the time derivative of the scaled entropy and, by using (2.25), we obtain

$$\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} = -\beta^{-1}\frac{1}{N}\int\left(\nabla \ln\left(\frac{\rho_N(\mathbf{x},t)}{M_N(\mathbf{x})}\right)\right)^2\rho_N(\mathbf{x},t)\mathrm{d}\mathbf{x} \coloneqq -\beta^{-1}\mathcal{I}\left(\rho_N|M_N\right), \tag{2.35}$$

[3]All the examples of equilibrium systems we provide satisfy these assumptions.

where in the last equality we have defined the scaled relative Fisher information $\mathcal{I}(\rho_N|M_N)$. The quantity that determines the convergence towards the Gibbs measure in relative entropy is the N-particle Sobolev constant

$$\lambda_N := \inf_{\rho_N \notin \{M_N\}} \frac{\beta^{-1}\mathcal{I}(\rho_N|M_N)}{\mathcal{E}(\rho_N|M_N)}. \tag{2.36}$$

If the log Sobolev constant is strictly positive, $\lambda_N > 0$, the system converges exponentially in relative entropy towards the stationary state $M_N$for any initial condition. A classical criterion for assessing exponential convergence in relative entropy derives from the Bakry-Emery theory [BGL14, Hel02]: if there exists a radius $R > 1$ and a constant $\lambda > 0$ such that the Hessian of the Hamiltonian function (2.23), $D^2\mathcal{H}_N$, satisfies the following convexity condition

$$D^2\mathcal{H}_N(\mathbf{x}) > \lambda \mathbf{1}_{NM\times NM} \quad \text{for every } |\mathbf{x}| > R, \tag{2.37}$$

where $\mathbf{1}_{NM\times NM}$ is the NM dimensional identity matrix, then $\lambda_N > 0$ and there is exponential convergence to equilibrium. One fundamental question regarding the above convexity condition is whether it is valid uniformly with N, that is whether it still holds for the thermodynamic limit $N \to +\infty$. In this regard, we define the mean field log Sobolev constant λ_∞ in a similar way. In section 2.3 we have shown that, in these equilibrium settings, the McKean Vlasov equation has a gradient structure with respect to a free energy functional $F[\rho]$, see equation (2.28). Moreover, the Free energy functional, when evaluated on solutions of the McKean Vlasov equation, satisfies an equation that is reminiscent of (2.35), see (2.30), and can be written as

$$\frac{\mathrm{d}F[\rho]}{\mathrm{d}t} = -D[\rho] \quad \text{where } D[\rho] := \int_{\mathbb{R}^M} \rho(\mathbf{x},t)\left(\nabla\frac{\delta F}{\delta\rho}\right)^2 \mathrm{d}\mathbf{x}. \tag{2.38}$$

We can then define the mean field log Sobolev constant λ_∞ as

$$\lambda_\infty := \inf_{\rho\notin\mathcal{K}} \frac{D[\rho]}{F[\rho] - \inf F}. \tag{2.39}$$

Since, in general, the McKean Vlasov equation might have multiple stationary solutions at which the infimum of the free energy is attained, the infimum procedure in the definition of mean field log Sobolev equation is not taken with respect to such stationary solutions, that is we have defined

$$\mathcal{K} = \{\rho : F[\rho] = \inf F\}. \tag{2.40}$$

If $\lambda_\infty > 0$, there is exponential convergence in terms of free energy functional of the mean field system towards its stationary solution. It is of fundamental relevance to understand the relationship between the N-particle log Sobolev constant λ_N and its mean field equivalent λ_∞. The authors in [DGPS23] provide strong indication that their conjecture

$$\lim_{N\to+\infty} \lambda_N \stackrel{?}{=} \lambda_\infty \tag{2.41}$$

could indeed be true but have not given a direct proof of it. They can however show that

$$\limsup_{N\to+\infty} \lambda_N \leq \lambda_\infty. \tag{2.42}$$

Furthermore, they are able to fully characterise the consequences of the (non) degeneracy of the log Sobolev constant λ_N in the thermodynamic limit in terms of uniform in time propagation of chaos and fluctuations around the mean field limit.
In particular, they show that if[4] $\liminf_{N\to+\infty} \lambda_N > 0$, then

$$\bar{\mathrm{d}}_2 \left(\rho_N \left(\{\mathbf{x}^k\}, t\right), \prod_{k=1}^{N} \rho\left(\mathbf{x}^k, t\right) \right) \leq \frac{C}{N^\gamma} \quad \text{for all } t > 0, \tag{2.43}$$

where $\bar{\mathrm{d}}_2 = \frac{1}{\sqrt{N}}\mathrm{d}_2$ is a suitably scaled 2-Wasserstein distance between probability measures, see [DGP21, CCY19]. Moreover, $C > 0$ is a constant and $\gamma > 0$ is an exponent that depends on the properties of the system. The above result constitutes a quantitative version of the uniform propagation of chaos property for weakly interacting diffusions. It quantifies the degree of "closeness" of the N-particle distribution $\rho_N\left(\{\mathbf{x}^k\}, t\right)$ to the product of mean field solution $\prod_{k=1}^{N} \rho\left(\mathbf{x}^k, t\right)$ of equation (2.28). In particular, it states that if the N-particle log Sobolev constant does not degenerate in the thermodynamic limit, uniform in time propagation of chaos holds. We remark that the assumption $\liminf_{N\to+\infty} \lambda_N > 0$ implies that $\lambda_\infty > 0$. Indeed, we

[4]There are other "technical" hypotheses for the proof of this result, we refer the reader to [DGPS23]

can write that

$$\liminf_{N\to+\infty} \lambda_N \leq \limsup_{N\to+\infty} \lambda_N \leq \lambda_\infty, \tag{2.44}$$

where the first inequality derives from the definition of the limiting procedures and the second one derives from (2.42). This implies that the mean field dynamics admits only one unique stationary solution, see later discussion. The above result is compatible with the exponential converge in relative entropy obtained in [Mal01]. Another consequence of the non degeneracy of the N-particle log Sobolev constant is that the fluctuation process Y_N, defined in (2.33), is in the thermodynamic limit a centred Gaussian measure, meaning that the empirical measure X_N exhibits gaussian fluctuations around the unique solution of the mean field dynamics at all times $t > 0$.

All the above results do not hold when the free energy functional is not convex. If $F[\rho]$ admits a critical point $\rho^*(\mathbf{x})$ such that $\frac{\delta F[\rho]}{\delta \rho}|_{\rho=\rho^*} = 0$ and that is not a minimiser (see the red dots in Figure 2.1), the authors in [DGPS23] have shown that the mean field log Sobolev constant is zero $\lambda_\infty = 0$ and the N-particle log Sobolev constant degenerates in the thermodynamic limit, that is, $\lim_{N\to+\infty} \lambda_N = 0$. It is possible to find an upper bound for λ_N as

$$\lambda_N \leq \frac{C}{N}, \tag{2.45}$$

where $C > 0$ is a constant. The mean field log Sobolev constant λ_∞ captures global features of the free energy landscape, that is, it is not related to a specific stationary solution of the mean field dynamics. By inspecting Figure 2.1 in the low temperatures regime, $\beta > \beta_c$, it is easy to see that the $\lambda_\infty = 0$. As a matter of fact, the red dots correspond to measures ρ^* such that the numerator in the definition of λ_∞vanishes, $D[\rho^*] = 0$, while the denominator remains positive. Such behaviour is observed at the phase transition $\beta = \beta_c$ too. The previous argument still holds for a discontinuous phase transition since the non minimising critical points still exist, see bottom panel of Figure 2.1. The reason why the mean field log Sobolev constant vanishes in a continuous phase transition is more subtle, since, at $\beta = \beta_c$, there is only one critical point of the free energy functional. It is however possible to show that the condition $\lambda_\infty = 0$ at the phase transition corresponds to a loss of stability of this solution [DGPS23].

The above results are in agreement with the theory of relaxation phenomena for spin systems [Yos03]. In this context, one distinguishes between two regimes. The high temperature regime features a disordered state where thermal fluctuations dominate and the spins point up or down in an erratic fashion almost independently of each other. This state is characterised by an unique phase of the system. On the other hand, in the low temperature regime the thermal fluctuations are not as strong and the interaction among the spins tend to create macroscopic areas of spins pointing in the same direction. As a result, one can identify in the thermodynamic limit two equilibrium states, usually called pure states. In one of the pure states a great part of the spins point up, whereas in the other pure state the reverse is attained. Such pure states can be mathematically represented by infinite volume limit of Gibbs measures. The low temperature regime is also known as phase coexistence since the systems supports more than one pure states. It is a fundamental result of the theory of relaxation phenomena that the thermodynamic limit of the log Sobolev constant for the spin system vanishes [Yos03]. This behaviour is associated with *slow relaxation* of the dynamics, meaning that the time it takes to flip a macroscopic area of spins pointing in the same direction diverges in the thermodynamic limit. In the context of weakly interacting diffusions we can identify the pure states as the solutions of the McKean Vlasov equation (2.28). A phase coexistence regime is attained when multiple stationary solutions are supported, or, for equilibrium statistical mechanics systems, when the free energy is non convex. In the next section a Large Deviation approach will be employed to explain how transitions among such pure states happen on a timescale that scales exponentially with the number of the particles N and thus diverges in the thermodynamic limit. We also remark that at the phase transition point the log Sobolev constant vanishes because of the development of critical scaling for the fluctuation process Y_N, see discussion below equation (2.33). It is still an open research question to what extent the fluctuations given by Y_N exhibit universality properties at a (continuous) phase transition.

2.5 Large Deviations: finite size metastability

In the previous chapters a *Law of Large Numbers* for the empirical measure X_N of the system has been established in terms of the mean field dynamics given by the McKean Vlasov equation (2.12) as a result of propagation of chaos properties. In the previous section, fluctuations of the empirical measure around the mean field dynamics have been established through a *Central Limit Theorem* result for the fluctuation process Y_N. If the mean field dynamics supports a unique invariant measure, the long time asymptotic state of the N particle system is well approximated by such mean field invariant solution. Loosely speaking, the empirical measure fluctuates closely to the mean field dynamics for any time $t > 0$.

In this chapter we will adopt a Large Deviation perspective to investigate the behaviour of the system when the mean field dynamics support more than one invariant measure. In this situation, uniform in time propagation of chaos does not hold and one does not expect the empirical measure X_N to follow the invariant solutions for all times. In particular, for a finite dimensional setting one expects in typical situations, such as equilibrium statistical mechanical systems, that the system exhibits ergodic properties. On the other hand, the existence of two different invariant solutions for the mean field dynamics show that in the thermodynamic limit ergodicity is broken. The goal of *Large Deviation Theory* for weakly interacting diffusions is to study the behaviour of the empirical measure for a finite, but big, number of particles N. In this regime, one expects the N particle system to exhibit a separation of timescale and tunnelling phenomena. The empirical measure X_N will fluctuate around one of the mean field invariant solution for a long, N-dependent, time but then will transition on a faster timescale close to another mean field invariant solution through a large deviation.

Large deviations due to random perturbations for finite dimensional stochastic systems can be studied through the Freidlin and Wentzell theory [FW84, Tou18]. The quasipotential, or non equilibrium potential, framework [GHT91, HTG94, ZAAH12, ZL16] has been a valuable too to investigate noise induced transitions in high dimensional multistable non equilibrium systems, with numerous applications to natural systems including the climate system [LB17, LB20]. Most often these large deviation results are obtained in the low noise $\sigma \to 0$ regime where the concept of determinstic attractors is still applicable. In particular, escape processes from the deterministic attractors of the systems are found to obey Arrhenius-like formulas [FW84, Tou18] such that the mean escape time from an attractor scales exponentially with the inverse of (the square of) the noise intensity σ. For simple settings of the deterministic attractors such as fixed points or limit cycles it is possible to obtain sharper formulas for the expected escape times that includes sub exponential corrections [BEGK04, BR16]. We also remark that the connection between the quasipotential and the transition statistics is non-trivial if more than two deterministic attractors are present [MGLL21].

The transition statistics for weakly interacting diffusions depend both on the number of particles N and the strength of the noise σ. The low noise regime has been investigated in [BM21, BFG07a, BFG07b, BBM10]. On one hand, sharp results for the expected escape times, including uniform control on the number of particles N and sub exponential corrections, can be obtained in these settings. On the other hand, a low noise regime might not be appropriate for the investigation of metastability properties close to a phase transition, where such features, see later discussion, become very relevant.

In the following we will present a large deviation approach in the thermodynamic regime $N \to +\infty$ for a generic value of strength of the noise σ. A mathematically rigorous large deviation theory has been developed by Dawson and Gartner [DG87b, DG87a]: their theory encompasses both equilibrium and nonequilibrium systems with multiplicative, state dependent diffusion matrix. This large deviation theory represents an infinite dimensional generalisation of the aforementioned Freidlin-Wentzell theory. This analogy can be made stronger as one notices that the large deviation theory for weakly interacting diffusions in the limit $N \to +\infty$ can be interpreted as a *weak noise limit* of a diffusive Markov process in the space of probability measures on $\mathbb{R}^M$, rather than on the N-particle phase space $\mathbb{R}^{M\times N}$ [DG87b]. Loosely speaking, this large deviation theory provides the probability, in the thermodynamic limit, that the empirical measure X_N at any time t is close to a given measure $\mu(t)$ on $\mathbb{R}^M$. Below we follow [BGN16] where the authors of the paper develop a simpler and less rigorous approach to the development of a large deviation principle. On one side, the calculations below should be interpreted in a formal sense, since the mathematics of some of the formulas is still uncertain or known to be problematic in most cases. On the other side, their approach provides a more intuitive and physical interpretation of the theory and the loss of mathematical rigour can be justified by the fact that their conclusions are in agreement with Dawson and Gartner's theory. We assume here that the diffusion matrix is constant (thermal noise) so that the equations of motions of the N particle system are

$$\mathrm{d}\mathbf{x}^k = \mathbf{F}_\alpha(\mathbf{x}^k)\mathrm{d}t - \frac{\theta}{N}\sum_{l=1}^{N} \nabla\mathcal{U}\left(\mathbf{x}^k - \mathbf{x}^l\right)\mathrm{d}t + \sigma \mathrm{d}\mathbf{W}_k, \tag{2.46}$$

where $k = 1, \ldots, N$. Following [BGN16] we define the empirical density

$$\rho^N(\mathbf{x}, t) = \frac{1}{N}\sum_{k=1}^{N} \delta\left(\mathbf{x} - \mathbf{x}^k\left(t\right)\right), \tag{2.47}$$

where $\delta\left(\mathbf{x}\right)$ represents the Dirac-δ measure. We remark that the above definition is analogous to the definition of the empirical measure X_N, with the caveat that the above definition refers to probability densities rather than measures. By using Ito's formula and (2.46) it is possible to find an equation for the empirical density

$$\frac{\partial \rho^N}{\partial t} = -\nabla\cdot\left(\left(\mathbf{F}\left(\mathbf{x}\right) - \theta\nabla\mathcal{U}\star\rho^N\right)\rho^N\right) + \frac{\sigma^2}{2}\Delta\rho^N - \frac{\sigma}{N}\nabla\cdot\sum_{k=1}^{N}\delta\left(\mathbf{x} - \mathbf{x}^k(t)\right)\dot{\mathbf{W}}_k. \tag{2.48}$$

To conform to the notation in [BGN16] we have written the above stochastic equation in a non rigorous way by introducing the white noise in time $\dot{\mathbf{W}}_k$, with vanishing mean and covariance $\mathbb{E}\left(\dot{W}_{i,k}(t)\dot{W}_{j,k'}(t')\right) = \delta\left(t - t'\right)\delta_{ij}\delta_{kk'}$ We observe that the above equation is exact since no approximations have been made starting from the equations of motions. In fact, it is not a closed equation for the empirical density since it still contains the microscopic information of the state $\mathbf{x}^k$ of all the particles through the last noisy term. We also remark that, in the limit $N \to +\infty$, one might naively think that this last term vanishes. With this assumption the above equation becomes the McKean Vlasov equation we have introduced in section 2.2. Of course, a stronger analysis is provided in this section in terms of propagation of chaos properties and the *Law of Large Numbers* for the empirical measure X_N. Dean [Dea96] proposed a way of closing the above equation for the empirical density. If we consider the quantity

$$\frac{\sigma}{N}\sum_{k=1}^{N}\delta\left(\mathbf{x}-\mathbf{x}^k\right)\dot{W}_{i,k}, \tag{2.49}$$

as white noise in time with values in vector fields in $\mathbb{R}^M$ parametrised by the state of the particles $\mathbf{x}^k$, we can evaluate its covariance (the mean is zero by definition) as

$$\begin{aligned}&\mathbb{E}\left(\frac{\sigma}{N}\sum_{k=1}^{N}\delta\left(\mathbf{x}-\mathbf{x}^k\right)\dot{W}_{i,k}(t)\frac{\sigma}{N}\sum_{k'=1}^{N}\delta\left(\mathbf{y}-\mathbf{x}^{k'}\right)\dot{W}_{j,k'}(t')\right)=\\&=\frac{\sigma^2}{N^2}\delta_{ij}\delta(t-t')\sum_{k=1}^{N}\delta(\mathbf{x}-\mathbf{x}^k)\delta(\mathbf{y}-\mathbf{x}^k)=\\&=\frac{\sigma^2}{N}\delta_{ij}\delta(t-t')\delta(\mathbf{x}-\mathbf{y})\rho^N(\mathbf{x},t),\end{aligned} \tag{2.50}$$

where in the last equality we have used the distributional identity $\delta(\mathbf{x}-\mathbf{x}^k)\delta(\mathbf{y}-\mathbf{x}^k)=\delta(\mathbf{x}-\mathbf{y})\delta(\mathbf{x}-\mathbf{x}^k)$ and we have introduced the empirical density (2.47).

Now, we consider a different noise

$$\frac{\sigma}{\sqrt{N}}\sqrt{\rho(\mathbf{x})}\xi_i(\mathbf{x},t), \tag{2.51}$$

where $\rho(\mathbf{x})$ is an arbitrary density and $\xi_i(\mathbf{x},t)$ is space time white noise with vanishing mean and covariance $\mathbb{E}\left(\xi_i(\mathbf{x},t)\xi_j(\mathbf{y},t')\right)=\delta_{ij}\delta(t-t')\delta(\mathbf{x}-\mathbf{y})$. The newly introduce noisy process can be viewed as white noise in time with values in vector fields in $\mathbb{R}^M$ parametrised by the density $\rho(\mathbf{x})$ rather than by states of all the particles. Its covariance is

$$\mathbb{E}\left(\frac{\sigma}{\sqrt{N}}\sqrt{\rho(\mathbf{x})}\xi_i(\mathbf{x},t)\frac{\sigma}{\sqrt{N}}\sqrt{\rho(\mathbf{y})}\xi_j(\mathbf{y},t')\right)=\frac{\sigma^2}{N}\delta_{ij}\delta(t-t')\delta(\mathbf{x}-\mathbf{y})\rho(\mathbf{x}), \tag{2.52}$$

which coincides with (2.50) if $\rho(\mathbf{x})=\rho^N$. This observation led Dean to identify the two noises (2.49) and (2.51) and thus write a closed equation for the empirical density as

$$\begin{aligned}\frac{\partial\rho}{\partial t}&=-\nabla\cdot\left(\left(\mathbf{F}\left(\mathbf{x}\right)-\theta\nabla\mathcal{U}\star\rho\right)\rho\right)+\frac{\sigma^2}{2}\Delta\rho-\frac{\sigma}{\sqrt{N}}\nabla\cdot\left(\sqrt{\rho}\boldsymbol{\xi}(\mathbf{x},t)\right)=\\&:=\mathcal{L}_\rho\rho+\sqrt{\frac{2}{N}}\eta[\rho],\end{aligned} \tag{2.53}$$

where we dropped the apex from the empirical density, introduced the non linear operator $\mathcal{L}_\rho$ defined in (2.16) and written the white noise in time $\eta[\rho](\mathbf{x},t)$, parametrised by the density ρ,

$$\eta[\rho](\mathbf{x},t) = -\frac{\sigma}{\sqrt{2}}\nabla\cdot\left(\sqrt{\rho}\boldsymbol{\xi}(\mathbf{x},t)\right), \tag{2.54}$$

with vanishing mean and covariance

$$\mathbb{E}\left(\eta[\rho](\mathbf{x},t)\eta[\rho](\mathbf{y},t')\right) = \delta(t-t')\mathcal{C}[\rho](\mathbf{x},\mathbf{y}), \tag{2.55}$$

where

$$\mathcal{C}[\rho](\mathbf{x},\mathbf{y}) = \frac{\sigma^2}{2}\nabla_{\mathbf{x}}\cdot\nabla_{\mathbf{y}}\left(\rho\left(\mathbf{x}\right)\delta\left(\mathbf{x}-\mathbf{y}\right)\right). \tag{2.56}$$

The above stochastic partial differential equation is usually called Dean equation and is assumed to model macroscopic fluctuations in systems with interacting agents [DVE14, AR04, CSZ20, KLvR20], see also the recent review on Macroscopic Fluctuation Theory [BDSG$^+$15]. As mentioned at the beginning of the chapter, its mathematical status is unclear and the extensive literature theory for stochastic partial differential equations with white noise in time and space does not apply to the Dean equation. Not only it is difficult to give a mathematical meaning of the terms in the Dean equation in a function space that includes the empirical densities of weakly interacting diffusions, but also it is even harder to formulate a theory of solutions for such stochastic partial differential equation [BGN16]. Nevertheless, the Dean equation is thought to be a good representation of fluctuations of the empirical density ρ^N around the solutions of the McKean-Vlasov equation in the thermodynamic limit where the noisy term, proportional to the inverse of (the square root of) the number of particles, can be considered small. The physical insight contained in the Dean equation is that fluctuations, given the multiplicative nature of the noise proportional to $\sqrt{\rho}$, vanish in regions of the phase space where no particles are present. We remark that such equation is not expected to be a viable approach to describe the critical fluctuations at the phase transition, since the classical central limit theorem scaling $\propto \frac{1}{\sqrt{N}}$ is not expected to hold. Indeed, as explained in section 2.4, critical fluctuations are characterised by a bigger amplitude and by a *critical slowing down* phenomenon, that is their typical evolution timescale is an increasing function of the number of particles N.

We can associate to the Dean Equation a Fokker Planck equation for the probability distribution functional $\mathrm{P}[\rho]$ expressing, loosely speaking, the probability that the empirical density $\rho^N(\mathbf{x},t)$ is close to an arbitrary path $\rho(\mathbf{x},t)$ for any time $t_i \leq t \leq t_f$, where $t_{i,f}$ are two fixed times. Equivalently, one might interpret $\mathrm{P}[\rho]$ as the probability distributional functional of the empirical densities in the time period $[t_i, t_f]$. The Fokker-Planck equation associated to (2.53) is

$$\frac{\partial \mathrm{P}}{\partial t} = \mathcal{F}\mathrm{P}, \tag{2.57}$$

where $\mathcal{F} = \mathcal{G}^\dagger$ is the adjoint of the generator $\mathcal{G}$ of the Dean Equation, see equation below, and the adjoint is evaluated with the rule $\left(\frac{\delta}{\delta\rho}\right)^\dagger = -\frac{\delta}{\delta\rho}$

$$\mathcal{G}\psi[\rho] = \int \frac{\delta\psi[\rho]}{\delta\rho(\mathbf{x})}\mathcal{L}_\rho\rho \mathrm{d}\mathbf{x} + \frac{1}{N}\int\int \frac{\delta^2\psi[\rho]}{\delta\rho(\mathbf{x})\rho(\mathbf{y})}\mathcal{C}[\rho](\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}, \tag{2.58}$$

and $\psi[\rho]$ is a test functional of ρ. The main goal of a large deviation theory for weakly interacting diffusions is to find a large deviation principle for the probability distribution functional $\mathrm{P}[\rho]$ such that

$$\mathrm{P}[\rho] \asymp e^{-N\mathcal{A}[\rho]}. \tag{2.59}$$

The quantity $\mathcal{A}[\rho]$ is called rate (or action) functional and it gives the exponential weight to the probability $\mathrm{P}[\rho]$ of observing the empirical density close to a measure path ρ in the time interval $[t_i, t_f]$. In particular, the symbol $\asymp$ in equation (2.59) should be interpreted as the asymptotic relation

$$\mathcal{A}[\rho] = -\lim_{N\to+\infty}\frac{1}{N}\ln\left(P[\rho]\right). \tag{2.60}$$

In other words, the symbol $\asymp$ is used to stress the concept that in the limit $N \to +\infty$ the dominant behaviour of the probability $P[\rho]$ is given by the decaying exponential $e^{-N\mathcal{A}[\rho]}$. Alternatively, one might think of the symbol $\asymp$ as an equality relationship in logarithmic scale, that is

$$a_N \asymp b_N \iff \lim_{N\to+\infty}\frac{1}{N}\ln a_N = \lim_{N\to+\infty}\frac{1}{N}\ln b_N. \tag{2.61}$$

This theory only captures the dominant exponential behaviour. As a result, sub exponential corrections in N cannot be obtained through a large deviation perspective. Nevertheless, a large deviation principle as in (2.59) provides a suitable framework to study both small and large fluctuations of the empirical densities, including tunnelling phenomena between, if present, different invariant solutions of the mean field dynamics (2.12). Following [BGN16, JP79] it is possible to use the Martin-Siggia-Rose [MSR73] formalism to find a large deviation principle starting from (2.57). The result of such procedure provides a rate function

$$\mathcal{A}[\rho] = \min_{u \in \mathrm{U}} \frac{\left(\int_{t_i}^{t_f} \mathrm{d}t \int \mathrm{d}\mathbf{x} \left(\partial_t \rho - \mathcal{L}_\rho \rho\right) u(\mathbf{x}, t)\right)^2}{2\sigma^2 \int_{t_i}^{t_f} \mathrm{d}t \int \mathrm{d}\mathbf{x} \rho \left(\nabla u\right)^2}, \tag{2.62}$$

where $\mathrm{U} = \{u(\mathbf{x}, t), t \in [t_i, t_f] : \int u(\mathbf{x}, t)\mathrm{d}\mathbf{x} = 0\}$ and we have defined for simplicity $\partial_t \rho = \frac{\partial \rho}{\partial t}$. The above formula for the rate function agrees with what has been obtained by Dawson and Gartner [DG87b, DG87a] in the specific case of additive noise, where the diffusion matrix is constant.

The minimum procedure in the above formula for the rate function can be explicitly carried out and one obtains an expression for $\mathcal{A}[\rho]$ that is analogous to the Freidlin-Wentzell rate function

$$\mathcal{A}[\rho] = \frac{1}{4} \int_{t_i}^{t_f} \mathrm{d}t \int \int \left(\partial_t \rho - \mathcal{L}_\rho \rho\right)(\mathbf{x}, t) \mathcal{C}^{-1}[\rho](\mathbf{x}, \mathbf{y}) \left(\partial_t \rho - \mathcal{L}_\rho \rho\right)(\mathbf{y}, t) \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}. \tag{2.63}$$

The above expression provides a meaningful interpretation of the action functional. In fact, we observe that the action functional vanishes on solutions of the mean field dynamics, that is of the McKean Vlasov equation $\partial_t \rho = \mathcal{L}_\rho \rho$. The actional functional thus measures the difficulty for an empirical density to deviate from its mean field behaviour due to finite size fluctuations.

2.5.1 Quasipotential: Escape from a mean field attractor

In order to study noise induced escape processes from determistic attractors a dynamical quantity called quasi potential is introduced in the Freidlin Wentzell theory [FW84]. It turns out that it is possible, at least on a formal level, to define such concepts in this infinite dimensional setting. If we denote $\mathrm{P}_\infty[\rho]$ the stationary solution of the functional Fokker Planck equation (2.57), we define the *global* quasipotential $\mathcal{Q}[\rho]$ as

$$\mathrm{P}_\infty[\rho] \asymp e^{-N\mathcal{Q}[\rho]}. \tag{2.64}$$

The *global* quasipotential represents a generalisation of an energy landscape for this infinite dimensional, stochastic and non equilibrium settings. If we consider the ansatz $P[\rho] \approx e^{-N\mathcal{Q}[\rho]}$ in the functional Fokker Planck equation (2.57), it is possible to see that $\mathcal{Q}$ satisfies the following Hamilton-Jacobi equations

$$\int \frac{\delta \mathcal{Q}}{\delta \rho(\mathbf{x})} \mathcal{L}_\rho \rho \mathrm{d}\mathbf{x} + \int \int \frac{\delta \mathcal{Q}}{\delta \rho(\mathbf{x})} \mathcal{C}[\rho](\mathbf{x}, \mathbf{y}) \frac{\delta \mathcal{Q}}{\delta \rho(\mathbf{y})} \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} = 0. \tag{2.65}$$

It is possible to define a local notion of quasipotential. We assume that at time $t_i = 0$ the empirical density is equals to $\rho_0(\mathbf{x})$, an invariant solution of the mean field dynamics. We define the quasipotential relative to ρ_0 as

$$\mathcal{Q}_{\rho_0}[\rho] = \min_{\{\hat{\rho} | \hat{\rho}(\mathbf{x},0) = \rho_0, \hat{\rho}(\mathbf{x},t_f) = \rho\}} \mathcal{A}[\hat{\rho}]. \tag{2.66}$$

The meaning of the above definition is easily understood if we consider the transition probability from a state $\rho_0(\mathbf{x})$ at time $t_i = 0$ to a state $\rho(\mathbf{x})$ at time t_f in a path integral formulation [BGN16]

$$\mathrm{P}[\rho_0, t_i; \rho, t_f] \asymp \frac{1}{\mathcal{Z}} \int e^{-N\mathcal{A}[\hat{\rho}]} \mathcal{D}[\hat{\rho}], \tag{2.67}$$

over all paths $\hat{\rho}(\mathbf{x}, t)$ satisfying the boundary conditions given in equation (2.66). A simple saddle node argument shows that most of the contribution to the exponential decaying term in the above equation is given for $N \to +\infty$ by settings where $\mathcal{A} \approx \mathcal{Q}_{\rho_0}[\rho]$, that is

$$\mathrm{P}[\rho_0, t_i; \rho, t_f] \asymp e^{-N\mathcal{Q}_{\rho_0}[\rho]}. \tag{2.68}$$

Let us now consider a simple situation where the mean field dynamics admits only one invariant solution $\rho_0(\mathbf{x})$. We remark that in this setting, away from a phase transition, we expect uniform in time propagation of chaos and gaussian fluctuations. In this setting the quasipotential $\mathcal{Q}_{\rho_0}[\rho]$ coincides with the global one $\mathcal{Q}$. In fact, given the assumption on the uniqueness of the stationary state $\lim_{t_f\to+\infty} \mathrm{P}[\rho_0, t_i; \rho, t_f] = P_\infty[\rho]$ regardless of the initial condition at $t_i = 0$ and from (2.68) we have that $\mathcal{Q}_{\rho_0}[\rho] = \mathcal{Q}$. In the more interesting case where the mean field dynamics supports more than one invariant solutions ρ_i, the above is no longer true because one has to consider the more complicated layout of the basins of attractions of each ρ_i. In principle, one is still able to construct the global quasipotential $\mathcal{Q}$ in terms of the local quasipotentials $\mathcal{Q}_{\rho_i}[\cdot]$ but the complexity of the procedure highly increases [MGLL21]. It is interesting to observe that the quasipotential $\mathcal{Q}_{\rho_0}[\cdot]$ can be associated to a generalisation of the instanton or fluctuation dynamics of the Freidlin Wentzell theory. In particular, the minimisation procedure in (2.66) is attained for a measure $\hat{\rho}$ that satisfies the instanton dynamics [BGN16]

$$\frac{\partial \rho}{\partial t} = \mathcal{L}_\rho^{inst} \rho, \tag{2.69}$$

where

$$\mathcal{L}_\rho^{inst} \rho = \mathcal{L}_\rho \rho + 2 \int \mathcal{Q}[\rho](\mathbf{x}, \mathbf{y}) \frac{\delta \mathcal{Q}}{\delta \rho(\mathbf{y})} \mathrm{d}\mathbf{y}. \tag{2.70}$$

We remark that in nonequilibrium settings, the instanton dynamics differs in general from the (time reversed) mean field dynamics $\partial_t \rho = \mathcal{L}_\rho \rho$, that, in this large deviation framework, could be also be called relaxation dynamics.

Lifetime of a mean field attractor in finite systems

We now turn to the investigation of tunnelling phenomena between invariant solutions of the mean field dynamics. Let us consider the simple situation where there exist only two stable fixed-point-like invariant solutions $\rho_0(\mathbf{x})$ and $\rho_1(\mathbf{x})$ and an unstable saddle node $\bar{\rho}(\mathbf{x})$. We assume that the system is at time $t_{in} = 0$ in the basin of attraction $\mathcal{D}_0$ of $\rho_0(\mathbf{x})$, where the basin of attraction is taken with respect to the relaxation dynamics $\partial_t \rho = \mathcal{L}_\rho \rho$. Given a solution ρ of the Dean equation (2.53) we define the first exit time from $\mathcal{D}_0$ as

$$\tau_0[\rho] = \inf\{t \geq 0 : \rho(\mathbf{x}, t) \notin \mathcal{D}_0\}. \tag{2.71}$$

We remark that this setting is the infinite dimensional analogous to the classical set up in non equilibrium statistical mechanics of a noise induced escape process from a fixed point attractor of the deterministic dynamics.

However, given the above formal results, one might expect that the usual large deviation results apply. In particular, given the equation for the transition probability (2.68) , one can use standard arguments [FW84, BR16] to argue that a large deviation principle holds for the exit time

$$\lim_{N\to+\infty} \mathrm{P}\left(e^{N(\mathcal{Q}^*-\delta)} < \tau_0[\rho] < e^{N(\mathcal{Q}^*+\delta)}\right) = 1, \tag{2.72}$$

where $\delta > 0$ is an arbitrary constant and the probability $\mathrm{P}(\cdot)$ is given by equation (2.57). The quantity $\mathcal{Q}^*$ is the analogous of the lowest quasi potential height and is defined by

$$\mathcal{Q}^* = \inf_{\rho\in D_0} \inf_{t_{fin}\geq 0} \mathcal{Q}_{\rho_0}[\rho]. \tag{2.73}$$

Equation (2.72) states that the most likely exit time scales in the asymptotic limit as $\asymp e^{N\mathcal{Q}^*}$. Clearly, a Kramer's like formula follows for the expected exit time as

$$\mathbb{E}[\tau_0[\rho]] \asymp e^{N\mathcal{Q}^*}, \tag{2.74}$$

where, again, the expectation is taken with respect to the probability P. The above formula shows that the leading behaviour of the mean exit time depends exponentially both on the height difference of the quasi potential and the number of particles N.

Escape from an *equilibrium* mean field attractor

Dawson and Gartner [DG87a, DG89] have established an important link between dynamic properties, such as the quasipotentials, and static properties for a class of equilibrium statistical mechanics systems. As described in section 2.3, in such system it is possible to identify a free energy functional $F[\rho]$ whose convexity properties define the long time behaviour of the mean field dynamics. With the assumptions we have made in this section, the free energy functional $F[\rho]$ would have two minima, not necessarily symmetric, and a maximum, see for reference

Figure 2.2.

Dawson and Gartner proved [DG87a, DG89] that the minimisation procedure in (2.66) is attained for the path $\hat{\rho}$ that corresponds to the time reversal path of the relaxation dynamics $\partial\rho = \mathcal{L}_\rho \rho$ with initial condition ρ_0 and end point ρ. This corresponds to the generalisation of the well known property that relaxation and fluctuation paths are the same, up to time reversal, for finite dimensional equilibrium systems [Tou18]. Moreover, the quasipotential $\mathcal{Q}_{\rho_0}[\rho]$ can be written in terms of the free energy functional

$$\mathcal{Q}_{\rho_0}[\rho] = \beta \left(F[\rho] - F[\rho_0] \right), \tag{2.75}$$

where $\beta = \frac{2}{\sigma^2}$ is the inverse temperature and $\rho \in \mathcal{D}_0$. The quasipotential $\mathcal{Q}^*$describing the expected time of escape from $\mathcal{D}_0$ is thus given by

$$\mathcal{Q}^* = \beta \Delta F \coloneqq \beta \left(F[\bar{\rho}] - F[\rho_0] \right). \tag{2.76}$$

The expected escape time in equilibrium systems scales as $\mathbb{E}[\tau_0] \asymp e^{\beta N \Delta F}$, where ΔF represents the height barrier between the stable mean field attractor ρ_0 and the saddle point $\bar{\rho}$.

The above equation shows that metastability and tunnelling phenomena become relevant for settings near a transition point. In fact, as the transition is approached the free energy barrier decreases, see Figure 2.2, making the transitions exponentially more likely to happen.

2.6 The Desai-Zwanzig model: a paradigmatic example

The Desai-Zwanzig (DZ) model, introduced in [DZ78], has been used as a paradigmatic example of an equilibrium order-disorder (continuous) phase transition reminiscent of the Ising-like ferromagnetic transitions in spin systems [Daw83, Shi87]. More recently, it has also been employed for new purposes such as applications to systemic risk [GPY13] and system size stochastic resonance [PZ03, PZdlC02]. The DZ model describes an ensemble of one-dimensional subsystems $x^k \in \mathbb{R}$ whose dynamics is prescribed by the following equations

$$\begin{aligned} \mathrm{d}x^k &= -V'_\alpha(x^k)\mathrm{d}t - \frac{\theta}{N}\sum_{l=1}^{N}\left(x^k - x^l\right)\mathrm{d}t + \sigma \mathrm{d}W_k \\ &= -V'_\alpha(x^k)\mathrm{d}t - \theta\left(x^k - \bar{x}\right)\mathrm{d}t + \sigma \mathrm{d}W_k, \end{aligned} \tag{2.77}$$

where $k = 1, \ldots, N$. The confining potential $V_\alpha(x) = -\frac{\alpha}{2}x^2 + \frac{x^4}{4}$ is non convex, with a double well shape, for $\alpha > 0$. The diffusion matrix is constant, resulting in thermal fluctuations for the system. The interaction potential is given by a convex, quadratic function $\mathcal{U}(x) = \frac{x^2}{2}$, leading to a coupling among the agents that attempts to synchronise them by nudging them towards their centre of mass $\bar{x} = \frac{1}{N}\sum_{k=1}^{N} x^k(t)$. We observe that the centre of mass can be interpreted as the expectation over the empirical measure X_N of the observable $x \in \mathbb{R}$. Indeed

$$\int x X_N(t, \mathrm{d}x) = \int x\rho_N(t, x)\mathrm{d}x = \frac{1}{N}\sum_{k=1}^{N}\int x\delta(x - x^k(t))\mathrm{d}x = \frac{1}{N}\sum_{k=1}^{N} x^k(t) = \bar{x}. \tag{2.78}$$

In the absence of coupling, $\theta = 0$, the above equations describe the simple motion of a particle in a double well potential, subject to additive noise hopping around, for a sufficiently low σ, the two deterministic stable fixed points $x^* = \pm\sqrt{\alpha}$.

The presence of the coupling allows for a long range coordination of the system that in the thermodynamic limit $N \to +\infty$ results in a proper phase transition. In this regime, by varying the parameters (α, θ, σ), the system undergoes a continuous phase transition, similar to the pitchfork bifurcation diagram for the Ising model. In order to study the transition, we observe that the *Law of Large Numbers* for the empirical measure X_N results in the following McKean Vlasov equation for the one particle distribution

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\left(V'_\alpha(x) + \theta\left(x - \langle x\rangle\right)\rho\right) + \frac{\sigma^2}{2}\frac{\partial \rho}{\partial x}\right), \tag{2.79}$$

where the non linearity in the drift term is given by the first moment $\langle x\rangle = \int x\rho(x,t)\mathrm{d}x$ of the distribution ρ. The first moment $\langle x\rangle$ corresponds to a suitable order parameter for the system, see e.g. [DZ78, Shi87, Daw83] and it is usually referred to as the "magnetisation" of the system in analogy with spin systems.

Since the DZ model describes an equilibrium system it is possible to define a free energy functional $F[\rho]$

$$F[\rho] = \int V_\alpha(x)\rho(x)\mathrm{d}x + \frac{\theta}{4}\int\int \rho(x)\left(x-y\right)^2\rho(y)\mathrm{d}x\mathrm{d}y + \beta^{-1}\int \rho\ln\rho\mathrm{d}x, \tag{2.80}$$

where $\beta = \frac{2}{\sigma^2}$ is the inverse temperature of the system, such that the McKean Vlasov equation can be written as

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\rho\frac{\partial}{\partial x}\frac{\delta F}{\delta \rho}\right). \tag{2.81}$$

As explained in previous sections, invariant solutions of the mean field dynamics can be characterised by extremes of the free energy functional. We will adopt here a different but equivalent approach using a self consistency equation, reminiscent of the Kirkwood-Monroe approach (2.32)[5]. From (2.79) it is simple to see that stationary solutions $\rho_0(x)$ satisfy

$$\frac{\partial \ln\rho_0}{\partial x} = -\beta\left(V_\alpha'(x) + \theta\left(x - \langle x\rangle\right)\right), \tag{2.82}$$

which, after integrating, results in a family of a distributions $\rho_0(x;m)$

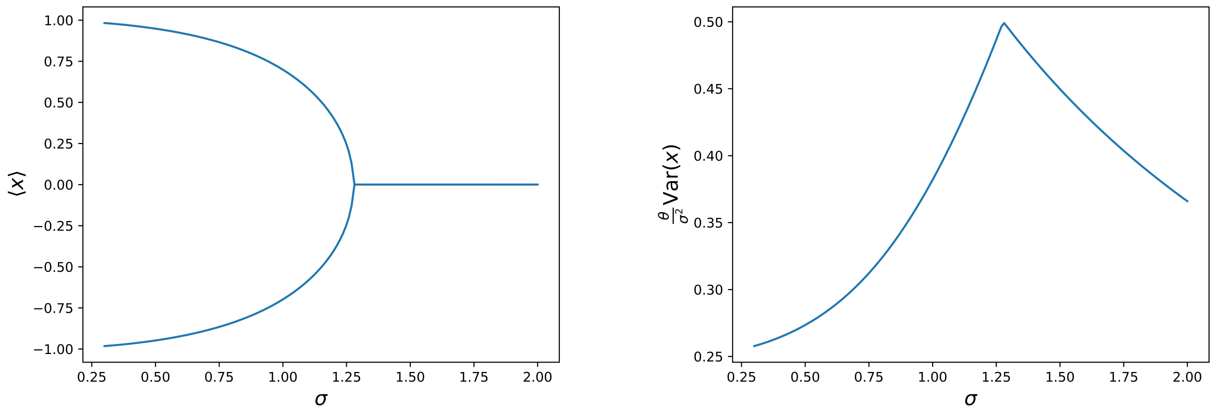
$$\rho_0(x,m) = \frac{1}{Z}e^{-\beta\left(V_\alpha(x)+\theta\left(\frac{x^2}{2}-xm\right)\right)}, \tag{2.83}$$

parametrised by a parameter m satisfying the constraint

$$m = R(m) := \int xp_0(x;m)\mathrm{d}x, \tag{2.84}$$

and $Z = Z(m) = \int e^{-\beta\left(V_\alpha(x)+\theta\left(\frac{x^2}{2}-xm\right)\right)}\mathrm{d}x$ is the mean field partition function. The above equation, commonly called as self consistency equation, plays a major role in determining the stationary properties of the system. Solutions $m^*$of the self consistency equation correspond to stationary measures $\rho_0(x,m^*)$, with magnetisation $\langle x\rangle = m^*$. Moreover, the free energy functional, being a Lyapunov function for the dynamics, see (2.30), provides stability properties of the stationary solutions. In particular, the free energy functional, evaluated at the stationary

[5] We will use the self consistency equation approach for a more general problem in chapter 6



(a) (b)

Figure 2.3: Phase diagram for the continuous phase transition for the McKean Vlasov equation (2.79). Left panel: order parameter $\langle x \rangle$ as a function of σ. Right panel: rescaled variance as a function of σ. The mean field rescaled variance, at the transition, is finite and equals to $1/2$ as given by (2.87). These results have been obtained through the numerical analysis of the self consistency equation (2.84).

solution $\rho_0(x; m)$ can be written as [GKPY19]

$$F[\rho] = F(m) = -\beta^{-1} \ln Z + \frac{\theta}{2} m^2. \tag{2.85}$$

The first derivative of the free energy yields the self consistency equation $F'(m) = \theta \left(m - \langle x \rangle \right)$ as a consequence of the fact that extreme points of the free energy correspond to invariant solutions of the McKean Vlasov equation. Stability properties are related to convexity features of the free energy functional: if the second variation is positive, $\delta^2 F > 0$, for any perturbation $\delta\rho$ around ρ_0, then the stationary solution ρ_0 is stable. If there exists a perturbation $\delta\rho$ such that the above does not hold, the stationary solution is unstable [Shi87, Fra05]. Evaluating the second derivative of the free energy we get

$$F''(m) = -\theta \left(\beta\theta \left(\langle x^2 \rangle - m^2 \right) - 1 \right). \tag{2.86}$$

The investigation of the solutions of the self consistency equation and the stability analysis provided by the free energy makes it possible to identify two regimes of the dynamics [Shi87]. Fixed, the parameter α and the strength of the coupling θ, there exists a high temperatures regime $\beta \to 0$, or equivalently $\sigma \to +\infty$, such that the self consistency equation has the unique (stable) solution $m^* = 0$ corresponding to a disordered state with vanishing magnetisation. As the temperature is decreased, the disordered state loses its stability and bifurcates through a pitchfork bifurcation in two stable and symmetric states of opposite magnetisation $\langle x \rangle = \pm m^* \neq 0$, see Figure 2.3. This corresponds to a continuous phase transition, characterised by the condition $F''(0) = 0$, that is [Shi87]

$$\frac{2\theta}{\sigma^2}\langle x^2 \rangle_0 = 1, \tag{2.87}$$

where the expectation value is taken with respect to the stationary measure $\rho_0(x;0)$. It is possible to obtain the critical hyperplane in the parameter space (α, θ, σ) where the phase transition takes place [Daw83]

$$\frac{D_{-3/2}\left(\frac{\theta-\alpha}{\sigma}\right)}{D_{-1/2}\left(\frac{\theta-\alpha}{\sigma}\right)} = \frac{\sigma}{\theta}, \tag{2.88}$$

where $D_\nu(z) = \frac{e^{-\frac{z^2}{4}}}{\Gamma(-\nu)} \int_0^{+\infty} e^{-zt-\frac{t^2}{2}} t^{-\nu-1} \mathrm{d}t$ with $\nu < 0$ is a parabolic cylinder function and $\Gamma(z)$ is the standard Gamma function. In particular, fixed (α, θ), the above condition yields the critical noise strength value σ_c (or equivalently inverse temperature β_c) at which the phase transition is attained.

