PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à École Normale Supérieure

Macroeconomic Agent-based Models: A Statistical Physics perspective

Soutenue par
Dhruv SHARMA
Le 26 Novembre 2020

École doctorale n°ED614
Physique en Île-de-France

Spécialité
Physique Statistique

Composition du jury :

Leticia CUGLIANDOLO Professeure, LPTHE	*Présidente du Jury*
Tiziana Assenza Professeure, Toulouse School of Economics	*Rapporteure*
Matteo MARSILI Professeur, ICTP	*Rapporteur*
Isabelle SALLE Researcher, Bank of Canada	*Examinatrice*
Alan KIRMAN Directeur, EHESS	*Examinateur*
Francesco ZAMPONI Directeur de Recherche,LPENS	*Directeur de thèse*
Jean-Philippe BOUCHAUD Chairman,CFM	*Co-encadrant*

1794
ENS | PSL

Macroeconomic Agent-Based Models: A Statistical Physics Perspective

Dhruv Sharma

ABSTRACT

Agent-based models (ABMs) have emerged as a complementary paradigm for modeling macroeconomic phenomena. Compared to other, more established models such as DSGE (Dynamic Stochastic General Equilibrium) models, ABMs provide a flexible framework for understanding the complexity of the macroeconomy while at the same time taking into account the heterogeneous nature of economic actors, institutions and markets without making overly restrictive assumptions. ABMs take a "bottom-up" approach towards macroeconomic modeling by simulating the behavior of each individual agent in the economy and then aggregating to reveal emergent phenomena such as endogenous business cycles or flash crashes.

The object of this thesis is to advance a methodology commonly used in statistical physics and apply it to the study of two macroeconomic agent-based models. In both models studied here, we first determine the "phase-diagram" of the model to identify the relevant macroscopic regimes to develop an intuitive understanding of the macro-dynamics using a small subset of parameters.

The first ABM presented here builds upon the paradigm of constraint satisfaction problems (CSPs) and integrates it within the model's behavioral rules via agents' budgetary constraints. These constraints, similar to the well-studied perceptron CSP, reveal the existence of three regimes and underscore the importance of debt for macroeconomic stability: at low-levels of debt, the economy remains structure-less with frequent bankruptcies while high debt leads to *endogenous* business cycles. Between these two extremes, an intermediate regime of relative stability is found with low levels of bankruptcies for all times.

Within this ABM, agents' preferences, serving as the source of disorder in the CSP, evolve continuously in time. We thus study a simple dynamical scheme for the perceptron and discover that a rugged landscape can indeed exist with dynamic, annealed disorder.

Finally, we extend the Mark-0 ABM to simulate exogenous consumption and productivity shocks due to the Covid pandemic. Whereas standard approaches design a model to understand a particular outcome, this model can generate a variety of scenarios after a Covid-like shock. Furthermore, we also investigate the efficacy of several policies, including the much-debated "helicopter money" drop, in avoiding economic collapse. We thus highlight the importance of ABMs as multi-purpose "scenario generators",

for producing outcomes that are difficult to foresee due to the intrinsic complexity of macro-economic dynamics.

RÉSUMÉ

Les modèles à agents (Agent-Based Models ou ABMs) sont apparus comme un paradigme complémentaire pour la modélisation des phénomènes macro-économiques. Par rapport à d'autres modèles plus établis, tels que les modelés DSGE (Dynamic Stochastic General Equilibrium), les ABMs offrent un cadre flexible pour comprendre la complexité de la macroéconomie tout en prenant en compte la nature hétérogène des acteurs économiques, des institutions et des marchés, sans faire d'hypothèses trop restrictives. Ces modelés adoptent une approche "bottom-up" de la modélisation macro-économique en simulant le comportement de chaque agent individuel dans l'économie puis en s'agrégeant pour révéler des phénomènes émergents tels que les cycles économiques endogènes ou les crashs soudains.

L'objet de cette thèse est de faire progresser une méthodologie communément utilisée en physique statistique et de l'appliquer à l'étude de deux modèles macro-économiques. Dans les deux modèles étudiés ici, nous déterminons d'abord le "diagramme de phase" du modèle pour identifier les régimes macroscopiques pertinents afin de développer une compréhension intuitive de la macro-dynamique en n'utilisant qu'un petit sous-ensemble de paramètres.

Le premier modèle présenté ici s'appuie sur le paradigme des problèmes de satisfaction des contraintes (de l'anglais Constraint Satisfaction Problems, CSPs) et l'intègre dans le cadre des règles de comportement du modèle via les contraintes budgétaires des agents. Ces contraintes, similaires à celles du perceptron, un CSP bien-etudié, révèlent l'existence de trois régimes et soulignent l'importance de la dette pour la stabilité macro-économique : à un faible niveau d'endettement, l'économie reste sans structure et les faillites sont fréquentes, alors qu'à un niveau élevé la dette conduit à des cycles économiques *endogènes*. Entre ces deux extrêmes, l'on trouve un régime intermédiaire de stabilité relative avec de faibles niveaux des faillites tout le temps.

Dans ce modèle, les préférences des agents, qui sont à l'origine du désordre dans le CSP, évoluent continuellement dans le temps. Nous étudions donc un schéma dynamique simple pour le perceptron et découvrons qu'un paysage rugueux peut en effet exister avec un désordre dynamique.

Enfin, nous généralisons l'ABM Mark-0 pour simuler les chocs exogènes de consommation et de productivité dus à la pandémie de COVID. Alors que les approches standards élaborent un modèle pour comprendre un résultat particulier, ce modèle

peut générer une variété de scénarios