The above analysis provides a complete overview of the mean field properties given by the McKean Vlasov equation, that is by the *Law of large numbers* for the empirical measure X_N. To our knowledge, the first available results for the critical fluctuations at the phase transition of X_N around its mean field limit were provided by Dawson for the DZ model [Daw83]. We assume that $\sigma > \sigma_c$ and that the system is in its (unique) stable steady state $\rho_0(x;0)$. As mentioned in section 2.4, we expect that the empirical measure will exhibit gaussian fluctuations around its mean field value. We define as in the previous chapters the empirical fluctuation process

$$Y_N = \sqrt{N}\left(X_N(t,\cdot) - \rho_0(x;0)\mathrm{d}x\right). \tag{2.89}$$

Dawson showed, see also [DGPS23] and references therein for a more general result for equilibrium systems, that Y_N will converge as $N \to +\infty$ to a stochastic process $Y(t,\cdot)$ whose deterministic dynamics is given by an operator $\tilde{\mathcal{L}}_{\rho_0}$

$$\tilde{\mathcal{L}}_{\rho_0} Y = \mathcal{L}_{\rho_0} Y - \theta \frac{\partial}{\partial x} \rho_0(x;0) \int y Y(t, \mathrm{d}y), \tag{2.90}$$

where

$$\mathcal{L}_{\rho_0}(\cdot) = \frac{\partial}{\partial x} \left((V'_\alpha(x) + \theta x) \cdot \right) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} \tag{2.91}$$

is the Fokker-Planck operator one obtains by setting the stationarity condition $\rho = \rho_0(x;0)$ in the McKean Vlasov operator defined in (2.79). For settings above the phase transitions, $\sigma > \sigma_c$, it is possible to show that $\tilde{\mathcal{L}}_{\rho_0}$is a stable operator, that is all its eigenvalues, but one[6], have negative real part. In this regime, the fluctuation process $Y(t, \cdot)$ reaches an asymptotic equilibrium given by a generalised Gaussian field. An important consequence of this *Central Limit Theorem* concerns the fluctuations of the observables of the N-particle system. Given any smooth function $\phi(x)$, we consider any observable of the form $\Phi = \frac{1}{N} \sum_{k=1}^{N} \phi\left(x_k\left(t\right)\right)$, e.g., the empirical magnetisation or centre of mass $\bar{x} = \frac{1}{N} \sum_{k=1}^{N} x_k(t)$. The above *Central Limit Theorem* implies that [Daw83]

$$\langle (\Phi - \langle \Phi \rangle)^2 \rangle_{\rho_N} \sim \frac{1}{\sqrt{N}}, \tag{2.92}$$

where $\langle \cdot \rangle_{\rho_N}$ denotes the expectation value over the N-particle distribution ρ_N solution of the stationary $N-$particle Fokker Planck equation (2.5).

On the other hand, at the phase transition ($\sigma = \sigma_c$) the operator $\tilde{\mathcal{L}}_{\rho_0}$ develops a new vanishing eigenvalue. The linearised dynamics (2.90) can no longer support an equilibrium state and does not represent a good approximation of the fluctuations. Ultimately, this is due to the fact that the critical fluctuations do not follow a scaling as in (2.89). Indeed, Dawson showed that critical fluctuations scale as

$$Y_N^c \coloneqq N^{1/4} \left(X_N \left(N^{1/2} t, \cdot \right) - \rho(x;0) \mathrm{d}x \right). \tag{2.93}$$

There are two meaningful features in the above scaling. Firstly, the scaling of the amplitude of the critical fluctuations is different with respect to non critical fluctuations. This means that

[6]The one corresponding to its stationary state, see [Daw83] for details

any observable Φ undergoes critical fluctuations of bigger amplitude scaling as

$$\langle(\Phi - \langle\Phi\rangle)^2\rangle_{\rho_N} \sim \frac{1}{N^{1/4}}. \tag{2.94}$$

This is due to the fact that at a phase transition the variables $\{x^k(t)\}$ are no longer weakly correlated, but exhibit strong correlations resulting in a macroscopic nature of critical fluctuations [CJL78]. Secondly, the rescaling of time in (2.93) is a manifestation of the critical slowing down phenomenon, meaning that fluctuations persist over longer timescale. It is possible to show that Y_N^c satisfies a *Central Limit Theorem* as one passes to the thermodynamic limit. Loosely speaking, this theorem says that the critical empirical measure can be written as [Daw83]

$$X_N\left(N^{1/2}t, \mathrm{d}x\right) \approx \frac{1}{N^{1/4}} x\rho_0(x;0)z(t)\mathrm{d}x, \tag{2.95}$$

where $z(t)$is a stationary stochastic process solution of

$$\mathrm{d}z = -cz^3(t)\mathrm{d}t + \sigma^*\mathrm{d}w, \tag{2.96}$$

where c and σ^* are two positive constants and $w(t)$ is a standard Wiener process. The above equations show another manifestation of the macroscopic nature of the fluctuations at the phase transition. In fact, as opposed to the non critical fluctuations described by a gaussian random field, critical fluctuations are here coherent, that is the entire empirical measure X_N is driven by the same process $z(t)$.

We conclude this section by briefly mentioning the main dynamical features of the system for the low temperatures $\sigma < \sigma_c$ regime. As previously mentioned, this corresponds to a bistable system where there exist two stable symmetric invariant solutions of the McKean Vlasov equation, corresponding to the two minima of the free energy $F(m)$. Moreover, the free energy has a maximum at $m = 0$, corresponding to the unstable disordered state, see Figure 2.4 . The empirical measure X_N will fluctuate closely to one of the two stable solutions until performing, through a large deviation, a transition to the other symmetric state. As described in section 2.5, the statistics of the transition can be described through a large deviation approach,.

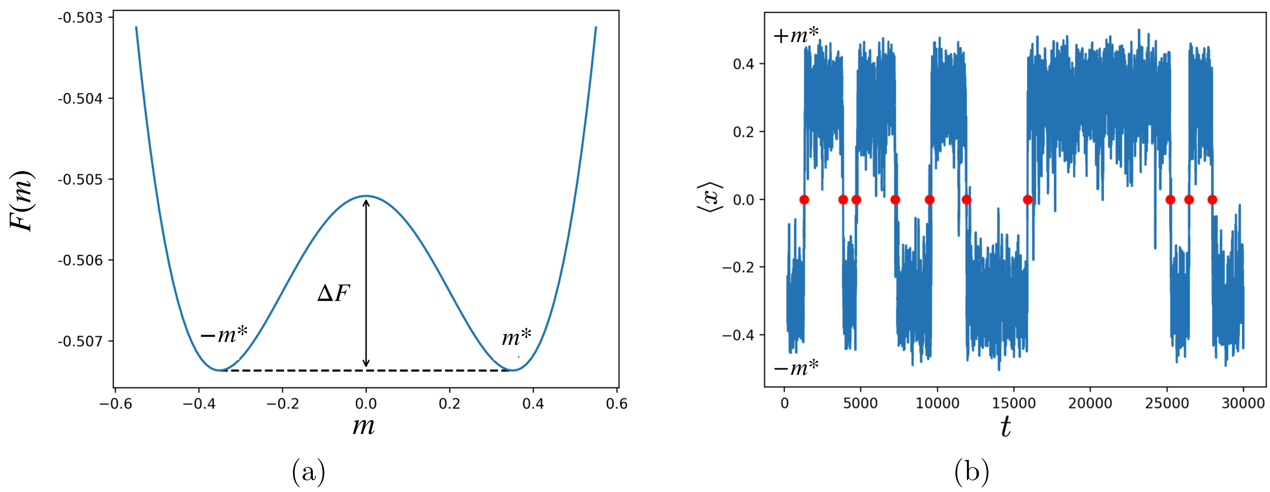


Figure 2.4: Metastability properties of the Desai Zwanzig model for a finite system. Left panel: Free energy $F(m)$, see equation (2.85). The two minima of F correspond to the invariant solutions of the McKean Vlasov equation, characterised by an order parameter $\langle x \rangle = \pm m^*$. Right panel: Typical transitions that are observed in a finite system obtained from a direct numerical integration of equations (2.77) on an ensemble of $N = 8000$ agents.

In particular the mean escape time can be written as

$$\mathbb{E}[\tau] \asymp e^{N\Delta F}, \tag{2.97}$$

where $\Delta F = F(0) - F(\pm m^*)$ is the free energy barrier between the stable and unstable states.

Chapter 3

Linear Response Theory

The aim of Response Theory is to predict how statistical properties of a system change as a result of arbitrary (weak) exogenous perturbations. In linear approximation, the Fluctuation Dissipation Theorem (FDT) holds for systems amenable to a statistical mechanics description. The FDT establishes a connection between spontaneous, unforced fluctuations and their relaxation properties with the response to external forcings. In particular, it allows to write the Green function, describing the linear response of such systems, as a time dependent correlation function of suitable observables evaluated in the unperturbed state. The FDT represents a very general and powerful tool to investigate the forced fluctuations of a system from the analysis of the unforced ones. Indeed, this result allows one to predict the response of (averages of) observables without applying any perturbation. As a matter of fact, one of the earliest applications of the FDT was the derivation of transport coefficient of liquids by simulating their equilibrium (unperturbed) molecular dynamics. The idea of relating dissipation properties to fluctuations can be traced back first to Einstein's work on Brownian motion [Ein05] and then to Nyquist's investigation of thermal agitation in electrical resistors [Nyq28]. A few years later Onsager formulated his regression hypothesis [Ons31] stating that the relaxation process of a nonequilibrium perturbation would follow the same physical laws governing the equilibrium fluctuations. These ideas are at the basis of the FDT developed by Callen and Welton [CW51] and Kubo [Kub57, Kub66]. In particular, the formulation of the latter resembles the more modern one, connecting the

Green function and correlation functions at equilibrium. For a historical perspective, see the very interesting review [MPRV08, Section 1-2]. Although such FDTs were originally applied to Hamiltonian systems near equilibrium, the validity of a generalised FDT extends to a vast class of systems featuring non equilibrium dissipative chaotic or stochastic dynamics . It turns out that in the latter systems the applicability of FDT and the derivation of response formulas are easier to justify than in deterministic settings, see later discussion. We refer the reader to [MPRV08, LV07, SV19] for a modern formulation of the generalised FDT. Furthermore, a generalisation of the FDT to non linear settings of the response can be found in [LC12]. Linear Response Theory and the FDT have been a fruitful conceptual framework and successful approach for a plethora of phenomena. Classic applications include the investigation of the physics of plasma [Nam76, Nam77], stochastic resonance [GHJM98], optical materials [LSPV05], galactic dynamics [BT08], turbulence [Kan20, BDLV01], Markov chains [SGLCG21, Luc16], optomechanical systems [MDN21], simple toy models of chaotic dynamics, [CS07, Luc09, Rei02], but also more recently developed fields such as neural networks [Ces19, CAC21, Lim21], financial markets [PTSSG$^+$21] and the climate system [Lei75, GL20]. In this regards, linear response theory has been successfully applied to a variety of systems, ranging from atmospheric toy models [LS11, AM07, CH11, AM08, MAG10, CEH13], barotropic models [Bel80], quasigeostrophic models [GL17, DG01], atmospheric models [NBH93, CVS04, GB07, GBM08, RP08, GBD02] to coupled climate models [LA05, KD09, FSH15, RLL16, LLR20, LRL17]. In modern terms, the goal is to define practical ways to reconstruct the measure supported time-dependent pullback attractor [CSG11] of the climate by studying the response of a suitably defined reference climate state [GL20]. More applications of response theory can be found in [Ött05] and the recent special issue edited by Gottwald [Got20].

The incredible success of the FDT and linear response theory in predicting the change of statistical properties of complex systems, as exemplified by the aforementioned applications, has led to the wrong assumption that linear response theory holds for general deterministic chaotic systems. As it happens, it is possible to show that simple chaotic dynamical systems violate linear response [BS08, BS10, BBS15, Bal00, Bal14]. In order to understand how linear response can fail, it is instructive to provide a general formulation of the problem. We consider a system, either stochastic or deterministic, that supports a unique ergodic invariant measure μ. We also assume that this measure is physical, i.e. its basin of attraction has positive Lebesgue measure. In other words, if the systems admits an ergodic physical measure there exists a set of initial conditions of non zero Lebesgue measure such that long time averages of observables Φ are equivalent to phase space averages over this measure μ

$$\lim_{T\to+\infty}\frac{1}{T}\int_0^T \Phi(\mathbf{x}(t))\mathrm{d}t = \int \Phi(\mathbf{x})\mu(\mathrm{d}\mathbf{x}), \tag{3.1}$$

where $\mathbf{x}(t)$ represents the trajectory stemming from one of these initial conditions. Moreover, we assume that the system depends on a parameter ε and that, for each fixed ε, the relative ergodic physical measure μ_ε is unique. The parameter ε corresponds to a, possibly time and phase space dependent, perturbation of the unperturbed system one obtains for e.g. $\varepsilon = 0$, whose unperturbed statistics are given by the measure μ_0. Linear response theory aims to predict how the expected value

$$\langle\Phi\rangle_\varepsilon = \int \Phi(\mathbf{x})\mu_\varepsilon(\mathrm{d}\mathbf{x}) \tag{3.2}$$

changes as the parameter ε is varied. In particular, we say that the (unperturbed) system exhibits linear response if the derivative

$$\langle\Phi\rangle' = \frac{\partial\langle\Phi\rangle_\varepsilon}{\partial\varepsilon}|_{\varepsilon=0} \tag{3.3}$$

exists. Thus, in linear response theory approximation, the average of the observable in the perturbed state is expressed as

$$\langle\Phi\rangle_\varepsilon \approx \langle\Phi\rangle_0 + \varepsilon\langle\Phi\rangle'. \tag{3.4}$$

Response formulas allow one to write the derivative $\langle\Phi\rangle'$ in terms of properties of the unperturbed measure μ_0. A FDT can be established if $\langle\Phi\rangle'$ depends only on correlation functions evaluated in the unperturbed state described by μ_0.

A sufficient condition for the derivative (3.3) to exist is that the invariant measure μ_ε is differentiable with respect to perturbation parameter ε. In fact, smoothness and differentiability properties of the invariant measures are at the core of the validity of linear response theory. By and large, in stochastic systems the noise has a regularising effect that provides a smooth invariant measure. Consequently, linear response theory can be very often justified for stochastic systems [HM10, DD10, Hä78]. In these settings, the response and the unforced fluctuations of a system satisfy satisfies a FDT [MPRV08] . We remark that these results hold for a system that does not feature critical transition (tipping points) nor phase transitions. We refer the reader to the next sections of this chapter for the investigation of response properties close to critical transitions. We remark that one of the contributions of this thesis, see next chapters, is to extend Response Theory for the thermodynamic limit of interacting systems, corresponding to an infinite dimensional system described by a nonlinear, nonlocal Fokker Planck equation.

On the other hand, dissipative chaotic deterministic systems do not exhibit smooth invariant probability measures. In fact, such systems live in a nonequilibrium steady state (NESS) characterised by an average contraction of phase space volumes and production of entropy. The natural mathematical framework to describe statistical properties of nonequilibrium chaotic systems is represented by Sinai Ruelle Bowen (SRB) measures [ER85, Rue89, You02] that enjoy smoothness properties along unstable manifolds in the phase space but are rather singular otherwise.

Ruelle [Rue97, Rue98b, Rue09] has shown that a response theory applies for uniformly hyperbolic Axiom A systems. Interestingly, the relationship between forced and unforced variability, i.e., between response and unperturbed fluctuations, is more complex in these settings due to the inherently different properties of the tangent space in its stable and unstable directions. As a matter of fact, the response of a uniformly hyperbolic system described by a SRB measure can be split in two different contributions [CS07, Rue98b]. The first one, deriving from the unstable manifold, can be expressed in terms of a correlation function evaluated in the unperturbed state in accordance to the FDT. The other contribution depends on the dynamics along the stable manifold and does not correspond to unperturbed correlation functions. Consequently, the numerical implementation of Ruelle's response formulas is nontrivial for deterministic chaotic systems [AM07]. Very recently, contributions based on adjoint and shadowing methods provide a promising way forward in this regards [Wan13, CW20, Ni21, SW22]. This new contribution solely derives from the dissipative nonequilibrium nature of the dynamics. Small perturbations around equilibrium steady states satisfy the classic FDT since the invariant measure of the system is absolutely continuous with respect to the Lebesgue measure [Rue09]. The extent of the effect of the new non-canonical term in Ruelle's response formulas on the applicability of the FDT is still a matter of research. The authors in [CRV12] show that for low dimensional deterministic systems the FDT can be applied most of the times, except for very peculiar situations where the unperturbed dynamics exhibits carefully oriented manifolds and for carefully chosen perturbations. Nevertheless, this non-canonical term is not always negligible [CS07, AM07]. On the other hand, in the context of high dimensional chaotic deterministic systems, the validity of the FDT is usually expected for smooth observables representing projection operators from high-dimensional spaces to lower dimensional ones. In fact, one might expect the projection operation to smooth out the singularities of the invariant measure of the whole system [CRV12, MPRV08]. We also remark that, although unjustified in deterministic settings, the presence of a small degree of noise provides a simple reason for the validity of the FDT in general systems.

As previously mentioned, a generic, non uniformly hyperbolic system might support an invariant measure with a rough dependence on the perturbation parameter ε. In such case, a linear response theory does not exist and neither does a FDT. It is however common belief in the scientific community that typical deterministic high dimensional system composed of multiple interacting degrees of freedom exhibit a smooth linear response. In most cases, such a behaviour is justified by invoking a strong chaoticity assumption on the dynamics, the"chaotic hypothesis" of Gallavotti and Cohen [GC95b, GC95a, Gal14, Gal20], according to which a high dimensional system can be considered for all practical purposes as a uniformly hyperbolic Axiom A system. Invoking the *chaotic hypothesis* to justify the existence of linear response theory is however unjustified: even if the hypothesis is true, it does not address how the equivalent Axiom A systems of the unperturbed and the perturbed system relate to each other, that is, it does not provide any information on the family of measures $\{\mu_\varepsilon\}_\varepsilon$ which is crucial for any statement on response theory [WG18].

In this regards, a stronger argument for the validity has been obtained by Wormell and Gottwald [WG19]. The authors study high dimensional deterministic system composed of mean-field coupled microscopic agents, that, individually, do not obey linear response theory. They observe that macroscopic observables of such systems satisfy linear response if the macroscopic dynamics features an effective noisy behaviour through self generated noise and the distribution of the microscopic parameters is smooth. Interestingly, the connection between microscopic and macroscopic dynamics is not as straightforward. They also show that, when the macroscopic observable exhibits a non trivial dynamics, linear response can fail at the macroscopic level even if the individual components possess a smooth response. Non validity of the "chaotic hypothesis" for high dimensional systems has also been observed in [Wor22b] where the existence of non-hyperbolic large-scale dynamical structures is established in a mean-field coupled deterministic system. Moreover, a recent work points in the direction that the emergence of linear response theory in high dimensional systems highly depends on the properties of SRB measures conditioned on generic manifolds, that is not necessarily stable or unstable, and, specifically, on the validity of *conditioned decay of correlations* properties [Wor22a].

We are here interested in the application of linear response to the context of weakly interacting diffusions where it is natural to assume a stochastic component of the dynamics. In this setting, we will not be concerned with these issues since the smoothness of the invariant measure of the stochastic system ensures the derivation of response formulas and the validity of the FDT away from critical settings.

Before delving into the investigation of response properties of stochastic systems, we want to remark that nonlinear response theories and nonlinear FDTs have also been established near both equilibrium and nonequilibrium dynamics [Rue98a, LSPV05, LC12, LCSZ08b, LCSZ08a, BB05]. Moreover, linear response formulas have been extended to encompass the situation of perturbations of finite amplitude, where the state vector is perturbed at one instance of time by a (not necessarily small) perturbation and thereafter left evolved through the unperturbed dynamics [Abr19].

Having established the close link given by the FDT between response and unperturbed correlations, it is relevant to present, in the next section, a theory of mixing properties for finite dimensional stochastic systems. In the context of non equilibrium systems, the authors in [SV19] remark that the unperturbed correlations appearing in FDTs involve complicated quantities that do not appear at equilibrium. The determination of such quantities is technically difficult since it requires the knowledge of the invariant measure of the system, which is, as opposed to equilibrium systems, not known a priori and heavily system-dependent [Der07]. Moreover, there is still no consensus on the interpretation of the physical meaning of such terms, see [SV19] and references therein.

For deterministic systems, the theory of the Ruelle-Pollicott (RP) resonances and the transfer operator formalism represent a rigorous approach to the investigation of mixing properties of the system, establishing decomposition formulas of correlation functions and power spectra [Bal00, BER89, Fro97, MG08, GLP13, LM94, BMM12]. This approach provides a more general framework to study time dependent correlation functions between arbitrary observables of the system by decomposing the dynamics in terms of universal (as opposed to observable-dependent) features. More recently it has been shown that a similar formalism, based on the theory of Markov semigroups and their spectral theory, can be applied to stochastic systems as well [CTDN20, TLD18]. The theory of RP resonances characterises other relevant properties of the evolution of deterministic and stochastic dynamical systems such as coherent structures [FPET07, Fro13, FPG14], metastability [MS81, SS13] [Pav14, Chapter7]. Stochastic RP resonances have also been successfully applied to the design of stochastic parameterizations to solve challenging closure problems issued from geophysical turbulence [SGLCG21, KCB18, CLM21]. In the following we will be interested in the role of RP resonances in the context of critical transitions [TvdBD15, TLLD18, TLD18] and sensitivity to perturbations [CNK$^+$14, Luc16, SGL20]. In the next section we will illustrate the fundamental results of the theory of stochastic RP resonances and mixing properties of the system.

Furthermore, in section 3.2 we will establish a link, along the same lines as [CNK$^+$14, GL22], between linear response theory and RP resonances. One of the fundamental results of this approach is the decomposition in terms of the modes and corresponding eigenvalues of the Koopman operator [Mez05, BMM12], governing the time evolution of observables of the system, of the Green functions of the systems. In particular, this will allow us to identify critical scenarios, usually called critical transitions or tipping points, of high dimensional complex systems where the linear response of the system breaks down and a rough dependence on the parameters of the system is expected. As a last remark, it is worth pointing out that the evaluation of the RP resonances is non trivial for highly non linear, complex, high dimensional systems. Most of the methods that have been developed to compute RP resonances from finite-dimensional data-driven approximations of infinite-dimensional transfer operators are somehow related to the Ulam's approach. There, one seeks a suitable partition of the phase space and approximates the underlying transfer operator by finite dimensional Markov matrices describing the transition probabilities between the finite number of states given by the partition. It is easy to realise that these methods suffer from the curse of dimensionality, the complexity of such procedures dramatically increases with the dimension of the phase space [FLQ10, KNK+18, TvdBD15, CTDN20, TCND19]. Alternative data-informed methods, including some deep learning approaches, exist but they suffer from practical limitations in their implementation [SC22]. Promising results in this regards have been obtained in [SC22], where the authors present a flexible *simulation-free* approach based on Neural Networks architectures able to find solutions of eigenvalue problems in moderately large dimensions.

3.1 The theory of stochastic Ruelle Pollicott resonances

We consider a finite dimensional system described by the following stochastic differential equations

$$\mathrm{d}\mathbf{x} = \mathbf{F}(\mathbf{x})\mathrm{d}t + \mathbf{s}(\mathbf{x})\mathrm{d}\mathbf{W}, \quad \mathbf{x} \in \mathbb{R}^d. \tag{3.5}$$

The drift part of the dynamics is given in terms of a generally non linear smooth vector field $\mathbf{F} : \mathbb{R}^d \to \mathbb{R}^d$, whereas the diffusion dynamics is determined by the state dependent matrix $\mathbf{s} : \mathbb{R}^d \to \mathbb{R}^{d\times q}$ and by the $\mathbb{R}^q-$valued Wiener process $\mathbf{W} = (W_1, \dots, W_q)$. In particular, the $i - th$ component of the diffusion part is

$$[\mathbf{s}(\mathbf{x})\mathrm{d}\mathbf{W}]_i = \sum_{j=1}^{q} \mathbf{s}_{ij}(\mathbf{x})\mathrm{d}W_j, \quad i = 1, \dots, d. \tag{3.6}$$

We note that $q \leq d$ is not necessarily equal to d and it is not guaranteed that the diffusion matrix $\mathbf{\Sigma} = \mathbf{s}\mathbf{s}^T$ is full rank, meaning that one could potentially find a change of variables where noise act only on a subset of such dynamical variables. It is well known that the evolution of the probability density functions associated with (3.5) is described by a Fokker-Planck equation

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= \mathcal{L}\rho := -\nabla \cdot \left(\mathbf{F}\left(\mathbf{x}\right)\rho(\mathbf{x},t)\right) + \frac{1}{2}\mathrm{D}^2 : \left(\mathbf{\Sigma}\left(\mathbf{x}\right)\rho(\mathbf{x},t)\right) \\ \rho(\mathbf{x},0) &= \rho_{in}(\mathbf{x}),\end{aligned} \tag{3.7}$$

where $\mathcal{L}$ is the Fokker Planck operator. Furthermore D^2 is the matrix of second derivatives $\mathrm{D}^2_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}$ and : denotes the following operation between matrices $A : B = \sum_{i,j=1}^{d} A_{ij}B_{ij}$. The probability distribution $\rho(\mathbf{x},t)$ represents the probability that the stochastic process $\mathbf{x}(t)$ originating from (3.5), originally distributed according to $\rho_{in}(\mathbf{x})$, is equal to $\mathbf{x}$ at time t. Such probability can be written as

$$\rho(\mathbf{x},t) = \int_{\mathbb{R}^d} \rho(\mathbf{x},t|\mathbf{y},0)\rho_{in}(\mathbf{y})\mathrm{d}\mathbf{y}, \tag{3.8}$$

where $\rho(\mathbf{x},t|\mathbf{y},0)$ represents the transition probability from state $\mathbf{y}$ at time $t = 0$ to state $\mathbf{x}$ at time t. In the introduction we have highlighted the relevance of the smoothness properties of the invariant probability distributions for the derivation of fluctuation dissipation theorems. Such smoothing effects derive from a sufficiently strong spreading of the noise in the evolution of all dynamical variables. In some applications, a uniform ellipticity assumption on the diffusion matrix is valid. This amounts to requiring that $q = d$ and that the diffusion matrix is uniformly positive definite, i.e. , there exists a constant $\lambda > 0$ such that

$$\langle \mathbf{y}, \mathbf{\Sigma}(\mathbf{x})\mathbf{y}\rangle \geq \lambda|\mathbf{y}|^2, \quad \forall \mathbf{y} \in \mathbb{R}^d, \tag{3.9}$$

uniformly for any $\mathbf{x} \in \mathbb{R}^d$, where $\langle \cdot, \cdot \rangle$ and $|\cdot|^2$ represents the usual scalar product and the (squared) norm of vectors in $\mathbb{R}^d$. The uniform ellipticity assumption guarantees the existence of smooth probability densities [Pav14, Chapter 4] but requires that the noise acts according to (3.9) on all dynamical variables. Most applications deal with the case of degenerate noise, where the external source of randomness only acts on a subset of dynamical variables, see [Pav14, Chapter 6] for examples regarding Langevin dynamics of particles including both position and momentum variables or the appendix of [CTDN20] for examples from fluid dynamics. In this situation, smooth densities exist provided that the interactions among the different dynamical variables allow for a sufficiently strong transmission of the noise. Mathematically, this is represented by the hypoellipticity property of the generator of the diffusion stochastic process [Pav14, Chapter 6]. Hörmander condition [Ḧ67] provides a useful criterion to check whether the structure of the noise given by $\mathbf{\Sigma}(\mathbf{x})$ generates a hypoelliptic process. Loosely speaking, Hörmander's condition states that at any point of the phase space $\mathbf{x} \in \mathbb{R}^d$ the directions generated by the second differential operator in $\mathcal{L}$ stemming from a suitable combination of the noise and the drift fully span $T_{\mathbf{x}}\mathbb{R}^d$, i.e. the tangent linear space of $\mathbb{R}^d$ at point $\mathbf{x}$. General smoothness and boundedness assumptions on the drift $\mathbf{F}(\mathbf{x})$ and diffusion $\mathbf{\Sigma}(\mathbf{x})$ coefficients together with Hörmander's condition suffice to guarantee the existence of smooth solutions of (3.7).

3.1.1 Markov Semigroup theory

The properties of the Fokker Planck equation (3.7) can be conveniently investigated through the Markov semigroup formalism that we introduce below. Such formalism encompasses more general Markov processes than the continuous path ones that the Fokker Planck equation describes. We state the theory in its generality first and apply it to systems described by Stochastic Differential Equations at a later stage. Given any function $\Phi(\mathbf{x}) \in \mathcal{C}_b\left(\mathbb{R}^d\right)$, where $\mathcal{C}_b\left(\mathbb{R}^d\right)$ denotes the space of bounded and continuous function of the phase space, we define a one parameter family of linear operators P_t, $t \geq 0$, defined by

$$\left(P_t \Phi\right)(\mathbf{y}) := \mathbb{E}\left[\Phi\left(\mathbf{x}\left(t\right)\right)|\mathbf{x}\left(0\right) = \mathbf{y}\right] = \int_{\mathbb{R}^d} \Phi\left(\mathbf{x}\right)\rho\left(\mathbf{x}, t|\mathbf{y}, 0\right)\mathrm{d}\mathbf{x}. \tag{3.10}$$

P_t is a linear operator and is called the Markov semigroup associated to the stochastic process. It is associated to statistical properties since it is defined through an expectation value over all realisations of the noise, namely over all possible values of final states at time t compatible with the stochastic dynamics. It is possible to show that P_t is equipped with a semigroup structure

$$P_0 = \mathrm{Id}, \quad P_{t+s} = P_t \circ P_s, \quad t, s \geq 0, \tag{3.11}$$

where Id is the identity operator. The first equation is easily proved since

$$(P_0\Phi)(\mathbf{y}) = \int_{\mathbb{R}^d} \Phi(\mathbf{x})\,\rho(\mathbf{x},0|\mathbf{y},0)\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d} \Phi(\mathbf{x})\,\delta(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{x} = \Phi(\mathbf{y}). \tag{3.12}$$

The second semigroup property derives from the Markovianity of the process described, in particular from the Chapman Kolmogorov equation [Pav14, Chapter 2]. In particular, a breakdown of such property indicates that the stochastic process exhibits some non-Markovian features. We define by $\mathcal{D}(\mathcal{K})$ the set of all bounded and continuous observables $\Phi \in \mathcal{C}_b(\mathbb{R}^d)$ for which the following strong limit

$$\mathcal{K}\Phi := \lim_{t\to 0} \frac{P_t\Phi - \Phi}{t} \tag{3.13}$$

exists. The operator $\mathcal{K} : \mathcal{D}(\mathcal{K}) \to \mathcal{C}_b(\mathbb{R}^d)$ is called the (infinitesimal) generator of the semigroup operator P_t and is usually referred to as the generator of the Markov process. If the dynamics is deterministic the transition probability in (3.10) is actually a Dirac delta function selecting only the final state $\mathbf{x} = S(t)\mathbf{y}$, where $S(t)$ is the deterministic flow describing the evolution of trajectories in the phase space. In a deterministic setting, the Markov semigroup operator is usually called Koopman operator and is given by

$$(P_t\Phi)(\mathbf{y}) = \Phi(S(t)\mathbf{y}). \tag{3.14}$$

The definition of the generator $\mathcal{K}$ and the semigroup properties imply that at least at a formal level we can write

$$P_t = e^{t\mathcal{K}}. \tag{3.15}$$

The physical relevance of the Markov semigroup operator P_t (or its generator $\mathcal{K}$) is represented by the fact that they govern the dynamical evolution of expectation values of observables. To further illustrate this fundamental property of P_t, we define the function $u(\mathbf{y},t) := (P_t\Phi)(\mathbf{y})$ and evaluate its time derivative

$$\frac{\partial u}{\partial t} = \frac{\mathrm{d}}{\mathrm{d}t}(P_t\Phi) = \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{t\mathcal{K}}\Phi\right) = \mathcal{K}\left(e^{t\mathcal{K}}\Phi\right) = \mathcal{K}u. \tag{3.16}$$

Given that at the initial time $u(\mathbf{y},0)=P_0\Phi=\Phi(\mathbf{y})$, we can state that the expectation values of observables $\Phi\in\mathcal{C}_b(\mathbb{R}^d)$ satisfy the following initial value problem

$$\begin{aligned}\frac{\partial u}{\partial t}&=\mathcal{K}u,\\ u(\mathbf{y},0)&=\Phi(\mathbf{y}).\end{aligned}\tag{3.17}$$

The above equation is usually called, in stochastic settings, the *backward Kolmogorov equation*. An alternative approach to the investigation of statistical properties of stochastic systems is to study the evolution of probability measures of the systems, rather than studying the evolution of its observables. The Fokker Planck equation (3.7) represents in fact such an approach. It turns out that the Fokker Planck equation constitutes a dual approach to the Markov semigroup operator defined above. We define the adjoint semigroup $P_t^\dagger$ that acts on (not necessarily smooth) probability measures μ in the following way

$$P_t^\dagger\mu\,(\Gamma)=\int_{\mathbb{R}^d}\rho(\Gamma,t|\mathbf{x},0)\mu(\mathrm{d}\mathbf{x}),\tag{3.18}$$

where Γ is a measurable set in $\mathbb{R}^d$and $\rho(\Gamma,t|x,0)$ is the transition probability from state $\mathbf{x}$ at time 0 to the set Γ at time t associated to equation (3.7). In particular, given the hypoellipticity assumption on the stochastic process the transition probability $\rho(\Gamma,t|x,0)$ is smooth with respect to the Lebesgue measure and can be written as

$$\rho(\Gamma,t|x,0)=\int_\Gamma\rho(\mathbf{y},t|\mathbf{x},0)\mathrm{d}\mathbf{y}.\tag{3.19}$$

The operator $P^\dagger$ is formally the adjoint of the Markov semigroup P_t, namely

$$\int_{\mathbb{R}^d}P_t\Phi(\mathbf{x})\mathrm{d}\mu(\mathbf{x})=\int_{\mathbb{R}^d}\Phi(\mathbf{x})\mathrm{d}(P^\dagger\mu)(\mathbf{x}).\tag{3.20}$$

In a deterministic context, the operator $P^\dagger$ is commonly known with different names as transfer operator, Ruelle operator or Perron-Frobenius operator. We can define the generator of $P^\dagger$ in a similar way as before and write

$$P^\dagger=e^{t\mathcal{L}},\tag{3.21}$$

where the operator $\mathcal{L}$ is the flat L^2-adjoint of the generator of the process $\mathcal{K}$

$$\int_{\mathbb{R}^d} (\mathcal{K}f)\, h\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d} f\,(\mathcal{L}h)\,\mathrm{d}\mathbf{x}. \tag{3.22}$$

Now, given a stochastic process $\mathbf{x}(t)$ with generator $\mathcal{K}$, initially distributed according to a measure μ , that is $\mathbf{x}(0) \sim \mu$, we define the *law* of the Markov process as

$$\mu(t) \coloneqq P_t^{\dagger}\mu. \tag{3.23}$$

The evolution of probability measures according to the underlying Markov process is determined by the operator $\mathcal{L}$. In fact, deriving with respect to time the above equation and using similar steps that led to (3.17) we obtain

$$\begin{aligned} \frac{\partial \mu(t)}{\partial t} &= \mathcal{L}\mu(t), \\ \mu(0) &= \mu. \end{aligned} \tag{3.24}$$

The above equation is called *forward Kolmogorov equation* and determines the time evolution of probability measures. When $\mu(t)$ is absolutely continuous with respect to the Lebesgue measure $\mu(t)(\mathrm{d}\mathbf{x}) = \rho(\mathbf{x}, t)\mathrm{d}\mathbf{x}$, the same equation holds for the probability distribution $\rho(\mathbf{x}, t)$. When we consider Markovian stochastic evolutions featuring drift and diffusion but with no jumps processes, the generator $\mathcal{K}$ and $\mathcal{L}$ featuring in the *backward* and *forward Kolmogorov equations* respectively can be explicitly evaluated [Pav14, Chapter 2]. In particular, the *forward Kolmogorov equation* (3.24) is simply the Fokker-Planck equation (3.7). The generator $\mathcal{L}$ of the adjoint semigroup is given by the Fokker-Planck operator

$$\mathcal{L}(\cdot) = -\nabla\cdot(\mathbf{F}(\mathbf{x})\,\cdot\,) + \frac{1}{2}\mathrm{D}^2 : (\boldsymbol{\Sigma}(\mathbf{x})\,\cdot\,) = -\sum_{i=1}^{d}\frac{\partial}{\partial x_i}(F_i(\mathbf{x})\,\cdot\,) + \frac{1}{2}\sum_{ij=1}^{d}\frac{\partial^2}{\partial x_i \partial x_j}(\Sigma_{ij}(\mathbf{x})\,\cdot\,)\,. \tag{3.25}$$

We remark that in deterministic setting, $\mathbf{\Sigma} = 0$, the above operator is usually called the Liouville operator. The generator of the continuous stochastic process (3.5) is then easily evaluated as

$$\mathcal{K} = \mathbf{F}(\mathbf{x}) \cdot \nabla + \frac{1}{2}\mathbf{\Sigma}(\mathbf{x}) : \mathrm{D}^2 = \sum_{i=1}^{d} F_i(\mathbf{x}) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{ij}^{d} \Sigma_{ij}(\mathbf{x}) \frac{\partial^2}{\partial x_i \partial x_j}. \tag{3.26}$$

3.1.2 Spectral properties of the Markov semigroup

In general, one is interested in the stationary measures that describe the asymptotic statistical behaviour of the solutions of the Stochastic Differential equation (3.5). In this respect, ergodic measures are central for the characterisation of long time statistical properties of the system. We here introduce the notion of ergodic measures in the context of the theory of Markov semigroups and show how spectral properties of these operators in suitable function spaces play a fundamental role. We say that the process described by (3.5) is ergodic if $\lambda = 0$ is a simple eigenvalue of the generator $\mathcal{K}$. This is equivalent to the condition that the equation

$$\mathcal{K} g = 0 \tag{3.27}$$

has only constant solutions $g =$ const. Ergodic properties can be investigated by looking at the null space of the generator of the stochastic process. It is possible to characterise ergodic properties in terms of the Markov semigroup operator P_t rather than its generator. Given the definition of the generator and (3.27), we can say that the stochastic process is ergodic if the equation

$$P_t g = g \tag{3.28}$$

has only constant solutions for all $t \geq 0$. Equivalent, more familiar, versions of ergodicity are related to the adjoint semigroup perspective. Given the law of the process (3.23), an invariant measure μ is defined as a fixed point of the adjoint semigroup

$$P_t^{\dagger} \mu = \mu. \tag{3.29}$$

This is in fact the L^2-adjoint of (3.28). If there is a unique μ satisfying the above equation, the Markov process is ergodic with respect to μ. In terms of the adjoint generator, and assuming for simplicity that $\mu(\mathrm{d}\mathbf{x}) = \rho(\mathbf{x})\mathrm{d}\mathbf{x}$, the ergodic invariant density $\rho(\mathbf{x})$ satisfies

$$\mathcal{L}\rho = 0, \tag{3.30}$$

which is simply the stationary Fokker-Planck equation. If the Markov process is ergodic with respect to μ, then the long term dynamics is described by μ, since

$$\lim_{t\to+\infty} P_t^\dagger \mu_0 = \mu, \tag{3.31}$$

where μ_0 is the initial distribution of the dynamical variables $\mathbf{x}(0)$. Finally, it is possible to show that the "physical" definition of ergodicity holds

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T \Phi\left(\mathbf{x}\left(t\right)\right)\mathrm{d}t = \int_{\mathbb{R}^d} \Phi(\mathbf{x})\mu(\mathrm{d}\mathbf{x}), \tag{3.32}$$

where $\Phi \in C_b(\mathbb{R}^d)$ is an observable of the system. The above equation states that time averages on an infinitely long stochastic trajectory are equivalent, provided that the system is ergodic with respect to μ, to phase space averages weighted by the measure μ. Proving that a stochastic process is ergodic is known to be a very difficult task for general systems. In practice, one usually tries to establish the existence of a unique invariant measure, i.e. a unique solution of (3.29). A classical and powerful approach is to show that the Markov semigroup is Strong Feller and irreducible. The irreducibility property expresses the idea that any neighborhood of any point of $\mathbb{R}^d$ can be reached with a positive probability. The Strong Feller property is related instead to smoothing properties of the Markov semigroup, namely that it maps bounded measurable functions into bounded *continuous* functions. We observe that the hypoellipticity assumption mentioned in 3.1 guarantees that the Markov semigroup is Strong Feller. In the following, we will not be concerned with proving that a unique ergodic measure exists and will always assume that this is indeed the case. Further details on such issues and Markov semigroups can be found in [DPZ96, EN00, LB07]. The spectral properties of the generator of the process carry information not only on the ergodic properties of the stochastic process, as we have just illustrated above, but also on the mixing properties of the system.

3.1.3 Spectral Decomposition of Correlation Functions

We here consider a stochastic process of type (3.5) and assume that an invariant measure μ exists. The existence of an invariant measure allows one to show [CTDN20] that the Markov semigroup P_t associated to the stochastic process is a strongly continuous semigroup [EN00] in $L^p_\mu(\mathbb{R}^d)$ for any $p \geq 1$. In this section, we will consider a class of observables $\Phi(\mathbf{x})$, often arising in applications, that belong to $L^p_\mu(\mathbb{R}^d)$, namely the functions that are p-integrable with respect to the invariant measure μ

$$\Phi \in L^p_\mu(\mathbb{R}^d) \Longleftrightarrow \int_{\mathbb{R}^d} |\Phi(\mathbf{x})|^p \mu\,(\mathrm{d}\mathbf{x}) < \infty. \tag{3.33}$$

In particular we will mostly be interested in $\Phi \in L^2_\mu(\mathbb{R}^d)$. The Markov semigroup theory described in the previous section has been developed in the functional space $\mathcal{C}_b(\mathbb{R}^d)$ of bounded continuous functions. However, it can be extended to the space $L^p_\mu(\mathbb{R}^d)$, see for example [GZ03]. Below, we will briefly recall the spectral theory of strongly continuous semigroups that will lead to the fundamental result of this section, the spectral decomposition of the Markov semigroup. The spectrum of the generator of a strongly continuous semigroup can be essentially divided in two qualitatively different sets, a discrete part $\sigma(\mathcal{K})$ composed of isolated eigenvalues and a continuous part, usually called essential spectrum. We assume here a typical setting for the stochastic process under investigation, that is, we assume that the invariant measure μ is unique and ergodic. The measure is thus stable meaning that the spectrum of the generator $\mathcal{K}$ is all included in the left side of the complex plane, with eigenvalues with non positive real part. An important class of semigroup operators, the compact semigroups, do not exhibit any essential spectrum. The characterisation of the essential spectrum of a generic semigroup $\mathcal{P} = (P_t)_{t\geq 0}$ is usually carried out by comparing it to compact operators. In particular, one defines the essential growth bound as

$$\omega_{ess}(\mathcal{P}) = \inf_{t>0} \frac{1}{t} \ln ||P_t||_{ess}, \tag{3.34}$$

where $||\cdot||_{ess}$ measures the distance of P_t to the set of linear and compact operators of $L^p_\mu(\mathbb{R}^d)$. Clearly, compact semigroups have $\omega_{ess}(\mathcal{P}) \equiv -\infty$. We can also define an *eventually compact* semigroup $\mathcal{P}$ if there exists $t_0 > 0$ such that P_{t_0} is compact. A general class of stochastic processes that we will often consider in the following is represented by *quasi-compact* semigroups, that is semigroups that approach compact operators in the asymptotic long time limit, namely $||P_t||_{ess} \to 0$ as $t \to +\infty$. Quasi compact semigroups are characterised by an essential growth bound $-\infty < \omega_{ess}(\mathcal{P}) < 0$. Spectral components with modulus larger than $e^{-|\omega_{ess}|t}$ are represented by eigenvalues of finite multiplicity. We also remark that the Spectral Mapping Theorem [EN00] provides a way to relate eigenvalues of $\mathcal{K}$ and its associated semigroup $\mathcal{P}$. In particular if λ_j is an eigenvalue of the generator $\mathcal{K}$, with $\mathbf{Re}\lambda_j > \omega_{ess}$, relative to an eigenfunction ψ_j, so is the eigenvalue $e^{\lambda_j t}$ of $P_t = e^{t\mathcal{K}}$ relative to the same eigenfunction. We define the stochastic Ruelle Pollicott resonances $\{\lambda_j\}_{j=1}^n$ of a strongly continuous Markov semigroup as the isolated discrete eigenvalues of the generator such that their real part is greater than the essential growth bound.

$$\{\lambda_j\}_{j=1}^n \text{ are stochastic RP resonances } \iff \lambda_j \in \sigma(\mathcal{K}) \text{ and } \lambda_j > \omega_{ess}(\mathcal{P}) \quad \forall j = 0, \dots, n. \tag{3.35}$$

In other terms, the RP resonances are the point spectrum of the generator $\mathcal{K}$. We remark that the number n is not necessarily finite. For simplicity of notation, we also assume that the RP resonances are ordered such that $\lambda_0 = 0 > \mathbf{Re}\lambda_1 \geq \mathbf{Re}\lambda_1 \geq \dots$. The simple RP resonance $\lambda_0 = 0$ corresponds to the eigenvalue relative to the existence of the stationary measure μ, see previous section. We remark that the resonance in 0 is simple because we assume that the measure μ is unique and ergodic. We can finally state the spectral decomposition of the Markov semigroup in $L^2_\mu(\mathbb{R}^d)$ [LM94, EN00, EN06]

$$P_t = e^{t\mathcal{K}} = \sum_{j=0}^{n} T_j(t) + \mathcal{R}_n(t), \tag{3.36}$$

where

$$T_j(t) = \left[\sum_{k=0}^{m_j-1} \frac{t^k}{k!} \left(\mathcal{K} - \lambda_j \mathrm{Id}\right)^k\right] e^{\lambda_j t}\Pi_j. \tag{3.37}$$

The operators $T_j(t)$ represent the contribution originating from the RP resonance λ_j. In particular, we have denoted as m_j the (finite) algebraic multiplicity of λ_j and with Π_j the spectral projection operator onto the eigenspace relative to λ_j. On the other hand, $\mathcal{R}_n$ is the operator representing the role of the essential spectrum. It is possible to bound such operator at any point in time, in fact given any $\omega > \omega_n^*$ there always exists $M > 0$ such that

$$||\mathcal{R}_n(t)|| \leq Me^{\omega t}, \quad t \geq 0, \tag{3.38}$$

where $\omega_n^* = \sup\left(\{\omega_{ess}(\mathcal{P})\} \cup \{\mathbf{Re}\lambda : \lambda \in \sigma(\mathcal{P})\backslash\{\lambda_1, \cdots, \lambda_n\}\}\right)$. For quasi-compact operators, the essential growth bound is negative, $\omega_{ess} < 0$, thus the contribution of the essential spectrum decreases at least exponentially with time. In fact, we can write

$$\mathcal{P} \text{ is quasi-compact } \Longrightarrow ||\mathcal{R}_n(t)|| \leq Me^{-|\omega|t}, \quad t \geq 0, \tag{3.39}$$

for any $|\omega| < |\omega_n^*|$. The physical relevance of the RP resonances becomes evident when the spectral decomposition is applied to the determination of correlation functions of observables of the system. In this regards, we define the correlation function between Φ and $\Psi \in L^2_\mu(\mathbb{R}^d)$ in the stationary state of the stochastic process described by the invariant measure μ as[1]

$$C_{\Phi,\Psi}(t) = \langle \Phi(\mathbf{x}(t))\Psi(\mathbf{x}(0))\rangle_\mu = \int_{\mathbb{R}^d} \left(e^{t\mathcal{K}}\Phi(\mathbf{x})\right)\Psi(\mathbf{x})\mu(\mathrm{d}\mathbf{x}). \tag{3.40}$$

We observe that the above definition in terms of the Markov semigroup is analogous, given a smooth invariant measure $\mu(\mathrm{d}\mathbf{x}) = \rho(\mathbf{x})\mathrm{d}\mathbf{x}$, to the one given in (2.22). In fact, given the

[1]Without loss of generality, we will assume that the variables have vanishing expectation value in the stationary state, that is $\int_{\mathbb{R}^d} \Phi\mu(\mathrm{d}\mathbf{x}) = \int_{\mathbb{R}^d} \Psi\mu(\mathrm{d}\mathbf{x}) = 0$

definition of the Markov semigroup (3.10) we can write

$$\begin{aligned} C_{\Phi,\Psi}(t) &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}\Phi(\mathbf{y})\rho\left(\mathbf{y},t|\mathbf{x},0\right)\mathrm{d}\mathbf{y}\right)\Psi(\mathbf{x})\rho(\mathbf{x})\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\Phi(\mathbf{y})\Psi(\mathbf{x})\rho(\mathbf{y},t|\mathbf{x},0)\rho(\mathbf{y}) = \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\Phi(\mathbf{y})\Psi(\mathbf{x})e^{t\mathcal{L}}\delta(\mathbf{y}-\mathbf{x})\rho(\mathbf{y}) = \int_{\mathbb{R}^d}\mathrm{d}\mathbf{y}\Phi(\mathbf{y})e^{t\mathcal{L}}\int\mathrm{d}\mathbf{x}\Psi(\mathbf{x})\delta(\mathbf{y}-\mathbf{x})\rho(\mathbf{x}) \\ &= \int_{\mathbb{R}^d}\Phi(\mathbf{y})e^{t\mathcal{L}}\Psi(\mathbf{y})\rho(\mathbf{y})\mathrm{d}\mathbf{y}, \end{aligned} \tag{3.41}$$

where we have used similar arguments for the time evolution of the transition probability as for (2.22). Considering the decomposition of the Markov operator (3.36), the correlation function $C_{\Phi,\Psi}(t)$ admits the following spectral decomposition of correlation functions

$$C_{\Phi,\Psi}(t) = \sum_{j=1}^{n}\sum_{k=0}^{m_j-1}\frac{t^k}{k!}e^{\lambda_j t}\left(\int_{\mathbb{R}^d}\Psi(\mathbf{x})\left(\mathcal{K}-\lambda_j\mathrm{Id}\right)^k\Pi_j\Phi(\mathbf{x})\mu\left(\mathrm{d}\mathbf{x}\right)\right) + \mathcal{Q}_n(t), \tag{3.42}$$

where $\mathcal{Q}_n(t)$ represents the contribution to the correlation function stemming from $\mathcal{R}_n(t)$. The inequality (3.38) on the essential spectrum operator yields a bound on the quantity $\mathcal{Q}_n(t)$

$$\mathcal{Q}_n(t) \leq M||\Psi||_{L^2_\mu}||\Phi||_{L^2_\mu}e^{\omega t}, \tag{3.43}$$

where $||\cdot||_{L^2_\mu}$ is the L^2_μ norm and $\omega > \omega_n^*$. Analogously to the previous discussion, if the Markov semigroup is quasi-compact, the above bound provides the exponentially decaying rate of $\mathcal{Q}_n(t)$. Indeed, considering (3.39) and (3.43) we have the following result

$$\mathcal{P}\text{ is quasi-compact } \Longrightarrow |\mathcal{Q}_n(t)| \leq M||\Psi||_{L^2_\mu}||\Phi||_{L^2_\mu}e^{-|\omega|t}, \quad t \geq 0. \tag{3.44}$$

The spectral decomposition of correlation functions (3.42) simplifies when one assumes that the RP resonances are simple eigenvalues , i.e. $m_j = 1\ \forall j = 0, 1, \dots$. In fact, we can write

$$\begin{aligned} C_{\Phi,\Psi}(t) &= \sum_{j=1}^{n} e^{\lambda_j t}\left(\int_{\mathbb{R}^d} \Psi(\mathbf{x})\Pi_j\Phi(\mathbf{x})\mu(\mathrm{d}\mathbf{x})\right) + \mathcal{Q}_n(t) = \\ &= \sum_{j=1}^{n} e^{\lambda_j t}\int_{\mathbb{R}^d} \Psi(\mathbf{x})\langle\psi^*, \Phi\rangle_\mu \psi_j(\mathbf{x})\mu(\mathrm{d}\mathbf{x}) + \mathcal{Q}_n(t) = \\ &= \sum_{j=1}^{n} e^{\lambda_j t}\langle\psi_j^*, \Phi\rangle_\mu \int_{\mathbb{R}^d} \Psi(\mathbf{x})\psi_j(\mathbf{x})\mu(\mathrm{d}\mathbf{x}) + \mathcal{Q}_n(t) = \\ &= \sum_{j=1}^{n} e^{\lambda_j t}\langle\psi_j^*, \Phi\rangle_\mu \langle\Psi, \psi_j\rangle_\mu + \mathcal{Q}_n(t), \end{aligned} \tag{3.45}$$

where we have introduced the L^2_μ-eigenfunction ψ_j associated with the eigenvalue λ_j of $\mathcal{K}$ and ψ_j^* the corresponding eigenfunction of $\mathcal{L} = \mathcal{K}^\dagger$. Furthermore, $\langle\cdot,\cdot\rangle_\mu$ represents the usual L^2_μ inner product. If we assume that the stochastic dynamics is generated by a quasi-compact semigroup and we consider (3.44), we can neglect with an exponentially decreasing in time error the contribution from the essential spectrum and write

$$C_{\Phi,\Psi}(t) = \sum_{j=1}^{n} e^{\lambda_j t}\langle\psi_j^*, \Phi\rangle_\mu \langle\Psi, \psi_j\rangle_\mu. \tag{3.46}$$

The above equation provides a meaningful interpretation of the RP resonances. First, we observe that the sum in all the previous expression starts from $j = 1$ rather than $j = 0$. Indeed, the spectral projector Π_0 relative to the simple RP resonance $\lambda_0 = 0$ projects onto the invariant measure. From (3.27) we can infer that the eigenfunction $\psi_0 = \text{const}$, so that $\langle\Psi, \psi_0\rangle_\mu = 0$ and the first non vanishing contribution to the correlation functions stems from $j = 1$. Each RP resonance is associated with an exponentially decaying contribution to the correlation function, with the $\mathbf{Re}\lambda_j$ determining its (inverse) exponential rate. In particular, the closer a RP resonance is to the imaginary axis, the slower the decay. The dominant mode in the spectral decomposition is given by $\gamma := |\mathbf{Re}\lambda_1|$ which is commonly known as spectral gap of the generator. When the spectral gap vanishes, one or more RP resonances touch the imaginary axis, preventing the decay of correlation functions and thus leading to a divergence of mixing properties of the system, analogously to deterministic systems [Rue09]. The imaginary part, $\mathbf{Im}\lambda_j$, of the RP resonances determine the angular frequency of the oscillations of the j-th mode. As a result, when the spectral gap shrinks to zero, the correlation functions exhibit a non decaying oscillatory behaviour. It is fundamental to point out that the spectral decomposition of correlation functions provides some degree of universality. It shows that the exponential (and oscillatory) behaviour of correlation functions is not related to the specific details of the chosen observables since it solely depends on the spectrum of the operators describing the stochastic process. We should also remark that the coefficients in the weights of the spectral decomposition (3.46) depend on the projection of the observables onto the eigenspaces of the RP resonances. Symmetries of the system and carefully chosen observables might lead to vanishing projections onto ψ_j, resulting in a null contribution from the RP resonance λ_j.

Representing somewhat the statistical dynamical skeleton of the evolution of the system, RP resonances can also be associated to power spectra computed along stochastic paths of the stochastic evolution (3.5) [CTDN20]. Given an observable $\Phi \in L^2_\mu$ its correlation spectrum is defined through the Wiener-Khintchine theorem [Wie30, Khi34] as the (one-sided) Fourier Transform of its autocorrelation function, that is

$$S_\Phi(\omega) = \int_0^{+\infty} \mathcal{C}_{\Phi,\Phi}(t) e^{i\omega t} \mathrm{d}t. \tag{3.47}$$

It is possible to relate the correlation spectrum to spectral properties of the generator of the stochastic process by introducing the resolvent $R(z,\mathcal{K}) = (z\mathrm{Id} - \mathcal{K})^{-1}$ of the Markov semigroup $P_t = e^{t\mathcal{K}}$. Given our assumptions on the existence of a unique invariant ergodic measure, the resolvent $R(z,\mathcal{K})$ is a well defined linear operator. In particular, the RP resonances can be identified as the poles of the resolvent. Recalling the characterisation of the resolvent of a strongly continuous semigroup in terms of its Laplace Transform [EN00]

$$R(z,\mathcal{K})\Phi = \int_0^{+\infty} e^{-zt} P_t \Phi \mathrm{d}t, \tag{3.48}$$

where $\Phi \in L^2_\mu$, we can write the correlation spectrum as

$$S_\Phi(\omega) = \int_{\mathbb{R}^d} \Phi(\mathbf{x}) \left[R(-i\omega; \mathcal{K})\Phi\right](\mathbf{x})\mu(\mathrm{d}\mathbf{x}). \tag{3.49}$$

We remark that the minus sign in the argument of the resolvent derives from our convention of the Fourier Transform in (3.47). The RP resonances, being poles of the resolvent, lead to a singular behaviour of the correlation spectrum at the complex frequencies $\omega_j = i\lambda_j = -\mathbf{Im}\lambda_j + i\mathbf{Re}\lambda_j = -\mathbf{Im}\lambda_j - i|\mathbf{Re}\lambda_j|$ contained in the lower half of the complex plane. The Power Spectral Density (PSD) is defined as $|S_\Phi(\omega)|$, where $\omega \in \mathbb{R}$ is taken to be real. The RP resonances closer to the imaginary axis manifest themselves as peaks of the PSD at real frequency values equals to $-\mathbf{Im}\lambda_j$, with the width of the peaks being related to $|\mathbf{Re}\lambda_j|$. When a RP resonance approaches the imaginary axis, its corresponding peak in the PSD develops into a diverging singularity. This is analogous to the critical behaviour of correlation functions when the spectral gap of the generator shrinks to zero.

If we assume that the dynamics is described by a quasi-compact Markov semigroup and that the RP are non degenerate ($m_j = 1 \forall j = 1, \dots$), we can find a simple decomposition of the correlation spectrum in terms of the RP [CTDN20] by applying a Fourier Transform to (3.46)

$$S_\Phi(\omega) = -\frac{1}{\pi} \sum_{j=1}^{n} \frac{\mathbf{Re}\lambda_j}{(\omega + \mathbf{Im}\lambda_j)^2 + (\mathbf{Re}\lambda_j)^2}. \tag{3.50}$$

The above formula provides a decomposition of the correlation spectrum in terms of Lorentzian functions, a functional form typical of applications in spectroscopy phenomena due to resonances [NZ15, LSPV05, GKWG09].

3.2 Linear Response Formulas

The goal of this section is to investigate the role of the stochastic RP resonances for the linear response properties of stochastic systems. In deterministic settings, the relationship between RP resonances and the response to perturbations of uniformly hyperbolic systems is well known, see for example [Rue09]. Along the lines of [GL22], we will below present a (mostly linear) response theory for general stochastic system with respect to both deterministic and stochastic forcings. We will consider first a deterministic perturbation of equation (3.5) as

$$\mathrm{d}\mathbf{x} = \left(\mathbf{F}\left(\mathbf{x}\right) + \varepsilon \mathbf{X}\left(\mathbf{x}\right) T\left(t\right)\right) \mathrm{d}t + \mathbf{s}(\mathbf{x})\mathrm{d}\mathbf{W}, \quad \mathbf{x} \in \mathbb{R}^d, \tag{3.51}$$

where $\varepsilon \in \mathbb{R}$ is a real number. We assume, as before, that the generator $\mathcal{K}_0$ of the unperturbed system $\varepsilon = 0$ generates an ergodic stochastic process with respect to the unique invariant measure μ_0. We assume here that the Hörmander's condition is satisfied for $\mathcal{K}_0$, thus leading to a hypoelliptic diffusion process and a smooth invariant probability measure $\mu_0(\mathrm{d}\mathbf{x}) = \rho_0(\mathbf{x})\mathrm{d}\mathbf{x}$. When $\varepsilon \neq 0$ a time modulated, through the bounded function $T(t)$, state dependent perturbation $\mathbf{X}(\mathbf{x})$ is applied to the system, making it non autonomous. The evolution of the probability distribution $\rho(\mathbf{x}, t)$ associated to (3.51) satisfies the following Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = \mathcal{L}(t)\rho(\mathbf{x}, t), \tag{3.52}$$

where the Fokker-Planck operator can be split in an *unperturbed* operator $\mathcal{L}_0$ and a *perturbation operator* $\mathcal{L}_1$ as

$$\mathcal{L}(t) = \mathcal{L}_0 + \varepsilon T(t)\mathcal{L}_1. \tag{3.53}$$

The unperturbed operator $\mathcal{L}_0$ is the adjoint of the generator $\mathcal{K}_0$ of the unperturbed process and reads, see equation (3.25),

$$\mathcal{L}_0(\cdot) = -\nabla \cdot \left(\mathbf{F}\left(\mathbf{x}\right) \cdot\right) + \frac{1}{2}\mathrm{D}^2 : \left(\mathbf{\Sigma}\left(\mathbf{x}\right) \cdot\right). \tag{3.54}$$

The perturbation operator is instead

$$\mathcal{L}_1(\cdot) = -\nabla \cdot \left(\mathbf{X}\left(\mathbf{x}\right) \cdot\right). \tag{3.55}$$

Given the non autonomous feature of the dynamics, the operator $\mathcal{L}(t)$ is time dependent and the theory of Markov semigroup does not hold exactly as described above. In particular, the (adjoint) semigroup that $\mathcal{L}(t)$ generates is given in terms of a *time-ordered* exponential operator [Gil17], rather than a simple exponential form as in (3.21). It is also worth to point out that, being a non autonomous system, the probability distributions depends in general both on the "final" phase space variable $\mathbf{x}$ at time t *and* the "initial" phase space variable $\mathbf{y}$ at time s. However, the main goal of response theory is to investigate the perturbation, given by $\mathcal{L}_1$, on the stationary process described by $\mathcal{L}_0$. One can then usually fix the initial time $s = 0$ and assume that the system is initially prepared in the stationary state described by the invariant probability distribution $\rho_0(\mathbf{x})$. The probability distribution $\rho(\mathbf{x}, t)$ can then be interpreted as in (3.8)

$$\rho(\mathbf{x}, t) = \int_{\mathbb{R}^d} \rho(\mathbf{x}, t|\mathbf{y}, 0)\rho_0(\mathbf{y})\mathrm{d}\mathbf{y}. \tag{3.56}$$

Response theory adopts a perturbative approach to the investigation of the statistical properties of the perturbed system. In particular, one seeks a perturbative expansion of the probability distribution $\rho(\mathbf{x}, t)$ as below

$$\rho(\mathbf{x}, t) = \sum_{k=0}^{+\infty} \varepsilon^k \rho_k = \rho_0(\mathbf{x}) + \varepsilon\rho_1(\mathbf{x}, t) + O(\varepsilon^2). \tag{3.57}$$

The invariant measure $\rho_0(\mathbf{x})$ satisfies $\mathcal{L}_0\rho_0 = 0$ and is considered to be known. In a linear regime one wishes to obtain an expression of ρ_1 in terms of ρ_0. Inserting the expansion (3.57) in (3.52) and gathering the linear term in ε one obtains

$$\frac{\partial \rho_1}{\partial t} = \mathcal{L}_0\rho_1 + T(t)\mathcal{L}_1\rho_0. \tag{3.58}$$

Being a linear equation in ρ_1, it is possible to find an explicit expression, for example using the variation of parameters formula, for the first correction to the perturbed measure as

$$\rho_1(\mathbf{x}, t) = \int_0^t T(s)e^{(t-s)\mathcal{L}_0}\mathcal{L}_1\rho_0(\mathbf{x})\mathrm{d}s. \tag{3.59}$$

In a similar fashion one can obtain an above linear correction, e.g. the k-th correction, in (3.57) by gathering the corresponding ε^k terms in the Fokker Planck equation [Rue98b, Luc08]. Given the correction to the invariant measure due the applied perturbation, it is now possible to evaluate the change in the statistical properties of observables of the system. We consider an observable $\Psi \in L^2_{\mu_0}$ and write its perturbed expectation value as

$$\int_{\mathbb{R}^d} \Psi(\mathbf{x})\rho(\mathbf{x},t) = \sum_{k=0}^{+\infty} \delta^k \left[\Psi\right](t) = \int_{\mathbb{R}^d} \Psi(\mathbf{x})\rho_0(\mathbf{x})\mathrm{d}\mathbf{x} + \varepsilon \int_{\mathbb{R}^d} \Psi(\mathbf{x})\rho_1(\mathbf{x},t)\mathrm{d}\mathbf{x} + O(\varepsilon^2) \tag{3.60}$$

Considering equation (3.59), the first correction to the unperturbed expectation value can be written as

$$\begin{aligned} \delta^1 \left[\Psi\right](t) &= \varepsilon \int_0^t \mathrm{d}s T(s) \int \Psi(\mathbf{x}) e^{(t-s)\mathcal{L}_0} \mathcal{L}_1 \rho_0 \mathrm{d}\mathbf{x} = \\ &= \varepsilon \int_0^t \mathrm{d}s T(s) \int \mathcal{L}_1 \rho_0 e^{(t-s)\mathcal{K}_0} \Psi(\mathbf{x}) \mathrm{d}\mathbf{x} = \\ &= \varepsilon \int_0^t \mathrm{d}s T(s) g(t-s), \end{aligned} \tag{3.61}$$

where we have introduced the function

$$g(t) = \int \mathcal{L}_1 \rho_0 e^{t\mathcal{K}_0} \Psi(\mathbf{x}) \mathrm{d}\mathbf{x}. \tag{3.62}$$

It is important to mention that the function $g(t)$ has a non-negative support because of *causality* [Rue09, Luc18], meaning that the response of the system is zero before the perturbation gets applied at time $t = 0$. In particular, one enforces causality by writing

$$G(t) = \Theta(t) g(t) = \Theta(t) \int \mathcal{L}_1 \rho_0 e^{t\mathcal{K}_0} \Psi(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{3.63}$$

where $\Theta(t)$ is the Heavyside function. Furthermore, given that the unperturbed generator $\mathcal{K}_0$ describes a stable stochastic process, its eigenvalues are contained in the left side of the complex plane. One might then assume that the function $G(t)$ decays sufficiently quick so that the time integration in (3.61) can be extended to infinite positive times. Considering both *causality* and this last remark one can then write

$$\delta^1 \left[\Psi\right](t) = \varepsilon \int_{-\infty}^{+\infty} \mathrm{d}s T(s) G(t-s) = \varepsilon \left(T \star G\right)(t), \tag{3.64}$$

where $\star$ is the convolution product. From the previous expression it is clear that the function $G(t)$, describing the response of the system, would take the role of a Green function [Luc08]. We remark that the knowledge of the Green Function would provide complete information about the response of the system, regardless the time modulation $T(t)$. It is worth to point out that (3.63) represents a version of the Fluctuation Dissipation Theorem for a generic non equilibrium stochastic system. Indeed, considering (3.40), we can write

$$G(t) = \Theta(t)\int \mathcal{L}_1\rho_0 e^{t\mathcal{K}_0}\Psi(\mathbf{x})\mathrm{d}\mathbf{x} = \Theta(t)\int \frac{\mathcal{L}_1\rho_0}{\rho_0(\mathbf{x})} e^{t\mathcal{K}_0}\Psi(\mathbf{x})\mu_0\left(\mathrm{d}\mathbf{x}\right) = \Theta(t)C_{\Psi,\psi}(t). \tag{3.65}$$

The Green function can be written as a correlation function between the considered observable $\Psi(\mathbf{x})$ and a suitable function $\psi(\mathbf{x}) = \frac{\mathcal{L}_1\rho_0}{\rho_0}$. We remark that in order to obtain this result it was fundamental to assume that the unperturbed invariant measure is smooth, i.e. $\mu_0(\mathrm{d}\mathbf{x}) = \rho_0(\mathbf{x})\mathrm{d}\mathbf{x}$. Below we want to show how the same formalism is able to encompass situations where the perturbation field acts on the diffusive part of equation (3.5) rather than on its deterministic part. We consider the following stochastic differential equation

$$\mathrm{d}\mathbf{x} = \mathbf{F}\left(\mathbf{x}\right)\mathrm{d}t + \left(\mathbf{s}(\mathbf{x}) + \varepsilon\mathbf{\Gamma}\left(\mathbf{x}\right)T\left(t\right)\right)\mathrm{d}\mathbf{W}, \quad \mathbf{x}\in\mathbb{R}^d, \tag{3.66}$$

where $\mathbf{\Gamma}:\mathbb{R}^d\to\mathbb{R}^{d\times q}$ is a perturbation to the matrix $\mathbf{s}(\mathbf{x})$ giving the structure to the noise feature of the dynamics. Evaluating the associated Fokker Planck equation as before one obtains

$$\frac{\partial\rho}{\partial t} = \mathcal{L}(t)\rho, \tag{3.67}$$

where $\mathcal{L}(t) = \mathcal{L}_0 + \varepsilon T(t)\mathcal{L}_1$. The unperturbed operator $\mathcal{L}_0$ is again given by (3.54), whereas the perturbation operator is

$$\mathcal{L}_1\left(\cdot\right) = \frac{1}{2}\mathrm{D}^2 : \left(\left(\mathbf{s}\mathbf{\Gamma}^T + \mathbf{\Gamma}\mathbf{s}^T\right)\cdot\right). \tag{3.68}$$

We remark that the above equation derives from the first order expansion of the perturbed diffusion matrix $\left(\mathbf{s}(\mathbf{x}) + \varepsilon\mathbf{\Gamma}\left(\mathbf{x}\right)\right)\left(\mathbf{s}(\mathbf{x}) + \varepsilon\mathbf{\Gamma}\left(\mathbf{x}\right)\right)^T = \mathbf{s}\mathbf{s}^T + \varepsilon T(t)\left(\mathbf{s}\mathbf{\Gamma}^T + \mathbf{\Gamma}\mathbf{s}^T\right) + O(\varepsilon^2)$. All the previous results, including the definition of the Green function $G(t)$ (3.63), the convolution expression (3.64) and the fluctuation dissipation theorem (3.65), are still valid. One has to carefully choose the correct perturbation operator $\mathcal{L}_1$ depending on what part of the dynamics is being perturbed. We also observe that the observable ψ in the fluctuation dissipation theorem (3.65) is different depending whether one is perturbing the deterministic or stochastic part of the dynamics. Moreover, we note that in the case of a stochastic perturbation $\mathbf{\Gamma}$ such that $\mathbf{s}\mathbf{\Gamma}^T + \mathbf{\Gamma}\mathbf{s}^T = 0$, in particular in the case of an unperturbed deterministic dynamics ($\mathbf{s} = 0$), the first correction to the response of the system is at second order ε^2 [Luc12, Abr17]. As a last remark, we mention that, given the linearity of the problem, a sort of superposition principle holds where one can separately study the effect of a deterministic and stochastic forcing and then add them up to to evaluate the impact of a combined perturbation [GL22].

3.3 Spectral Decomposition of Susceptibilities

Historically, the scientific context in which Linear Response Theory was developed was related to optics phenomena. For evident reasons, the response of the system was not investigated in the time domain by looking at the Green function $G(t)$ but rather in the dual frequency domain. In this framework, emphasis is given to very general *causality* properties of the system leading to integral dispersion relations, the Kramers-Kronig relations that link the real and imaginary part of the response in the frequency domain to external radiation. We will not here report these classical results and we refer the interested read to [PVA99, LSPV05]. Furthermore, in chapter 4 we will investigate a modified version of the Kramers-Kronig relations for interacting systems exhibiting phase transitions. Nevertheless, the analysis of the response of the system in the frequency domain represents still a very powerful and general tool to study the effect of time modulated perturbations to both equilibrium and nonequilibrium systems. Along these lines we define the (dynamic) susceptibility $\chi(\omega)$ of the system as the Fourier transform of the Green Function

$$\chi(\omega) = \int_{-\infty}^{+\infty} G(t) e^{i\omega t} \mathrm{d}t. \tag{3.69}$$

The first order correction in the frequency domain to the expectation value of the observable Ψ is then, by taking the Fourier Transform of (3.64),

$$\delta^1[\Psi](\omega) = \varepsilon\chi(\omega)T(\omega), \tag{3.70}$$

where we have used, for simplicity, an "abuse" of notation such that $f(\omega) = \int_{-\infty}^{+\infty} f(t)e^{i\omega t}\mathrm{d}t$ represents the Fourier Transform of the function f. The above formula represents a fundamental result of Linear Response Theory and establishes that the change of the expectation value of the observable Ψ at frequency ω is simply given by the input signal at frequency ω, $T(\omega)$, modulated by the susceptibility $\chi(\omega)$ of the system. In particular, it is well known that in a linear response regime there is no creation of new harmonics in the output signal $\delta^1[\Psi]$. We remark that the susceptibility fully determines the linear response properties of the system in the frequency domain. The goal of this section is to show that a spectral decomposition of the susceptibility $\chi(\omega)$ in terms of RP resonances holds. First, we will derive an analogous result for the Green Function. We recall the spectral decomposition of the Markov semigroup (3.36) and (3.37) and apply it to the unperturbed semigroup $e^{t\mathcal{K}_0}$ in (3.63), yielding

$$G(t) = \Theta(t)\sum_{j=1}^{n}\sum_{k=0}^{m_j-1}\alpha_j^{(k)}\frac{t^k}{k!}e^{\lambda_j t}, \tag{3.71}$$

where the coefficients

$$\alpha_j^{(k)} = \int_{\mathbb{R}^d}\mathcal{L}_1\rho_0\left(\mathcal{K}_0 - \lambda_j\mathrm{Id}\right)^k\Pi_j\Psi(\mathbf{x})\mathrm{d}\mathbf{x} \tag{3.72}$$

stem from the projection of the observable Ψ on the RP eigenspaces of the unperturbed generator. Furthermore, the coefficients $\alpha_j^{(k)}$depends on the details of the forcing through the perturbation operator $\mathcal{L}_1$. In writing (3.71) we have assumed that continuous part of the spectrum does not contribute to the decomposition. We remark that this assumption is attained (with an exponentially small in time mistake, see (3.39)) for quasi compact semigroups $e^{t\mathcal{K}_0}$or if we assume that the observable has null projection on this part of the spectrum. Again, it is worth to point out that the number n is not necessarily finite and that the summation starts from $j = 1$ because of the same reasons described right below (3.46) . If we take the Fourier

Transform of (3.71) we obtain a spectral decomposition of the susceptibility of the system $\chi(\omega)$[2]

$$\chi(\omega) = \sum_{j=1}^{n} \sum_{k=0}^{m_j-1} \frac{\alpha_j^{(k)}}{(i\omega + \lambda_j)^{k+1}} \tag{3.73}$$

As in a deterministic setting [Rue86], the function $\chi(\omega)$ can be meromorphically extended (that is extended to a holomorphic function apart from isolated poles) in the lower half in the complex plane in a strip given by $\omega_{ess} < \mathbf{Im}\omega \leq 0$. The RP resonances of the generator of the unperturbed generator $\mathcal{K}_0$ introduce singularities $\omega_j = i\lambda_j$ in the susceptibility of the system. Such resonances manifest themselves in an enhanced response of the system to those (real) frequencies located around $\omega \approx \mathbf{Im}\lambda_j$ for those RP resonances that are close enough to the imaginary axis, that is for the RP resonances such that $\mathbf{Re}\lambda_j$ is small enough. Instead, the RP resonances far away from the imaginary axis, that is with very negative real part, will provide a continuum background for the susceptibility, thus leading to a very mild and smooth response. The contribution coming from the essential spectrum would yield a similar effect and would be difficult to untangle its contribution from the one stemming from RP resonances far away from the imaginary axis [CTDN20]. We remark that the spectral decomposition of the susceptibility(3.73) indicates that the resonances in the linear response properties of the system uniquely depend on universal properties of the dynamics, such as the spectral properties of the Markov Semigroup (Transfer Operator). In particular, the details of the applied forcing (or of the arbitrarily chosen observable) do not affect the resonant behaviour of the response, as they only contribute to the weights $\alpha_j^{(k)}$of the spectral decomposition. It is worth to point out that, in the presence of symmetries in the dynamics and for carefully chosen observables and applied forcings, some weights could identically vanish, $\alpha_j^{(k)} \equiv 0$, thus masking the effect of the relative resonance λ_j.

[2]We remark that this formula slightly differs from [GL22] because of our definition of the Fourier Transform adopts the opposite sign.

3.3.1 Critical Transitions

The Ruelle Pollicott resonances constitute the relevant dynamical features of both deterministic and stochastic systems as they determine statistical properties such as correlation functions (3.46), power spectra (3.49) and dynamic susceptibilities (3.73). Assuming that the weight α_1^k does not vanish, the dominant exponential mode in the spectral decomposition of such statistical properties derives from λ_1. In other words, the spectral gap $\gamma = |\mathbf{Re}\lambda_1|$ determines both the rate of mixing properties and the sensitivity to perturbations of the system. Indeed, if we assume that a δ-like in time perturbation $T(t) = \delta(t)$, i.e. a perturbation that excites all frequencies, is applied to the system, the response of the observable Ψ is given to first order, see (3.64) and (3.46),

$$\delta^1 [\Psi] (t) = \varepsilon G(t) \approx \varepsilon \Theta(t) \alpha_1 e^{\lambda_1 t} + c.c. = \varepsilon \Theta(t) \alpha_1 e^{\gamma t} e^{i \mathbf{Im} \lambda_1 t} + c.c. \tag{3.74}$$

where we assumed $m_1 = 1$ and that the contribution of the other RP resonances can be neglected since they are further away from the imaginary axis. If the spectral gap γ is very small, the linear response $\delta^1 [\Psi] (t)$ will take a very long time, of the order of $\tau \approx \frac{1}{\gamma}$, to decay to zero for any observable Ψ of the system. In physical terms, a small spectral gap correspond to a system where the negative feedbacks are very weak, leading to a low resilience to external stimuli. High dimensional complex systems might incur in critical transitions, commonly characterised by vast, sudden and potentially dire changes of the state of the system. The mathematical theory of the RP resonances we have developed above provides a suitable framework to investigate such critical phenomena for finite dimensional systems, both deterministic and stochastic. Critical transitions can be associated to settings where the spectral gap of the generator $\mathcal{K}_0$ of the unperturbed system becomes vanishingly small as a result of the RP resonances touching the imaginary axis. Consequently, correlation properties and power spectra acquire a singular behaviour due to the loss of mixing properties as the transition is approached [TCND19, TLLD18, GNPT95]. Since there is a one-to-one correspondence between the radius of expansion of linear response theory and the spectral gap of the transfer operator [LG06, Luc16], near critical transitions the linear response breaks down and one finds rough dependence of the system properties on its parameters [CNK$^+$14, TLD18]. Indeed, settings where the RP resonances touch the imaginary axis correspond to settings where the poles of the susceptibility $\chi(\omega)$ approach, from below, the real axis of the complex frequency plane. At a critical transition one observes a divergence of the response of the system due to the negative feedbacks becoming more and more inefficient as shown by (3.74) when the spectral gap $\gamma \to 0^+$. This situation correspond to an infinite relaxation time of the response of the system to perturbations.

Chapter 4

Linear Response Theory for McKean Vlasov Equation

4.1 Introduction

In chapter 2 we have introduced the notion of weakly interacting diffusions, namely ensembles of identical and exchangeable agents interacting with each other. Such class of multiagent systems includes a very rich variety of applications, ranging from cooperation [Daw83], synchronization [ABPV^{+}05] to systemic risk [GPY13] and consensus opinion formation [WLEC17, GPY17]. See also section 2.1 for a more complete review of applications of weakly interacting diffusions in the natural and social sciences. It is well known that one can pass to the limit as the number of agents goes to infinity, i.e. the thermodynamic or *mean field* limit. In particular, in this limit the evolution of the empirical measure is described by a nonlinear, nonlocal Fokker-Planck equation, the *McKean-Vlasov Equation*. An important feature of weakly interacting diffusions is that in the thermodynamic limit they can exhibit phase transitions [CP10, Tam84]. In this framework, phase transitions are characterized in terms of exchange of stability of non-unique stationary states for the McKean-Vlasov equation at critical settings of the parameters of the system. In the case of equilibrium systems, such stationary states are associated with critical points of a suitably defined energy landscape, see equation (2.28). As a paradigmatic example, for the Kuramoto model of nonlinear oscillators, the uniform distribution on the torus, corresponding to a non synchronised state, becomes unstable at the critical noise strength and stable localized stationary states emerge (phase-locking), leading to a synchronization phase transition [PKRK03].

In Chapter 3 we have formulated a general Response Theory for finite dimensional systems by using mathematical tools deriving from a functional analysis approach to the study of partial differential equations, such as the Fokker-Planck equation. In particular, we have illustrated how the concept of Ruelle Pollicott resonances, firstly introduced for deterministic systems [Rue98a, Rue09] can be extended to stochastic settings. This framework allows to investigate the statistical properties of physical systems by studying spectral properties of suitable transfer operators that govern the time evolution of expectation values of observables of the system, or, equivalently, the time evolution of probability measures.

In particular, critical transitions appear when the spectral gap of the transfer operator of the unperturbed system becomes vanishingly small, as a result of the Ruelle-Pollicott resonances touching the imaginary axis. Since there is a one-to-one correspondence between the radius of expansion of linear response theory and the spectral gap of the transfer operator, near critical transitions the linear response breaks down. Indeed, the susceptibility of the system acquires a singular behaviour due to poles approaching the real axis of frequencies, see (3.73) and (3.74). At a critical transition, singularly diverging resonances appear in the response of the system for real frequencies, leading to a blow up of response properties.

The main goal of this chapter is to investigate phase transitions for weakly interacting diffusions by looking at the response of the infinite dimensional mean field dynamics to weak external perturbations. In particular, we associate the nearing of a phase transition with the setting where a very small cause leads to very large effects. In other words, as in the case of critical transitions for finite dimensional systems, we associate phase transitions of the thermodynamic limit of the interacting agents to the breakdown of linear response properties and the development of non analytical behaviour of response functions (susceptibilities).

We remark that linear response theory has long been studied in detail for diffusion processes and very strong rigorous results have been obtained in this direction [HM10, DD10]. These results can be applied to the McKean-Vlasov equation *in the absence of phase transitions* to justify rigorously linear response theory and to establish fluctuation-dissipation results.

As mentioned in Chapter 1, phase transitions are usually defined by

a. identifying an order parameter, i.e. a suitable observable of the system able to capture some degree of collective behaviour of the possibly high dimensional system

b. verifying that in the thermodynamic limit, for some value of the parameter of the system, the properties of such an order parameter undergo a sudden change

It should be emphasized, however, that, it is not always possible to identify an order parameter, in particular for high dimensional nonequilibrium systems with no particular symmetries. The way we define phase transitions comes from a somewhat complementary viewpoint, which aims at clarifying analogies and differences with respect to the case of critical transitions. We remark that, by adopting a response theory perspective, we can identify critical settings in terms of universal properties of the dynamics of the system, stemming from the spectral properties of transfer operators, rather than from the specific details of the dynamics, applied forcings or of the observables under investigation.

Moreover, the approach we adopt here is, in spirit, along the same lines of the work of Sornette and collaborators that have highlighted the importance of separating the effects of endogeneous vs. exogenerous processes in determining the dynamics of a complex system and, especially, in defining the conditions conducive to crises [Sor06], and have proposed multiple applications in the natural- see, e.g. [HSG03] - as well as the social - see, e.g., [Sor03] - sciences. The existence of a relationship between the response of the system to exogeneous perturbations and the decorrelation due to endogenous dynamics is interpreted as resulting from a fluctuation-dissipation relation-like properties which is at the core of a linear response theory approach, see chapter 3. Finally, Sornette and collaborators have also emphasized the importance of memory effects especially in the context of endogenous dynamics [SH03, WSS18], which we also find by investigating the response of the infinite dimensional system of interacting agents, see (4.11).

While our viewpoint and methods are different from theirs, what we pursued here shares similar goals and delves into closely related concepts.

The main results of this chapter can be summarized as follows:

- We derive linear response formulas for the thermodynamic limit of coupled identical systems and Kramers-Kronig relations and sum rules for the related susceptibilities;
- We state conditions leading to phase transitions as opposed to the classical scenario of critical transitions for finite dimensional systems in terms of spectral properties of suitable operators;
- We derive the corrections to the standard Kramers-Kronig relations and sum rules occurring at a phase transition transition setting;
- We clarify, through the use of functional analytical arguments, why no divergence of the integrated autocorrelation time is expected for microscopic degrees of freedom in the case of phase transitions, whereas the opposite holds in the case of critical transitions. We will re-examine this fundamental issue in chapter 5 by providing a link between a *microscopic* and *macroscopic* observables.

4.2 Linear Response Formulas

We consider a system of weakly interacting diffusions, i.e. an ensemble of N exchangeable interacting M-dimensional systems whose dynamics is described by the following stochastic differential equations:

$$\mathrm{d}x_i^k = F_{i,\alpha}(\mathbf{x}^k)\mathrm{d}t - \frac{\theta}{N}\sum_{l=1}^{N}\partial_{x_i^k}\mathcal{U}\left(\mathbf{x}^k - \mathbf{x}^l\right)\mathrm{d}t + \sigma\hat{s}_{ij}(\{\mathbf{x}^l\}_{l=1}^{N})\mathrm{d}W_j, \quad k = 1,\dots,N \quad i = 1,\dots,M$$
$$\mathbf{x}^k(t=0) \sim \rho_{in}(\mathbf{x}^k), \tag{4.1}$$

where $\mathbf{F}_\alpha(\mathbf{x})$ is a smooth vector field, possibly depending on a parameter (or a set of parameters) α. Additionally, $\mathrm{d}W_i$, $i = 1, \ldots, p$ are independent Brownian motions (the Ito convention is used); $\hat{s}_{ij}$ is the N-particle volatility matrix defined in equation (2.2) and the parameter $\sigma > 0$ controls the intensity of the stochastic forcing. We consider here a chaotic initial condition for the weakly interacting diffusions, namely each of the agents is distributed, at time $t = 0$, according to the same initial distribution ρ_{in}. Additionally, the N particles undergo a global, all-to-all coupling given by the pairwise interaction potential $\mathcal{U}(\mathbf{x})$. In the following chapters we will consider quadratic (Curie-Weiss) interactions, i.e. a quadratic potential $\mathcal{U}(\mathbf{x}) = |\mathbf{x}|^2/2$. The coefficient θ modulates the intensity of such a coupling, which attempts at synchronising all systems by nudging them to the center of mass $\bar{x} = \frac{1}{N}\sum_{k=1}^{N} \mathbf{x}^k$. In particular, if $\theta = 0$, the N systems are decoupled. We remark that, while the specific details of our linear response calculations, such as (4.24) and (4.25), depend on the assumption of a quadratic potential, the Markov semigroup approach to stability properties and linear response features of weakly interacting diffusions is far more general. Indeed, an interesting link between the existence of phase transitions and the spectral properties of suitable mean field operators has indeed been established for weakly interacting diffusions in statistical equilibrium settings [DGPS23]. We are interested in investigating the thermodynamic limit of the ensemble of interacting agents(4.1). Therefore we consider the empirical measure

$$X_N\left(t; A\right) = \frac{1}{N}\sum_{k=1}^{N} \mathbf{1}_A\left(\mathbf{x}^k\left(t\right)\right), \tag{4.2}$$

where $\mathbf{1}_A(\cdot)$ is the indicator function associated to a measurable set $A \subset \mathbb{R}^M$, where $\mathbb{R}^M$ is the single-particle phase space. It is possible to use martingale techniques [Daw83, Szn89, Oel84, DG87b], see section 2.4 for further details on *propagation of chaos* properties, to show that the empirical measure converges weakly to a measure $\mu(\mathrm{d}\mathbf{x}) = \rho(\mathbf{x}, t)\mathrm{d}\mathbf{x}$ where the smooth probability density distribution $\rho(\mathbf{x}, t)$ satisfies the following McKean Vlasov equation, which is a nonlinear and nonlocal Fokker-Planck equation

$$\begin{aligned}\frac{\partial \rho(\mathbf{x},t)}{\partial t} &= -\nabla \cdot \left[\rho(\mathbf{x},t)\left(\mathbf{F}_\alpha(\mathbf{x}) - \theta \nabla \mathcal{U} \star \rho\right)\right] + \frac{\sigma^2}{2}\tilde{\Delta}\rho(\mathbf{x},t) = \\ &= -\nabla \cdot \left[\rho(\mathbf{x},t)\left(\mathbf{F}_\alpha(\mathbf{x}) - \theta\left(\mathbf{x} - \langle \mathbf{x}\rangle (t)\right)\right)\right] + \frac{\sigma^2}{2}\tilde{\Delta}\rho(\mathbf{x},t) = \\ &:= \mathcal{L}^{\alpha,\theta,\sigma}_{\langle \mathbf{x}\rangle}\rho(\mathbf{x},t),\end{aligned} \tag{4.3}$$

with $\langle \mathbf{x}\rangle = \langle \mathbf{x}\rangle(t) = \int \mathbf{x}\rho(\mathbf{x},t)\mathrm{d}\mathbf{x}$ and $\star$ denoting the convolution operator. Additionally, we have that $\tilde{\Delta}$ is a linear diffusion operator such that $\tilde{\Delta}\rho(\mathbf{x},t) = \mathrm{D}^2 : (\Sigma(\mathbf{x})\rho(\mathbf{x},t)) = \sum_{i=1}^M \sum_{j=1}^M \partial_{x_i}\partial_{x_j}\left(\Sigma_{ij}(\mathbf{x})\rho(\mathbf{x},t)\right)$, which coincides with the standard M-dimensional Laplacian ($\tilde{\Delta} = \Delta$) if the single particle diffusion matrix $\mathbf{\Sigma}(\mathbf{x}) = \mathbf{s}(\mathbf{x})\mathbf{s}(\mathbf{x})^T$ is the identity matrix, where $\mathbf{s}(\mathbf{x})$ represents the single particle volatility matrix, that is the diagonal blocks of the matrix $\hat{\mathbf{s}}$ defined in (2.2). If $\sigma = 0$, we are considering a nonlinear Liouville equation. In general, we assume that $\sigma > 0$ and that the system of stochastic differential equations (4.1) describes a hypoelliptic diffusion process so that the probability distributions (including the stationary distributions) given by (4.1) are smooth. We observe that the operator $\mathcal{L}^{\alpha,\theta,\sigma}_{\langle \mathbf{x}\rangle}$ in the last line of (4.3) is the nonlinear integro-differential McKean-Vlasov operator $\mathcal{L}_{\rho(\mathbf{x},t)}$ defined in (2.16). For this choice of interaction potential the operator $\mathcal{L}_{\rho(\mathbf{x},t)}$ depends on the probability distribution $\rho(\mathbf{x},t)$ only through its first moment $\langle \mathbf{x}\rangle$, hence the notation for the subscript. We also observe that $\mathcal{L}^{\alpha,\theta,\sigma}_{\langle \mathbf{x}\rangle}$ depends on the parameters of the system. However, in order to avoid cumbersome notation, we will write $\mathcal{L}^{\alpha,\theta,\sigma}_{\langle \mathbf{x}\rangle} = \mathcal{L}_{\langle \mathbf{x}\rangle}$ and reverse to the full notation once such parameter dependence will be relevant.

We denote with $\rho_0(\mathbf{x}) = \rho_0(\mathbf{x};\alpha,\theta,\sigma)$ a reference invariant measure of the system. Since we are considering a system with an infinite number of particles, such an invariant measure needs not be unique. Specifically, in statistical equilibrium settings, it is possible to obtain a characterisation of invariant measures of the system in terms of convexity properties of the dynamics. Indeed, if the one particle volatility matrix $\mathbf{s}(\mathbf{x})$ is proportional to the identity and $\mathbf{F}_\alpha(\mathbf{x}) = -\nabla V_\alpha(\mathbf{x})$ and $V_\alpha(\mathbf{x})$ is not convex, thus allowing for more than one local minimum, for a given value of θ the system undergoes a phase transition for sufficiently weak noise. We remark that the invariant measure depends on the values of α, θ, σ and is characterised by a constant first moment $\langle \mathbf{x}\rangle(t) = \langle \mathbf{x}\rangle_0^{\alpha,\theta,\sigma} = \langle \mathbf{x}\rangle_0$. As a result, in the stationary state described by $\rho_0(\mathbf{x})$ we can define the "parametrised" operator

$$\mathcal{L}_{\langle \mathbf{x} \rangle_0}\psi = -\nabla \cdot \left[\psi(\mathbf{x})\left(\mathbf{F}_\alpha(\mathbf{x}) - \theta\left(\mathbf{x} - \langle \mathbf{x} \rangle_0\right)\right)\right] + \frac{\sigma^2}{2}\tilde{\Delta}\psi = 0, \tag{4.4}$$

where $\psi = \psi(\mathbf{x})$ is a smooth function. $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$ is a linear differential (Fokker-Planck) operator, with the invariant measure $\rho_0(\mathbf{x})$ being its eigenvector with vanishing eigenvalue, i.e. $\mathcal{L}_{\langle \mathbf{x} \rangle_0}\rho_0 = 0$. Such operator can be interpreted as the adjoint of the generator - in the stationary state $\rho_0(\mathbf{x})$- of time translation of smooth observables in the mean field limit, see the discussion at the end of section 2.2 and [Fra04]. In other words, the operator $\mathcal{L}^\dagger_{\langle \mathbf{x} \rangle_0} = \mathcal{K}_{\langle \mathbf{x} \rangle_0}$ determines the single particle correlation properties in the mean field stationary state $\rho_0(\mathbf{x})$.

Taking inspiration from [OY12, PR14], we now study the impact of perturbations on the invariant measure $\rho_0(\mathbf{x})$ of the McKean Vlasov equation. In particular, we follow and generalise the results presented in [Fra04]. We consider a perturbation of the deterministic part of the dynamics in (4.3) by setting $\mathbf{F}_\alpha(\mathbf{x}) \to \mathbf{F}_\alpha(\mathbf{x}) + \varepsilon\mathbf{X}(\mathbf{x})T(t)$, where, as in Chapter 3, we consider a time modulated state dependent perturbation and $\varepsilon \in \mathbb{R}$ is a real number. We investigate the change in the statistical properties of the non-autonomous perturbed system by considering an expansion of the probability distribution as

$$\rho(\mathbf{x}, t) = \sum_{k=0}^{+\infty} \varepsilon^k \rho_k = \rho_0(\mathbf{x}) + \varepsilon\rho_1(\mathbf{x}, t) + O(\varepsilon^2). \tag{4.5}$$

Inserting the above perturbative expansion in (4.3) we obtain, at linear order in ε, an equation for the first correction to the unperturbed invariant measure

$$\begin{aligned}
\frac{\partial \rho_1(\mathbf{x}, t)}{\partial t} &= \mathcal{L}_{\langle \mathbf{x} \rangle_0}\rho_1 + T(t)\mathcal{L}_1\rho_0 - \theta\nabla \cdot \left(\rho_0(\mathbf{x}) \int \mathbf{y}\rho_1(\mathbf{y}, t)\mathrm{d}\mathbf{y}\right) = \\
&= \mathcal{L}_{\langle \mathbf{x} \rangle_0}\rho_1 + T(t)\mathcal{L}_1\rho_0 - \theta\nabla \cdot \left(\langle \mathbf{x} \rangle_1(s)\rho_0(\mathbf{x})\right) = \\
&:= \tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}\rho_1 + T(t)\mathcal{L}_1\rho_0,
\end{aligned} \tag{4.6}$$

where we have introduced the perturbation operator

$$\mathcal{L}_1 = -\nabla \cdot (\mathbf{X}(\mathbf{x}) \cdot\ ) \tag{4.7}$$

that depends on the state dependent forcing $\mathbf{X}(\mathbf{x})$. We remark that the linear integro-differential operator $\tilde{\mathcal{L}}_{\langle\mathbf{x}\rangle_0}$ acting on $\rho_1(\mathbf{x}, t)$ defined in the last line of the previous equation is not the operator whose zero eigenvector is the unperturbed invariant measure. In particular, this operator does not determine the mean field single particle correlation properties of the system, at variance with the response theory developed in section 3.2 . We observe that the correction to the unperturbed operator $\mathcal{L}_{\langle\mathbf{x}\rangle_0}$ is proportional to θ and emerges as a result of the nonlinearity of the McKean-Vlasov equation due to the interactions among the agents. In particular, for a decoupled system $\theta = 0$, one would recover the response properties of the single agents. We will further discuss the operator $\tilde{\mathcal{L}}_{\langle\mathbf{x}\rangle_0}$ in section 4.2.1 below. From (4.3) one obtains a formal solution for the perturbation to the invariant measure as

$$\begin{aligned}\rho_1(\mathbf{x}, t) &= \int_0^t T(s) e^{(t-s)\mathcal{L}_{\langle\mathbf{x}\rangle_0}} \mathcal{L}_1 \rho_1(\mathbf{x}) \mathrm{d}s - \theta \int_0^t e^{(t-s)\mathcal{L}_{\langle\mathbf{x}\rangle_0}} \nabla \cdot (\rho_0(\mathbf{x}) \langle\mathbf{x}\rangle_1(s))\, \mathrm{d}s = \\ &= \int_0^t T(s) e^{(t-s)\mathcal{L}_{\langle\mathbf{x}\rangle_0}} \mathcal{L}_1 \rho_1(\mathbf{x}) \mathrm{d}s - \theta \sum_{k=1}^{M} \int_0^t \langle x_k\rangle_1(s) e^{(t-s)\mathcal{L}_{\langle\mathbf{x}\rangle_0}} \partial_{x_k} \rho_0(\mathbf{x}) \mathrm{d}s.\end{aligned} \tag{4.8}$$

The statistical properties of any observable $\Psi(\mathbf{x})$ of the system are determined by

$$\begin{aligned}\int \Psi(\mathbf{x}) \rho(\mathbf{x}, t) \mathrm{d}\mathbf{x} = \sum_{k=0}^{+\infty} \delta^k [\Psi] &= \int \Psi(\mathbf{x}) \rho_0(\mathbf{x}) \mathrm{d}\mathbf{x} + \varepsilon \int \Psi(\mathbf{x}) \rho_1(\mathbf{x}, t) \mathrm{d}\mathbf{x} + O(\varepsilon^2) = \\ &:= \langle\Psi\rangle_0 + \varepsilon \langle\Psi\rangle_1 + O(\varepsilon^2),\end{aligned} \tag{4.9}$$

where we have defined $\langle\cdot\rangle_0$ and $\langle\cdot\rangle_1 = \langle\cdot\rangle_1(t)$ as the expectation value with respect to $\rho_0(\mathbf{x})$ and $\rho_1(\mathbf{x}, t)$ respectively. In the following, we will consider the observable $\Psi = x_i$ for $i = 1, \dots, M$. We will show that critical phenomena for weakly interacting diffusions can be related to spectral properties of either $\mathcal{L}_{\langle\mathbf{x}\rangle_0}$ or $\tilde{\mathcal{L}}_{\langle\mathbf{x}\rangle_0}$ and, as such, do not depend on the choice of the selected observable, leading to some degree of universality. In a linear response regime the shift of the expected value of observable x_i due to the perturbation is

$$
\begin{aligned}
\langle x_i \rangle_1 &= \int_0^t \mathrm{d}s T(s) \int x_i e^{(t-s)\mathcal{L}_{\langle \mathbf{x} \rangle_0}} \mathcal{L}_1 \rho_0(\mathbf{x}) \mathrm{d}\mathbf{x} - \theta \sum_{k=1}^{M} \int_0^t \mathrm{d}s \langle x_k \rangle(s) \int x_i e^{(t-s)\mathcal{L}_{\langle \mathbf{x} \rangle_0}} \partial_{x_k} \rho_0(\mathbf{x}) \mathrm{d}\mathbf{x} = \\
&= \int_0^t \mathrm{d}s T(s) \int \mathcal{L}_1 \rho_0(\mathbf{x}) e^{(t-s)\mathcal{K}_{\langle \mathbf{x} \rangle_0}} x_i \mathrm{d}\mathbf{x} - \theta \sum_{k=1}^{M} \int_0^t \mathrm{d}s \langle x_k \rangle(s) \int \partial_{x_k} \rho_0(\mathbf{x}) e^{(t-s)\mathcal{K}_{\langle \mathbf{x} \rangle_0}} x_i \mathrm{d}\mathbf{x},
\end{aligned}
\tag{4.10}
$$

where we remark again that $\mathcal{K}_{\langle \mathbf{x} \rangle_0}$ is the adjoint of $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$. By considering *causality* and assuming some decay properties of the integrals, see discussion after equation (3.61), we can rewrite the previous expression as:

$$
\langle x_i \rangle_1(t) = \int_{-\infty}^{\infty} \mathrm{d}s T(s) G_i(t-s) + \theta \sum_{k=1}^{M} \int_{-\infty}^{\infty} \mathrm{d}s \langle x_k \rangle_1(s) Y_{ik}(t-s), \tag{4.11}
$$

where we have introduced the *microscopic* Green functions

$$
\begin{aligned}
G_i(t) &= \Theta(t) \int \mathcal{L}_1 \rho_0(\mathbf{x}) e^{(t-s)\mathcal{K}_{\langle \mathbf{x} \rangle_0}} x_i \mathrm{d}\mathbf{x}, \\
Y_{ik}(t) &= -\Theta(t) \int \partial_{x_k} \rho_0(\mathbf{x}) e^{(t-s)\mathcal{K}_{\langle \mathbf{x} \rangle_0}} x_i \mathrm{d}\mathbf{x}.
\end{aligned}
\tag{4.12}
$$

Notwithstanding the Markovianity of the dynamics, the second term on the right hand side of (4.11) describes a memory effect in the response of the observable $\mathbf{x}$. Such a term emerges in the thermodynamic limit and effectively imposes a condition of self-consistency between forcing and response; see different yet related results obtained by Sornette and collaborators [SH03, Sor06, WSS18]. We remark that such effect is a manifestation of the aforementioned difference between the operators $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$ and $\tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}$ and is not present for finite dimensional systems. Note also that if $\mathbf{X}(\mathbf{x}) = \hat{\mathbf{v}}_k$, where $\hat{\mathbf{v}}_k$ is the unit vector in the k^{th} direction, then $G_i(t) = Y_{ik}(t)$. The *microscopic* Green functions can be written in a Fluctuation Dissipation form as a correlation functions in the unperturbed state of mean field correlation functions between the observable

x_i and other suitable observables. In particular, we can write

$$\begin{aligned} G_i(t) &= \Theta(t)\int \mathcal{L}_1\rho_0(\mathbf{x})e^{(t-s)\mathcal{K}_{\langle\mathbf{x}\rangle 0}}x_i\mathrm{d}\mathbf{x} = \Theta(t)\int \Psi(\mathbf{x})e^{(t-s)\mathcal{K}_{\langle\mathbf{x}\rangle 0}}x_i\mu_0(\mathrm{d}\mathbf{x}) = \Theta(t)C_{x_i,\Psi}(t), \\ Y_{ik}(t) &= -\Theta(t)\int \partial_{x_k}\rho_0(\mathbf{x})e^{(t-s)\mathcal{K}_{\langle\mathbf{x}\rangle 0}}x_i\mathrm{d}\mathbf{x} = \Theta(t)\int \Phi_k(\mathbf{x})e^{(t-s)\mathcal{K}_{\langle\mathbf{x}\rangle 0}}x_i\mu_0(\mathrm{d}\mathbf{x}) = \Theta(t)C_{x_i,\Phi}(t). \end{aligned} \tag{4.13}$$

where $\mu_0(\mathrm{d}\mathbf{x}) = \rho_0(\mathbf{x},t)\mathrm{d}\mathbf{x}$ is the unperturbed measure and we have introduced the observables $\Psi(\mathbf{x}) = \frac{\mathcal{L}_1\rho_0(\mathbf{x})}{\rho_0(\mathbf{x})}$ and $\Phi_k(\mathbf{x}) = -\partial_{x_k}\ln\rho_0(\mathbf{x})$. In the next sections of this Chapter we will show that the *microscopic* susceptibilities do not define, separately, the total response of the thermodynamic limit of the system, but one could interpret them as a measure of the resilience of the single agents.

4.2.1 Spectral Decomposition of the Response

We here seek a decomposition of the response properties of the system in terms of the spectral features of the unperturbed Markov semigroup generated by $\mathcal{K}_{\langle\mathbf{x}\rangle_0}$. We recall that we can always decompose a Markov semigroup that supports an invariant measure in terms of its point and essential spectrum as

$$e^{t\mathcal{K}_{\langle\mathbf{x}\rangle 0}} = \sum_{j=0}^{n} T_j(t) + \mathcal{R}_n(t), \tag{4.14}$$

where

$$T_j(t) = \left[\sum_{l=0}^{m_j-1}\frac{t^l}{l!}\left(\mathcal{K}-\lambda_j\mathrm{Id}\right)^l\right]e^{\lambda_j t}\Pi_j. \tag{4.15}$$

The summation in (4.14) runs over all the n, with n possibly infinite, stochastic Ruelle Pollicott resonances $\lambda_j \in \mathbb{C}$ with $\mathbf{Re}\lambda_j \leq 0$ and $\lambda_0 = 0 \geq \mathbf{Re}\lambda_1 \geq \mathbf{Re}\lambda_2 \geq \dots$ representing the point spectrum. The simple eigenvalue $\lambda_0 = 0$ corresponds to the unperturbed invariant distribution $\rho_0(\mathbf{x})$ and the other eigenvalues have in general a multiplicity $m_j \geq 1$. Π_j is the spectral projector onto the eigenspace relative to the Ruelle Pollicott resonance λ_j, and, in particular, Π_0 projects onto the eigenspace relative to the invariant distribution. Furthermore, the operator $\mathcal{R}_n(t)$ is the residual operator associated with the essential spectrum. The norm of this operator is related to the distance of the essential spectrum from the imaginary axis in the complex plane.

Inserting the above spectral decomposition in (4.12) yields the following decomposition of the *microscopic* Green Functions

$$
\begin{aligned}
G_i(t) &= \Theta(t) \sum_{j=1}^{n} \sum_{l=0}^{m_j-1} \alpha_{i,j}^{(l)} \frac{t^l}{l!} e^{\lambda_j t} + \mathcal{Q}_G(t), \\
Y_{ik}(t) &= \Theta(t) \sum_{j=1}^{n} \sum_{l=0}^{m_j-1} \beta_{i,j,k}^{(l)} \frac{t^l}{l!} e^{\lambda_j t} + \mathcal{Q}_Y(t),
\end{aligned}
\tag{4.16}
$$

where the coefficients are given by

$$
\begin{aligned}
\alpha_{i,j}^{(l)} &= \int \mathcal{L}_\infty \rho_0(\mathbf{x}) \left(\mathcal{K}_0 - \lambda_j \mathrm{Id}\right)^l \Pi_j x_i \mathrm{d}\mathbf{x}, \\
\beta_{i,j,k}^{(l)} &= \int \partial_{x_k} \ln \rho_0(\mathbf{x}) \left(\mathcal{K}_0 - \lambda_j \mathrm{Id}\right)^k \Pi_j x_i \mathrm{d}\mathbf{x}.
\end{aligned}
\tag{4.17}
$$

We remark that in the case the Ruelle Pollicott resonances are simple eigenvalues only the term $l = 0$ survives in the summation over all the multiplicities, and the coefficients $\alpha_{i,j}^{(l)} = \alpha_{i,j}^{(0)} = \alpha_{i,j}$ can be written as

$$
\alpha_{i,j} = \langle \psi_j^*(\mathbf{x}), x_i \rangle_0 \langle \frac{\mathcal{L}_1 \rho_0(\mathbf{x})}{\rho_0(\mathbf{x})}, \psi_j(\mathbf{x}) \rangle_0, \tag{4.18}
$$

where $\langle \cdot, \cdot \rangle_0$ is the usual $L^2(\mathbb{R}^M, \rho_0)$ inner product weighted with the invariant distribution $\rho_0(\mathbf{x})$ and ψ_j (ψ_j^*) are the $L^2(\mathbb{R}^M, \rho_0)$ eigenfunctions of the (adjoint of the) generator $\mathcal{K}_{\langle \mathbf{x} \rangle_0}$. Similarly, we can write

$$
\beta_{i,j,k} = \langle \psi_j^*(\mathbf{x}), x_i \rangle_0 \langle \partial_{x_k} \ln \rho_0, \psi_j(\mathbf{x}) \rangle_0. \tag{4.19}
$$

Note that the decomposition (4.16) does not include the term relative to the simple eigenvalue $\lambda_0 = 0$ since the coefficient

$$
\begin{aligned}
\alpha_{i,0} &= \langle \psi_0^*(\mathbf{x}), x_i \rangle_0 \langle \frac{\mathcal{L}_1 \rho_0(\mathbf{x})}{\rho_0(\mathbf{x})}, \psi_0(\mathbf{x}) \rangle_0 = \\
&= \langle \psi_0^*(\mathbf{x}), x_i \rangle_0 \langle \frac{\mathcal{L}_1 \rho_0(\mathbf{x})}{\rho_0(\mathbf{x})}, \mathbf{1}_{\mathbb{R}^M} \rangle_0 = \langle \psi_0^*(\mathbf{x}), x_i \rangle_0 \langle \frac{\mathcal{L}_1 \rho_0(\mathbf{x})}{\rho_0(\mathbf{x})} \rangle_0 = 0
\end{aligned}
\tag{4.20}
$$

vanishes for any perturbation $\mathbf{X}(\mathbf{x})$. We observe that in the above formula we have used the fact that the eigenfunction ψ_0 (the dual of the invariant measure) is a constant function, see [CTDN20] and (3.27). We now investigate the response properties in frequency space and apply the Fourier transform to (4.11), obtaining

$$\sum_{k=1}^{M} P_{ik}(\omega)\langle x_k\rangle_1(\omega) = \chi_i(\omega)T(\omega), \quad P_{ik}(\omega) = \delta_{ik} - \theta\Upsilon_{ij}(\omega), \tag{4.21}$$

where we have used the (standard) abuse of notation $f(\omega) = \int_{-\infty}^{+\infty} f(t)e^{i\omega t}\mathrm{d}t$ in defining the Fourier transform of $T(t)$ and $\langle x_j\rangle_1(t)$ and have defined the *microscopic* susceptibilities

$$\chi_i(\omega) = \int_{-\infty}^{+\infty} G_i(t)e^{i\omega t}\mathrm{d}t = \sum_{j=1}^{n}\sum_{l=0}^{m_j-1} \frac{\alpha_{i,j}^{(l)}}{(i\omega+\lambda_j)^{l+1}}, \tag{4.22}$$

and

$$\Upsilon_{ik}(\omega) = \int_{-\infty}^{+\infty} Y_{ik}(t)e^{i\omega t}\mathrm{d}t = \sum_{j=1}^{n}\sum_{l=0}^{m_j-1} \frac{\beta_{i,j,k}^{(l)}}{(i\omega+\lambda_j)^{l+1}}. \tag{4.23}$$

The functions $\chi_i(\omega)$ and $\Upsilon_{ik}(\omega)$ can be meromorphically extended (that is extended to a holomorphic function apart from isolated poles) in the lower half in the complex plane in a strip given by[1] $\omega_{ess} < \mathbf{Im}\omega \leq 0$. The stochastic Ruelle Pollicott resonances of the generator $\mathcal{K}_{\langle\mathbf{x}\rangle_0}$ of the unperturbed generator $\mathcal{K}_0$ introduce poles $\omega_j = i\lambda_j$ in the susceptibility of the system. Additionally, if ω_j is a pole, so is $-\omega_j^*$ (since, correspondingly, λ_j comes together with λ_j^*), where * denotes the complex conjugate. We remark that the susceptibilities given in (4.22) and (4.23) are instead holomorphic in the upper half complex $\omega-$plane if $\mathbf{Re}\lambda_j < 0,\ j = 1,\ldots,n$. As previously mentioned, the spectral decomposition indicates that all susceptibilities, regardless of the observable and the forcing considered, share the same poles located at $\omega_j = i\lambda_j$, since the Ruelle Pollicott resonances solely depends on the unperturbed generator $\mathcal{K}_{\langle\mathbf{x}\rangle_0}$. We also observe that in writing (4.22) and (4.23) we have supposed that the essential spectrum does not contribute to the spectral decomposition. Such a condition is exactly attained when the generator $\mathcal{K}_{\langle\mathbf{x}\rangle_0}$ generates a compact or eventually compact semigroup or when the observable x_i has no projection on the essential spectrum. Otherwise, it is asymptotically attained (with an error decreasing exponentially with time) for quasi-compact semigroups. From (4.21) we

[1] ω_{ess} represents the essential growth bound of the unperturbed Markov semigroup, a relevant spectral property related to the essential spectrum, see section 3.1.3

obtain our final result for the response of the system

$$\langle x_i\rangle_1(\omega) = \sum_{j=1}^{M} \mathbf{P}^{-1}_{ij}(\omega)\chi_j(\omega)T(\omega) = \tilde{\chi}_i(\omega)T(\omega), \tag{4.24}$$

where $\mathbf{P}^{-1}_{ij}(\omega) = (\mathbf{P}^{-1})_{ij}$ is a shorthand notation for the ij-th element of the inverse matrix $\mathbf{P}^{-1}$ and we have defined a *macroscopic susceptibility* that determines the full response of the system in frequency space

$$\tilde{\chi}_i(\omega) = \sum_{j=1}^{M} \mathbf{P}^{-1}_{ij}(\omega)\chi_j(\omega). \tag{4.25}$$

The *macroscopic* susceptibility

We observe that (4.24) describing the full response of the system generalises previous findings presented in [TKP01]. If the coupling is absent, so that $\theta = 0$, we obtain the same result as in the case of a single particle system where$\langle x_i\rangle_1(\omega) = \chi_{i,\theta=0}(\omega)T(\omega)$, since $\tilde{\chi}_{i,\theta=0}(\omega) = \chi_{i,\theta=0}(\omega)$. The effect of switching on the coupling, $\theta > 0$, is two-fold in terms of response:

- First, the microscopic function $\chi_i(\omega)$ is modified, because the unperturbed operator $\mathcal{L}^0_{\theta,\langle\mathbf{x}\rangle_0}$depends both explicitly and through the unperturbed invariant measure $\rho_0(\mathbf{x})$ on θ. Indeed, changes in the value of θ impact expectation values and mean field correlation properties. This susceptibility can be interpreted as the frequency response of the single agents, when the interaction among them is taken into account in a parametrised, static way by setting in a linear response framework $\langle\mathbf{x}\rangle(t) = \langle\mathbf{x}\rangle_0$ in the definition of the operator $\mathcal{L}_{\langle\mathbf{x}\rangle_0}$.
- More importantly, the presence of a non-vanishing value of θ introduces a memory effect, see equation (4.10), that results in a non-trivial correction with respect to the identity to the matrix $P_{ij}(\omega)$. We can interpret the function $\tilde{\chi}_i(\omega)$ as the macroscopic susceptibility, which takes fully into account, in a self-consistent way, the interaction between the systems.

We observe that (4.24) generalises the frequency-dependent version of the well-known Clausius-

Mossotti relation [Jac75, LSPV05, TT13], which connects the macroscopic polarizability of a material and the microscopic polarizability of its elementary components and provides a meaningful analogy of the susceptibility $\chi_i(\omega)$ and $\tilde{\chi}_i(\omega)$.

Below we show that the *macroscopic* susceptibility is related to spectral properties of the operator $\tilde{\mathcal{L}}_{\langle\mathbf{x}\rangle_0}$. In fact, from the last line of (4.6) we see that we can write the perturbation to the invariant distribution as

$$\rho_1(\mathbf{x}, t) = \int_0^t T(s) e^{(t-s)\tilde{\mathcal{L}}_{\langle\mathbf{x}\rangle_0}} \mathcal{L}_1 \rho_0(\mathbf{x}) \mathrm{d}s. \tag{4.26}$$

We can evaluate the change in the expectation value of observable x_i as

$$\langle x_i \rangle_1(t) = \int_{-\infty}^{+\infty} T(s) \tilde{G}_i(t-s) \mathrm{d}s = \left(T \star \tilde{G}_i\right)(t), \tag{4.27}$$

where the *macroscopic* Green Function is

$$\tilde{G}_i(t) = \Theta(t) \int \mathcal{L}_1 \rho_0(\mathbf{x}) e^{t\tilde{\mathcal{K}}_{\langle\mathbf{x}\rangle_0}} x_i \mathrm{d}\mathbf{x}. \tag{4.28}$$

Note that $\tilde{\mathcal{K}}_{\langle\mathbf{x}\rangle_0}$ cannot be interpreted as the generator of time translation for mean field smooth observables and, as such, $\tilde{G}_i(t)$ cannot be written in a Fluctuation Dissipation form. From (4.27) we can infer that the Fourier Transform of $\tilde{G}_i(t)$ is the *macroscopic* susceptibility (4.25)

$$\tilde{\chi}_i(\omega) = \int \tilde{G}_i(t) e^{i\omega t} \mathrm{d}t \tag{4.29}$$

The benefit of deriving the expression of $\tilde{\chi}_i(\omega)$ as done in the previous section lies in the possibility of separating the effect of the poles of $\chi(\omega)$ and the zeros of $P_{ij}(\omega)$. On the other hand, (4.28) represents the basis for the spectral decomposition of the *macroscopic* susceptibility of the system. Indeed, similarly to (4.14), we can write:

$$e^{t\tilde{\mathcal{K}}_{\langle\mathbf{x}\rangle_0}} = \sum_{j=0}^{n} \tilde{T}_j(t) + \tilde{\mathcal{R}}_n(t), \tag{4.30}$$

with

$$\tilde{T}_j(t) = \left[\sum_{l=0}^{m_j-1} \frac{t^l}{l!} \left(\tilde{\mathcal{K}} - \tilde{\lambda}_j \mathrm{Id}\right)^l\right] e^{\tilde{\lambda}_j t} \tilde{\Pi}_j, \tag{4.31}$$

where the symbols have the same interpretation as in the discussion after (4.14) but now refer to the operator $\tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}$. Using similar arguments as in the previous section, the *macroscopic* Green function can be written as

$$\tilde{G}_i(t) = \Theta(t) \sum_{j=1}^{n} \sum_{l=0}^{m_j-1} \tilde{\alpha}_j^{(l)} \frac{t^l}{l!} e^{\tilde{\lambda}_j t} + \mathcal{Q}_{\tilde{G}}, \tag{4.32}$$

where the coefficients $\tilde{\alpha}_j^{(l)}$ can be obtained from (4.28) and (4.31). We now apply the Fourier transform to (4.32)and obtain:

$$\tilde{\chi}_i(\omega) = \sum_{j=1}^{n} \sum_{l=0}^{m_j-1} \frac{\tilde{\alpha}_j^{(l)}}{\left(i\omega + \tilde{\lambda}_j\right)^{l+1}}, \tag{4.33}$$

where we have neglected the contribution of the essential spectrum. Comparing (4.33) and (4.25), it is clear that the poles $\tilde{\omega}_j = i\tilde{\lambda}_j$ of $\tilde{\chi}_i(\omega)$ are those of $\chi_i(\omega)$ plus those of the matrix $\mathbf{P}_{ij}^{-1}(\omega)$.

4.3 Critical phenomena for interacting systems

In this Thesis we identify critical phenomena for the ensemble of interacting agents as settings where response properties of the system break down. Equivalently, critical phenomena are characterised by the development of a singularity of the *macroscopic* susceptibility $\tilde{\chi}_i(\omega)$ for a real value of the frequency $\omega \in \mathbb{R}$. Equation (4.25), that we report below,

$$\tilde{\chi}_i(\omega) = \sum_{j=1}^{M} \mathbf{P}_{ij}^{-1}(\omega) \chi_j(\omega). \tag{4.34}$$

suggests that the singular behaviour of the *macroscopic* susceptibility $\tilde{\chi}(\omega)$ can arise either from a singularity in the microscopic susceptibility $\chi_i(\omega)$ or from zeros of $P_{ij}(\omega)$, leading to poles of $\mathbf{P}_{ij}^{-1}(\omega)$. Remarkably, these correspond to two different physical critical phenomena. We will call the former scenario *critical transition* whereas we will refer to the latter as *phase transitions*. The characteristics of both critical phenomena can be linked to the following two spectral conditions

1. The Ruelle Pollicott resonances for the unperturbed Markov semigroup have strictly negative real part, i.e. $\mathbf{Re}\lambda_j < 0$ for $j = 1, \dots, n$

2. The matrix $P_{ij}(\omega)$ is invertible and, additionally, has no zeros in the upper complex ω-plane.

We remark that if both conditions 1 and 2 are satisfied then the following condition is satisfied (and viceversa)

3. The spectral gap of $\tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}$ is finite, meaning that all the Ruelle Pollicott resonances associated with this operator have strictly negative real part

Condition 3 corresponds to non critical settings for the ensemble of interacting agents.

Critical transitions are instead characterised by a breakdown of condition 1 for a set of parameters $(\alpha, \theta, \sigma) = (\bar{\alpha}, \bar{\theta}, \bar{\sigma})$. This breakdown is due to the presence of a vanishing spectral gap for the operator $\mathcal{L}^{\bar{\alpha},\bar{\theta},\bar{\sigma}}_{\langle \mathbf{x} \rangle_0}$, and, a fortiori, for the operator $\tilde{\mathcal{L}}^{\bar{\alpha},\bar{\theta},\bar{\sigma}}_{\langle \mathbf{x} \rangle_0}$. In such a scenario, the functions $\chi_{i,\bar{\alpha},\bar{\theta}}(\omega)$ and $\tilde{\chi}_{i,(\bar{\alpha},\bar{\theta},\bar{\sigma})}(\omega)$ feature one or more poles $\omega_j \in \mathbb{R}$ in the real $\omega-$axis. In other terms, the linear response blows up for forcings having non-vanishing spectral power $|T(\omega_j)|^2$ at the corresponding frequencies. In this case the blow-up of the linear susceptibilities corresponds to an ultraslow decay of mean field correlations leading to a singularity in the integrated decorrelation time. In other terms, this type of criticality conforms to the classic framework of the theory of critical transitions [Kue11, CNK+14, TLD18, GL20, Sch09b]. We remark once again that the presence of a divergence of correlation properties does not depend on the specific functional form of the perturbation field $\mathbf{X}(\mathbf{x})$ nor on the choice of observable under investigation, whilst the properties of the response do depend in general from it.

On the other hand, phase transitions are characterised by settings where the breakdown of condition 2 for a set of parameters $(\alpha,\theta,\sigma)=(\tilde{\alpha},\tilde{\theta},\tilde{\sigma})$ is associated with the fact that the spectral gap of the operator $\tilde{\mathcal{L}}^{(\tilde{\alpha},\tilde{\theta},\tilde{\sigma})}_{\langle\mathbf{x}\rangle_0}$ vanishes, whilst the spectral gap of the operator $\mathcal{L}^{(\tilde{\alpha},\tilde{\theta},\tilde{\sigma})}_{\langle\mathbf{x}\rangle_0}$ remains finite. In this latter case, only the *macroscopic* susceptibilities $\tilde{\chi}_i^{\tilde{\alpha},\tilde{\theta},\tilde{\sigma}}(\omega)$ have one or more poles for real values of ω, whereas the functions $\chi_i^{\tilde{\alpha},\tilde{\theta},\tilde{\sigma}}(\omega)$ are holomorphic in the upper complex $\omega-$ plane. We remark that the non-invertibility of the P matrix depends on the presence of sufficiently strong coupling between the systems, which leads to them being coordinated. In this case, we do not observe a divergence of correlation properties of the single *microscopic* agents, since the divergence stems as an emergent behaviour due to the interactions among the agents. In fact, the non trivial structure of the matrix P_{ij} derives from the nonlinearity of (4.3) as a result of the thermodynamic limit $N\to+\infty$ and the presence of non vanishing interactions ($\theta\neq 0$) among the agents. Therefore, we interpret the singularities in the linear response resulting from the breakdown of condition 2 as being associated to a *phase transition* of the system. We will provide examples of phase transitions in Chapter 5.

4.3.1 Dispersion Relations far from Criticalities

We here assume that condition 3 is satisfied, that is, the system is in non critical settings. This implies that that all the Ruelle Pollicott resonances λ_j of the generator $\mathcal{K}_0$ of the unperturbed semigroup have strictly negative real part, $\mathbf{Re}\lambda_j<0$. In this situation, since $G_i(t)$ is causal, the function $\chi_i(\omega)$ is a holomorphic function in the upper complex ω-plane, with all the poles $\omega_j=i\lambda_j=-i|\mathbf{Re}\lambda_j|-\mathbf{Im}\lambda_j$ being all contained in the lower complex ω-plane far from the real axis of frequencies. We first consider the short-time behaviour $t\to 0^+$ of the Green Function $G_i(t)$. Using (4.12) and the definition of the adjoint operator for Markov Semigroups (3.26), we derive:

$$G_i(t)=\Theta(t)\left(\langle X_i(\mathbf{x})\rangle_0+\left(\sum_{k=1}^{M}\langle X_k(\mathbf{x})\partial_{x_k}F_i(\mathbf{x})\rangle_0-\theta\langle X_i(\mathbf{x})\rangle_0\right)t+o(t^2)\right) \tag{4.35}$$

As a result, the high-frequency behaviour of the susceptibility $\chi_i(\omega)$ can be written as:

$$\chi_i(\omega) = i\frac{\langle X_i(\mathbf{x})\rangle_0}{\omega} - \frac{\sum_{k=1}^{M}\langle X_k(\mathbf{x})\partial_{x_k}F_i(\mathbf{x})\rangle_0 - \theta\langle X_i(\mathbf{x})\rangle_0}{\omega^2} + o(\frac{1}{\omega^2}). \tag{4.36}$$

The causality of $G_i(t)$ implies that one can write the following "tautology" $G_i(t) = \Theta(t)G_i(t)$. By performing the Fourier transform of both sides of this identity, we obtain the following equation $\chi_i(\omega) = \frac{1}{2\pi}\left(\chi_i \star \hat{\Theta}\right)(\omega)$, where $\hat{\Theta}(\omega) = \int \Theta(t)e^{i\omega t}\mathrm{d}t = i\mathbf{P}(1/\omega) + \pi\delta(\omega)$ is the Fourier transform of $\Theta(t)$, and $\mathbf{P}\left(\frac{1}{\omega}\right)$ indicating the principal value distribution. By separating the real (**Re**) and imaginary (**Im**) parts of $\chi_i(\omega)$, the previous relations can be written in the standard Kramers-Kronig form:

$$\mathbf{P}\int_{-\infty}^{\infty}\frac{\mathbf{Re}\{\chi_i(\nu)\}}{\nu-\omega}\mathrm{d}\nu = -\pi\mathbf{Im}\{\chi_i(\omega)\}, \tag{4.37}$$

$$\mathbf{P}\int_{-\infty}^{\infty}\frac{\mathbf{Im}\{\chi_i(\nu)\}}{\nu-\omega}\mathrm{d}\nu = \pi\mathbf{Re}\{\chi_i(\omega)\}. \tag{4.38}$$

Since $G_i(t)$ is a real function of real argument t, its Fourier transform obeys the following conditions: $\chi_i(\omega) = (\chi_i(-\omega^*))^*$, where * denotes the complex conjugate. Hence, for real values of ω we have $\mathbf{Re}\{\chi_i(\omega)\} = \mathbf{Re}\{\chi_i(-\omega)\}$ and $\mathbf{Im}\{\chi_i(\omega)\} = -\mathbf{Im}\{\chi_i(-\omega)\}$. From this it follows an alternative form of the Kramers-Kronig relations [LSPV05]:

$$\mathbf{P}\int_{0}^{\infty}\frac{\mathbf{Re}\{\chi_i(\nu)\}}{\nu^2-\omega^2}\mathrm{d}\nu = -\frac{\pi}{2\omega}\mathbf{Im}\{\chi_i(\omega)\}, \tag{4.39}$$

$$\mathbf{P}\int_{0}^{\infty}\frac{\nu\mathbf{Im}\{\chi_i(\nu)\}}{\nu^2-\omega^2}\mathrm{d}\nu = \frac{\pi}{2}\mathbf{Re}\{\chi_i(\omega)\}. \tag{4.40}$$

It is then possible to derive the following sum rules:

$$\int_{0}^{\infty}\mathbf{Re}\{\chi_i(\nu)\}\mathrm{d}\nu = \lim_{\omega\to\infty}\left(\frac{\pi}{2}\omega\mathbf{Im}\{\chi_i(\omega)\}\right) = \frac{\pi}{2}\langle X_i(\mathbf{x})\rangle_0, \tag{4.41}$$

$$\int_{0}^{\infty}\frac{\mathbf{Im}\{\chi_i(\nu)\}}{\nu}\mathrm{d}\nu = \lim_{\omega\to 0}\left(\frac{\pi}{2}\mathbf{Re}\{\chi_i(\omega)\}\right) = \frac{\pi}{2}\tau_{G_j}G_i(0^+), \tag{4.42}$$

where

$$\tau_{G_i} = \frac{\int_0^{\infty} G_i(t)\mathrm{d}t}{G_i(0^+)}, \tag{4.43}$$

if $G_i(0^+) \neq 0$, is a measure of the decorrelation of the system. Additionally, if $\langle X_i(\mathbf{x})\rangle_0 = 0$, so that the imaginary part of the susceptibility decreases asymptotically at least as fast as ω^{-3} the following additional sum rules holds:

$$\int_0^\infty \nu \mathbf{Im}\{\chi_i(\nu)\}\mathrm{d}\nu = \lim_{\omega\to\infty}\left(-\frac{\pi}{2}\omega^2\mathbf{Re}\{\chi_i(\omega)\}\right) = \frac{\pi}{2}\sum_{k=1}^{M}\langle X_k(\mathbf{x})\partial_{x_k}F_i(\mathbf{x})\rangle_0. \tag{4.44}$$

We turn now our attention to the asymptotic properties for large values of ω of the matrix $P_{ij}(\omega)$. We proceed as above and consider the short time behaviour of $Y_{ij}(t)$,:

$$Y_{ij}(t) = \Theta(t)\left(\delta_{ij} + o(t)\right). \tag{4.45}$$

As a result, for large values of ω, we have that the *microscopic* susceptibility can be written as

$$\Upsilon_{ij}(\omega) = \frac{i}{\omega}\delta_{ij} + o(\omega^{-1}), \tag{4.46}$$

so that $P_{ij}(\omega) = \delta_{ij}\left(1 - i\frac{\theta}{\omega}\right) + o(\omega^{-2})$ and $\mathbf{P}^{-1}_{ij}(\omega) = \delta_{ij}\left(1 + i\frac{\theta}{\omega}\right) + o(\omega^{-2})$. The asymptotic behaviour for large values of frequencies of the *macroscopic* susceptibility can then be written as

$$\tilde{\chi}_i(\omega) = i\frac{\langle X_i(\mathbf{x})\rangle_0}{\omega} - \frac{\sum_{k=1}^{M}\langle X_k(\mathbf{x})\partial_{x_k}F_i(\mathbf{x})\rangle_0}{\omega^2} + o(\frac{1}{\omega^2}). \tag{4.47}$$

Comparing the above expression with (4.36), we observe that there is a correction in the asymptotic behaviour of the *macroscopic* susceptibility with respect to the *microscopic* one. Nonetheless, since also condition 2 is satisfied away from criticality, the matrix $P_{ij}(\omega)$ can be inverted for all values of ω in the upper complex ω-plane and the Kramers-Kronig relations (4.39)-(4.40) and the sum rules (4.41)-(4.44) apply also for $\tilde{\chi}_i(\omega)$.

4.3.2 Dispersion relations for *Phase Transitions*

In what follows, we focus on the criticalities associated with condition 2, that is *phase transitions*. We are interested here in assessing how dispersion relations change due to onset of a macroscopic collective behaviour of the system. In the following, for simplicity and with no loss of generality, we consider the parameter σ as fixed. We assume that for some reference values for $\alpha = \alpha_0$ and $\theta = \theta_0$ the system is stable, meaning that the *macroscopic* Green function $\tilde{G}_i^{\alpha_0,\theta_0}(t)$ that takes into account all the interactions among the identical systems has only positive support and is a smoothly decaying function with relevant timescales given by the Ruelle Pollicott resonances $\tilde{\lambda}_j$, see (4.32). Correspondingly, the *macroscopic* susceptibilities $\tilde{\chi}_i^{\alpha_0,\theta_0}(\omega)$, just like the *microscopic* ones, are holomorphic in the upper complex $\omega-$ plane. This implies that the entries of the matrix $\Pi_{ij}^{\alpha_0,\theta_0}(\omega)$ do not have poles in the upper complex $\omega-$ plane.

Let's now consider the following modulation of the system. We consider the protocol $(\alpha_s, \theta_s) = (\alpha_0 + \delta_\alpha(s), \theta_0 + \delta_\theta(s))$ and assume for $0 \leq s < \tilde{s}$ the system retains stability. For $(\alpha_{\tilde{s}}, \theta_{\tilde{s}}) = (\tilde{\alpha}, \tilde{\theta})$, the system loses stability as R poles ω_l with $l = 1, \ldots, R$ cross into the upper complex ω-plane (with $\mathbf{Im}\{\omega_l\} = 0$, $l = 1, \ldots, R$) for themacroscopic susceptibilities $\tilde{\chi}_i^{\tilde{\alpha},\tilde{\theta}}(\omega)$, whilst the mean field susceptibilies $\chi_i^{\tilde{\alpha},\tilde{\theta}}(\omega)$ are holomorphic in the upper complex ω-plane. The setting $(\tilde{\alpha}, \tilde{\theta})$ corresponds to a *phase transition* point.

We have that $P_{ij}^{\tilde{\alpha},\tilde{\theta}}$ does *not* have full rank for $\omega = \omega_l$, $l = 1, \ldots, R$. For such real frequencies ω_l the *macroscopic* susceptibility diverge. Indeed, we remark that the invertibility conditions of the matrix $P_{ij}(\omega)$ is intrinsic and does not depend on the applied external forcing $\mathbf{X}$, which enters, instead, only in the definition of the *microscopic* susceptibility $\chi_i(\omega)$. We interpret this as the fact that the divergence of the response is due to eminently endogenous, rather than exogeneous, processes. We also remark that $P_{ij}(\omega) = \delta_{ij} - \Upsilon_{ij}(\omega)$, where $\Upsilon_{ij}(\omega)$ can be seen as *microscopic* susceptibility for the expectation value of x_i associated with an infinitesimal change of the value of the j^{th} component of $\langle \mathbf{x} \rangle_0$, see (4.6) and (4.12). This supports the idea that $\langle \mathbf{x} \rangle$ is an appropriate order parameter for the system. We assume, for simplicity, that only simple poles are present, i.e. $m_l = 1$ for $l = 1, \ldots, R$. . In order to study the effect of phase transitions on the Kramers-Kronig relations, we then decompose the matrix $\mathbf{\Pi}^{\tilde{\alpha},\tilde{\theta}}(\omega)$ in the upper complex $\omega-$ plane as follows:

$$\Pi_{ij}^{\tilde{\alpha},\tilde{\theta}}(\omega) = \Pi_{h;ij}^{\tilde{\alpha},\tilde{\theta}}(\omega) + \sum_{l=1}^{R} \frac{\mathrm{Res}(\Pi_{ij}^{\tilde{\alpha},\tilde{\theta}}(\omega))_{\omega=\omega_l}}{\omega - \omega_l}, \tag{4.48}$$

where we have separated the holomorphic component $\Pi_{h;ij}^{\tilde{\alpha},\tilde{\theta}}(\omega)$ from the singular contributions coming from the poles ω_l, $l = 1, \ldots, R$; note that $\mathrm{Res}(f(\omega))_{\omega=\nu}$ indicates the residue of the function f for $\omega = \nu$. Note that if ω_l is a pole on the real axis, $-\omega_l$ is also a pole. Additionally, $\mathrm{Res}(f(\omega))_{\omega=\omega_l} = -\mathrm{Res}(f(\omega))^*_{\omega=-\omega_l}$, so that if $\omega_l = 0$ the residue has vanishing real part. Building on equation (4.48), the macroscopic susceptibility can then be written as:

$$\tilde{\chi}_i^{\tilde{\alpha},\tilde{\theta}}(\omega) = \Pi_{ij}^{\tilde{\alpha},\tilde{\theta}}(\omega)\chi_i^{\tilde{\alpha},\tilde{\theta}}(\omega) = \Pi_{h;ij}^{\tilde{\alpha},\tilde{\theta}}(\omega)\chi_i^{\tilde{\alpha},\tilde{\theta}}(\omega) + \sum_{l=1}^{R} \frac{\mathrm{Res}(\Pi_{ij}^{\tilde{\alpha},\tilde{\theta}}(\omega))_{\omega=\omega_l}}{\omega - \omega_l}\chi_{i,\tilde{\alpha},\tilde{\theta}}(\omega_l), \tag{4.49}$$

where the Kramers-Kronig relations given in (4.37) are then modified as follow, taking into account the extra poles along the real ω-axis:

$$\mathbf{P}\int_{-\infty}^{\infty} \mathrm{d}\nu \frac{\tilde{\chi}_i^{\tilde{\alpha},\tilde{\theta}}(\nu)}{\nu - \omega} = i\pi\tilde{\chi}_{i,\tilde{\alpha},\tilde{\theta}}(\omega) + i\pi \sum_{l=1}^{R} \frac{\mathrm{Res}(\Pi_{ij}^{\tilde{\alpha},\tilde{\theta}}(\omega))_{\omega=\omega_l}}{\omega_l - \omega}\chi_i^{\tilde{\alpha},\tilde{\theta}}(\omega_l). \tag{4.50}$$

By taking the limit $\omega \to \infty$ we can generalise the sum rule given in (4.41):

$$\int_0^{\infty} \mathrm{d}\nu \mathbf{Re}\{\tilde{\chi}_i^{\tilde{\alpha},\tilde{\theta}}(\nu)\} = \frac{\pi}{2}\langle X_i(\mathbf{x})\rangle_0 - \frac{\pi}{2}\mathbf{Im}\left\{\sum_{l=1}^{R} \mathrm{Res}(\Pi_{ij}^{\tilde{\alpha},\tilde{\theta}}(\omega))_{\omega=\omega_l}\chi_{i,\tilde{\alpha},\tilde{\theta}}(\omega_l)\right\}. \tag{4.51}$$

Instead, by taking the limit $\omega \to 0$ we can generalise the sum rule given in (4.42) as follows:

$$\int_0^{\infty} \mathrm{d}\nu \frac{\mathbf{Im}\{\tilde{\chi}_i^{\tilde{\alpha},\tilde{\theta}}(\nu)\}}{\nu} = \lim_{\omega\to 0}\left(\frac{\pi}{2}\mathbf{Re}\{\tilde{\chi}_i^{\tilde{\alpha},\tilde{\theta}}(\omega)\}\right) + \frac{\pi}{2}\mathbf{Re}\left\{\sum_{\omega_l \neq 0} \frac{\mathrm{Res}(\Pi_{ij}^{\tilde{\alpha},\tilde{\theta}}(\omega))_{\omega=\omega_l}}{\omega_l}\chi_i^{\tilde{\alpha},\tilde{\theta}}(\omega_l)\right\}. \tag{4.52}$$

where we note that the zero-frequency poles do not contribute to the second term on the right hand side.

4.3.3 Two Scenarios of Phase Transitions

In the discussion above, we are assuming that for $(\alpha, \theta) = (\tilde{\alpha}, \tilde{\theta})$ we have that $\det(P_{ij}(\omega))$ vanishes for R real values of ω, namely $\det(P_{ij}(\omega_l)) = 0$ for $l = 1, \cdots R$. Since $P_{ij}(\omega) = (P_{ij}(-\omega^*))^*$, we have that $\det(P_{ij}(\omega)) = (\det(P_{ij}(-\omega^*)))^*$. Therefore, the solutions to the equation $\det(P_{ij}(\omega)) = 0$ come in conjugate pairs if they are complex. Generically, we can assume that as we tune the parameter s to the critical value $\tilde{s}$ such that $(\alpha_{\tilde{s}}, \theta_{\tilde{s}}) = (\tilde{\alpha}, \tilde{\theta})$ either one real solution or the real part of one pair of solutions crosses to positive values. We then consider the following two scenarios for the poles ω_l, $l = 1, \ldots, R$:

- $\omega_1 = 0$, $R = 1$; or
- $\omega_1 = -\omega_2 > 0$, $R = 2$.

Indeed, we wish to consider the two qualitatively different cases of either i) a single pole with zero frequency; or ii) a pair of poles with nonvanishing and opposite frequencies emerging at $(\alpha, \theta) = (\tilde{\alpha}, \tilde{\theta})$. Of course, more than two poles could simultaneously emerge $(\alpha, \theta) = (\tilde{\alpha}, \tilde{\theta})$, but we consider this as a non-generic case.

- If $\omega_l = 0$ is a pole, then we have a static phase transition, associated with a breakdown in the linear response describing the parametric modulation of the measure of the system. While such a statement applies for rather general systems and perturbations, this situation can be better understood by considering the specific perturbation $\mathbf{X}(\mathbf{x}) = \langle \mathbf{x} \rangle_0 - \mathbf{x}$ with $T(t) = 1$, which amounts to studying, within linear approximation, how the measure of the system changes as the value of θ is changed to $\theta + \varepsilon$. This phase transition corresponds to a insulator-metal phase transition in condensed matter, because the electric susceptibility $\chi_{ij}^{elect}(\omega)$ of a conductor diverges as $i\sigma_{ij}/\omega$ for small frequencies, where σ is a real tensor and describes the static electric conductivity, which is vanishing for an insulator [LSPV05].
- If, instead, we have a pair of poles located at $\pm\omega_l \neq 0$, we have a dynamic phase transition activated by a forcing with non-vanishing spectral power at the frequency $\pm\omega_l$. In this case, a limit cycle emerges corresponding to self-sustained oscillation, which is made possible by the feedback encoded in the nonlinearity of the McKean-Vlasov equation.

4.4 Equilibrium Phase Transitions: Gradient Systems

In this section we will investigate *equilibrium* phase transitions for weakly interacting diffusions using the linear response formalism developed above and the spectral theory for compact

Markov semigroups. Firstly, we will recall some of the results on equilibrium phase transitions for weakly interacting diffusions that we have introduced in section 2.3. When the local force can be written as a gradient of a potential $\mathbf{F}_\alpha(\mathbf{y}) = -\nabla V_\alpha(\mathbf{y})$ and the diffusion matrix is the identity matrix $s_{ij} = \delta_{ij}$, equations (4.1) describe an equilibrium system. In particular, the N particles system has a unique ergodic invariant measure when the potential satisfies suitable confining properties [Tam84, Pav14], namely the Gibbs measure

$$M_N\left(\{\mathbf{x}^k\}\right) = \frac{e^{-\beta \mathcal{H}_N}}{Z_N}, \tag{4.53}$$

where Z_N is the partition function of the N-particle system and $\mathcal{H}_N$ is the Hamiltonian function of the system defined in terms of the local and interaction potentials as

$$\mathcal{H}_N\left(\{\mathbf{x}^k\}\right) = \sum_{k=1}^{N} V_\alpha(\mathbf{x}^k) + \frac{\theta}{2N}\sum_{k,l=1}^{N} \mathcal{U}(\mathbf{x}^k - \mathbf{x}^l). \tag{4.54}$$

Equivalently, the generator $\mathcal{K}_N$, the adjoint operator of the N-particle Fokker Planck operator $\mathcal{L}_N$, see equation (2.4), of the finite particle stochastic process has purely discrete spectrum, a nonzero spectral gap and the system converges exponentially fast to the unique equilibrium state, both in the L^2 space weighted by the invariant measure and in relative entropy [Pav14, Chapter 4.6]. On the other hand, in the thermodynamic limit, the system is described by the McKean Vlasov equation (4.3) whose stationary measures are solutions of the Kirkwood-Monroe equation (2.32):

$$\rho_0(\mathbf{x}) = \frac{1}{Z} e^{-\frac{2}{\sigma^2}(V(\mathbf{x}) + \mathcal{U}\star\rho_0(\mathbf{x}))}, \quad Z = \int e^{-\frac{2}{\sigma^2}(V(\mathbf{x}) + \mathcal{U}\star\rho_0(\mathbf{x}))} \mathrm{d}\mathbf{x}. \tag{4.55}$$

When the confining and interaction potentials are strongly convex and convex, respectively, then it is well known that (4.55) has only one solution, corresponding to the unique steady state of the McKean-Vlasov dynamics. In addition, the dynamics converges exponentially fast, in relative entropy, to the stationary state and the rate of convergence to equilibrium can be quantified [Mal01]. However, when the confining potential is not convex, e.g. is bistable, then more than one stationary states can exist, at sufficiently low noise strength (equivalently, for sufficiently strong interactions). The loss of uniqueness of the invariant measure can thus be interpreted as a continuous (or discontinuous) phase transition due to the interactions between the agents and arising strictly in the thermodynamic limit. For the quadratic interaction potential $\mathcal{U}(\mathbf{x}) = \frac{|\mathbf{x}|^2}{2}$ we are considering in this chapter, the equilibrium stationary measures (4.55) can be written as

$$\rho_0(\mathbf{x}) = \frac{1}{Z} e^{-\frac{2}{\sigma^2}\hat{V}(\mathbf{x})}, \quad Z = \int e^{-\frac{2}{\sigma^2}\hat{V}(\mathbf{x})} \mathrm{d}\mathbf{x}, \tag{4.56}$$

where we have introduced the modified potential $\hat{V}(\mathbf{x}) = V(\mathbf{x}) + \theta(\frac{|\mathbf{x}|^2}{2} - \langle \mathbf{x} \rangle_0 \cdot \mathbf{x})$, with the term proportional to θ arising from the interactions between the subsystems and $\langle \mathbf{x} \rangle_0 = \int \mathbf{x} \rho_0(\mathbf{x}) \mathrm{d}\mathbf{x}$ is the first moment of the invariant distribution. The linear Fokker-Planck operator associated to the stationary Mc-Kean Vlasov equation 4.3 describing the mean field equilibrium properties relative to (4.56) reads

$$\mathcal{L}_{\langle \mathbf{x} \rangle_0}(\cdot) = \nabla \cdot \Big(\nabla \hat{V}(\mathbf{x}) \ \cdot \ \Big) + \frac{\sigma^2}{2} \Delta \cdot . \tag{4.57}$$

It is known that Markov semigroups describing gradient dynamics in $\mathbb{R}^M$ are compact [LB07] , provided some growth conditions on the potential are satisfied. Indeed, the adjoint $\mathcal{K}_{\langle \mathbf{x} \rangle_0}$ of the operator $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$ generates a semigroup $e^{t\mathcal{K}_{\langle \mathbf{x} \rangle_0}}$that is associated with a gradient dynamics with deterministic part given by the potential $\hat{V}(\mathbf{x})$. If $\hat{V}$ satisfies the confining property [Pav14, Sect. 4.5]

$$\lim_{|\mathbf{x}| \to +\infty} \left(\frac{|\nabla \hat{V}|^2}{2} - \Delta \hat{V} \right) = +\infty, \tag{4.58}$$

then the following results hold:

- the semigroup $e^{t\mathcal{K}_{\langle \mathbf{x} \rangle_0}}$ is compact and its generator $\mathcal{K}_{\langle \mathbf{x} \rangle_0}$ is self adjoint in $L^2(\mathbb{R}^M, \rho_0)$, the space of square integrable functions in $\mathbb{R}^M$ weighted with by the invariant density. [Pav14, Chapters 4.6-4.7]

- the generator $\mathcal{K}_{\langle \mathbf{x} \rangle_0}$(or equivalently the Fokker Planck operator $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$) has a spectral gap in $L^2_{\rho_0}$

The above properties have important consequences on the spectral decomposition of the microscopic Green Functions $G_i(t), Y_{ik}(t)$ and their corresponding susceptibilities. In particular, a compact semigroup does not feature an essential spectrum so that $\mathcal{Q}_G$ and $\mathcal{Q}_Y$ in (4.12) identically vanish. Moreover, the number of Ruelle Pollicott resonances for such semigroups is infinite, $n \to +\infty$, and, since $\mathcal{K}_{\langle \mathbf{x} \rangle_0}$ is self adjoint, they are real, with $\lambda_0 = 0 > \lambda_1 \geq \lambda_2 \geq \dots$ and $\lambda_j \to -\infty$ as $j \to +\infty$. The existence of a spectral gap $\gamma = |\lambda_1| > 0$ of the generator guarantees that the microscopic constituents of the dynamics are robust and resilient to external and prevents the system from undergoing a critical transition, i.e a breakdown of condition 1 described in section 4.3 . However, we remark that the existence of a spectral gap of the generator $\mathcal{K}_{\langle \mathbf{x} \rangle_0}$ does not preclude the possibility that the system will experience a phase transition, that is a breakdwon of condition 2.

For gradient dynamics, it is possible to obtain a formula for the microscopic susceptibility $G_i(t)$ in terms of the time derivative of *known* correlation functions. Let us consider, for simplicity, a uniform forcing $\mathbf{X} = \hat{\mathbf{v}}_k$, with $\hat{\mathbf{v}}_k$ being the unit vector in the $k-th$ direction. The microscopic susceptibility can be written as

$$G_i(t) = \Theta(t) \int x_i e^{t\mathcal{L}_{\langle \mathbf{x} \rangle_0}} \mathcal{L}_1 \rho_0 \mathrm{d}\mathbf{x} = Y_{ik}(t) = -\Theta(t) \int x_i e^{t\mathcal{L}_{\langle \mathbf{x} \rangle_0}} \partial_{x_k} \rho_0 \mathrm{d}\mathbf{x}. \tag{4.59}$$

Since the system is at equilibrium, the stationary probability density can be written as in (4.55), $\partial_{x_k} \rho_0 = -\frac{2}{\sigma^2} \rho_0(\mathbf{x}) \partial_{x_k} \hat{V}$, physically representing the fact that the probability current associated to the invariant measure vanishes at equilibrium. Furthermore, using (4.57) it is easy to verify the following identity $\mathcal{L}_{\langle \mathbf{x} \rangle_0} (x_k \rho_0(\mathbf{x})) = -\rho_0(\mathbf{x}) \partial_{x_k} \hat{V}$. The mean field susceptibility can then be written as

$$G_i(t) = -\frac{2}{\sigma^2} \Theta(t) \int x_i e^{t\mathcal{L}_{\langle \mathbf{x} \rangle_0}} \mathcal{L}_{\langle \mathbf{x} \rangle_0} x_k \rho_0 \mathrm{d}\mathbf{x} = -\frac{2}{\sigma^2} \Theta(t) \frac{\mathrm{d}}{\mathrm{d}t} \int x_i e^{t\mathcal{L}_{\langle \mathbf{x} \rangle_0}} x_k \rho_0 \mathrm{d}\mathbf{x} = \tag{4.60}$$

$$= -\frac{2}{\sigma^2} \Theta(t) \frac{\mathrm{d}}{\mathrm{d}t} \langle x_i(t) x_k(0) \rangle_0 = -\frac{2}{\sigma^2} \Theta(t) \frac{\mathrm{d}}{\mathrm{d}t} \langle z_i(t) z_k(0) \rangle_0 = \tag{4.61}$$

$$= -\frac{2}{\sigma^2} \Theta(t) \frac{\mathrm{d}}{\mathrm{d}t} C_{z_i, z_k}(t), \tag{4.62}$$

where in the last equation we have introduced the fluctuation variables $z_i = x_i - \langle x_i \rangle_0$ and $C_{z_i,z_k}(t)$ represents the mean field correlation function between variables z_i and z_k. Equation (4.62) shows that the *microscopic* susceptibility is closely related to equilibrium correlation functions. It is then possible to associate to each correlation function the mean field integrated correlation time

$$\tau_{ij} = \frac{\int_0^{+\infty} \langle z_i(t) z_j(0) \rangle_0 \mathrm{d}t}{\langle z_i z_j \rangle_0} = \frac{\int_0^{+\infty} C_{z_i,z_j}(t) \mathrm{d}t}{C_{z_i,z_j}(0)}. \tag{4.63}$$

Note that this time scale differs from the one introduced in (4.42), which in this case can be written as

$$\tau_{G_i} = \frac{\int_0^{\infty} G_i(t) \mathrm{d}t}{G_i(0^+)} = -\frac{\langle z_i z_j \rangle_0}{\lim_{t \to 0^+} \frac{\mathrm{d}}{\mathrm{d}t} \langle z_i(t) z_j(0) \rangle_0}. \tag{4.64}$$

By comparing the expressions of τ_{G_i} and τ_{ij} and by considering (4.16), one understands that τ_{G_i} and τ_{ij} correspond to two differently weighted averages of the timescales associated with each subdominant mode of the operator $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$. Usually, the singular behaviour of correlation properties has been used as an indicator of critical transitions [Sch09a]. However, let us remark again that, being related to the spectrum of the operator $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$, in our case neither τ_{G_i} nor τ_{ij} show any critical behaviour at a phase transition, while they both diverge in the case of critical transitions corresponding to the breakdown of 1 above.

Chapter 5

Detecting Phase Transitions through Linear Response in finite dimensional systems

5.1 Introduction

In the previous chapter we have derived linear response formulas for the thermodynamic limit of a system of weakly interacting diffusions described a McKean Vlasov equation and have explicitly identified two qualitatively different scenarios for the breakdown of the linear response, namely critical transitions and phase transitions. Critical transitions occur for any finite (possibly infinite too) dimensional system and are characterised by the spectral gap of the transfer operator governing the time evolution of observables shrinking to zero. Consequently, a critical transition is accompanied by a divergence of the mixing properties of the (microscopic) dynamical variables of the system, and, as a result of Fluctuation Dissipation formulas, one expects a singular behaviour of the response properties and a rough dependence of the system on changes in its parameters. On the other hand, phase transitions are a genuine thermodynamic phenomenon, where the divergence of the response stems from the coordination taking place, in suitable conditions, because of the coupling between the infinite number of agents composing the total system. Moreover, the coupling among the subsystems results in a memory effect, see (4.10), that leads to obtaining the total response function of the system as a macroscopic version of its microscopic counterpart, see (4.24), with formal similarities with the well-known Clausius-Mossotti relation for the polarizability of an electric medium [LSPV05, Jac75, TT13].

In this chapter we focus on the latter case, investigating with much greater detail on the relationship between the occurrence of phase transitions and the non-analyticity of the susceptibility of the system describing the frequency-dependent response of an observable to a given perturbation in the upper complex frequency plane. Such singular behaviour manifests itself as a pole that crosses the real axis of the frequency variable, leading to a diverging resonance of the system. We consider two paradigmatic models of equilibrium and nonequilibrium (respectively) interacting diffusions, namely the Desai Zwanzig model, that we have already introduced in section 2.6 and the Bonilla-Casado-Morillo. In particular, our aim is twofold.
Firstly, we prove that the susceptibility of the thermodynamic limit of interacting diffusions, described by a McKean Vlasov equation, develops a singular behaviour at the phase transition. We are able to fully characterise the location of the pole reaching the real axis of frequencies and, for the Desai Zwanzig model, we provide an explicit expression of its residue in terms of mean field correlation functions. We also remark that in chapter 6 we prove a similar result for a nonequilibrium version of the Desai Zwanzig model.
Secondly, we adopt a numerical perspective that mirrors spectroscopic techniques that are used for investigating the frequency dependence of the optical properties of materials [LSPV05]. In particular, by studying how the real and imaginary part of the susceptibility of the systems depend on the number of agents, we are able to predict the position of the pole and the associated residue, fully characterising the emergence of the singularity of the susceptibility in the thermodynamic limit. We verify that the position of the pole depends on the considered model, but, instead, that for a given model the loss of analyticity depends neither on the choice of the observable, nor on the applied perturbation, and is, in this sense, an universal feature of the system, since it is related to spectral properties of suitable operators.
In this chapter we also clarify the link between the breakdown of response properties of the McKean Vlasov equation and the phenomenon of *critical slowing down*, see section 2.4. This

phenomenon stems from the fact that, near critical settings, the negative feedbacks of the system become increasingly ineffective, resulting in arbitrarily large, usually non-Gaussian, fluctuations and a divergence of correlation properties of the system. The McKean Vlasov equation, by definition, does not capture such singular behaviour since it originates from *propagation of chaos* properties that, at the phase transition, do not hold uniformly in time. Response theory, as first observed in [Shi87], allows one to investigate such singular behaviour and provides a bridge between microscopic and macroscopic features of the interacting agents.

5.2 Examples: a mathematical analysis

5.2.1 Equilibrium Phase Transition: the Desai Zwanzig Model

The Desai Zwanzig model [DZ78] has a paradigmatic value as it features an equilibrium thermodynamic phase transition, characterised by a pitchfork bifurcation of the infinite dimensional invariant measure of the system, arising from the interaction between the agents. This model can be seen as a stochastic model of key importance for elucidating order-disorder phase transitions [Fra13]. More details about the model and its applications can be found in section 2.6. We here provide a few key features that will be relevant for our analysis of response properties. Each of the interacting systems can be interpreted as a particle, moving in one dimension, $M = 1$, in a double well potential $V_\alpha(x) = -\frac{\alpha}{2}x^2 + \frac{x^4}{4}$, interacting with the other particles via a quadratic interaction $\mathcal{U}(x) = \frac{x^2}{2}$. The N-particle system is described by the equations

$$\mathrm{d}x^k = F_\alpha(x^k)\mathrm{d}t - \frac{\theta}{N}\sum_{l=1}^{N}(x^k - x^l)\mathrm{d}t + \sigma \mathrm{d}W^k, \tag{5.1}$$

where $k = 1, \dots, N$. The local force is $F_\alpha = -V'_\alpha$ and the one particle volatility matrix is the identity matrix $s_{ij} = \delta_{ij}$ resulting in thermal noise. Furthermore, V_α is double well shaped when $\alpha > 0$, otherwise it has a unique global minimum. In the thermodynamic limit $N \to \infty$, the one particle density satisfies the McKean-Vlasov equation (4.3) and it is possible to show that the infinite particle system undergoes a continuous phase transition, with $\langle x \rangle$ being a suitable order parameter. The phase transition is characterised by the condition (2.87) that we report here for simplicity

$$\frac{2\theta}{\sigma^2}\langle x^2\rangle_0 = 1, \tag{5.2}$$

where $\langle\cdot\rangle_0$ is the expectation over ρ_0, the invariant measure solution of the McKean Vlasov equation. We remark that at the phase transition $\langle x\rangle_0 = 0$, so that $\langle x^2\rangle_0$ represents the variance of the observable x. Moreover, critical values of the parameters at which the phase transition is attained satisfy the following property

$$\frac{D_{-3/2}\left(\frac{\theta-\alpha}{\sigma}\right)}{D_{-1/2}\left(\frac{\theta-\alpha}{\sigma}\right)} = \frac{\sigma}{\theta}, \tag{5.3}$$

where $D_\nu(z)$ is a parabolic cylinder functions. To assess the existence of the phase transitions for a sufficiently big finite system, we below report the results of the numerical integration of (5.1) by adopting a simple Euler-Maruyama scheme [KP11]. In particular, we have tested the convergence of our results in the thermodynamic limit $N \to \infty$ by looking at increasing values of the number N of particles. We present in Figs. 5.1a-5.1b-5.1c the results obtained with $N = 5000$ for $0.2 \leq \theta \leq 1.0$ and $0.4 \leq \sigma \leq 1.0$. The stationary, unperturbed expectation values $\langle f\rangle_0$ of a generic function $f(x) : \mathbb{R} \to \mathbb{R}$ are evaluated as a time average, at stationarity, of the empirical mean over all the particles of f, namely

$$\langle f\rangle_0 = \int f(x)\rho_0(x)\mathrm{d}x \approx \frac{1}{T}\int_{t^*}^{t^*+T} \bar{f}(t)\mathrm{d}t, \tag{5.4}$$

where T is a long time interval and t^* is a suitably long initial time such that stationarity is attained for all $t \geq t^*$ and the bar denotes averaging over all the agents,

$$\bar{f}(t) = \frac{1}{N}\sum_{k=1}^{N} f(x_k(t)). \tag{5.5}$$

Moreover, mean field correlations functions, see definition (2.22), between observable f and g are estimated as an average over all particles of the single-particle correlation functions

$$C_{fg}(t) = \int f(\mathbf{x})e^{\mathcal{L}_{\langle x\rangle_0}t}g(\mathbf{x})\rho_0(\mathbf{x})\mathrm{d}\mathbf{x} \approx \frac{1}{N}\sum_{k=1}^{N} C_{f(\mathbf{x}_k)g(\mathbf{x}_k)}(t) \tag{5.6}$$

where the one particle correlation function $C_{f(\mathbf{x}_k)g(\mathbf{x}_k)}(t)$ with $k = 1, \dots, N$ can be evaluated from the two time series $f(\mathbf{x}_k(t))$ and $g(\mathbf{x}_k(t))$. The relevant expectation values and correlations have been evaluated considering averages performed over $T = 2.5 \times 10^3$ time units. Figures 5.2a-5.2b portray two sections performed approximately in the middle of the domain of the heat maps provided in figures 5.1a-5.1b-5.1c, with the goal of clarifying the obtained results. The order parameter $\langle x \rangle_0$ clearly indicates a continuous phase transition. The (re-scaled) variance of the fluctuations, being related to the operator $\mathcal{L}_{\langle x \rangle_0}$, see equation (4.4), is finite and equal to $\frac{1}{2}$ at the transition point, in agreement with (5.2). The re-scaled correlation time $\hat{\tau} = \theta \times \tau$, where τ is defined in (4.63) as the integrated correlation time associated to the mean field autocorrelation function $C_{fg}(t)$ with $f = g = x$, is also non-singular, as opposed to a critical transition scenario.

Singularity of the susceptibility

Given the simplicity of this model, it is possible to explicitly evaluate all the relevant quantities that characterise the phase transition for the Desai Zwanzig model. As explained in greater detail in section 4.2 we perturb the deterministic part of the McKean Vlasov equation by considering $F_\alpha(x) \to F_\alpha(x) + \varepsilon T(t) X(x)$. We consider here a purely temporal perturbation, that is $X(x) = 1$. In this case the *microscopic* Green Functions (4.12) coincide[1], $Y(t) = G(t)$, and (4.21), describing the response of the observable x in the frequency domain, can be written as

$$P(\omega)\langle x \rangle_1(\omega) = \chi(\omega)T(\omega), \tag{5.7}$$

where the 1×1 matrix is $P(\omega) = 1 - \theta\chi(\omega)$. The macroscopic susceptibility is then obtained as

$$\tilde{\chi}(\omega) = P^{-1}(\omega)\chi(\omega) = \frac{\chi(\omega)}{1 - \theta\chi(\omega)}. \tag{5.8}$$

Furthermore, this is a gradient system satisfying all the assumptions that have been made in

[1] Being a one dimensional problem we drop here the indices from the Green Functions and Susceptibilities

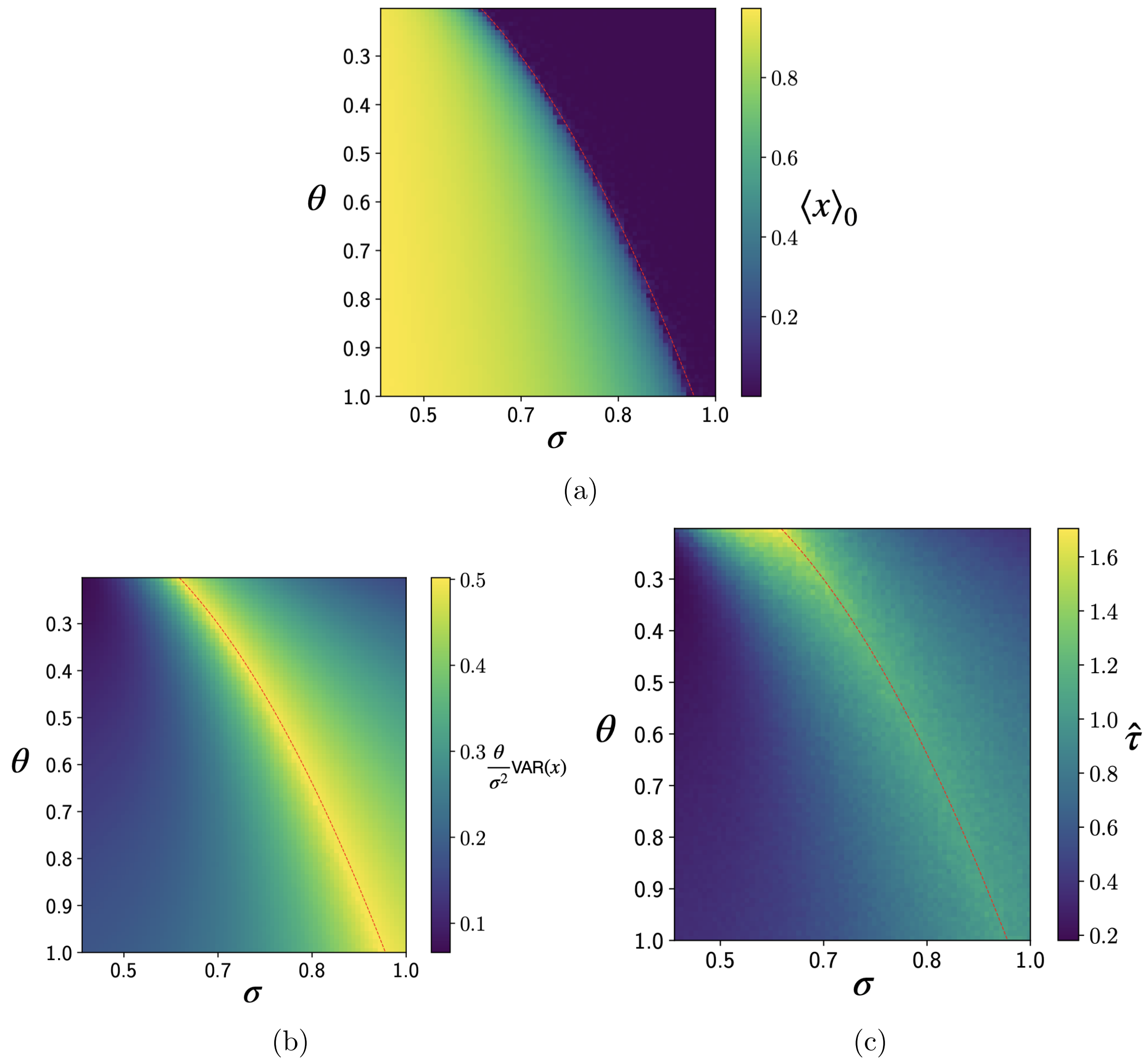


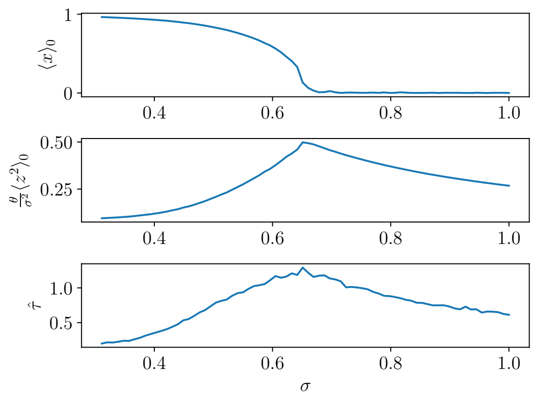
Figure 5.1: Results of numerical simulations of equations (5.1) with $\alpha = 1$. Heat maps of the order parameter $\langle x\rangle_0$ (panel a); of the re-scaled variance $\frac{\theta}{\sigma^2}\mathrm{VAR}(x) = \frac{\theta}{\sigma^2}\left(\langle x^2\rangle_0 - \langle x\rangle_0^2\right)$ (panel b); and of the rescaled correlation time $\hat{\tau} = \theta \times \tau$ (panel c). The dotted red line shows the transition line (5.3). See text for details.

section, so that the *microscopic* susceptibility can be written as

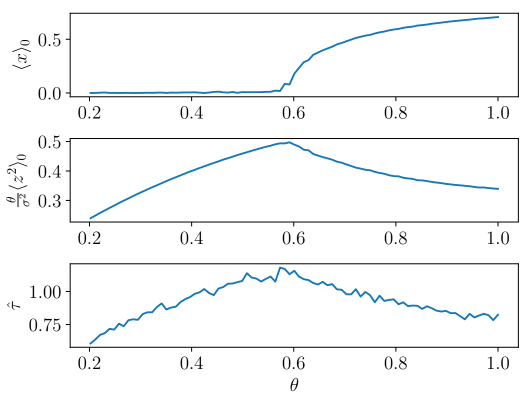
$$G(t) = -\Theta(t)\frac{2}{\sigma^2}\frac{\mathrm{d}}{\mathrm{d}t}C_{z,z}(t), \tag{5.9}$$

where $z(t) = x - \langle x\rangle_0$. Taking the Fourier transform results in

$$\chi(\omega) = \frac{2}{\sigma^2}\left[C_{z,z}(0) + i\omega\gamma(\omega)\right], \tag{5.10}$$



(a) From top to bottom: order parameter, rescaled variance, and rescaled integrated autocorrelated time as a function of the strength of the noise. Here $\theta \approx 0.4$.



(b) From top to bottom: order parameter, rescaled variance, and rescaled integrated autocorrelated time as a function of the strength of the coupling. Here $\sigma \approx 0.78$.

Figure 5.2: A horizontal (left) and a vertical (right) section of the heat maps shown in figures 5.1a-5.1c.

where $\gamma(\omega) = \int_0^\infty C_{z,z}(t) e^{i\omega t} \mathrm{d}t$ is the (one-sided) Fourier transform of the correlation function. As previously mentioned, $\chi(\omega)$ can be written in terms of the spectrum of the operator $\mathcal{L}_{\langle x \rangle_0}$ which in this specific example reads, see (4.57),

$$\mathcal{L}_{\langle x \rangle_0} = -\hat{V}'(x)\partial_x + \frac{\sigma^2}{2}\partial_{xx}, \tag{5.11}$$

where the modified potential is $\hat{V} = V_\alpha + \theta(\frac{x^2}{2} - \langle x \rangle_0 x)$. It can be proven [Daw83] that the above operator is self-adjoint and has a pure point spectrum $\{\lambda_j\}$ with $0 = \lambda_0 > \lambda_1 > \lambda_2 > \dots$, with the vanishing eigenvalue corresponding to the stationary distribution ρ_0. In fact, it is easy to show that condition (4.58) holds, with $e^{t\mathcal{K}_{\langle x \rangle_0}}$ being thus a stable compact with a non vanishing spectral gap, see section 4.4. The operator $\tilde{\mathcal{L}}_{\langle x \rangle_0}$ describing the response of the invariant measure ρ_0 of the McKean Vlasov equation, see (4.3), is instead

$$\tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}\rho_1 = \mathcal{L}_{\langle \mathbf{x} \rangle_0}\rho_1 - \theta \langle x \rangle_1(t)\partial_x \rho_0. \tag{5.12}$$

Dawson [Daw83] proved that, away from the transition point - in particular, above it, where $\langle x \rangle_0 = 0$ - the operator $\tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}$ has similar spectral properties to $\mathcal{L}_{\langle x \rangle_0}$. Consequently the system has a smooth response to perturbations, characterised by a *macroscopic* susceptibility $\tilde{\chi}(\omega)$

with Ruelle Pollicott poles contained in the lower half of the complex frequency ω-plane. At the transition, though, $\tilde{\mathcal{L}}_{\langle x\rangle_0}$ shows a vanishing spectral gap, with the operator developing a null eigenvalue. This situation corresponds to a phase transition scenario, as expressed by a breakdown of condition 2 described in section 4.3, in which the *microscopic* susceptibility $\chi(\omega)$ is holomorphic in the upper complex ω-plane, while the *macroscopic* $\tilde{\chi}_(\omega)$ develops a pole, arising from the non invertibility of $P(\omega) = 1 - \theta\chi(\omega)$. Let us observe again that this implies that at the transition there is no divergence of the integrated autocorrelation time τ, because the spectral gap of the operator $\mathcal{L}_{\langle x\rangle_0}$ does not shrink to zero. This is clearly shown in the two-dimensional map shown in figure 5.1c and in the two sections shown in figures 5.2a-5.2b. We can fully characterise the singular behaviour of the *macroscopic* susceptibility $\tilde{\chi}(\omega)$ at the transition. Indeed, the transition point is characterised by the condition (5.2) that we can write

$$1 - \frac{2\theta}{\sigma^2} C_{z,z}(0) = 0, \tag{5.13}$$

so that the *macroscopic* susceptibility (5.8) becomes, considering (5.10) and the equation above,

$$\tilde{\chi}(\omega) = -\frac{1}{\theta} + \frac{i}{\omega} \frac{C_{z,z}(0)}{\theta\gamma(\omega)}. \tag{5.14}$$

We remark that $\gamma(\omega)$ is related to properties of the *mean field* correlation functions, and, as such, it is a holomorphic function in the upper complex plane, with no poles on the real axis. Consequently, the above expression shows that at the transition point $\tilde{\chi}(\omega)$ develops a simple pole in $\omega_0 = 0$, with residue

$$\mathrm{Res}\left(\tilde{\chi}(\omega)\right)_{\omega=0} = i\frac{C_{z,z}(0)}{\theta\gamma(0)} = \frac{i}{\theta\tau_z}, \tag{5.15}$$

where τ is the *mean field* integrated correlation time associated to the fluctuation variable $z(t) = x(t) - \langle x\rangle_0$. We observe that the residue is completely imaginary as it is expected from a static phase transition characterised by a pole $\omega_0 = 0$, see discussion in section 4.3.

5.2.2 Nonequilibrium Phase Transition: the Bonilla-Casado-Morillo model

In this section we will study the response features of the Bonilla-Casado-Morrillo model [BCM87] and elucidate the properties of a non equilibrium self-synchronization phase transition for an ensemble of nonlinear oscillators, by looking at the divergence of the *macroscopic* susceptibility $\tilde{\chi}(\omega)$. We anticipate that the susceptibility develops a pair of symmetric poles $\omega_1 = -\omega_2 > 0$ at the transition point, thus following the scenario of a dynamic phase transition discussed in section 4.3. The model consists of N two-dimensional ($M = 2$) non linear oscillators $\mathbf{x}^k = (x_1^k, x_2^k)$, interacting via a quadratic interaction potential $\mathcal{U}(\mathbf{x}) = \frac{|\mathbf{x}|^2}{2}$ and subjected to thermal noise

$$\mathrm{d}x_i^k = F_{i,\alpha}(\mathbf{x}^k)\mathrm{d}t - \frac{\theta}{N}\sum_{l=1}^{N}\partial_{x_i^k}U(\mathbf{x}^k - \mathbf{x}^l)\mathrm{d}t + \sigma\mathrm{d}W_i^k, \quad k = 1, \ldots, N. \tag{5.16}$$

The local force is not conservative, giving rise to a non equilibrium process, and reads $\mathbf{F}_\alpha(\mathbf{x}) = (\alpha - |\mathbf{x}|^2)\,\mathbf{x}{+}\mathbf{x}^+$where $\mathbf{x}^+ = (-x_2, x_1)$. This term corresponds to a rotation, which is divergence-free with respect to the (Gibbsian) invariant measure and, therefore, does not change the stationary state, but it makes it a non-equilibrium one [LNP13, DPZ17, DLP16]. The parameter $\alpha > 0$ controls the amplitude of the oscillations of the individual non linear oscillators. In fact, when $\theta = \sigma = 0$, each subsystem oscillates as $\mathbf{x}^j(t) = \sqrt{\alpha}\,(\cos(t + \beta_j), \sin(t + \beta_j))$ where $\beta_j = \tan(x_2^j(0)/x_1^j(0))$. The coupling tries to synchronise the subsystems by attracting them towards the center of mass $\frac{1}{N}\sum_{j=1}^{N}\mathbf{x}^j$. In the thermodynamic limit, the system is described by a McKean-Vlasov equation

$$\partial_t\rho(\mathbf{x}, t) = \mathcal{L}_{\langle\mathbf{x}\rangle}\rho = -\nabla\cdot\left[\left(\hat{\mathbf{F}}(\mathbf{x}) + \theta\langle\mathbf{x}\rangle(t)\right)\rho(\mathbf{x}, t)\right] + \frac{\sigma^2}{2}\Delta\rho(\mathbf{x}, t), \tag{5.17}$$

where $\hat{\mathbf{F}}(\mathbf{x}) = \mathbf{F}_\alpha(\mathbf{x}) - \theta\mathbf{x}$, the last term representing the mean field contribution of the coupling to the local force. The authors in [BCM87] prove that the infinite particle system undergoes a phase transition, with a stationary measure $\rho_0(\mathbf{x})$ losing stability to a time dependent probability measure $\bar{\rho} = \bar{\rho}(\mathbf{x}, t)$. Physically, this phenomenon can be interpreted as a process of synchronization. In fact, $\rho_0(\mathbf{x})$ represents a disordered state, with the oscillators moving out of phase, while $\bar{\rho}$ describes a state of collective organisation with the oscillators moving in an organised rhythmic manner. The transition can be investigated via the order parameter $\langle \mathbf{x} \rangle$ which vanishes in the asynchronous state, $\langle \mathbf{x} \rangle_0 = 0$, and is different from zero and oscillating in time in the synchronous state. In particular, the stationary measure $\rho_0(\mathbf{x})$ can be written as

$$\rho_0(\mathbf{x}) = \frac{1}{Z} e^{-\phi(\mathbf{x})}, \quad \phi(\mathbf{x}) = \left(\theta - \alpha + \frac{1}{2} |\mathbf{x}|^2 \right) \frac{|\mathbf{x}|^2}{\sigma^2}, \tag{5.18}$$

and satisfies the stationary McKean-Vlasov equation $\mathcal{L}_{\langle \mathbf{x} \rangle_0} \rho_0(\mathbf{x}) = \mathcal{L}_{\langle \mathbf{x} \rangle_0} \rho_0(\mathbf{x}) = 0$ where

$$\mathcal{L}_0 = -\nabla \cdot \left(\hat{F}(\mathbf{x}) \cdot \right) + \frac{\sigma^2}{2} \Delta. \tag{5.19}$$

We can perform a linear response theory around this stationary state ρ_0 by replacing $\mathbf{F}_\alpha \to \mathbf{F}_\alpha + \varepsilon \mathbf{X}(\mathbf{x}) T(t)$ and studying the perturbation ρ_1 of the measure defined via $\rho(\mathbf{x}, t) = \rho_0(\mathbf{x}) + \varepsilon \rho_1(\mathbf{x}, t)$. As previously outlined, $\rho_1(\mathbf{x}, t)$ satisfies (4.6) from which the whole linear response theory follows. However, to conform to the notation in [BCM87] we will here define $\rho_1(\mathbf{x}, t) = \rho_0^{1/2} q(\mathbf{x}, t)$ and write the corresponding equation for $q(\mathbf{x}, t)$. After some algebra, it is possible to write that

$$\begin{aligned} \partial_t q(\mathbf{x}, t) &= \mathcal{M}_0(q) - T(t) \rho_0^{-1/2} \nabla \cdot (\mathbf{X}(\mathbf{x}) \rho_0) + \theta \rho_0^{1/2} \langle \rho_0^{1/2} \mathbf{y}, q(\mathbf{y}, t) \rangle \cdot \nabla \phi(\mathbf{x}) = \\ &= \tilde{\mathcal{M}}_0(q) - T(t) \rho_0^{-1/2} \nabla \cdot (\mathbf{X}(\mathbf{x}) \rho_0) , \end{aligned} \tag{5.20}$$

where we have introduced the usual L^2 inner product $\langle f, g \rangle = \int f(\mathbf{x}) g(\mathbf{x}) \mathrm{d}\mathbf{x}$ and defined the linear differential operator

$$\mathcal{M}_0 q = \frac{\sigma^2}{4} \left[\Delta \phi - \frac{1}{2} |\nabla \phi|^2 \right] q + \left[-\mathbf{x}^+ \cdot \nabla + \Delta \right] q, \tag{5.21}$$

and the linear integro-differential operator

$$\tilde{\mathcal{M}}_0 q = \mathcal{M}_0 q + \theta \rho_0^{1/2} \langle \rho_0^{1/2} \mathbf{y}, q(\mathbf{y}, t) \rangle \cdot \nabla \phi(\mathbf{x}). \tag{5.22}$$

We mention that this operator has the structure of a Schrödinger operator in a magnetic field [Pav14, Sec. 4.9]. Furthermore, let us observe that $\langle \rho_0^{1/2}\mathbf{y}, q(\mathbf{y},t)\rangle = \int \rho_0^{1/2}\mathbf{y}q(\mathbf{y},t)\mathrm{d}\mathbf{y} = \int \mathbf{y}\rho_1(\mathbf{y},t)\mathrm{d}\mathbf{y} = \langle \mathbf{y}\rangle_1$. It is then clear that $\mathcal{M}_0$ ($\tilde{\mathcal{M}_0}$) is the analogous, in a "Schrödinger picture", of the operator $\mathcal{L}_0$ ($\tilde{\mathcal{L}_0}$). In the following we show the equivalent response formulas using these new operators. A formal solution of the above equation is

$$q(\mathbf{x},t) = \int_{-\infty}^{t} e^{(t-s)\mathcal{M}_0}\left(-T(s)\rho_0^{-1/2}\nabla\cdot(\mathbf{X}(\mathbf{x})\rho_0) + \theta\rho_0^{1/2}\langle \rho_0^{1/2}\mathbf{y}, q(\mathbf{y},s)\rangle\cdot\nabla\phi(\mathbf{x})\right)\mathrm{d}s, \quad (5.23)$$

which is analogous to (4.8). Using the above expression we can evaluate the response of the observable x_i as

$$\begin{aligned}\langle x_i\rangle_1 &= \langle \rho_0^{1/2}x_i, q(\mathbf{x},t)\rangle = \\ &= \int \mathrm{d}\mathbf{x}\int_{-\infty}^{t}\mathrm{d}s\; x_i e^{(t-s)\mathcal{M}_0}\left[-T(s)\rho_0^{-1/2}\nabla\cdot(\mathbf{X}(\mathbf{x})\rho_0) + \theta\rho_0^{1/2}\langle \rho_0^{1/2}\mathbf{y}, q(\mathbf{y},s)\rangle\cdot\nabla\phi(\mathbf{x})\right].\end{aligned} \quad (5.24)$$

The above equations show that an analogous response theory holds for the operators $\mathcal{M}_0$ and $\tilde{\mathcal{M}_0}$. In particular, their spectrum is related to the Fourier transform of the *microscopic* susceptibility $\chi(\omega)$ and *macroscopic* susceptibility $\tilde{\chi}(\omega)$ (respectively) through equations similar to (4.22) and (4.33). The authors in [BCM87] have studied the spectrum of both these operators in order to perform a stability analysis of the stationary distribution $\rho_0(\mathbf{x})$. In particular, they observe that the operator $\mathcal{M}_0$ can be written as $\mathcal{M}_0 = \mathcal{M}_H + \mathcal{M}_A$ where

$$\mathcal{M}_H(q) = \frac{\sigma^2}{4}\left[\Delta\phi - \frac{1}{2}|\nabla\phi|^2\right]q + \frac{\sigma^2}{2}\Delta q, \quad (5.25)$$

and

$$\mathcal{M}_A(q) = -\mathbf{x}^{+}\cdot\nabla q, \quad (5.26)$$

with vanishing commutator $[\mathcal{M}_H, \mathcal{M}_A] = 0$. The operator $\mathcal{M}_H$ is related to the conservative part of the local force. As a matter of fact, it is a self-adjoint (Hermitian) operator with real eigenvalues. $\mathcal{M}_A$ is instead anti-Hermitian, with purely imaginary eigenvalues (describing oscillations) given by the non conservative part of $\mathbf{F}$. Furthermore, $\mathcal{M}_H$ has only one zero eigenvalue corresponding to the ground state $\sqrt{\rho_0}$ while all the remaining eigenvalues are negative, meaning that a critical transition, according to which the spectral gap of the mean field operator $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$ vanishes, see discussion in section 4.3, cannot take place in this setting. In particular, microscopic correlation properties do not diverge. Phase transitions can, instead, take place according to the scenario where $\tilde{\chi}(\omega)$ develops a singular behaviour for real frequencies whereas $\chi(\omega)$ is a holomorphic function in the upper complex plane. Infact, the spectral gap of the operator $\tilde{\mathcal{M}}_0$ vanishes [BCM87] at surface in the (α, σ, θ) parametric space defined by the following equation:

$$A = \frac{\delta^2}{2} \left[1 - \frac{1}{\delta} \exp\left(-\frac{A^2}{\delta^2} \right) \left[\int_{-\frac{A}{\delta}}^{\infty} e^{-r^2} dr \right]^{-1} \right], \tag{5.27}$$

where $A = \frac{\alpha}{\theta} - 1$ and $\delta = \frac{\sqrt{2\sigma^2}}{\theta}$. In particular, they are able to prove that the eigenvalues associated to eigenfunctions of $\tilde{\mathcal{M}}_0$ which are orthogonal to the subspace of $L^2(\mathbb{R}^2)$ spanned by $\sqrt{\rho_0}$ and $\mathbf{n} \cdot \mathbf{x}\sqrt{\rho_0}$, $\mathbf{n} \in \mathbb{R}^2$ being any unit vector, are always negative. Nevertheless, $\tilde{\mathcal{M}}_0$ can become unstable from eigenfunctions which are not orthogonal to $\mathbf{n} \cdot \mathbf{x}\sqrt{\rho_0}$. As a matter of fact, it is possible to identify the eigenfunctions that at the transition point yield eigenvalues with vanishing real part. In particular, at the transition line (5.27), the eigenfunction $\Omega(\mathbf{x}) = (0,1) \cdot \mathbf{x}\sqrt{\rho_0} + i(1,0) \cdot \mathbf{x}\sqrt{\rho_0}$ gives an eigenvalue $\tilde{\lambda} = i$, with $\Omega(\mathbf{x})^*$ corresponding to the complex conjugate eigenvalue $\tilde{\lambda}^* = -i$. The *macroscopic* susceptibility (4.33) consequently develops a pair of symmetric poles in $\omega = \pm 1$, corresponding to the Ruelle Pollicott poles $\omega = i\lambda$ and $\omega = i\lambda^*$. The development of real poles corresponds to a dynamic phase transition, giving rise to a Hopf-like bifurcation yielding the time dependent state $\bar{\rho}(\mathbf{x}, t)$ that defines the synchronized state. As a result, near the transition, the order parameter $\langle x \rangle_{\bar{\rho}}$, where the expectation value is computed using the measure $\bar{\rho}(\mathbf{x}, t)$, oscillates at frequency $\omega = 1$ with amplitude $A_1(\alpha, \sigma, \theta)$. Instead, since it is a quadratic quantity, the rescaled variance $\theta/\sigma^2 \langle z^2 \rangle$, where $z = x - \langle x \rangle$, oscillates at frequency $\omega = 2$ with amplitude $A_2(\alpha, \sigma, \theta)$ around the value $B_2(\alpha, \sigma, \theta)$.

We have investigated this non equilibrium transition through numerical integration of (5.16) via an Euler-Maruyama scheme. Again, the convergence of our results to the thermodynamic limit has been tested by looking at increasing values of the number of agents. The expectation

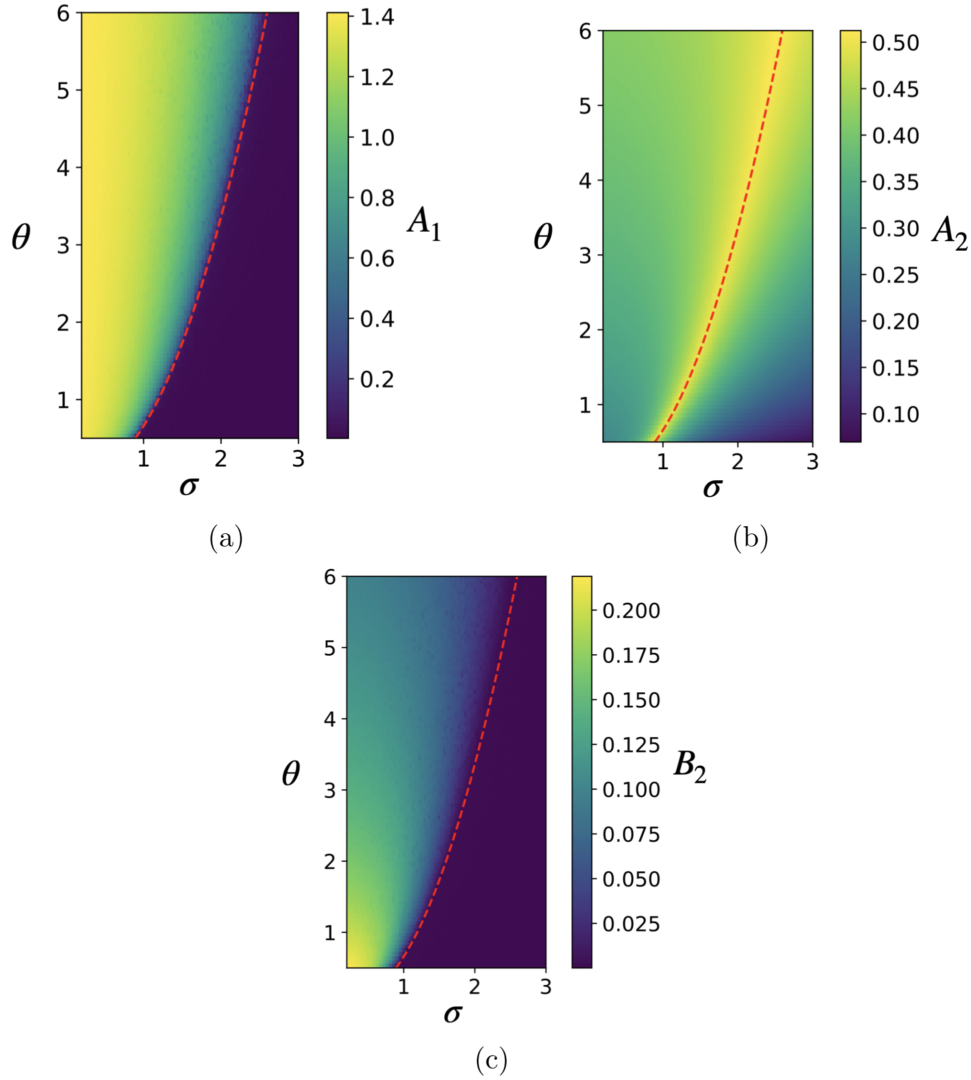
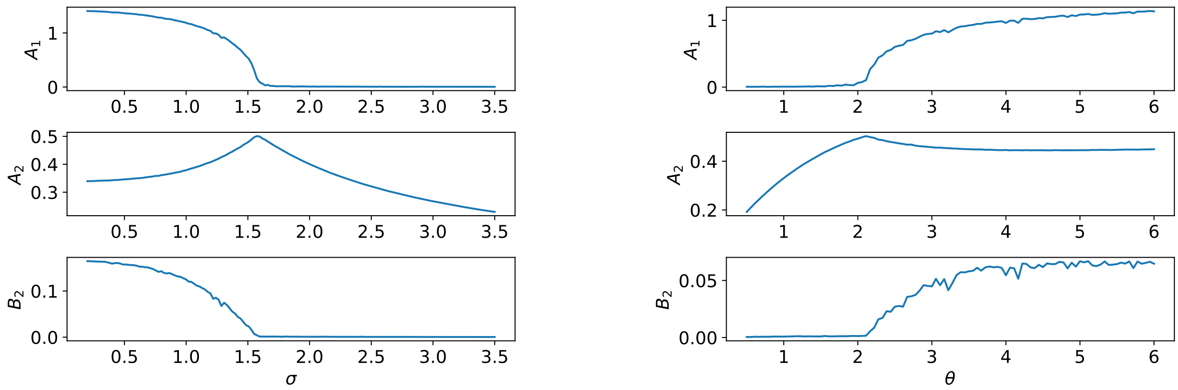


Figure 5.3: Results of numerical simulations of equation (5.16) with $\alpha = 2$. Heat maps of the amplitude A_1 of the oscillations of the variable x (panel a), of the amplitude A_2 of the oscillations of the re-scaled variance $\frac{\theta}{\sigma^2}\langle z^2\rangle$ (panel b), and of the time mean value B_2 of $\frac{\theta}{\sigma^2}\langle z^2\rangle$ (panel c). The red dotted line represents the transition line given by equation (5.27);see [BCM87]. See text for details.

values are evaluated as in the previous section, see (5.4). We display here the results by taking $N = 5000$ and choosing $\alpha = 2$. Figure 5.3 shows the value of A_1 (panel a), A_2 (panel b), and B_2 (panel c) for $\alpha = 2$ in the parametric region $0.2 \leq \sigma \leq 3$, $0.5 \leq \theta \leq 6$ of the two dimensional parameter space (σ, θ). For the sake of clarity, we also provide in figure 5.4 a snapshot of a horizontal and vertical section of the heat plots.

These numerical experiments confirm that the system indeed undergoes a continuous phase transition, with a collective synchronisation stemming from a disordered state as the system passes through the transition line given by (5.27) for $\alpha = 2$. Let us remark again that the fluctuations of the microscopic quantities, being related to the spectrum of $\mathcal{L}_{\langle x \rangle_0}$, are always finite, see figure 5.4.



(a) From top to bottom: A_1, A_2, and B_2. Here $\alpha = \theta = 2$

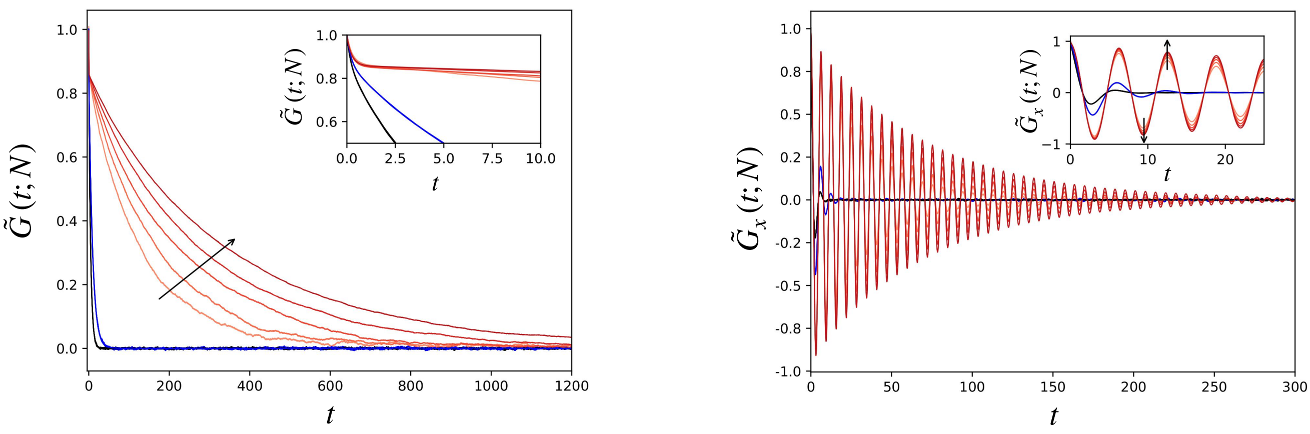
(b) From top to bottom: A_1, A_2, and B_2.Here $\alpha = 2, \sigma = 1.6$.

Figure 5.4: Horizontal (left) and vertical (right) sections of the heat plots 5.3a-5.3c.

5.3 Spectroscopy of phase transitions

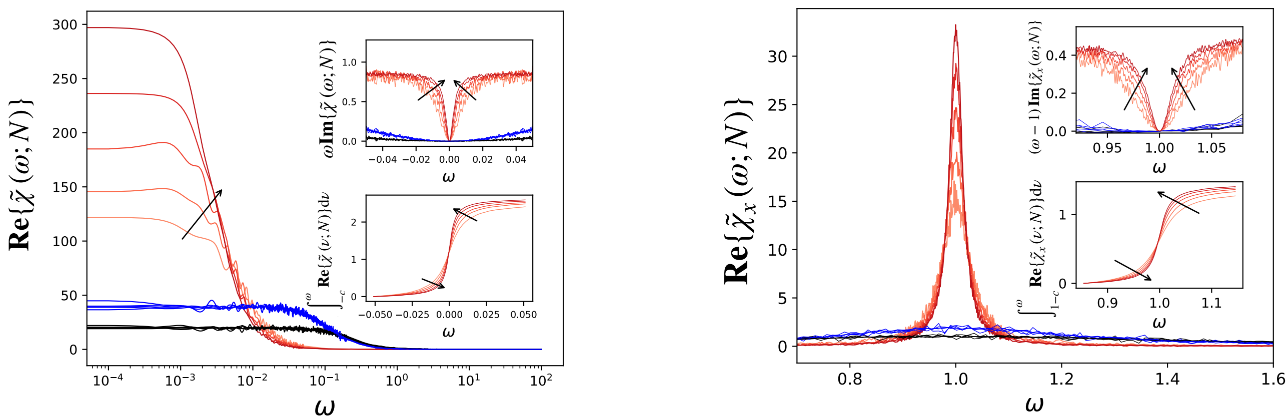
The goal of this section is to characterise the phase transitions of the previous two examples by investigating the response properties of the finite systems described by (5.1) and (5.16) as the number of the agents tend to infinity, reaching thus the thermodynamic limit. In particular, we repeat the response experiments, see discussion below, for various choices of N, in order to detect the emergence of singularities for the combination of the parameters corresponding to phase transitions. Here, we keep fixed the values of the internal parameter α and the coupling strength θ. Both models undergo a phase transition at the transition line $\tilde{\sigma} = \sigma(\theta, \alpha)$ in the parameter space (σ, θ, α) given by (5.3) for the Desai Zwanzig model and (5.27). Following [MPRV08], we perform n simulations where the initial conditions are chosen according to the unperturbed invariant measure $\rho_0(\mathbf{x})$ and where at time[2] $t = 0$ apply a perturbation proportional to a time-Dirac $T(t) = \delta(t)$ function. It is important to remark that we apply such perturbation to all the agents composing the system, thus modifying the dynamics of the whole system in a coherent way. We recall here that in the thermodynamic limit where the McKean Vlasov equation is

[2]We set here the origin of time when the perturbation is applied



(a) (b)

Figure 5.5: macroscopic response functions as a function of time t. Panel a: response $\tilde{G}(t;N)$ for the one dimensional order parameter of the DZ model. Panel b: response $\tilde{G}_x(t;N)$ of the first component of the bi-dimensional order parameter for the BCM model. Black and blue lines correspond to non critical values of the strength of the noise σ. Red lines correspond to response functions at the transition point. For each value of σ, there are five lines corresponding to different values of N, namely $N = 2^k \times 10^3$ with $k = 1, \ldots, 5$. The arrows and the colour gradient indicate the direction of increasing N.



(a) (b)

Figure 5.6: Macroscopic susceptibilities as a function of the frequency ω. Panel a: susceptibility $\tilde{\chi}(\omega)$ for the one dimensional order parameter of the DS model. Panel b: susceptibility $\tilde{\chi}_x(\omega)$ for the first component of the two dimensional order parameter for the BCM model. The parameter in the lower extreme of the integral is for both cases $c = 0.05$. Blue and black lines in panel 5.6b have been multiplied by a factor 5 for visualisation purposes. Colour code and plotting conventions as in Figure 5.5.

valid, the response of the variable x_i $i = 1, \dots, M$ after a perturbation of the deterministic part of the dynamics can be written in a linear response regime as, see (4.27),

$$\langle x_i \rangle_1 = \int_{-\infty}^{+\infty} T(s)\tilde{G}_i(t-s)\mathrm{d}s = \left(T \star \tilde{G}_i\right)(t). \tag{5.28}$$

Now, given the previous equation, it is clear that the response of the mean field observable x_i to a delta-like perturbation yields the *macroscopic* Green Function $\langle x_i \rangle_1(t) = \tilde{G}_i(t)$. Performing the aforementioned response experiments at a fixed value of number of particles N and looking at the response of the empirical order parameter $\bar{x} = \frac{1}{N}\sum_{k=1}^{N}$ over the n simulations gives an estimate $\tilde{G}_i(t;N)$ of the macroscopic response function. Figure 5.5 shows the response functions $\tilde{G}_i(\tau;N)$ for an additive perturbation $\mathbf{X}(\mathbf{x}) = 1$ for the DZ model (left panel) and $\mathbf{X}(\mathbf{x}) = (1,0)$ for the BCM model (right panel). The two response functions are qualitatively different because, by and large, the one for the DZ model describes a monotonic decay, whereby the system relaxes towards the unperturbed state, while the one for the BCM combines the decay with an oscillatory behaviour taking place at the natural frequency $\tilde{\omega} = 1$. In the DZ model, the response functions initially undergo a fast and substantial decay, both far from and at the phase transition, associated with a time scale of order 1. However, at the phase transition, a new, much longer, timescale appears. This timescale increases monotonically with N. The same is observed in the case of the BCM model if one considers the envelope of the response function rather than the response function itself: at the transition the decay of the oscillations towards the stable invariant measure becomes slower and slower as N increases.
The origin of the new timescales resides in the appearance of simple pole at $\omega = \omega_0$ in the *macroscopic* susceptibility $\tilde{\chi}(\omega)$, the Fourier transform of the response function. By applying the Fourier Transform to $\tilde{G}_i(\tau;N)$ we define the finite size *macroscopic* susceptibility

$$\tilde{\chi}_i(\omega;N) = \int \tilde{G}_i(t;N)e^{i\omega t}\mathrm{d}t. \tag{5.29}$$

Below we show that the response experiments clearly show the existence of a pole at the phase transition located at $\omega_0 = 0$ for the DZ model and at $\omega_0 = \tilde{\omega} = 1$ for the BCM model. When considering finite values of N, the susceptibilities describing the response of (virtually) any observable to (virtually) any external perturbation have a contribution of the form

$$\tilde{\chi}_i(\omega; N) = \frac{\tilde{\alpha}}{\omega - \omega_0 + i\gamma(N)} + r(\omega), \tag{5.30}$$

where $\tilde{\alpha}$ represents the residue of the pole and $\gamma(N) \to 0^+$ as $N \to +\infty$ and $r(\omega)$ is the regular part (for real frequencies) of the susceptibility, that, near a phase transition, becomes progressively negligible with respect to the resonant behaviour stemming from the pole ω_0. We observe that the singularity arises in the thermodynamic limit as

$$\lim_{N\to\infty} \frac{\tilde{\alpha}}{\omega - \omega_0 + i\gamma(N)} = -i\pi\tilde{\alpha}\delta(\omega - \omega_0) + \tilde{\alpha}\mathcal{P}\left(\frac{1}{\omega - \omega_0}\right). \tag{5.31}$$

Comparing (5.30) and the decomposition of the *macroscopic* susceptibility (4.33), we note that equation (5.30) takes into account the approaching of a Ruelle Pollicott resonance towards the point on the imaginary axis $\lambda = i\omega_0$ as the number of agents tend to infinity. We remark that the asymptotic property does not depend on how fast the function $\gamma(N)$ vanishes for increasing values of N. The residue $\tilde{\alpha} \in \mathbb{C}$ depends on the choice of observable and of the perturbation as given by (4.33) and Eq. (4.25), whereas the location of the pole solely depends on the spectral properties of the operator $\tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}$, or, more specifically by the invertibility properties of the matrix $P_{ij}(\omega)$, see (4.21). Figure 5.6 shows the properties of the finite size *macroscopic* susceptibilities. When $\sigma \neq \tilde{\sigma}$, the susceptibilities do not show any singularity nor any remarkable dependence on N, thus indicating that the thermodynamic limit has been reached to a good approximation and the response of the system is smooth. As N increases, for both the DZ model (left panel) and the BCM model (right panel) the resonance at $\omega = \omega_0$ of the real part of the susceptibility approaches the limiting Dirac function $\pi k\delta(\omega - \omega_0)$ with coefficient $k > 0$. This singular behaviour is clear from the plot of the primitive function of the real part of the susceptibility (bottom inset) that tends to step function, that is a constant times the Heaviside distribution $\Theta(\omega - \omega_0)$. We remark that the position of the poles that is observed in the finite size *macroscopic* susceptibility agrees with the analytical calculations shown in section 5.2

For both models, the residue $\tilde{\alpha} = \mathrm{i}|\tilde{\alpha}|$ is an imaginary number. Indeed, the imaginary part of the susceptibility behaves exactly as the Cauchy principal value distribution and can be used

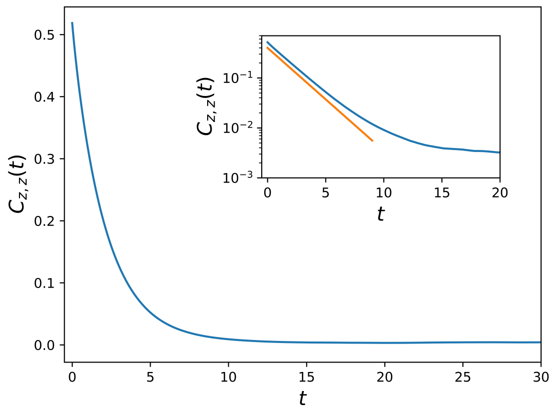


Figure 5.7: Mean field correlation $C_{z,z}(t)$ of the fluctuation variable $z = x - \langle x \rangle_0$ as a function of time. The orange line in the inset corresponds to an exponentially decaying function $y = 0.45e^{-t/\tau_z}$ where $\tau_z \approx 2.1$. The parameters of the model refer to the phase transition setting for the Desai Zwanzig model, see main text. The simulations have been performed on an ensemble of $N = 10^4$ agents.

to get easily a quantitative estimate of k. The top insets of Figure 5.6 shows the function $(\omega - \omega_0)\mathbf{Im}\{\tilde{\chi}_i(\omega)\}$. As $N \to \infty$, this function converges to k everywhere except for $\omega = \omega_0$. We can compare the value of the residue $\tilde{\alpha}$ in the case of the DZ model with the exact analytical result, see (5.15), $|\tilde{\alpha}| = \frac{1}{\theta \tau_z}$, where τ_z is the mean field autocorrelation time of the fluctuations $z(t) = x(t) - \langle x \rangle_0$. The timescale τ_z has been estimated from a time integral of the *mean field* correlation $C_{z,z}(t)$, see figure 5.7. A cutoff has been imposed on the time integral when the noisy signal takes over the exponential decay of the correlation function. As a result of the numerical simulations, we obtain $|\tilde{\alpha}| \approx 0.86$, which agrees within $\approx 2\%$ with the one resulting from the limiting behaviour of the susceptibility, thus validating our results. In the case of the BCM model, our procedure allows one to derive a direct estimate $|\tilde{\alpha}| \approx 0.44$; in this case no expression for the residue is available in the literature and, following [BCM87, LPZ20], its evaluation seems cumbersome. We here observe that, by evaluating the susceptibility for finite values of N, we are able to predict the residue of the pole at $\omega = \omega_0$, which appears, instead, only in the thermodynamic limit. As discussed earlier, the singular behaviour of the susceptibility has some degree of universality. By this we mean that while for a given model the value of the residue is forcing- and observable-dependent, its position is a fundamental property of the model itself. In order to show that the critical behaviour of the response does not depend on the type of perturbation, modulo a potential degenerate class of perturbations that have zero projection on the invariant measure $\rho_0(\mathbf{x})$, we report below the investigation the response of the DZ model for a state dependent perturbation $X(x) = x^2$, see figure 5.8. The *macroscopic* response function $\tilde{G}_i(\tau; N)$, both away and at the phase transition, has a rapid initial decay with a timescale that is different from the response function relative to a uniform perturbation. As a matter of fact, the timescale associated to the dominant mode of the response function for $t \to 0^+$ does in general depend on the applied perturbation, see section 4.3.1. As expected, the response function at the phase transition develops a much longer timescale that increases as the number of particle increases. A more accurate comparison with the result shown in the main text can only be performed in the frequency domain. Figure 5.8 (right panel) shows that, away from the transition, the susceptibilities have a smooth behaviour and no evident dependence on N. At the phase transition, the susceptibility develops the expected singular behaviour $\frac{\tilde{\alpha}}{\omega-\omega_0+i\gamma(N)}$, where $\gamma(N) \to 0^+$ as $N \to +\infty$ due to the appearance of a simple pole $\omega_0 = 0$. The residue $\tilde{\alpha}$ is purely imaginary and its magnitude can be inferred by visual inspection of the top inset representing the function$(\omega - \omega_0)\mathbf{Im}\{\tilde{\Gamma}(\omega)\}$ to be just less than 0.29. As a last remark, we provide some details about the numerical response experiments. The finite size *macroscopic* Green functions away from the transitions are estimated on an ensemble of $n = 10^5$ simulations, while the critical Green functions with $n = 10^6$ for Desai Zwanzig and $n = 7 \times 10^6$ for Bonilla-Casado-Morillo. Furthermore we investigate the response up to time $t = 5 \times 10^3$. The parameters $\theta = \alpha = 2$ have been fixed and the colour code for the figures is given by:

- black lines Desai Zwanzig $\sigma \approx 1$, Bonilla-Casado-Morillo $\sigma \approx 2$.
- blue lines Desai Zwanzig $\sigma \approx 0.87$, Bonilla-Casado-Morillo $\sigma \approx 1.8$
- red lines Desai Zwanzig $\tilde{\sigma} \approx 0.75$, Bonilla-Casado-Morillo $\tilde{\sigma} \approx 1.59$.

Linear Response as a way to investigate Critical Slowing Down in mean field models

The response theory we have developed in Chapter 4 and the numerical experiments reported above show that at a phase transition the *mean field* correlation functions (5.6) do not diverge,

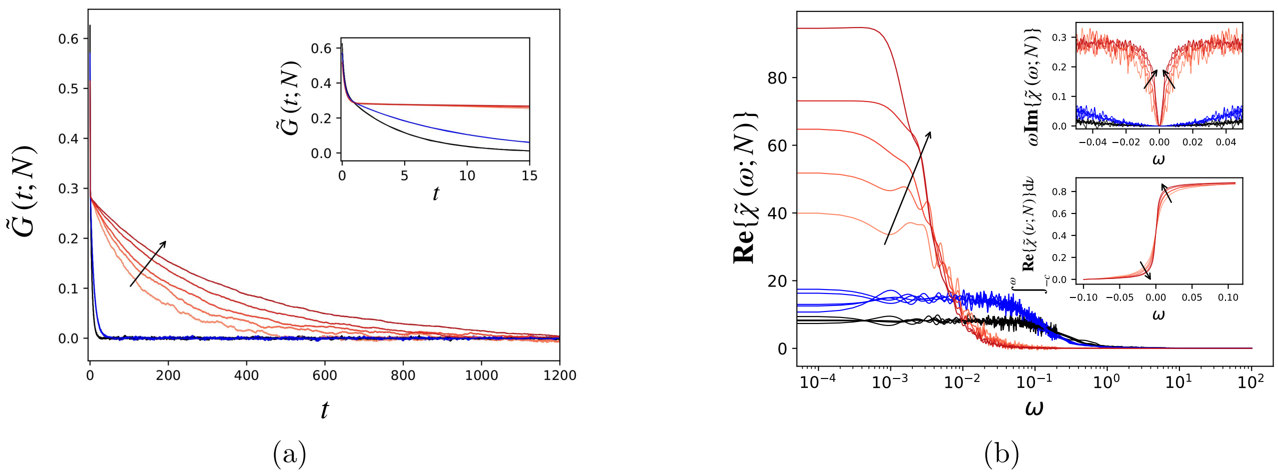


Figure 5.8: Panel a: Response function $\tilde{G}(t;N)$ and Panel b: susceptibility $\tilde{\chi}(\omega;N)$ for a spatially dependent perturbation $\mathbf{X}(\mathbf{x}) = x^2$ for the one dimensional order parameter for the DZ model. Colour code and plotting conventions as in Figure 5.5. The lower extreme of the integral of the bottom panel is $c = 0.05$.

since the Ruelle Pollicott resonances λ_j, $j = 1, \ldots, n$, of the operator $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$ are far away from the imaginary axis. On the other hand, in section 2.4 we have shown that a phase transition is characterised by *critical slowing down*, that is by large scale (in time and space) correlations among the particles, leading to a macroscopic emergent behaviour in the macroscopic observables of the system. Critical slowing down is observed as diverging correlation features of macroscopic observables when the thermodynamic limit $N \to +\infty$ is approached. From a mathematical standpoint, critical slowing down arises when one considers a *Central Limit Theorem* in the number of particles N around the McKean Vlasov equation, that can be considered in this respect a *Law of Large Numbers*. As observed in section 2.6, in order to study the critical fluctuations of the empirical measure for the Desai Zwanzig model one has to consider a scaling of the typical timescale of the fluctuations process of order $\tau_N = t/\sqrt{N}$, meaning that fluctuations persist, fixed N, over a longer timescale than the underlying microscopic dynamics. On the other hand, the McKean Vlasov equation can be essentially associated to a mathematically justified mean field ansatz, see equation (2.14), for the thermodynamic limit of the N-particle Fokker Planck equation where one considers all the particles to be statistically independent thus neglecting correlations among them. We remark that the *mean field* correlation functions (5.6) do not consider correlations among the agents and, for this reason, stay finite at the phase transition. On the contrary, if one were to consider macroscopic correlation functions of the macroscopic observable $\bar{x}(t) = \frac{1}{N}\sum_{k=1}^{N} x_k(t)$ over the full N-particle distribution ρ_N, namely

$$\langle \bar{x}(t)\bar{x}(0)\rangle_{\rho_N} = \frac{1}{N^2}\sum_{i,j}\langle x_i(t)x_j(0)\rangle_{\rho_N}, \tag{5.32}$$

then the timescale τ_N would manifest itself. For fixed N, the Response Theory developed in Chapter 3 establishes that, through the Fluctuation Dissipation Theorem, correlation properties are closely linked to sensitivity to external perturbations and, as such, a breakdown of response theory is expected at the phase transition as a result of the generator of the N-particle system developing a vanishing spectral gap $\gamma_N \to 0$ as $N \to +\infty$. We observe that the spectral gap γ_N is closely related to the log-Sobolev constant, see section 2.4, and that a vanishing spectral gap $\gamma_N \to 0$ implies that at a phase transition the log-Sobolev constant degenerates. The results we have presented in this chapter confirm and generalise what was originally proposed by Shiino [Shi85] for the Desai Zwanzig model, that is the fact that a Linear Response approach for the mean field model is a powerful way to investigate critical slowing down properties. The success of the response theory for the McKean Vlasov equation in predicting not only the location of the poles of the susceptibility but also the value of the residue in terms of *mean field* correlation functions, see also figure 6.3 for another example, is a striking phenomenon if one considers that the McKean Vlasov equation does not capture the correlations among the agents. This essentially boils down to two physical reasons. Firstly, equations (4.11) and (4.12) show that a Fluctuation Dissipation does not exist between the response of the system and the *mean field* correlation functions. A phase transition is rather characterised by a non trivial resonance, driven by the coupling among the agents, of the *mean field* correlation functions with the response of the system at time t and all previous times $s \leq t$, see (4.11), manifesting itself as a non invertibility of the matrix $P_{ij}(\omega)$. Secondly, the success of the theory is also due to our perturbation procedure of the N-particle system described in section 5.3. We recall that we perturb the deterministic part of the dynamics of *every* particle as $\mathbf{F}_\alpha(\mathbf{x}_k) \to \mathbf{F}_\alpha(\mathbf{x}_k) + \varepsilon T(t)\mathbf{X}(\mathbf{x}_k)$ with $k = 1,\ldots,N$, thus creating a macroscopic coherent perturbation of the system and projecting the Green Function of the N-particle system onto the eigenspaces of the operator $\tilde{\mathcal{L}}_{\langle\mathbf{x}\rangle_0}$ conducive to the phase transition, that is the ones corresponding to the Ruelle Pollicott resonances $\tilde{\lambda}_j$ with $\mathbf{Re}\{\tilde{\lambda}_j\} \to 0$ as the phase transition is approached. Different perturbation protocols of the N particle system could lead to different results. We conjecture that as long as a considerable fraction $\propto N$ of the agents is perturbed, similar results are to be expected as $N \to +\infty$, whereas this would not be the case if one were to perturb a fixed number $\bar{N}$ of agents. This interesting question is left for future work.

Chapter 6

Dimension Reduction for noisy interacting systems

6.1 Introduction

In this Chapter we investigate a different yet closely related feature regarding interacting agent systems, namely their dimension reduction properties in the thermodynamic limit. In this regime, such models often exhibit phase transitions as a result of the complex interplay between the interacting dynamics and the noise, see Chapter 2. Singularities associated to phase transitions, such as the divergence of correlation properties, the so called *critical slowing down*, and the breakdown of linear response properties, see Chapters 4 and 5 can only be observed in the thermodynamic limit. Consequently their investigation involves the study of the nonlinear and nonlocal McKean Vlasov equation or a *brute force* approach, i.e. extensive numerical simulations of very large ensemble of agents as elucidated in Chapter 5. Reduction of complexity can be achieved by defining collective variables (reaction coordinates) able to accurately describe the full dynamics in a low dimensional space. Nonetheless, while order parameters like magnetization can in many cases be easily deduced for equilibrium systems using, e.g. symmetry arguments, the definition of reaction coordinates for nonequilibrium system is far more challenging [MD05, BLP06, Rog21].

The goal of this Chapter is to present a *model reduction* approach for the study of such infinite systems based on a systematic approximation of the full infinite dimensional dynamics in terms of a *low* number of ODEs. We provide a systematic dimension reduction methodology for constructing low dimensional, reduced-order dynamics based on the cumulants of the probability distribution of the infinite system. We show that the low dimensional dynamics returns the correct *diagnostic* properties since it produces a quantitatively accurate representation of the stationary phase diagram of the system that we compare with exact analytical results and numerical simulations. Moreover, we prove that the reduced order dynamics yields the *prognostic*, i.e., time dependent properties too as it provides the correct response of the system to external perturbations. On one hand, this validates the use of our complexity reduction methodology since it retains information not only of the invariant measure of the system but also of the transition probabilities and time dependent correlation properties of the stochastic dynamics. On the other hand, the breakdown of linear response properties is a key signature of a phase transition phenomenon. We show that the reduced response operators capture the correct diverging resonant behaviour by quantitatively assessing the singular nature of the susceptibility of the system and the appearance of a pole for real value of frequencies. Consequently, this methodology can be interpreted as a low dimensional, reduced order approach to the investigation and detection of critical phenomena in high dimensional interacting systems in settings where order parameters are not known. In particular, we recall that here we will provide examples referring to quadratic interactions as in Chapter 4 but our methodology includes more general classes of interactions with numerous applications including synchronisation of nonlinear, possibly chaotic, oscillators [BCM87, PKRK03] and emergent phenomena in neural networks and life sciences [CDPF15, DP19].

Dimension reduction techniques for high dimensional systems usually refer to the Mori-Zwanzig formalism [Mor65, Zwa61] leading to a Generalised Langevin Equation that describes, on a formal level, the effect of the neglected degrees of freedom on the resolved dynamical variables one wishes to include in their dimension reduction problem. Practical approximations of the Generalised Langevin equation have been investigated by using two complementary approaches, namely data-driven and top-down methods. Successful data driven methods, such as Dynamic

Mode Decomposition [SCH10] rely on the use of the eigenvalues and eigenvectors of the whole system transfer operator as suitably selected basis functions, that are able to capture the modes of variability of the underlying dynamics. Other types of mode decomposition selection exist and have been successfully applied to the prediction of complex phenomena [GAD+02, CK17]. We also mention that machine learning approaches, such as variational autoencoders [KW14], have also been used to construct a surrogate, low dimensional representation of the system. Other data driven techniques, such as empirical model reduction [KCG15], can be used to obtain closure methods from partial observations of the system. The resulting closure structure is given in terms of multilayer stochastic systems whose relevance and robustness has also been highlighted from an alternative, theory-informed parametrization perspective [SGLCG21]. We also mention that for chaotic and turbulent system an approach based on Unstable Periodic Orbits (UPOs) theory has been successful in determining approximations of transfer operators governing the dynamics of the system [CAM+05]. By preferring a UPOs partition of the phase space, rather than a blind Ulam's partition, one is able to exploit the intrinsic topology of the flow to approximate in a systematic way transfer operators. Methods based on UPOs theory have been a powerful tool for the investigation of mixing properties of quasi-invariant sets of the phase space [MLG22] and for obtaining coarse grained, low dimensional descriptions of the dynamics [YHB21].

On the other hand, top-down approaches usually rely on approximations of the dynamical equations governing the evolution of the system to construct a stochastic, possibly non-Markovian, parametrization [WL12, WL13, CLW15]. Our dimension reduction methodology fits in this latter class of techniques and is based on a suitable closure method of the infinite hierarchy of equations for the moments or, equivalently, cumulants of the probability distribution of the infinite dimensional system. Such closure method results in a deterministic parametrization of the full dynamics in terms of a low number of cumulants. The focus in this chapter is not about establishing the link between the Generalised Langevin Equation and our methodology but rather providing a suitable low dimensional approximation of the transfer and response operator that govern the dynamics of the system in the thermodynamic limit. On one side, we show that the stationary properties of the reduced dynamics yield a quantitatively accurate representation of the stationary phase diagram of the system by comparing it with exact analytical results and numerical simulations. On the other side, we test the accuracy of the reduced dynamics in representing the response of the system to perturbations and show that phase transitions are associated with the breakdown of the reduced linear response operators and the emergence of poles in the susceptibility for real values of the frequency. This validates our methodology as a suitable low dimensional approach to detecting phase transitions in high dimensional systems by looking at resonances and singularities of reduced order response operators. As validation case studies, we apply our dimension reduction methodology to investigate the nonequilibrium continuous phase transition in a model featuring noise-induced stabilisation phenomena [VdBPAHM94] and a model featuring an equilibrium discontinuous transition [GKPY19].

6.2 The class of models

We consider a system of exchangeable weakly interacting one-dimensional diffusions whose dynamics is governed by the following Stratonovich SDE[1]

$$\mathrm{d}x_i = \left[F_\alpha(x_i) - \frac{\theta}{N} \sum_j^N \mathcal{U}'\left(x_i - x_j\right) \right] \mathrm{d}t + \sigma(x_i) \circ \mathrm{d}W_i, \tag{6.1}$$

with initial condition $x_i \sim \rho_{in}(x)$ and $i = 1, \ldots, N$. Each agent undergoes an internal dynamics given by the vector field $F_\alpha(x)$, depending on a set of parameters α, and is coupled with all the other agents through a symmetric interaction potential $\mathcal{U}(x) = \mathcal{U}(-x)$, with θ denoting the interaction strength. Furthermore, $\mathrm{d}W_i$, $i = 1, \ldots, N$, are independent Brownian motions and $\sigma(x) > 0$ $\forall x \in \mathbb{R}$ is a multiplicative, state dependent, diffusion coefficient. The main assumption here is that $F(x), \mathcal{U}(x)$ and the one particle diffusion matrix $\Sigma(x) = \sigma^2(x)$ all have a polynomial functional form. For simplicity, we consider quadratic interactions, $\mathcal{U}(x) = \frac{x^2}{2}$, as in the previous chapters. This corresponds to cooperative interactions among the agents

[1] At variance with the previous chapters, we here consider a Stratonovic prescription for the noise. As it is well known, it is possible to change prescriptions by suitably modifying the drift term of a stochastic differential equation, see also section 6.2.1 .

that attempt to synchronise them towards their common centre of mass $\bar{x}(t) = \frac{1}{N}\sum_i^N x_i(t)$. We remark that this methodology applies to generic polynomial interaction potentials and can also be easily generalised to higher dimensions. We are interested in the thermodynamic limit $N \to +\infty$ of equations (6.1). It is known, see Chapter 2, that the one particle distribution of the N-particle system converges to the distribution $\rho(x,t)$ satisfying the McKean Vlasov partial differential equation (2.12), that, according to our setting, can be written as

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\frac{\sigma^2(x)}{2}\rho\frac{\partial}{\partial x}\left(f_{\langle x\rangle}(x) + \ln\rho\right)\right) = \mathcal{L}_{\langle x\rangle}\rho(x,t), \tag{6.2}$$

where $\rho(x,0) = \rho_{in}(x)$ and

$$f_{\langle x\rangle}(x) = 2\int^x \frac{-\hat{F}_\alpha(y) + \theta\left(y - \langle x\rangle\right)}{\sigma^2(y)}\mathrm{d}y + \ln\sigma^2(x), \tag{6.3}$$

where $\langle x\rangle$ represents the first moment of the distribution $\rho(x,t)$ and $\hat{F}_\alpha(x) = F_\alpha(x) + \frac{1}{2}\sigma(x)\sigma'(x)$. We recall that (6.2) exhibits in general non-uniqueness of stationary solutions, where exchanges of stability and appearance/disappearance of stationary solutions can be interpreted as phase transitions .

We can characterise the stationary solutions of (6.2) as a one parameter family of distributions

$$\rho_0(x;m) = \frac{e^{-f_m(x)}}{\int_{\mathbb{R}} e^{-f_m(x)}\mathrm{d}x} \equiv \frac{e^{-f_m(x)}}{Z(m)}, \tag{6.4}$$

where the parameter m satisfies the selfconsistency equation

$$m = R(m) \equiv \int_{\mathbb{R}} x\rho_0(x;m)\mathrm{d}x \tag{6.5}$$

and $Z(m) > 0$ denotes the partition function. Equation (6.5) plays a major role in determining the stationary properties of the system. Solutions $m^\star$ of (6.5) correspond to stationary measures $\rho_0(x;m^\star)$ with first moment $\langle x\rangle = m^\star$, a suitable order parameter of the system for this type of quadratic interactions, see also the examples in chapter 5. Partial information on the stability of the invariant measures can be obtained by the investigation of the slope of the selfconsistency equation $R'(m^\star) = \frac{\mathrm{d}R(m)}{\mathrm{d}m}|_{m^\star}$. In particular, if $R'(m^\star) > 1$, the stationary solution $\rho(x;m^\star)$ is unstable.

6.2.1 Two paradigmatic examples

Below we provide some details on the two models we will investigate in this chapter. The first model (model A) was introduced in [VdBPAHM94] to study the effect of multiplicative noise on spatially extended systems. We consider the Desai Zwanzig model, see section 2.6, settings where the local dynamics $F(x) = -V'_\alpha(x)$ is given by a double well potential $V_\alpha(x) = \frac{x^4}{4} - \alpha\frac{x^2}{2}$ and the noise is additive $\sigma(x) = \sigma$. The equations for motions are given by, considering the quadratic coupling,

$$\mathrm{d}x_i = \left[\alpha x_i - x_i^3 - \theta\left(x_i - \bar{x}\right)\right]\mathrm{d}t + \sigma \mathrm{d}W_i, \tag{6.6}$$

where, given the additivity of the noise, the prescription for the noise is no longer a relevant issue. We assume now that the parameter α is not known exactly but rather erratically fluctuates in time, i.e. $\alpha \to \alpha + \sigma_m \mathrm{d}\xi$ where $\mathrm{d}\xi$ is another, uncorrelated, Brownian motion. This results in a set of equations for the N interacting agents that reads

$$\mathrm{d}x_i = \left[-V'(x_i) - \theta\left(x_i - \bar{x}\right)\right]\mathrm{d}t + \sigma_m x_i \circ^\nu \mathrm{d}\xi + \sigma \mathrm{d}W_i, \tag{6.7}$$

where the symbol $\circ^\nu$ stands for a generic (not necessarily Ito) prescription for the equations. It is convenient to write the above set of equations in the equivalent, in law, form

$$\mathrm{d}x_i = \left[-V'(x_i) - \theta\left(x_i - \bar{x}\right)\right]\mathrm{d}t + \sigma(x_i) \circ^\nu \mathrm{d}W_i, \tag{6.8}$$

where $\sigma(x) = \sqrt{\sigma^2 + \sigma_m^2 x^2}$ is a state dependent stochastic term. It is well known that the presence of multiplicative noise introduce a modelling issue, since it is not clear, a priori, what prescription should be given to the stochastic integral defining the stochastic equation [Pav14, Kli90, vK81]; see also discussion in [GL22]. We interpret equations (6.8) as a generic one parameter family of stochastic integrals parametrised by a parameter $\nu \in [0, 1]$. Different values of ν correspond to different prescription of the SDEs. In particular, $\alpha = 0, 1/2, 1$ correspond to the Ito, Stratonovich and Klimontovich prescription respectively. Different conventions of the stochastic integral lead to different stability properties of the SDE. Remarkably, the convention for a given system might also vary depending on the operational conditions [PMH+13]. In the following we choose a Stratonovich convention $\nu = \frac{1}{2}$. It is known that a generic SDE can be transformed into an Ito-SDE by suitably modifying the drift coefficient as $F_\alpha(x) \to F_{\alpha,\nu}(x) = F_\alpha(x) + \nu\sigma(x)\sigma'(x)$ [Pav14]. Since it is more convenient to work with the Ito prescription, we apply this transformation to equations (6.8) and obtain

$$\mathrm{d}x_i = \left[-V_\nu(x_i) - \theta\left(x_i - \bar{x}\right)\right]\mathrm{d}t + \sigma(x_i)\mathrm{d}W_i, \tag{6.9}$$

where $V_\nu(x) = V_\alpha(x) + \nu\sigma_m^2\frac{x^2}{2} = \frac{x^4}{4} - \left(\alpha + \nu\sigma_m^2\right)\frac{x^2}{2}$.

The introduction of a fluctuating parameter in the drift term corresponds to applying an external, state-dependent noise that breaks the detailed balance condition, thus driving the $N-$particle system to an out of equilibrium state. Equation (6.3) is for this model

$$\begin{aligned} f_{\langle x\rangle}(x) = &-\frac{\alpha - \theta + (\nu-1)\,\sigma_m^2 + \frac{\sigma^2}{\sigma_m^2}}{\sigma_m^2}\ln\left(1 + \left(\frac{\sigma_m}{\sigma}x\right)^2\right) + \\ &+ \frac{x^2}{\sigma_m^2} - 2\frac{\theta\langle x\rangle}{\sigma\sigma_m}\arctan\left(\frac{\sigma_m}{\sigma}x\right). \end{aligned} \tag{6.10}$$

The analysis of the self consistency equation (6.5) provides insightful information on the stationary phase diagram of the model. In particular, symmetries of the problem force the system to always have the trivial solution $m^\star = 0$, corresponding to disordered state $\rho_0(x;0)$ of vanishing order parameter. This can be easily shown by observing that $R(-m) = -R(m)$ since stationary distributions satisfy $\rho_0(x;m) = \rho_0(-x;-m)$, see equations (6.3) and (6.4). Moreover, if $m^\star$ is a solution of the self consistency equation, so is $-m^\star$. We thus expect that two symmetric branches of stable solutions will arise as soon as the disordered state loses stability. The disordered state becomes unstable as soon as $R'(0) = 1$, that can be written as

$$\frac{\theta}{\sigma\sigma_m}\langle x\arctan\left(\frac{\sigma_m}{\sigma}x\right)\rangle_0 = \frac{1}{2}, \tag{6.11}$$

where the expectation value $\langle\cdot\rangle_0$ is taken with respect to the stationary distribution $\rho_0(x;0)$. Since the order parameter at the transition point is known to be $m^* = 0$, the above equation

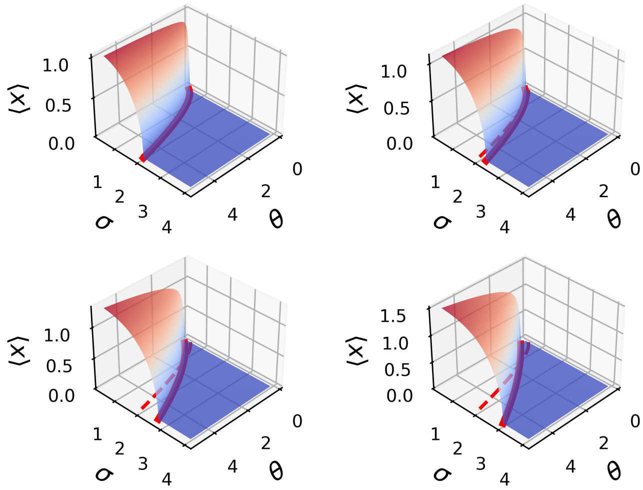


Figure 6.1: Order parameter $\langle x \rangle$ as a function of (σ, θ) obtained via the self consistency equation analysis. The red dashed line and the continuous red line represent the exact transition curve for $\sigma_m = 0$ and $\sigma_m \neq 0$, see equation (6.11). The other parameters of the model are fixed and equal to $\alpha = 1, \sigma_m = 1.5, \nu = 1/2$.

yields, fixed all the other parameters, the critical value $\sigma_c = \sigma_c(\alpha, \theta, \sigma_m)$ of the strength of the additive noise. We recall that (6.11) represents a generalisation to this nonequilibrium setting of equation (5.2) for the Desai Zwanzig model. Figure 6.1 shows the multiplicative noise induced stabilisation phenomenon we mentioned before. Indeed, the multiplicative noise has a rectifying effect, pushing, for strong enough coupling θ, the transition point to higher and higher values of σ. Moreover, the amplitude of the order parameter gets magnified, since it exceeds the maximum value $\sqrt{\alpha}$, the minimum point of the potential $V_\alpha(x)$, that is attained in the low noise regime ($\sigma \to 0$) when $\sigma_m = 0$.

The second model (model B) we have investigated features a discontinuous phase transition and is obtained by breaking the symmetry $x \to -x$ of the Desai Zwanzig model through a *tilted* potential as $V_{\alpha,\mu} = V_\alpha + \mu x$, with $\mu > 0$. Moreover, the system is subject to thermal noise $\sigma(x) = \sigma$. The pitchfork bifurcation of invariant solutions one obtains for $\mu = 0$ disappears in this case. In particular, there exists a smooth, stable branch of negative order parameter $\langle x \rangle$ for all values of the strength of the noise σ. However, decreasing σ, a pair of solutions appear through a saddle node bifurcation, yielding another branch of stable $\langle x \rangle > 0$, with the other one being unstable, see panel (b) of figure 6.2. The saddle node bifurcation is characterised by the condition $R'(m_c) = 1$ that reads

$$\frac{\theta}{\sigma^2}\langle (x - m_c)^2 \rangle_0 = \frac{1}{2}, \tag{6.12}$$

where m_c is the value of the positive order parameter at the transition point and the expectation value is taken with respect to the stationary distribution $\rho_0(x; m_c)$. Since m_c is not known a priori and has to be evaluated numerically by solving the self consistency equation, the above equation does not directly provide the value of the critical noise σ_c at which the saddle node bifurcation takes place. Nevertheless, it provides a criterion to assess how close the critical point evaluated numerically is to the exact one by evaluating the slope $R'(m_c)$ and comparing it to the exact value 1.

6.3 Reduced order dynamics

In order to construct the reduced order dynamics, we multiply (6.2) by x^n, $n \in \mathbb{N}$, and integrate over $\mathbb{R}$. Given our assumptions on the drift and diffusion terms, this procedure results in an infinite hierarchy of equations for the moments $M_n = \langle x^n \rangle$ of the probability distribution $\rho(x, t)$. In order to further elucidate on this, we will first consider model A defined in the previous section. We recall that model A features a continuous phase transition, see figure 6.1 or panel (a) of figure 6.2. The aforementioned procedure yields the following equations for the moments M_n

$$\begin{aligned}\frac{dM_n}{\mathrm{d}t} = n\left(\alpha - \theta + \frac{n}{2}\sigma_m^2\right) M_n - nM_{n+2} + \\ + \frac{n(n-1)}{2}\sigma^2 M_{n-2} + n\theta M_1 M_{n-1},\end{aligned} \tag{6.13}$$

with $M_0 = 1$, $M_{-1} \equiv 0$ being the (lower) boundary conditions. Firstly, we observe that the global coupling among the agents gives rise to an interaction term between the order parameter $\langle x \rangle = M_1$ and all the other moments M_n. Secondly, the nonlinear features of the

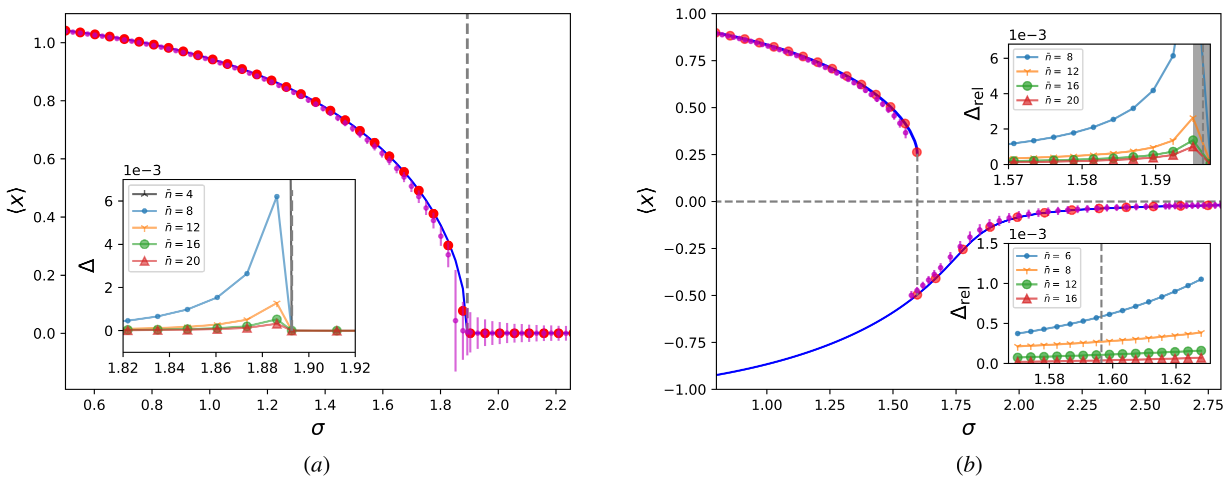


Figure 6.2: Phase diagram, $\langle x \rangle = \langle x \rangle(\sigma)$. The continuous blue line refers to the selfconsistency equation, the red dots to the reduced order dynamics ($\bar{n} = 4$) and the magenta dots (with errorbars) to the numerical integration. Panel (a): continuous transition given by model A. The inset at the bottom shows the absolute error Δ between the reduced order dynamics and the selfconsistency approach. The error Δ for $\bar{n} = 4$ is out of scale and peaks at a value $\Delta \approx 0.1$. The vertical dashed line refers to the critical condition $R'(0) = 1$. Fixed parameters are $(\alpha, \theta, \sigma_m) = (1, 4, 0.8)$. Panel (b): discontinuous phase transition of model B. The insets show the relative error Δ_{rel} between the reduced order dynamics and the selfconsistency approach. The inset at the top (bottom) refers to the upper (lower) branch of the phase diagram. The vertical dashed line is obtained numerically through the selfconsistency approach and its value has been consistently checked to yield a slope $R'(m)$ such that $R'(m) - 1 \approx 10^{-4}$. Fixed parameters are $(\alpha, \theta, \mu) = (1, 4, 0.02)$.

dynamics given by $V_\alpha(x)$ couple lower moments with higher degree ones. The infinite hierarchy of moment equations (6.13) is equivalent to (6.2) and no reduction in the level of complexity of the mathematical description has been accomplished yet. On a practical level, the necessity of finding appropriate closure schemes for the hierarchy arises. Were we to truncate the system of equations (6.13) at a specific level $\bar{n}$, a closure scheme for $M_{\bar{n}+1}$, $M_{\bar{n}+2}$ in terms of M_n with $n < \bar{n}$ is needed.

Inspired by [DZ78], we implement a cumulant truncation scheme [WB70, SS90, Bov78]. We introduce the cumulants k_n as the coefficients of the Taylor expansion of the cumulant generating function

$$\sum_{n=1}^{+\infty} k_n(t) \frac{\lambda^n}{n!} = \ln \int_{\mathbb{R}} \rho(x,t) e^{\lambda x} \mathrm{d}x. \tag{6.14}$$

The truncation scheme consists of imposing the condition $k_{\bar{n}+1} = k_{\bar{n}+2} = 0$. This procedure provides a closure relations for $\bar{M}_{\bar{n}+1} = \bar{M}_{\bar{n}+1}(M_1, \ldots, M_{\bar{n}})$ and $\bar{M}_{\bar{n}+2} = \bar{M}_{\bar{n}+2}(M_1, \ldots, M_{\bar{n}})$. Alternatively, one can obtain from (6.14) and (6.2) an infinite hierarchy of equations for the cumulants

$$\frac{\mathrm{d}k_n}{\mathrm{d}t} = G_n(k_1, \ldots, k_n, k_{n+1}, k_{n+2}), \tag{6.15}$$

where the nonlinear function $G_n(\cdot)$ is written in appendix B. Equation (6.15) indicates that the cumulant truncation scheme corresponds to a parametrization of the dynamics given by (6.1), in the limit $N \to +\infty$, in terms of a finite number $\bar{n}$ of cumulants. It is well known that such a scheme is inconsistent, since a function with a finite cumulant expansion cannot be positive if the order of the highest cumulant is larger than two [Gre71]. However, a parametrization in terms of cumulants is expected to perform better than parametrizations in terms of (central) moments based on the observation that a Gaussian distribution has vanishing cumulants $k_n = 0$ for $n > 2$, while all (central) moments are nonzero. For non-Gaussian distributions, one expects that neglected higher-order cumulants will be smaller than the corresponding (central) moments. For model A, equation (6.5) predicts that the stable solution $\langle x \rangle = 0$ bifurcates when $R'(0) = 1$, see (6.11), through a continuous phase transition in two symmetric, competing states with opposite order parameter. Panel (a) of figure 6.2 shows the continuous phase diagram for the state with positive order parameter, obtained with the exact selfconsistency equation and the reduced order dynamics (6.15). As soon as $\bar{n} = 4$ cumulants (main panel) are introduced, the reduced dynamics provides a very good approximation of the phase diagram. We observe that the reduced dynamics converges from below - the true transition point is underestimated by the reduced dynamics - to the exact phase diagram as higher truncation are considered. The accuracy of the reduced dynamics has been quantitatively assessed in terms of the absolute error Δ (shown in the inset) with respect to the selfconsistency approach. The reduced dynamics has also been compared to numerical simulations of an ensemble of $N = 12000$ agents described by equations (6.1). We have used the Milstein scheme [KP11], that has strong order of convergence 1, with time step $\Delta t = 0.01$ and estimated the order parameter as the time average, at stationarity, of the center of mass $\bar{x}(t)$. Moreover, the reduced order dynamics has been initialised with a Gaussian initial condition, such that $(k_1, k_2) = (0.1, 0.01)$ and all others cumulants set to zero. Very good agreement is observed between the reduced order and the full mean field dynamics. As for the behaviour of the finite size system ($N < \infty$) near the phase transition, one observes some discrepancies in the numerical simulations (not shown in the figure) with respect to the infinite size system. This is due to the fact that for values of the noise strength σ just below the transition point, the thermodynamic limit of the ensemble of agents is multistable as equation (6.5) supports two symmetric solutions. For finite systems, as observed in section 2.5, metastability phenomena are observed with the average exit time from one of the two symmetric solutions exponentially increasing with the number of particles. Fixed the number of particles N, the average exit time decreases exponentially as the transition point is approached. Consequently, noise-induced transitions among the two symmetric solutions become a relevant feature of the dynamics nearby the transition point and one should consider the *rectified* order parameter (shown in the figure), obtained as the time average of $\bar{x}(t)$ conditioned on the fact that the system is in the basin of attraction of the positive solution. Given the symmetry of the problem, one can also consider the quantity $|\bar{x}(t)|$ when the system fluctuates closely to the mean field invariant solutions, that is neglecting the transitions. We have also probed the validity of the cumulant based parametrization by investigating discontinuous phase transitions. Below we consider model B, which, we recall, is defined by a tilted potential $V_{\alpha,\mu} = V_\alpha(x) + \mu x$ and additive noise. Stationary properties of the reduced dynamics (6.15), with Gaussian initial condition $(k_1, k_2) = (1, 0.01)$, are in very good agreement with the other two approaches, see Panel (b). The insets show the relative error Δ_{rel} between the reduced dynamics and the selfconsistency equation. The top one, referring to the top branch of the phase diagram, shows that in the very close proximity, represented as a shaded area, of the transition point, Δ_{rel} jumps to higher values, due to the fact that the reduced dynamics' prediction for the transition point, depending on the level of truncation $\bar{n}$, underestimates the true one and approaches it from below as $\bar{n}$ increases. The bottom inset shows that Δ_{rel} for the bottom branch of the phase diagram is instead a smooth function that is not affected by the transition. This confirms that the reduced dynamics is able to track, as σ is parametrically changed, the disappearing attractor until jumping to the other stable, smoothly changing, attractor. Once again, noise-induced transitions are observed close to the phase transition in the finite system. Due to the asymmetry between the two competing states, the metastable lifetime of the state with $\langle x \rangle > 0$ decreases as the transition is approached and the system, after a short time, is driven to the other state of much longer lifetime. This can be more easily understood by considering the free energy functional, see section 2.3, associated to the mean field dynamics for model B

$$F[\rho] = \int V_{\alpha,\mu}(\mathbf{x})\rho(\mathbf{x})\mathrm{d}\mathbf{x} + \frac{\theta}{4}\int\int \rho(\mathbf{x})\left(\mathbf{x}-\mathbf{y}\right)^2\rho(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} + \frac{\sigma^2}{2}\int \rho(\mathbf{x})\ln\rho(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{6.16}$$

The expected escape time from the mean field attractor with positive order parameter $\rho_+(x)$ can be written in terms of the above functional as

$$\mathbb{E}[\tau_+] \asymp e^{\frac{2}{\sigma^2}N\Delta F}, \tag{6.17}$$

where $\asymp$ denotes an asymptotic relation valid in the thermodynamic limit $N \to +\infty$ and $\Delta F = F[\rho_0] - F[\rho_+]$ where ρ_0 is the unstable (saddle node) solution with vanishing order parameter, see section 2.5 for further details. An analogous formula holds for the expected escape time from the mean field invariant measure ρ_- with negative order parameter. We can then estimate the ratio between the escape times as

$$\frac{\mathbb{E}[\tau_+]}{\mathbb{E}[\tau_-]} \asymp e^{-\frac{2}{\sigma^2}N\delta F}, \tag{6.18}$$

where $\delta F = F[\rho_+] - F[\rho_-] > 0$. The escape time from ρ_+ is exponentially smaller than the one from ρ_- and, considering that δF does not vanish approaching a discontinuous phase transition, the finite size system will stay for a very short time in $\rho_+$and then transition to ρ_-, from which, given its exponentially longer life time, will not transition back if not on an exponentially longer time scale.

6.3.1 Reduced order response operators

So far we have illustrated the static properties of the system when perturbed with an adiabatic change of its parameters. Below, we investigate the nonautonomous, dynamical response to a time-dependent external perturbation. We report linear response properties of the reduced dynamics. We perturb a stable stationary state by modifying the drift term as $F(x) \to F(x) +$

$\varepsilon X(x)T(t)$, where ε is small. We here apply a uniform perturbation $X(x) = 1$, resulting in a one-cumulant perturbation $k_1^{(0)} \to k_1^{(0)} + \varepsilon$ for equations (6.15), where $k_1^{(0)}$ is the unperturbed order parameter. We then observe the *macroscopic* Green function $\tilde{G}(t)$, associated to the order parameter, defined in a linear response regime as

$$k_1(t) = k_1^{(0)} + \varepsilon \int_{-\infty}^{\infty} \tilde{G}(t-s)T(s)\mathrm{d}s. \tag{6.19}$$

As in the previous chapters, we choose as temporal modulation for the forcing a Dirac's δ, $T(t) = \delta(t)$, which corresponds to a broad band forcing in frequency space. We recall that in these settings, by observing the perturbed $k_1(t)$ we obtain the Green Function as $\tilde{G}(t) = \frac{k_1(t) - k_1^{(0)}}{\varepsilon}$. Convergence to the linear regime has been assessed evaluating the response for different values of ε. Panel (a) of Fig. 6.3 shows that, at the transition point (red lines), the Green function has an exponential decay (bottom inset) with an associated timescale that is order of magnitudes greater than what is observed in non-critical settings (blue and black lines). Moreover, such timescale is an increasing function of the level of truncation $\bar{n}$ of the reduced dynamics, whereas no dependence on $\bar{n}$ is observed for the non-critical Green functions (main inset). The critical behaviour is linked to the breakdown of linear response theory at the phase transition point, in the thermodynamic limit of equation (6.1) due to the agent-to-agent interactions, thus being associated with endogenous dynamical processes. This critical behaviour can be associate to spectral properties of the operator $\tilde{\mathcal{L}}_{\langle x \rangle_0}$ or invertibility properties of the matrix $P_{ij}(\omega)$, see chapter 4. As the number of agents N is increased, one observes an emerging singular behaviour in the *macroscopic* susceptibility $\tilde{\chi}(\omega)$, defined as the Fourier Transform of $\tilde{G}(t)$, signalled by a development of a pole ω_0 on the real axis of the frequencies as explained in chapter 5. We here observe an analogous scenario, in terms of $\bar{n}$ rather than N, see also equation (5.30), where the susceptibility can be written as

$$\tilde{\chi}(\omega) = \frac{\tilde{\alpha}}{\omega - \omega_0 + i\gamma(\bar{n})} + r(\omega), \tag{6.20}$$

where $\omega_0 = 0$ and $r(\omega)$ is an analytic function in the upper complex ω plane. As $\bar{n} \to +\infty$, $\gamma(\bar{n}) \to 0$ and the susceptibility develops a singular behaviour given by $\lim_{\bar{n}\to\infty} \chi(\omega) =$

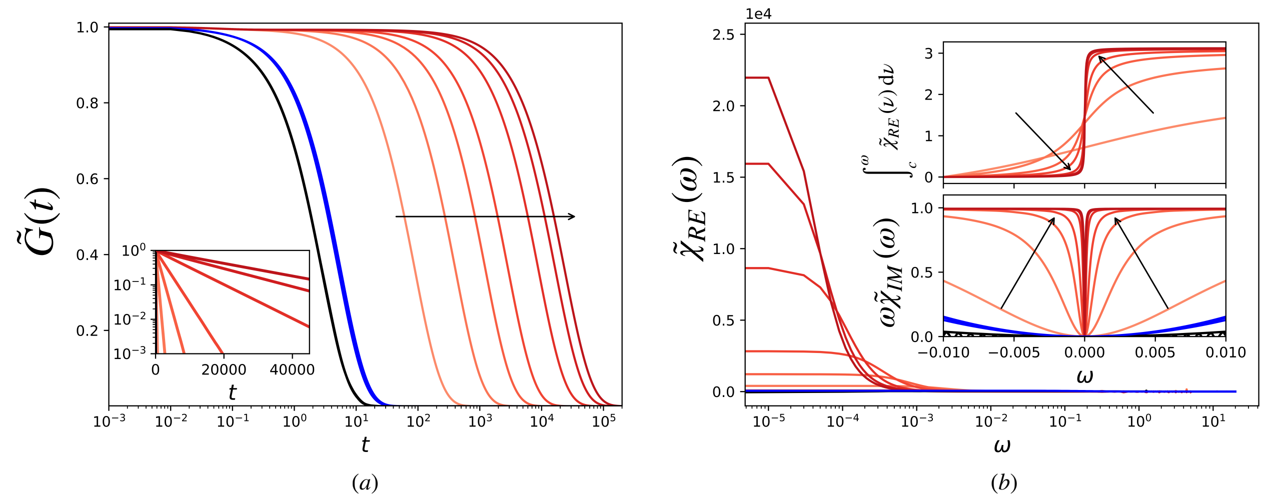


Figure 6.3: macroscopic Green function $\tilde{G}(t)$ (panel (a)) and macroscopic susceptibility $\tilde{\chi}(\omega)$ (panel (b)) for model A. The blue (black) lines refer to a non-critical setting 5% below (above) the transition point. Red lines refer to critical settings. Fixed parameters are as in Panel (a) of Fig. 6.2. The red color code and the arrows correspond to increasing values of $\bar{n} = 4, 6, 8, 10, 14, 18, 22$. In panel (b) black and blue lines have been multiplied by a scaling factor for graphical purposes.

$-i\pi\tilde{\alpha}\delta(\omega-\omega_0) + \tilde{\alpha}\mathcal{P}\left(\frac{1}{\omega-\omega_0}\right) + r(\omega)$. Panel (b) confirms the appearance of an emerging pole with an imaginary residue $\tilde{\alpha} = i|\tilde{\alpha}|$. The real part χ_{RE} (main panel) of the susceptibility clearly shows the resonant $\delta-$like behaviour for $\omega = \omega_0$. Alternatively, the top inset shows that the primitive function of χ_{RE} close to the pole ($c = -0.01$) converges accordingly to a Heaviside function. We observe that $\bar{n} = 4$ does not show a resonant behaviour, even though it is associated with a longer timescale. The imaginary part $\chi_{IM}(\omega)$ of the susceptibility (bottom inset), behaving like a Cauchy principal value distribution, yields a quantitative estimate $|\tilde{\alpha}| \approx 1$ for the residue of the pole.

An exact formula for the residue

It is possible to obtain a formula for the amplitude of the residue of the *macroscopic* susceptibility. Invariant measures $\rho_0(x)$ of the McKean Vlasov equation (6.2) satisfy the eigenvalue problem $\mathcal{L}_{\langle x\rangle_0}\rho_0(x) = 0$, where the linear differential operator $\mathcal{L}_{\langle x\rangle_0}$ is defined by

$$\mathcal{L}_{\langle x\rangle_0}\psi(x) = \frac{\partial}{\partial x}\left(\frac{\sigma^2(x)}{2}\psi\frac{\partial}{\partial x}\left(f_{\langle x\rangle_0}(x) + \ln\psi\right)\right), \tag{6.21}$$

where $\psi(x)$ is a smooth function and $f_{\langle x\rangle_0}(x)$ is defined in equation (6.3). The response of the observable x after a perturbation of the system can be written, see (4.24),

$$\langle x\rangle_1(\omega) = \tilde{\chi}(\omega)T(\omega), \tag{6.22}$$

where the *macroscopic* susceptibility $\tilde{\chi}(\omega)$ is

$$\tilde{\chi}(\omega) = P^{-1}(\omega)\chi(\omega) = \frac{\chi(\omega)}{1-\theta\chi(\omega)}. \tag{6.23}$$

The microscopic susceptibility $\chi(\omega)$ is related to microscopic correlation properties of the system in the unperturbed state described by ρ_0. In particular, $\chi(\omega)$ is the Fourier Transform of the microscopic response function $G(t)$ that can be written as a suitable correlation function, see (4.12). For gradient systems with thermal noise, it is possible to write $G(t)$ as a time derivative of suitable correlation properties. We remark that for general non equilibrium systems this is not always possible. However, given the structure of the problem, we are able find an analogous formula for $G(t)$. In fact, using (4.12) and the known form of the invariant measure (6.4) we have

$$G(t) = \Theta(t)\int x e^{t\mathcal{L}_{\langle x\rangle_0}}\rho_0(x)\frac{\partial}{\partial x}f_{\langle x\rangle_0}(x)\mathrm{d}x. \tag{6.24}$$

We now define the function $g(x) = -\frac{1}{\sigma\sigma_m}\arctan\left(\frac{\sigma_m}{\sigma}x\right)$ such that its derivative is $\frac{\partial g(x)}{\partial x} = -\frac{1}{\sigma^2(x)}$. We then evaluate the following expression

$$\begin{aligned}\mathcal{L}_{\langle x\rangle_0}(g\rho_0) &= \frac{\partial}{\partial x}\left(\frac{\sigma(x)^2}{2}g\rho_0\frac{\partial}{\partial x}\left(f_{\langle x\rangle_0}(x)+\ln\rho_0+\ln g\right)\right) = \\ &= \frac{\partial}{\partial x}\left(\frac{\sigma(x)^2}{2}g\rho_0\frac{\partial}{\partial x}\ln g\right) = \frac{\partial}{\partial x}\left(\frac{\sigma(x)^2}{2}\rho_0\frac{\partial}{\partial x}g\right) = \\ &= -\frac{1}{2}\frac{\partial}{\partial x}\rho_0 = +\frac{1}{2}\rho\frac{\partial}{\partial x}f_{\langle x\rangle_0}(x),\end{aligned} \tag{6.25}$$

where we have used the fact that $f_{\langle x\rangle_0}(x) + \ln\rho_0 = Z =$ constant. The microscopic response function can thus be written as

$$\begin{aligned} G(t) &= -\frac{2}{\sigma\sigma_m}\Theta(t)\int \mathrm{d}x x \exp\left(\mathcal{L}_{\langle x\rangle_0} t\right)\mathcal{L}_{\langle x\rangle_0}\arctan\left(\frac{\sigma_m}{\sigma}x\right)\rho_0(x) = \\ &= -\frac{2}{\sigma\sigma_m}\Theta(t)\frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}x x \exp\left(\mathcal{L}_{\langle x\rangle_0} t\right)\arctan\left(\frac{\sigma_m}{\sigma}x\right)\rho_0(x) = \\ &= -\frac{2}{\sigma\sigma_m}\Theta(t)\frac{\mathrm{d}}{\mathrm{d}t}C_{x,A}(t), \end{aligned} \tag{6.26}$$

where in the last line we have introduced the *mean field* correlation function between observable x and observable $A = \arctan\left(\frac{\sigma_m}{\sigma}x\right)$, defined according to equation (2.22). The microscopic susceptibility can thus be written as

$$\chi(\omega) = \int_{-\infty}^{+\infty} G(t)e^{i\omega t}\mathrm{d}t = \frac{2}{\sigma\sigma_m}\left(C_{x,A}(0) + i\omega\hat{C}_{x,A}(\omega)\right), \tag{6.27}$$

where $\hat{C}_{x,A}(\omega) = \int_0^{+\infty} e^{i\omega t}C_{x,A}(t)$ is the (one-sided) Fourier transform of the correlation function $C_{x,A}(t)$. We can then show that the macroscopic susceptibility $\tilde{\chi}(\omega)$ develops a singular behaviour for a real frequency $\omega_0 = 0$ at the phase transition. Let us observe that equation (6.11), that characterises the phase transition line, can be written as

$$\frac{\theta}{\sigma\sigma_m}C_{x,A}(0) = \frac{1}{2} \tag{6.28}$$

since $\langle x\rangle_0 = 0$ at the transition point. In conclusion, using all the above results, the *macroscopic* susceptibility at the phase transition $\tilde{\chi}(\omega)$ can be written as

$$\tilde{\chi}(\omega) = -\frac{1}{\theta} + i\frac{1}{\omega}\frac{\sigma\sigma_m}{\theta^2\hat{C}_{x,A}(\omega)} = -\frac{1}{\theta} + i\frac{1}{\omega}\frac{C_{x,A}(0)}{\theta\hat{C}_{x,A}(\omega)}. \tag{6.29}$$

Being related to the spectral properties of the operator $\mathcal{L}_{\langle x\rangle_0}$, see chapter 4, the quantity $\hat{C}_{x,A}(\omega)$ is an analytical function in the upper complex plane (including the real axis). Consequently, the above equation shows that linear response theory breaks down at the phase transition, with

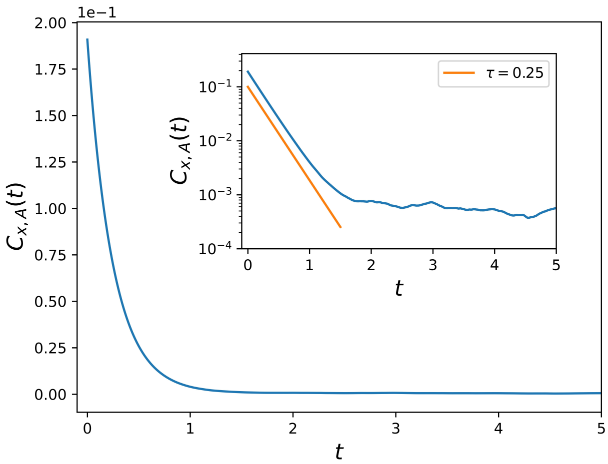


Figure 6.4: Correlation function $C_{x,A}(t)$ as a function of time. The orange line in the inset corresponds to an exponentially decaying function $y = 0.1e^{-t/\tau}$ where $\tau = 0.25$. The parameters of the model are the same as in Figure 6.3 of the main text.

the *macroscopic* susceptibility $\chi(\omega)$ developing a simple pole in $\omega = \omega_0 = 0$ with residue

$$\tilde{\alpha} = \operatorname*{Res}_{\omega=\omega_0} \chi(\omega) = \frac{i}{\theta} \frac{C_{x,A}(0)}{\hat{C}_{x,A}(0)} = \frac{i}{\theta \tau_{x,A}}, \tag{6.30}$$

where $\tau_{x,A}$ is the integrated auto-correlation time defined by

$$\tau_{x,A} = \frac{\hat{C}_{x,A}(0)}{C_{x,A}(0)} = \frac{\int_0^{+\infty} C_{x,A}(t)\mathrm{d}t}{C_{x,A}(0)}. \tag{6.31}$$

We observe that, as $\sigma_m \to 0$, the above equation is compatible with (5.15) obtained for the Desai Zwanzig model. We can now compare the value of the residue obtained from the reduced order dynamics and (6.31) where the *mean field* correlation functions are estimated from numerical simulations as described by equation (5.6). We have performed simulations over an ensemble of $N = 16000$ agents. The integrated correlation time $\tau_{x,A}$ has been estimated by imposing a cut off $T = 1.5$ on the time integral corresponding to the moment after which the noisy signal takes over the exponential decay of the correlation function (see inset of Figure 6.4). The resulting value is $\tau = 0.25091$ with corresponding amplitude, we recall $\theta = 4$, of the residue $|\tilde{\alpha}| = 0.99636$, which agrees to a very good approximation with what has been obtained through the reduced order dynamics, see Figure 6.3 in the main text. We remark that the existence of the pole ω_0 at the phase transition, as opposed to its residue $\tilde{\alpha}$, depends neither on the

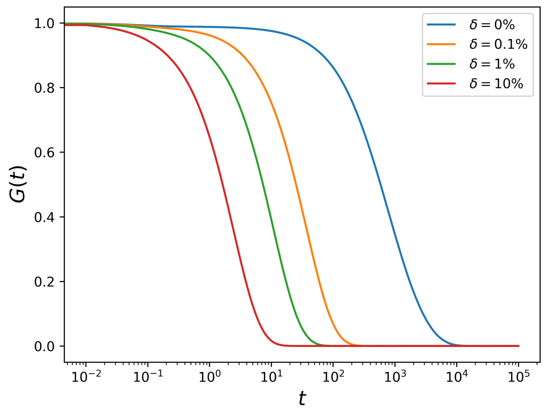


Figure 6.5: *macroscopic* Green Function $\tilde{G}(t)$ as a function of time for model B. δ represents the relative distance from the approximate phase transition point, see the discussion in the main text.

forcing $T(t)$ nor on the choice of the observable and can be related to spectral properties of suitably defined evolution operators, see chapter 4. This crucial property validates the use of our cumulant based reduced dynamics to settings where the order parameter is not known or cannot easily be written in terms of the cumulants.

Response for the reduced order dynamics for model B

The most interesting feature of model B is undoubtedly the response to a static modulation of its parameters corresponding to the jump characterising the discontinuous phase transition, see panel (b) of figure 6.2. Nevertheless one could study the dynamical response of the system as the transition point is approached (from below) on the top branch. Since it is associated with the loss of stability of the invariant measure, we expect similar results to hold for this model as well. We evaluate $\tilde{G}(t)$ associated to the order parameter $\langle x \rangle$ using a perturbation protocol analogous to the previous section. Figure 6.5 shows that for settings near the phase transition, one of the Ruelle Pollicott resonances of the reduced response operator associated to the reduced cumulants dynamics is approaching the imaginary axis. As a result, one exponential mode of the spectral decomposition of the macroscopic Green function is associated to a timescale that is orders of magnitude bigger than any exponential mode in non critical settings. Alternativey, such critical behaviour can be related to the development of a static

$\omega_0 = 0$ pole of the *macroscopic* susceptibility of the system. We remark that the figure refers to a level of truncation of $\bar{n} = 22$. One could also perform an analysis, similar to figure 6.3, by looking at different values of $\bar{n}$. We expect to obtain similar results. However, such analysis is more complicated here due to the discontinuous feature of the transition. Firstly, the reduced dynamics transition point depends on $\bar{n}$ and the analysis becomes increasingly hard very close to the transition point, see shaded area in panel (b) of Figure 6.2. Secondly, Figure 6.5, clearly shows that the timescale associated to the Green function is highly sensitive to small deviations, as small as $\delta = 0.1\%$, from the transition point.

To conclude, in this chapter we have obtained a reduced low-dimensional system for the moments of the probability distribution function of the mean field dynamics. We showed that such approximate dynamics provides an accurate representation of the stationary phase diagram, even for a very low number (e.g. 4) of cumulants. This indicates that the cumulants act as effective reaction coordinates, which are able to capture the essential properties of the system with moderate loss of information due to the cumulant truncation. Additionally, the linear response properties of the projected dynamics agrees with that of the full system, and the breakdown of the corresponding linear response operators can be used to characterise the phase transition occurring in the system. Hence, our methodology seems useful for performing linear stability analysis for a large class of interacting multiagent systems, and for predicting their response to forcings of general nature. We remark that, for polynomial drift and diffusion coefficients, the (untruncated) moments system is exact.

Chapter 7

Conclusions

In this thesis we have investigated critical phenomena for the thermodynamic limit of weakly interacting diffusions, i.e. ensembles of identical, exchangeable interacting agents. By looking at the response of the system to (weak) external perturbations we have developed a linear response theory describing the induced change of statistical properties of observables of the infinite system of agents. In particular, we associated the development of a critical phenomenon to the settings of a non smooth, singular response of the system and the appearance of resonant poles for the complex valued susceptibility describing the linear response of the system.
More specifically, we have shown that the response in frequency space of the macroscopic system to any time modulated forcing is given in terms of a *macroscopic* susceptibility $\tilde{\chi}(\omega)$ that captures not only the response of the single agents composing the system, given by the *microscopic* susceptibility $\chi(\omega)$, but also the effect of the interactions among them.
By using functional analysis methods and the spectral theory of Markov semigroups, we have shown that the properties of the response can be characterised in terms of suitable operators governing the time evolution of the unperturbed system. In particular, the *microscoscopic* susceptibility $\chi(\omega)$ is intimately linked to spectral properties of $\mathcal{L}_{\langle \mathbf{x} \rangle_0}$, the generator of the time evolution of microscopic observables in the unperturbed state. On the contrary, the response of the full system at time t is not only associated to correlation properties of microscopic observables but also to memory effect between the previous-time response, i.e. the response at

any time $s \leq t$, originating from the coupling among the system. As such, the susceptibility $\tilde{\chi}(\omega)$ is related to a more complicated operator $\tilde{\mathcal{L}}_{\langle \mathbf{x} \rangle_0}$ that describes the response of the system and that cannot be interpreted as generating mixing properties of the microscopic observables.

On one hand, we have fully analysed the smooth properties of the response and developed dispersion (Kramers-Kronig) relations and sum rules, deriving essentially from *causality* properties, for the macroscopic system far away from critical settings.
On the other hand, we have proved that, at criticality, the response of the system breaks down, signalled by a non analytical behaviour of the macroscopic susceptibility given by the development of a pole on the real axis of frequencies. We have been able to identify two different mechanisms leading to such a pole corresponding to two different scenarios of criticality, *critical transitions* and *phase transitions.* For simplicity we report here equation (4.25)

$$\tilde{\chi}_i(\omega) = \sum_{j=1}^{M} \mathbf{P}_{ij}^{-1}(\omega)\chi_j(\omega). \tag{7.1}$$

that links the properties of the response of the full system, given by $\tilde{\chi}(\omega)$, with the response of the single microscopic agents, expressed by $\chi(\omega)$, through the matrix $P_{ij(\omega)}$.

1. If the non analytical behaviour derives from a pole in the *microscopic* susceptibility $\chi_i(\omega)$, the system undergoes a critical phenomenon that conforms to the classic scenario of *critical transitions* for finite dimensional systems where one expects a divergence of correlation properties of the microscopic degrees of freedom due to the spectral gap of the transfer operator of the unperturbed system becoming vanishingly small.

2. When the pole derives from non invertibility properties of the renormalisation matrix $P_{ij}(\omega)$, the system undergoes a qualitatively different critical scenario that can be interpreted as a *phase transition.* This critical behaviour is not accompanied by a divergence of correlation properties of the single agents but arises, only in the thermodynamic limit, as a result of the interactions among them.

Through the spectral decomposition of Markov semigroups in terms of stochastic Ruelle Pollicott resonances we have established a link between the singular properties of the response at criticality and the spectral properties of dynamical operators. Focusing on phase transitions, we have shown through spectroscopic numerical experiments that there is a clear signature of the development of a pole on the real axis of frequencies when the thermodynamic limit is approached for both equilibrium and nonequilibrium systems. For finite size systems, the pole manifests itself as a resonance of the response with a finite width, which decreases as the thermodynamic limit is approached. We are also able to fully characterise the singular behaviour of the susceptibility by evaluating the value of the residue of the pole in terms of microscopic correlation functions in the unperturbed state.

The response theory we have developed, being linked to spectral properties of suitable response operators, shows that the singular behaviour of the susceptibility of the full system does not depend on the applied forcings nor on the observable under investigation, providing some degree of universality. Classical approaches to the investigation of *phase transitions* require the identification of order parameters, i.e. suitable observables of the system able to capture some degree of macroscopic behaviour of the system, which is most often a non trivial task to achieve, especially in nonequilibrium settings. The response theory perspective we have adopted in this thesis provides an alternative approach that is able to bypass the problem of the detection of order parameters.
In this regard, in the last chapter of the thesis we provide a dimension reduction methodology, based on a cumulant expansion of the probability distribution of the system in the thermodynamic limit, that allows to create a reduced order dynamics in terms of a (low) number of cumulants, that act as effective reaction coordinates for the system. By performing response experiments on the reduced dynamics, we have shown that the exact phase transition can be detected through the resonant, singular behaviour of the reduced order susceptibility of the system.

In conclusion, the work we have developed in this thesis indicates that response theory for identical interacting systems is a very powerful conceptual tool to the detection and investigation of critical phenomena, framed in the mathematical framework of spectral properties of dynamical operators. There are multiple directions of future research in this area, such as

- the analysis of more complicated interaction patterns among the agents. In particular, the spectral theory of selfadjoint operators (in suitable Hilbert spaces) would allow one to study general multidimensional equilibrium reversible interacting systems featuring phase transitions for generic interaction potentials.
- the designing of early warning indicators for critical phenomena alternative to the extensive literature that exists for critical transitions.
- the systematic investigation of the optimal class of observables for which the divergence of response properties would yield the best detection of the critical phenomena. Clearly, this fundamental problem is closely related to the identification of order parameters for such systems.
- the interpretation, in a response theory perspective, of phenomena of collective behaviour (synchronisation, cooperation, consensus, etc...) in both natural and social sciences.
- the designing of reduced order models based on the dimension reduction methodology performed in the last chapter for more complicated systems of weakly interacting diffusions, with applications to neural networks and life sciences [DP19, CDPF15].

Appendix A

Cumulant Truncation Scheme

In this appendix we provide the algebra to perform a cumulant truncation scheme at any generic order n. We remark that the equations of moments approach, and the cumulant truncation scheme, can be easily generalised to higher dimensions. Firstly, we observe that the relationship between cumulants and moments of a probability distribution is

$$k_n = \sum_{l=1}^{n} (-1)^{l-1}(l-1)! B_{nl}(M_1, \ldots, M_{n-l+1}) \tag{A.1}$$

where $B_{nl}(M_1, \ldots, M_{n-l+1})$ are partial (incomplete) Bell polynomials. In particular, these polynomials are given by

$$B_{nl}(M_1, \ldots, M_{n-l+1}) = \sum \frac{n!}{j_1! j_2! \ldots j_{n-l+1}!} \left(\frac{M_1}{1!}\right)^{j_1} \left(\frac{M_2}{2!}\right)^{j_2} \ldots \left(\frac{M_{n-l+1}}{(n-l+1)!}\right)^{j_1}$$

where the sum is taken over all the sequences $j_1 j_2 \ldots j_{n-l+1}$ of non negative integers such that the following two conditions hold

$$j_1 + j_2 + \ldots j_{n-l+1} = l$$

$$j_1 + 2j_2 + \cdots + (n-l+1) j_{n-l+1} = n$$

Moreover, we will make extensive use of the following two properties of the Bell polynomials

$$B_{n1}(M_1,\ldots,M_n) = M_n \tag{A.2}$$

$$B_{n2}(M_1,\ldots,M_{n-1}) = \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k} M_k M_{n-k} \tag{A.3}$$

The closure approximation $\bar{M}_{\bar{n}+1}$ can be easily found by separating the term $l = 1$ from equation A.1 and using A.2,

$$k_n = M_n + \sum_{l=2}^{n}(-1)^{l-1}(l-1)! B_{nl}(M_1,\ldots,M_{n-l+1}) \tag{A.4}$$

In fact, evaluating the above equation for $n = \bar{n}+1$ and imposing the condition $k_{\bar{n}+1} = 0$ results in

$$\bar{M}_{\bar{n}+1} = -\sum_{l=2}^{\bar{n}+1}(-1)^{l-1}(l-1)! B_{\bar{n}+1,l}(M_1,\ldots,M_{\bar{n}+2-l}) \tag{A.5}$$

The evaluation of $\bar{M}_{\bar{n}+2}$ requires more care since it involves $\bar{M}_{\bar{n}+1}$ as well. Let us first observe that the cumulant $k_{\bar{n}+2}$ can be written as, see equation A.1,

$$\begin{aligned} k_{\bar{n}+2} = {} & M_{\bar{n}+2} - B_{\bar{n}+2,2}(M_1,\ldots,M_{\bar{n}+1}) + \\ & + \sum_{l=1}^{n}(-1)^{l-1}(l-1)! B_{\bar{n}+2,l}(M_1,\ldots,M_{\bar{n}+3-l}) \end{aligned} \tag{A.6}$$

Using equation A.3 we can write

$$B_{\bar{n}+2,2}(M_1,\ldots,M_{\bar{n}+1}) = (\bar{n}+2)M_{\bar{n}+1}M_1 + \sum_{k=2}^{\bar{n}}\binom{\bar{n}+2}{k} M_k M_{\bar{n}+2-k} \tag{A.7}$$

where we have separated the term $k = 1$ and $k = \bar{n}+1$ from the total sum.

Finally, by imposing the condition $k_{\bar{n}+2} = 0$ and consistently estimating $M_{\bar{n}+1}$ as $\bar{M}_{\bar{n}+1}$ we obtain the approximated value for $M_{\bar{n}+2}$ as

$$\begin{aligned}\bar{M}_{\bar{n}+2} = (\bar{n}+2)\bar{M}_{\bar{n}+1}M_1 + \frac{1}{2}\sum_{k=2}^{\bar{n}} \binom{\bar{n}+2}{k} M_k M_{\bar{n}+2-k} - \\ - \sum_{l=3}^{\bar{n}+2} (-1)^{(l-1)}(l-1)! B_{\bar{n}+2,l}(M_1, \ldots, M_{\bar{n}+3-l})\end{aligned} \tag{A.8}$$

In conclusion, the cumulant truncation scheme consists in the finite set of equations 6.13 with $n = 1, \ldots, \bar{n}$ along with the boundary conditions $M_0 = 1$ and $M_{\bar{n}+1} = \bar{M}_{\bar{n}+1}$, $M_{\bar{n}+2} = \bar{M}_{\bar{n}+2}$ as given by equations A.5 and A.8 respectively.

Appendix B

Equations for the Cumulants

In this section we will provide a few more details on how to obtain the dynamical evolution of the cumulants of the distribution of the infinite system $\rho(x,t)$. Our analysis follows closely the one in [DZ78]. The calculations below refer to model A, since similar results hold for model B. As explained in the main text, $\rho(x,t)$ satisfies a non linear and non local fokker planck equation that we write here in an alternative way as

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\left(V'_{\alpha,\nu}(x) + \theta\left(x - \langle x\rangle\right)\right)\rho\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2(x)\rho\right) \tag{B.1}$$

The cumulants k_n are defined by the cumulant generating function $G(\lambda,t) = \ln g(\lambda,t)$

$$\sum_{n=1}^{\infty} k_n(t)\frac{\lambda^n}{n!} = \ln\int \rho(x,t)e^{\lambda x}\mathrm{d}x \equiv \ln g(\lambda,t) \tag{B.2}$$

Equation B.1 yields an evolution equation for the cumulant generating function

$$\frac{dG}{dt} = -\frac{\lambda}{g}\int \mathrm{d}x\left(x^3 - \left(\alpha - \theta + \nu\sigma^2 x^2\right) - \theta\langle x\rangle\right)\rho e^{\lambda x} + \frac{\lambda^2}{2g}\int \mathrm{d}x\left(\sigma^2 + \sigma_m^2 x^2\right)\rho e^{\lambda x}$$

By separating the different powers of the variable x we can write the above equation in terms of G, its derivative $G'(\lambda, t) = \frac{\partial G}{\partial \lambda}$ and higher order derivatives as

$$\begin{aligned}\frac{dG}{dt} &= \lambda\theta\langle x\rangle + \frac{\lambda^2\sigma^2}{2} + \lambda(\alpha - \theta + \nu\sigma^2)G' + \frac{\lambda^2\sigma_m^2}{2}\left(G'^2 + G''\right) - \\ &- \lambda\left(G'G'^2 + 3G'G'' + G'''\right)\end{aligned} \tag{B.3}$$

Using the definition of the cumulants given in equation B.2 and comparing same powers of λ one finally obtains the equations for the cumulants

$$\begin{aligned}\frac{1}{n}\frac{\mathrm{d}k_n}{\mathrm{d}t} &= \theta k_1\delta_{1n} + \frac{\sigma^2}{2}\delta_{n2} + \left(\alpha - \theta + \sigma_m^2\left(\nu + \frac{n-1}{2}\right)\right)k_n - k_{n+2} + \\ &+ \sigma_m^2(1-\delta_{n1})\frac{(n-1)!}{2}\sum_{i=1}^{n-1}\frac{k_i k_{n-i}}{(i-1)!(n-i-1)!} - \\ &- 3(n-1)!\sum_{i=1}^{n}\frac{k_i k_{n-i+2}}{(i-1)!(n-i)!} - \\ &- (n-1)!\sum_{i=1}^{n}\sum_{j=1}^{n-i+1}\frac{k_i k_j k_{n+2-i-j}}{(i-1)!(j-1)!(n-i-j+1)!}\end{aligned} \tag{B.4}$$

Bibliography

[ABPV+05] Juan A. Acebrón, L. L. Bonilla, Conrad J. Pérez Vicente, Félix Ritort, and Renato Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.*, 77:137–185, Apr 2005.

[Abr17] Rafail V. Abramov. Leading order response of statistical averages of a dynamical system to small stochastic perturbations. *J. Stat. Phys.*, 166(6):1483–1508, 2017.

[Abr19] Rafail V. Abramov. A theory of average response to large jump perturbations. *Chaos*, 29(8):083128, 19, 2019.

[AM07] R.V. Abramov and A.J. Majda. Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems. *Nonlinearity*, 20(12):2793, 2007.

[AM08] Rafail V. Abramov and Andrew J. Majda. New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems. *J. Nonlinear Sci.*, 18(3):303–341, 2008.

[AR04] Andrew J. Archer and Markus Rauscher. Dynamical density functional theory for interacting Brownian particles: stochastic or deterministic? *J. Phys. A*, 37(40):9325–9333, 2004.

[Arn92] V. I. Arnold. *Catastrophe theory.* Springer-Verlag, Berlin, third edition, 1992. Translated from the Russian by G. S. Wassermann, Based on a translation by R. K. Thomas.

[AWVC12] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox. Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1962):1166–1184, 2012.

[Bal00] V. Baladi. *Positive Transfer Operators and Decay of Correlations.* World Scientific, Singapore, 2000.

[Bal14] Viviane Baladi. Linear response, or else. In *ICM Seoul 2014, Proceedings*, volume III, page 525–545, 2014.

[BB05] Jean-Philippe Bouchaud and Giulio Biroli. Nonlinear susceptibility in glassy systems: A probe for cooperative dynamical length scales. *Phys. Rev. B*, 72:064204, Aug 2005.

[BBM10] Florent Barret, Anton Bovier, and Sylvie Méléard. Uniform Estimates for Metastable Transition Times in a Coupled Bistable System. *Electronic Journal of Probability*, 15(none):323 – 345, 2010.

[BBS15] Viviane Baladi, Michael Benedicks, and Daniel Schnellmann. Whitney-Hölder continuity of the SRB measure for transversal families of smooth unimodal maps. *Invent. Math.*, 201(3):773–844, 2015.

[BCM87] Luis L. Bonilla, JoséM. Casado, and Manuel Morillo. Self-synchronization of populations of nonlinear oscillators in the thermodynamic limit. *Journal of Statistical Physics*, 48(3):571–591, 1987.

[BDLV01] L. Biferale, I. Daumont, G. Lacorata, and A. Vulpiani. Fluctuation-response relation in turbulent systems. *Phys. Rev. E*, 65:016302, Dec 2001.

[BDSG+15] Lorenzo Bertini, Alberto De Sole, Davide Gabrielli, Giovanni Jona-Lasinio, and Claudio Landim. Macroscopic fluctuation theory. *Rev. Modern Phys.*, 87(2):593–636, 2015.

[BEGK04] Anton Bovier, Michael Eckhoff, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.

[Bel80] Thomas L. Bell. Climate sensitivity from fluctuation dissipation: Some simple model tests. *Journal of Atmospheric Sciences*, 37(8):1700 – 1707, 1980.

[BER89] V. Baladi, J.-P. Eckmann, and D. Ruelle. Resonances for intermittent systems. *Nonlinearity*, 2(1):119–135, 1989.

[BFG07a] Nils Berglund, Bastien Fernandez, and Barbara Gentz. Metastability in interacting nonlinear stochastic differential equations: I. from weak coupling to synchronization. *Nonlinearity*, 20(11):2551–2581, oct 2007.

[BFG07b] Nils Berglund, Bastien Fernandez, and Barbara Gentz. Metastability in interacting nonlinear stochastic differential equations: II. large-nbehaviour. *Nonlinearity*, 20(11):2583–2614, oct 2007.

[BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[BGN16] Freddy Bouchet, Krzysztof Gawędzki, and Cesare Nardini. Perturbative calculation of quasi-potential in non-equilibrium diffusions: A mean-field example. *Journal of Statistical Physics*, 163(5):1157–1210, 2016.

[BKPP21] A. Borovykh, N. Kantas, P. Parpas, and G. A. Pavliotis. On stochastic mirror descent with interacting particles: convergence properties and variance reduction. *Phys. D*, 418:Paper No. 132844, 21, 2021.

[BLP06] Giovanni Bussi, Alessandro Laio, and Michele Parrinello. Equilibrium free energies from nonequilibrium metadynamics. *Phys. Rev. Lett.*, 96:090601, Mar 2006.

[BM21] K. Bashiri and G. Menz. Metastability in a continuous mean-field model at low temperature and strong interaction. *Stochastic Processes and their Applications*, 134:132–173, 2021.

[BMM12] Marko Budišić, Ryan Mohr, and Igor Mezić. Applied Koopmanism. *Chaos*, 22(4):047510, 33, 2012.

[Bov78] D.C.C Bover. Moment equation methods for nonlinear stochastic systems. *Journal of Mathematical Analysis and Applications*, 65(2):306–320, 1978.

[BR16] Freddy Bouchet and Julien Reygner. Generalisation of the Eyring-Kramers transition rate formula to irreversible diffusion processes. *Ann. Henri Poincaré*, 17(12):3499–3532, 2016.

[BS08] Viviane Baladi and Daniel Smania. Linear response formula for piecewise expanding unimodal maps. *Nonlinearity*, 21(4):677–711, 2008.

[BS10] Viviane Baladi and Daniel Smania. Alternative proofs of linear response for piecewise expanding unimodal maps. *Ergodic Theory Dynam. Systems*, 30(1):1–20, 2010.

[BT08] J. Binney and S. Tremaine. *Galactic Dynamics*. Princeton University Press, Princeton, second edition, 2008.

[BW22] Erhan Bayraktar and Ruoyu Wu. Stationarity and uniform in time convergence for the graphon particle system. *Stochastic Processes and their Applications*, 150:532–568, 2022.

[CAC21] Bruno Cessac, Ignacio Ampuero, and Rodrigo Cofré. Linear response of general observables in spiking neuronal network models. *Entropy*, 23(2):Paper No. 155, 30, 2021.

[CAM+05] Predrag Cvitanovic, Roberto Artuso, Ronnie Mainieri, Gregor Tanner, Gábor Vattay, Niall Whelan, and Andreas Wirzba. Chaos: classical and quantum. *ChaosBook. org (Niels Bohr Institute, Copenhagen 2005)*, 69:25, 2005.

[CCH14] José Antonio Carrillo, Young-Pil Choi, and Maxime Hauray. *The derivation of swarming models: Mean-field limit and Wasserstein distances*, pages 1–46. Springer Vienna, Vienna, 2014.

[CCY19] José A. Carrillo, Katy Craig, and Yao Yao. *Aggregation-Diffusion Equations: Dynamics, Asymptotics, and Singular Limits*, pages 65–108. Springer International Publishing, Cham, 2019.

[CD22] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. i. models and methods, 2022.

[CDG20] Fabio Coppini, Helge Dietert, and Giambattista Giacomin. A law of large numbers and large deviations for interacting diffusions on erdős–rényi graphs. *Stochastics and Dynamics*, 20(02):2050010, 2020.

[CDPF15] Francesca Collet, Paolo Dai Pra, and Marco Formentin. Collective periodicity in mean-field models of cooperative behavior. *Nonlinear Differential Equations and Applications NoDEA*, 22(5):1461–1482, 2015.

[CE88] F. Comets and Th. Eisele. Asymptotic dynamics, non-critical and critical fluctuations for a geometric long-range interacting model. *Communications in Mathematical Physics*, 118(4):531–567, 1988.

[CEH13] F. C. Cooper, J. G. Esler, and P. H. Haynes. Estimation of the local response to a forcing in a high dimensional system using the fluctuation-dissipation theorem. *Nonlinear Processes in Geophysics*, 20(2):239–248, 2013.

[Ces19] Bruno Cessac. Linear response in neuronal networks: from neurons dynamics to collective response. *Chaos*, 29(10):103105, 24, 2019.

[CF05] Garnet Kin-Lic Chan and Reimar Finken. Time-dependent density functional theory of classical fluids. *Phys. Rev. Lett.*, 94:183001, May 2005.

[CGPS20] J. A. Carrillo, R. S. Gvalani, G. A. Pavliotis, and A. Schlichting. Long-time behaviour and phase transitions for the mckean–vlasov equation on the torus. *Archive for Rational Mechanics and Analysis*, 235(1):635–690, 2020.

[CH11] Fenwick C. Cooper and Peter H. Haynes. Climate sensitivity via a nonparametric fluctuation–dissipation theorem. *Journal of the Atmospheric Sciences*, 68(5):937 – 953, 2011.

[Cha08] Pierre-Henri Chavanis. Hamiltonian and brownian systems with long-range interactions: V. stochastic kinetic equations and theory of fluctuations. *Physica A: Statistical Mechanics and its Applications*, 387(23):5716–5740, 2008.

[Cha14] Pierre-Henri Chavanis. The brownian mean field model. *The European Physical Journal B*, 87(5):120, 2014.

[CJL78] M. Cassandro and G. Jona-Lasinio. Critical point behaviour and probability theory. *Advances in Physics*, 27(6):913–941, 1978.

[CK17] Mickaël D. Chekroun and Dmitri Kondrashov. Data-adaptive harmonic spectra and multilayer stuart-landau models. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(9):093110, 2017.

[CLM21] Mickaël D. Chekroun, Honghu Liu, and James C. McWilliams. Stochastic rectification of fast oscillations on slow manifold closures. *Proc. Natl. Acad. Sci. USA*, 118(48):Paper No. e2113650118, 9, 2021.

[CLW15] M. D. Chekroun, H. Liu, and S. Wang. *Stochastic Parameterizing Manifolds and Non-Markovian Reduced Equations*. SpringerBriefs in Mathematics. Springer International Publishing, Cham, 2015.

[CNK+14] M. D. Chekroun, J. D. Neelin, D. Kondrashov, J. C. McWilliams, and M. Ghil. Rough parameter dependence in climate models and the role of Ruelle-Pollicott resonances. *Proceedings of the National Academy of Sciences*, 111(5):1684–1690, 2014.

[Cop22] Fabio Coppini. Long time dynamics for interacting oscillators on graphs. *The Annals of Applied Probability*, 32(1):360 – 391, 2022.

[CP10] L. Chayes and V. Panferov. The McKean-Vlasov equation in finite volume. *J. Stat. Phys.*, 138(1-3):351–380, 2010.

[CP17] Pau Clusella and Antonio Politi. Noise-induced stabilization of collective dynamics. *Phys. Rev. E*, 95:062221, Jun 2017.

[CRV12] Matteo Colangeli, Lamberto Rondoni, and Angelo Vulpiani. Fluctuation-dissipation relation for chaotic non-hamiltonian systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(04):L04002, apr 2012.

[CS07] Bruno Cessac and Jacques-Alexandre Sepulchre. Linear response, susceptibility and resonances in chaotic toy models. *Physica D: Nonlinear Phenomena*, 225(1):13 – 28, 2007.

[CSG11] M. D. Chekroun, E. Simonnet, and M. Ghil. Stochastic climate dynamics: Random attractors and time-dependent invariant measures. *Physica D: Nonlinear Phenomena*, 240(21):1685–1700, 2011.

[CSZ20] Federico Cornalba, Tony Shardlow, and Johannes Zimmer. From weakly interacting particles to a regularised Dean-Kawasaki model. *Nonlinearity*, 33(2):864–891, 2020.

[CTDN20] Mickaël D. Chekroun, Alexis Tantet, Henk A. Dijkstra, and J. David Neelin. Ruelle–pollicott resonances of stochastic systems in reduced state space. part i: Theory. *Journal of Statistical Physics*, 2020.

[CVS04] I. Cionni, G. Visconti, and F. Sassi. Fluctuation dissipation theorem in a general circulation model. *Geophysical Research Letters*, 31(9), 2004.

[CW51] Herbert B. Callen and Theodore A. Welton. Irreversibility and generalized noise. *Phys. Rev.*, 83:34–40, Jul 1951.

[CW20] Nisha Chandramoorthy and Qiqi Wang. A computable realization of Ruelle's formula for linear response of statistics in chaotic systems. *arXiv e-prints*, page arXiv:2002.04117, February 2020.

[Daw83] Donald A. Dawson. Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *Journal of Statistical Physics*, 31(1):29–85, 1983.

[DD10] A. Dembo and J.-D. Deuschel. Markovian perturbation, response and fluctuation dissipation theorem. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(3):822–852, 2010.

[Dea96] David S. Dean. Langevin equation for the density of a system of interacting Langevin processes. *J. Phys. A*, 29(24):L613–L617, 1996.

[Der07] Bernard Derrida. Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07023–P07023, jul 2007.

[DG87a] D. A. Dawson and J. Gärtner. Long-time fluctuations of weakly interacting diffusions. In Hans Jürgen Engelbert and Wolfgang Schmidt, editors, *Stochastic Differential Systems*, pages 1–10, Berlin, Heidelberg, 1987. Springer Berlin Heidelberg.

[DG87b] D. A. Dawson and Jürgen Gärtner. Large deviations from the mckean-vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[DG89] D. A. Dawson and J. Gärtner. Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions. *Mem. Amer. Math. Soc.*, 78(398):iv+94, 1989.

[DG01] V. P. Dymnikov and A. S. Gritsoun. Climate model attractors: chaos, quasi-regularity and sensitivity to small perturbations of external forcing. *Nonlinear Processes in Geophysics*, 8(4/5):201–209, 2001.

[DGP21] Matias G. Delgadino, Rishabh S. Gvalani, and Grigorios A. Pavliotis. On the diffusive-mean field limit for weakly interacting diffusions exhibiting phase transitions. *Archive for Rational Mechanics and Analysis*, 2021.

[DGPS23] Matías G. Delgadino, Rishabh S. Gvalani, Grigorios A. Pavliotis, and Scott A. Smith. Phase transitions, logarithmic sobolev inequalities, and uniform-in-time propagation of chaos for weakly interacting diffusions. *Communications in Mathematical Physics*, 2023.

[DLP16] A. B. Duncan, T. Lelièvre, and G. A. Pavliotis. Variance reduction using nonreversible langevin samplers. *Journal of Statistical Physics*, 163(3):457–491, 2016.

[DP19] Paolo Dai Pra. Stochastic mean-field dynamics and applications to life sciences. In Giambattista Giacomin, Stefano Olla, Ellen Saada, Herbert Spohn, and Gabriel Stoltz, editors, *Stochastic Dynamics Out of Equilibrium*, pages 3–27, Cham, 2019. Springer International Publishing.

[DPZ96] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[DPZ17] A. B. Duncan, G. A. Pavliotis, and K. C. Zygalakis. Nonreversible Langevin Samplers: Splitting Schemes, Analysis and Implementation. *arXiv e-prints*, page arXiv:1701.04247, January 2017.

[DSvN+08] Vasilis Dakos, Marten Scheffer, Egbert H. van Nes, Victor Brovkin, Vladimir Petoukhov, and Hermann Held. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, 105(38):14308–14312, 2008.

[DVE14] Aleksandar Donev and Eric Vanden-Eijnden. Dynamic density functional theory with hydrodynamic interactions and fluctuations. *The Journal of Chemical Physics*, 140(23):234115, 2014.

[DZ78] Rashmi C. Desai and Robert Zwanzig. Statistical mechanics of a nonlinear stochastic model. *Journal of Statistical Physics*, 19(1):1–24, 1978.

[Ein05] A. Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[EK16] Joep H. M. Evers and Theodore Kolokolnikov. Metastable states for an aggregation model with noise. *SIAM Journal on Applied Dynamical Systems*, 15(4):2213–2226, 2016.

[ELP17] Deniz Eroglu, Jeroen S. W. Lamb, and Tiago Pereira. Synchronisation of chaos and its applications. *Contemporary Physics*, 58(3):207–243, 2017.

[EN00] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[EN06] Klaus-Jochen Engel and Rainer Nagel. *A Short Course on Operator Semigroups*. Springer, 2006.

[ER85] J.-P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Modern Phys.*, 57(3, part 1):617–656, 1985.

[Feu08] Ulrike Feudel. Complex dynamics in multistable systems. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 18(6):1607–1626, 2008.

[FLQ10] Gary Froyland, Simon Lloyd, and Anthony Quas. Coherent structures and isolated spectrum for Perron-Frobenius cocycles. *Ergodic Theory Dynam. Systems*, 30(3):729–756, 2010.

[FPET07] Gary Froyland, Kathrin Padberg, Matthew H. England, and Anne Marie Treguier. Detection of coherent oceanic structures via transfer operators. *Phys. Rev. Lett.*, 98:224503, May 2007.

[FPG14] Gary Froyland and Kathrin Padberg-Gehle. Almost-invariant and finite-time coherent sets: directionality, duration, and diffusion. In *Ergodic theory, open dynamics, and coherent structures*, volume 70 of *Springer Proc. Math. Stat.*, pages 171–216. Springer, New York, 2014.

[FPS18] Ulrike Feudel, Alexander N. Pisarchik, and Kenneth Showalter. Multistability and tipping: From mathematics and physics to climate and brain—minireview and preface to the focus issue. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(3):033501, 2018.

[Fra01] T.D. Frank. A langevin approach for the microscopic dynamics of nonlinear fokker–planck equations. *Physica A: Statistical Mechanics and its Applications*, 301(1):52–62, 2001.

[Fra04] T.D. Frank. Fluctuation–dissipation theorems for nonlinear fokker–planck equations of the desai–zwanzig type and vlasov–fokker–planck equations. *Physics Letters A*, 329(6):475 – 485, 2004.

[Fra05] T. D. Frank. *Nonlinear Fokker-Planck Equations Fundamentals and Applications.* Springer, Berlin, Heidelberg, 2005.

[Fra13] T. Frank. Strongly nonlinear stochastic processes in physics and the life sciences. *ISRN Mathematical Physics*, 2013:1–28, 03 2013.

[Fro97] Gary Froyland. Computer-assisted bounds for the rate of decay of correlations. *Comm. Math. Phys.*, 189(1):237–257, 1997.

[Fro13] Gary Froyland. An analytic framework for identifying finite-time coherent sets in time-dependent dynamical systems. *Phys. D*, 250:1–19, 2013.

[FSH15] David Fuchs, Steven Sherwood, and Daniel Hernandez. An exploration of multivariate fluctuation dissipation operators and their response to sea surface temperature perturbations. *Journal of the Atmospheric Sciences*, 72(1):472 – 486, 2015.

[FW84] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1984. Translated from the Russian by Joseph Szücs.

[GAD+02] M. Ghil, M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. Advanced spectral methods for climatic time series. *Reviews of Geophysics*, 40(1):3–1–3–41, 2002.

[Gal14] G. Gallavotti. *Nonequilibrium and irreversibility*. Springer, New York, 2014.

[Gal20] Giovanni Gallavotti. Nonequilibrium and fluctuation relation. *J. Stat. Phys.*, 180(1-6):172–226, 2020.

[GB07] Andrey Gritsun and Grant Branstator. Climate response using a three-dimensional operator based on the fluctuation–dissipation theorem. *Journal of the Atmospheric Sciences*, 64(7):2558 – 2575, 2007.

[GBD02] A. S. Gritsoun, G. Branstator, and V. P. Dymnikov. Construction of the linear response operator of an atmospheric general circulation model to small external forcing. *Russian J. Numer. Anal. Math. Modelling*, 17(5):399–416, 2002.

[GBM08] A. Gritsun, G. Branstator, and A. J. Majda. Climate response of linear and quadratic functionals using the fluctuation-dissipation theorem. *J. Atmos. Sci.*, 65, 2008.

[GC95a] G. Gallavotti and E. G. D. Cohen. Dynamical ensembles in nonequilibrium statistical mechanics. *Phys. Rev. Lett.*, 74:2694–2697, Apr 1995.

[GC95b] G. Gallavotti and E. G. D. Cohen. Dynamical ensembles in stationary states. *Journal of Statistical Physics*, 80(5-6):931–970, September 1995.

[Gei18] Sylvie Geisendorf. Evolutionary climate-change modelling: A multi-agent climate-economic model. *Computational Economics*, 52(3):921–951, 2018.

[Geo11] Hans-Otto Georgii. *Gibbs Measures and Phase Transitions.* De Gruyter, 2011.

[GHJM98] Luca Gammaitoni, Peter Hänggi, Peter Jung, and Fabio Marchesoni. Stochastic resonance. *Rev. Mod. Phys.*, 70:223–287, Jan 1998.

[GHT91] R. Graham, A. Hamm, and T. Tél. Nonequilibrium potentials for dynamical systems with fractal attractors or repellers. *Phys. Rev. Lett.*, 66:3089–3092, Jun 1991.

[Gil17] Tepper L. Gill. The feynman-dyson view. *Journal of Physics: Conference Series*, 845:012023, may 2017.

[GINR20] A. Garbuno-Inigo, N. Nüsken, and S. Reich. Affine invariant interacting Langevin dynamics for Bayesian inference. *SIAM J. Appl. Dyn. Syst.*, 19(3):1633–1658, 2020.

[GKPY19] Susana N. Gomes, Serafim Kalliadasis, Grigorios A. Pavliotis, and Petr Yatsyshin. Dynamics of the desai-zwanzig model in multiwell and random energy landscapes. *Phys. Rev. E*, 99:032109, Mar 2019.

[GKWG09] Gruberbauer, M., Kallinger, T., Weiss, W. W., and Guenther, D. B. On the detection of lorentzian profiles in a power spectrum: a bayesian approach using ignorance priors. *A&A*, 506(2):1043–1053, 2009.

[GL17] Andrey Gritsun and Valerio Lucarini. Fluctuations, response, and resonances in a simple atmospheric model. *Physica D: Nonlinear Phenomena*, 349:62–76, 2017.

[GL20] Michael Ghil and Valerio Lucarini. The physics of climate variability and climate change. *Rev. Mod. Phys.*, 92:035002, Jul 2020.

[GL22] Manuel Santos Gutiérrez and Valerio Lucarini. On some aspects of the response to stochastic and deterministic forcings. *Journal of Physics A: Mathematical and Theoretical*, 55(42):425002, oct 2022.

[GLP13] P. Giulietti, C. Liverani, and M. Pollicott. Anosov flows and dynamical zeta functions. *Ann. of Math. (2)*, 178(2):687–773, 2013.

[GNPT95] P. Gaspard, G. Nicolis, A. Provata, and S. Tasaki. Spectral signature of the pitchfork bifurcation: Liouville equation approach. *Phys. Rev. E*, 51:74–94, Jan 1995.

[GNS$^+$12] Benjamin D. Goddard, Andreas Nold, Nikos Savva, Grigorios A. Pavliotis, and Serafim Kalliadasis. General dynamical density functional theory for classical fluids. *Phys. Rev. Lett.*, 109:120603, Sep 2012.

[Got20] Georg A. Gottwald. Introduction to focus issue: Linear response theory: Potentials and limits. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(2):020401, 2020.

[GP18] S.N. Gomes and G.A. Pavliotis. Mean field limits for interacting diffusions in a two-scale potential. *J. Nonlin. Sci.*, 28(3):905–941, 2018.

[GPK12] B. D. Goddard, G. A. Pavliotis, and S. Kalliadasis. The overdamped limit of dynamic density functional theory: Rigorous results. *Multiscale Modeling & Simulation*, 10(2):633–663, 2012.

[GPY13] J. Garnier, G. Papanicolaou, and T. Yang. Large deviations for a mean field model of systemic risk. *SIAM Journal on Financial Mathematics*, 4(1):151–184, 2013.

[GPY17] J. Garnier, G. Papanicolaou, and T. Yang. Consensus convergence with stochastic effects. *Vietnam Journal of Mathematics*, 45(1):51–75, 2017.

[Gra77] R. Graham. Covariant formulation of non-equilibrium statistical thermodynamics. *Z. Phys. B*, 26(4):397–405, 1977.

[Gre71] P. J. H. Green. Characteristic functions by e. lukacs. [second edition. pp. viii 350. london: Griffin, 1970, £5·50]. *Journal of the Institute of Actuaries*, 97(1):134–135, 1971.

[GRVE22] Marylou Gabrié, Grant M. Rotskoff, and Eric Vanden-Eijnden. Adaptive monte carlo augmented with normalizing flows. *Proceedings of the National Academy of Sciences*, 119(10):e2109420119, 2022.

[GZ03] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003.

[H6̈7] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[Hel02] B. Helffer. *Semiclassical analysis, Witten Laplacians, and statistical mechanics*, volume 1 of *Series in Partial Differential Equations and Applications.* World Scientific Publishing Co., Inc., River Edge, NJ, 2002.

[HF20] Lukas Halekotte and Ulrike Feudel. Minimal fatal shocks in multistable complex networks. *Scientific Reports*, 10(1):11783, 2020.

[HK64] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[HL84] W. Horsthemke and R. Lefever. *Noise-induced transitions*, volume 15 of *Springer Series in Synergetics.* Springer-Verlag, Berlin, 1984. Theory and applications in physics, chemistry, and biology.

[HM10] M. Hairer and A. J. Majda. A simple framework to justify linear response theory. *Nonlinearity*, 23(4):909–922, 2010.

[HSG03] AgnèS. Helmstetter, Didier Sornette, and Jean-Robert Grasso. Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws. *Journal of Geophysical Research (Solid Earth)*, 108(B1):2046, January 2003.

[HTG94] A. Hamm, T. Tél, and R. Graham. Noise-induced attractor explosions near tangent bifurcations. *Physics Letters A*, 185(3):313–320, 1994.

[HVF21] Lukas Halekotte, Anna Vanselow, and Ulrike Feudel. Transient chaos enforces uncertainty in the british power grid. *Journal of Physics: Complexity*, 2(3):035015, jul 2021.

[Hä78] P. Hänggi. Stochastic processes 2: Response theory and fluctuation theorems. *Helvetica Physica Acta*, 51, 1978.

[Jac75] John David Jackson. *Classical electrodynamics; 2nd ed.* Wiley, New York, NY, 1975.

[JM98] Benjamin Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Annales de l'I.H.P. Probabilités et statistiques*, 34(6):727–766, 1998.

[JP79] B. Jouvet and R. Phythian. Quantum aspects of classical and statistical fields. *Phys. Rev. A*, 19:1350–1355, Mar 1979.

[JPS22] Pierre-Emmanuel Jabin, David Poyato, and Juan Soler. Mean-field limit of non-exchangeable systems, 2022.

[KAB+14] Mikko Kivelä, Alex Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 07 2014.

[Kan20] Yukio Kaneda. Linear response theory of turbulence. *J. Stat. Mech. Theory Exp.*, 2020(3):034006, 18, 2020.

[KCB18] Dmitri Kondrashov, Mickaël D. Chekroun, and Pavel Berloff. Multiscale stuart-landau emulators: Application to wind-driven ocean gyres. *Fluids*, 3(1), 2018.

[KCG15] Dmitri Kondrashov, Mickaël D. Chekroun, and Michael Ghil. Data-driven non-Markovian closure models. *Physica D: Nonlinear Phenomena*, 297:33–55, 2015.

[KD09] D. B. Kirk-Davidoff. On the diagnosis of climate sensitivity using observations of fluctuations. *Atmospheric Chemistry and Physics*, 9(3):813–822, 2009.

[KFT16] Bálint Kaszás, Ulrike Feudel, and Tamás Tél. Death and revival of chaos. *Phys. Rev. E*, 94:062221, Dec 2016.

[Khi34] A. Khintchine. Korrelationstheorie der stationären stochastischen prozesse. *Mathematische Annalen*, 109(1):604—615, 1934.

[KjHGG11] David Kinny, Jane Yung jen Hsu, Guido Governatori, and Aditya K. Ghose, editors. *Agents in Principle, Agents in Practice 14th International Conference, PRIMA 2011, Wollongong, Australia, November 16-18, 2011, Proceedings*. Lecture Notes in Artificial Intelligence ; 7047. Springer-Verlag, Berlin, Heidelberg, 1st ed. 2011. edition, 2011.

[Kli90] Yu.L. Klimontovich. Ito, stratonovich and kinetic forms of stochastic equations. *Physica A: Statistical Mechanics and its Applications*, 163(2):515–532, 1990.

[KLvR20] Vitalii Konarovskyi, Tobias Lehmann, and Max von Renesse. On Dean-Kawasaki dynamics with smooth drift potential. *J. Stat. Phys.*, 178(3):666–681, 2020.

[KM41] John G. Kirkwood and Elizabeth Monroe. Statistical mechanics of fusion. *The Journal of Chemical Physics*, 9(7):514–526, 1941.

[KNK+18] Stefan Klus, Feliks Nüske, Péter Koltai, Hao Wu, Ioannis Kevrekidis, Christof Schütte, and Frank Noé. Data-driven model reduction and transfer operator approximation. *J. Nonlinear Sci.*, 28(3):985–1010, 2018.

[Kol10] Vassili N. Kolokoltsov. *Nonlinear Markov Processes and Kinetic Equations*. Cambridge Tracts in Mathematics. Cambridge University Press, 2010.

[KP11] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2011.

[KPP19] Nikolas Kantas, Panos Parpas, and Grigorios A. Pavliotis. The sharp, the flat and the shallow: Can weakly interacting agents learn to escape bad minima?, 2019.

[Kub57] Ryogo Kubo. Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. *J. Phys. Soc. Japan*, 12:570–586, 1957.

[Kub66] R. Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1):255–284, 1966.

[Kue11] Christian Kuehn. A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena*, 240(12):1020 – 1035, 2011.

[KW14] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In Yoshua Bengio and Yann LeCun, editors, *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.

[LA05] Peter L. Langen and Vladimir A. Alexeev. Estimating $2 \times$ co2 warming in an aquaplanet gcm using the fluctuation-dissipation theorem. *Geophysical Research Letters*, 32(23), 2005.

[LB07] Luca Lorenzi and Marcello Bertoldi. *Analytical methods for Markov semigroups*, volume 283 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2007.

[LB17] Valerio Lucarini and Tamás Bódai. Edge states in the climate system: exploring global instabilities and critical transitions. *Nonlinearity*, 30(7):R32–R66, 2017.

[LB20] Valerio Lucarini and Tamás Bódai. Global stability properties of the climate: Melancholia states, invariant measures, and phase transitions. *Nonlinearity*, 33(9):R59–R92, jul 2020.

[LC12] Valerio Lucarini and Matteo Colangeli. Beyond the linear fluctuation-dissipation theorem: the role of causality. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(05):P05013, may 2012.

[LCSZ08a] E. Lippiello, F. Corberi, A. Sarracino, and M. Zannetti. Nonlinear susceptibilities and the measurement of a cooperative length. *Phys. Rev. B*, 77:212201, Jun 2008.

[LCSZ08b] Eugenio Lippiello, Federico Corberi, Alessandro Sarracino, and Marco Zannetti. Nonlinear response and fluctuation-dissipation relations. *Phys. Rev. E*, 78:041120, Oct 2008.

[Lei75] C. E. Leith. Climate response and fluctuation dissipation. *J. Atmos. Sci.*, 32:2022, 1975.

[LG06] C. Liverani and S. Gouëzel. Banach spaces adapted to Anosov systems. *Ergodic Theory and Dynamical Systems*, 26:189–217, 2006.

[Lim21] Soon Hoe Lim. Understanding recurrent neural networks using nonequilibrium response theory. *J. Mach. Learn. Res.*, 22:Paper No. 47, 48, 2021.

[LLR20] Valerio Lembo, Valerio Lucarini, and Francesco Ragone. Beyond forcing scenarios: Predicting climate change through response operators in a coupled general circulation model. *Scientific Reports*, 10(1):8668, 2020.

[LM94] Andrzej Lasota and Michael C. Mackey. *Chaos, fractals and noise*. Springer, New York, 1994.

[LNP13] T. Lelievre, F. Nier, and G. A. Pavliotis. Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2): 237–274 , 2013.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[LPZ20] V. Lucarini, G. A. Pavliotis, and N. Zagli. Response theory and phase transitions for the thermodynamic limit of interacting identical systems. *Proc. R. Soc. A.*, 476, 2020.

[LRL17] V. Lucarini, F. Ragone, and F. Lunkeit. Predicting climate change using response theory: Global averages and spatial patterns. *Journal of Statistical Physics*, 166(3):1036–1064, February 2017.

[LS11] V. Lucarini and S. Sarno. A statistical mechanical approach for the computation of the climatic response to general forcings. *Nonlin. Processes Geophys*, 18:7–28, 2011.

[LSPV05] V. Lucarini, J. J. Saarinen, K.-E. Peiponen, and E. M. Vartiainen. *Kramers-Kronig relations in Optical Materials Research*. Springer, New York, 2005.

[Luc08] V. Lucarini. Response theory for equilibrium and non-equilibrium statistical mechanics: Causality and generalized kramers-kronig relations. *Journal of Statistical Physics*, 131:543–558, 2008. 10.1007/s10955-008-9498-y.

[Luc09] Valerio Lucarini. Evidence of dispersion relations for the nonlinear response of the Lorenz 63 system. *J. Stat. Phys.*, 134(2):381–400, 2009.

[Luc12] Valerio Lucarini. Stochastic perturbations to dynamical systems: a response theory approach. *J. Stat. Phys.*, 146(4):774–786, 2012.

[Luc16] Valerio Lucarini. Response operators for Markov processes in a finite state space: radius of convergence and link to the response theory for Axiom A systems. *J. Stat. Phys.*, 162(2):312–333, 2016.

[Luc18] V. Lucarini. Revising and extending the linear response theory for statistical mechanical systems: Evaluating observables as predictors and predictands. *Journal of Statistical Physics*, 173(6):1698–1721, December 2018.

[LV07] G. Lacorata and A. Vulpiani. Fluctuation-response relation and modeling in systems with fast and slow dynamics. *Nonlinear Processes in Geophysics*, 14(5):681–694, 2007.

[MA01] N. Martzel and C. Aslangul. Mean-field treatment of the many-body Fokker-Planck equation. *J. Phys. A*, 34(50):11225–11240, 2001.

[MAG10] Andrew J. Majda, Rafail Abramov, and Boris Gershgorin. High skill in low-frequency climate response through fluctuation dissipation theorems despite structural instability. *Proc. Natl. Acad. Sci. USA*, 107(2):581–586, 2010.

[Mal01] F. Malrieu. Logarithmic sobolev inequalities for some nonlinear pde's. *Stochastic Processes and their Applications*, 95(1):109 – 132, 2001.

[McK66] H. P. McKean. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences*, 56(6):1907–1911, 1966.

[MD05] Ao Ma and Aaron R. Dinner. Automatic method for identifying reaction coordinates in complex systems. *The Journal of Physical Chemistry B*, 109(14):6769–6779, 04 2005.

[MDN21] Ali Motazedifard, A. Dalafi, and M. H. Naderi. A Green's function approach to the linear response of a driven dissipative optomechanical system. *J. Phys. A*, 54(21):Paper No. 215301, 22, 2021.

[Mer65] N. David Mermin. Thermal properties of the inhomogeneous electron gas. *Phys. Rev.*, 137:A1441–A1443, Mar 1965.

[Mez05] Igor Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynam.*, 41(1-3):309–325, 2005.

[MG08] Ian Melbourne and Georg A. Gottwald. Power spectra for deterministic chaotic dynamical systems. *Nonlinearity*, 21(1):179–189, 2008.

[MGLL21] Georgios Margazoglou, Tobias Grafke, Alessandro Laio, and Valerio Lucarini. Dynamical landscape and multistability of a climate model. *Proc. A.*, 477(2250):Paper No. 20210019 , 28, 2021.

[MJ67] H.P McKean Jr. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations Lecture Series in Differential Equations, Session 7, Catholic Univ.*, 1967.

[MLG22] Chiara Cecilia Maiocchi, Valerio Lucarini, and Andrey Gritsun. Decomposing the dynamics of the lorenz 1963 model using unstable periodic orbits: Averages, transitions, and quasi-invariant sets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(3):033129, 2022.

[Mor65] H. Mori. Transport, collective motion, and Brownian motion. *Progress of Theoretical Physics*, 33(3):423–455, March 1965.

[MPRV08] U. Marini Bettolo Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani. Fluctuation-dissipation: Response theory in statistical physics. *Phys. Rep.*, 461:111, 2008.

[MS81] B. J. Matkowsky and Z. Schuss. Eigenvalues of the Fokker-Planck operator and the approach to equilibrium for diffusions in potential fields. *SIAM J. Appl. Math.*, 40(2):242–254, 1981.

[MSR73] P. C. Martin, E. D. Siggia, and H. A. Rose. Statistical dynamics of classical systems. *Phys. Rev. A*, 8:423–437, Jul 1973.

[Nam76] Mitsuhiro Nambu. Linear response theory of a turbulent plasma. *The Physics of Fluids*, 19(3):412–419, 1976.

[Nam77] Mitsuhiro Nambu. Linear response theory of a turbulent magnetized plasma. *The Physics of Fluids*, 20(3):459–465, 1977.

[NBH93] G.R. North, R.E. Bell, and J.W. Hardin. Fluctuation dissipation in a general circulation model. *Clim. Dyn.*, 8:259, 1993.

[Ni21] Angxiu Ni. Approximating linear response by nonintrusive shadowing algorithms. *SIAM Journal on Numerical Analysis*, 59(6):2843–2865, 2021.

[NPT10] G. Naldi, L. Pareschi, and G. Toscani. *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*. Birkhäuser Basel, 2010.

[Nyq28] H. Nyquist. Thermal agitation of electric charge in conductors. *Phys. Rev.*, 32:110–113, Jul 1928.

[NZ15] Stéphane Nonnenmacher and Maciej Zworski. Decay of correlations for normally hyperbolic trapping. *Invent. Math.*, 200(2):345–438, 2015.

[Oel84] Karl Oelschlager. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 05 1984.

[Ons31] Lars Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37:405–426, Feb 1931.

[OSBA02] Edward Ott, Paul So, Ernest Barreto, and Thomas Antonsen. The onset of synchronization in systems of globally coupled chaotic and periodic oscillators. *Physica D: Nonlinear Phenomena*, 173(1):29–51, 2002.

[Ött05] H.C. Öttinger. *Beyond Equilibrium Thermodynamics*. Wiley, Hoboken, 2005.

[OY12] Shun Ogawa and Yoshiyuki Y. Yamaguchi. Linear response theory in the Vlasov equation for homogeneous and for inhomogeneous quasistationary states. *Phys. Rev. E*, 85:061115, Jun 2012.

[Pav14] Grigorios A. Pavliotis. *Stochastic Processes and Applications*, volume 60. Springer, New York, 2014.

[PC15] Louis M. Pecora and Thomas L. Carroll. Synchronization of chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(9):097611, 2015.

[Pec98] L. M. Pecora. Synchronization conditions and desynchronizing patterns in coupled limit-cycle and chaotic systems. *Phys. Rev. E*, 58:347, 1998.

[PG16] M. Porter and J. Gleeson. *Dynamical Systems on Networks: A Tutorial.* Frontiers in Applied Dynamical Systems: Reviews and Tutorials. Springer, Cham, 2016.

[PH96] Paolo Dai Pra and Frank den Hollander. Mckean-vlasov limit for interacting random processes in random media. *Journal of Statistical Physics*, 84(3):735–772, 1996.

[PKRK03] A. Pikovsky, J. Kurths, M. Rosenblum, and J. Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences.* Cambridge Nonlinear Science Series. Cambridge University Press, 2003.

[PMH+13] Giuseppe Pesce, Austin McDaniel, Scott Hottovy, Jan Wehr, and Giovanni Volpe. Stratonovich-to-itôtransition in noisy systems with multiplicative feedback. *Nature Communications*, 4(1):2733, 2013.

[PR14] Aurelio Patelli and Stefano Ruffo. General linear response formula for non integrable systems obeying the vlasov equation. *The European Physical Journal D*, 68(11):329, 2014.

[PT13] L. Pareschi and G. Toscani. Interacting multiagent systems: kinetic equations and monte carlo methods, 2013.

[PTSSG+21] Antonio M. Puertas, Juan E. Trinidad-Segovia, Miguel A. Sánchez-Granero, Joaquim Clara-Rahora, and F. Javier de las Nieves. Linear response theory in stock markets. *Scientific Reports*, 11(1):23076, 2021.

[PVA99] Kai-Erik Peiponen, Erik M. Vartiainen, and Toshimitsu Asakura. *Dispersion, Complex Analysis and Optical Spectroscopy.* Springer, 1999.

[PZ03] A. Pikovsky and A. Zaikin. System size stochastic and coherence resonance. *AIP Conference Proceedings*, 665(1):561–568, 2003.

[PZdlC02] A. Pikovsky, A. Zaikin, and M. A. de la Casa. System size resonance in coupled noisy systems and in the ising model. *Phys. Rev. Lett.*, 88:050601, Jan 2002.

[Rei02] Christian H. Reick. Linear response of the Lorenz system. *Phys. Rev. E (3)*, 66(3):036103, 11, 2002.

[RH02] U. Krause R. Hegselmann. Opinion dynamics and bounded confidence: models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

[Ris89] H. Risken. *The Fokker-Planck equation*, volume 18 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, 1989.

[RLL16] Francesco Ragone, Valerio Lucarini, and Frank Lunkeit. A new framework for climate sensitivity and prediction: a modelling perspective. *Climate Dynamics*, 46(5):1459–1471, 2016.

[Rog21] Jutta Rogal. Reaction coordinates in complex systems-a perspective. *The European Physical Journal B*, 94(11):223, 2021.

[RP08] Michael J. Ring and R. Alan Plumb. The response of a simplified gcm to axisymmetric forcings: Applicability of the fluctuation–dissipation theorem. *Journal of the Atmospheric Sciences*, 65(12):3880 – 3898, 2008.

[Rue86] D. Ruelle. Resonances of chaotic dynamical systems. *Physical Review Letters*, 56:405–407, February 1986.

[Rue89] David Ruelle. *Chaotic evolution and strange attractors*. Lezioni Lincee. [Lincei Lectures]. Cambridge University Press, Cambridge, 1989. The statistical analysis of time series for deterministic nonlinear systems, Notes prepared and with a foreword by Stefano Isola.

[Rue97] David Ruelle. Differentiation of SRB states. *Comm. Math. Phys.*, 187(1):227–241, 1997.

[Rue98a] D. Ruelle. Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms. *Nonlinearity*, 11(1):5–18, January 1998.

[Rue98b] David Ruelle. General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium. *Phys. Lett. A*, 245(3-4):220–224, 1998.

[Rue09] D. Ruelle. A review of linear response theory for general differentiable dynamical systems. *Nonlinearity*, 22(4):855–870, April 2009.

[RVE18] G. M. Rotskoff and E. Vanden-Eijnden. Neural networks as interacting particle systems: Asymptotic convexity of the loss landscape and universal scaling of the approximation error, 2018.

[Sak00] Hidetsugu Sakaguchi. Phase transition in globally coupled rössler oscillators. *Phys. Rev. E*, 61:7212–7214, Jun 2000.

[SBB$^+$09] Marten Scheffer, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egbert H. van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.

[SC22] Eric Simonnet and Mickaël D. Chekroun. Deep spectral computations in linear and nonlinear diffusion problems, 2022.

[Sch09a] M. Scheffer. *Critical Transitions in Nature and Society*. Princeton Studies in Complexity. Princeton University Press, Princeton, 2009.

[Sch09b] Martin Scheffer. *Critical Transitions in Nature and Society*. Princeton University Press, 2009.

[SCH10] PETER J. SCHMID. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, 2010.

[SGL19] Jose Simmonds, Juan A. Gómez, and Agapito Ledezma. The role of agent-based modeling and multi-agent systems in flood-based hydrological problems: a brief review. *Journal of Water and Climate Change*, 10 2019. jwc2019108.

[SGL20] Manuel Santos Gutiérrez and Valerio Lucarini. Response and sensitivity using markov chains. *Journal of Statistical Physics*, 2020.

[SGLCG21] Manuel Santos Gutiérrez, Valerio Lucarini, Mickaël D. Chekroun, and Michael Ghil. Reduced-order models for coupled dynamical systems: Data-driven methods and the koopman operator. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(5):053116, 2021.

[SH03] D Sornette and A Helmstetter. Endogenous versus exogenous shocks in systems with memory. *Physica A: Statistical Mechanics and its Applications*, 318(3):577 – 591, 2003.

[Shi85] Masatoshi Shiino. H-theorem and stability analysis for mean-field models of non-equilibrium phase transitions in stochastic systems. *Physics Letters A*, 112(6):302 – 306, 1985.

[Shi87] M. Shiino. Dynamical behavior of stochastic systems of infinitely many coupled nonlinear oscillators exhibiting phase transitions of mean-field type: H theorem on asymptotic approach to equilibrium and critical slowing down of order-parameter fluctuations. *Phys. Rev. A*, 36:2393–2412, Sep 1987.

[Sor06] D. Sornette. Endogenous versus exogenous origins of crises. In Kantz H. Albeverio S., Jentsch V., editor, *Extreme Events in Nature and Society*, pages 95–119. , Berlin, Heidelberg, 20006.

[Sor03] D. Sornette. *Why Stock Markets Crash (Critical Events in Complex Financial Systems)*. Princeton Univeriyty Press, Princeton, 2003.

[SS90] Rüdiger Schack and Axel Schenzle. Moment hierarchies and cumulants in quantum optics. *Phys. Rev. A*, 41:3847–3852, Apr 1990.

[SS13] Christof Schütte and Marco Sarich. *Metastability and Markov state models in molecular dynamics*, volume 24 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2013. Modeling, analysis, algorithmic approaches.

[SV19] A. Sarracino and A. Vulpiani. On the fluctuation-dissipation relation in non-equilibrium and non-hamiltonian systems. *Chaos*, 29:083132, 2019.

[SW22] Adam A. Śliwiak and Qiqi Wang. A trajectory-driven algorithm for differentiating SRB measures on unstable manifolds. *SIAM J. Sci. Comput.*, 44(1):A312–A336, 2022.

[Szn89] A.S. Sznitman. *Topics in propagation of chaos.*, volume 1464 of *Hennequin PL. (eds) Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989. Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 1989.

[Tam84] Y. Tamura. On asymptotic behaviors of the solution of a nonlinear diffusion equation. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 31(1):195–221, 1984.

[TBDR05] Takayuki Tatekawa, Freddy Bouchet, Thierry Dauxois, and Stefano Ruffo. Thermodynamics of the self-gravitating ring model. *Phys. Rev. E*, 71:056111, May 2005.

[TCND19] Alexis Tantet, Mickaël D. Chekroun, J. David Neelin, and Henk A. Dijkstra. Ruelle–pollicott resonances of stochastic systems in reduced state space. part iii: Application to the cane–zebiak model of the el niño–southern oscillation. *Journal of Statistical Physics*, 2019.

[TKP01] Dmitri Topaj, Won-Ho Kye, and Arkady Pikovsky. Transition to coherence in populations of coupled chaotic oscillators: A linear response approach. *Phys. Rev. Lett.*, 87:074101, Jul 2001.

[TLD18] Alexis Tantet, Valerio Lucarini, and Henk A Dijkstra. Resonances in a Chaotic Attractor Crisis of the Lorenz Flow. *Journal of Statistical Physics*, 170(3):584–616, 2018.

[TLLD18] Alexis Tantet, Valerio Lucarini, Frank Lunkeit, and Henk A. Dijkstra. Crisis of the chaotic attractor of a climate model: a transfer operator approach. *Nonlinearity*, 31(5):2221–2251, 2018.

[Tou18] Hugo Touchette. Introduction to dynamical large deviations of Markov processes. *Phys. A*, 504:5–19, 2018.

[TT13] E. Talebian and M. Talebian. A general review on the derivation of clausius-mossotti relation. *Optik*, 124(16):2324 – 2326, 2013.

[Tug14] Julian Tugaut. Phase transitions of mckean–vlasov processes in double-wells landscape. *Stochastics*, 86(2):257–284, 2014.

[TvdBD15] Alexis Tantet, Fiona R. van der Burgt, and Henk A. Dijkstra. An early warning indicator for atmospheric blocking events using transfer operators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3):036406, 2015.

[Var91] S. R. S. Varadhan. Scaling limits for interacting diffusions. *Communications in Mathematical Physics*, 135(2):313 – 353, 1991.

[VdBPAHM94] C. Van den Broeck, J. M. R. Parrondo, J. Armero, and A. Hernández-Machado. Mean field model for spatially extended systems in the presence of multiplicative noise. *Phys. Rev. E*, 49:2639–2643, Apr 1994.

[vK81] N. G. van Kampen. Itô versus stratonovich. *Journal of Statistical Physics*, 24(1):175–187, 1981.

[VKP15] Vladimir Vlasov, Maxim Komarov, and Arkady Pikovsky. Synchronization transitions in ensembles of noisy oscillators with bi-harmonic coupling. *Journal of Physics A: Mathematical and Theoretical*, 48(10):105101, feb 2015.

[Wan13] Qiqi Wang. Forward and adjoint sensitivity computation of chaotic dynamical systems. *Journal of Computational Physics*, 235(0):1 – 13, 2013.

[WB70] Ralph M Wilcox and Richard Bellman. Truncation and preservation of moment properties for fokker-planck moment equations. *Journal of Mathematical Analysis and Applications*, 32(3):532–542, 1970.

[WG18] Caroline L. Wormell and Georg A. Gottwald. On the validity of linear response theory in high-dimensional deterministic dynamical systems. *Journal of Statistical Physics*, 172(6):1479–1498, 2018.

[WG19] Caroline L. Wormell and Georg A. Gottwald. Linear response for macroscopic observables in high-dimensional systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(11):113127, 2019.

[Wie30] Norbert Wiener. Generalized harmonic analysis. *Acta Mathematica*, 55(none):117 – 258, 1930.

[WL12] J. Wouters and V. Lucarini. Disentangling multi-level systems: averaging, correlations and memory. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(03):P03003, March 2012.

[WL13] J. Wouters and V. Lucarini. Multi-level dynamical systems: Connecting the Ruelle response theory and the Mori-Zwanzig approach. *Journal of Statistical Physics*, 151(5), March 2013.

[WLEC17] C. Wang, Q. Li, W. E, and B. Chazelle. Noisy hegselmann-krause systems: Phase transition and the 2r-conjecture. *Journal of Statistical Physics*, 166(5):1209–1225, 2017.

[Wor22a] Caroline L. Wormell. Decay of correlations after conditioning on generic submanifolds in deterministic chaos, 2022.

[Wor22b] Caroline L. Wormell. Non-hyperbolicity at large scales of a high-dimensional chaotic system. *Proc. A.*, 478(2261):Paper No. 20210808, 19, 2022.

[WSS18] Spencer Wheatley, Michael Schatz, and Didier Sornette. The ARMA Point Process and its Estimation. *arXiv e-prints*, page arXiv:1806.09948, June 2018.

[YHB21] Gökhan Yalnız, Björn Hof, and Nazmi Burak Budanur. Coarse graining the state space of a turbulent flow using periodic orbits. *Phys. Rev. Lett.*, 126(24):Paper No. 244502, 5, 2021.

[Yos03] N. Yoshida. Phase transition from the viewpoint of relaxation phenomena. *Rev. Math. Phys.*, 15(7):765–788, 2003.

[You02] Lai-Sang Young. What are SRB measures, and which dynamical systems have them? volume 108, pages 733–754. 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.

[YRMK21] Serhiy Yanchuk, Antonio C. Roque, Elbert E. N. Macau, and Jürgen Kurths. Dynamical phenomena in complex networks: fundamentals and applications. *The European Physical Journal Special Topics*, 230(14):2711–2716, 2021.

[ZAAH12] Joseph Xu Zhou, M. D. S. Aliyu, Erik Aurell, and Sui Huang. Quasi-potential landscape in complex multi-stable systems. *Journal of The Royal Society Interface*, 9(77):3539–3553, 2012.

[ZL16] Peijie Zhou and Tiejun Li. Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, a-type integral and beyond. *The Journal of Chemical Physics*, 144(9):094109, 2016.

[ZLP21] Niccolò Zagli, Valerio Lucarini, and Grigorios A. Pavliotis. Spectroscopy of phase transitions for multiagent systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(6):061103, 2021.

[ZLP23] Niccolò Zagli, Valerio Lucarini, and Grigorios A. Pavliotis. Response theory identifies reaction coordinates and explains critical phenomena in noisy interacting systems, 2023.

[ZPLA23] Niccolò Zagli, Grigorios A. Pavliotis, Valerio Lucarini, and Alexander Alecio. Dimension reduction of noisy interacting systems. *Phys. Rev. Res.*, 5:013078, Feb 2023.

[Zwa61] R. Zwanzig. Memory effects in irreversible thermodynamics. *Physical Review*, 124(4):983–992, 1961.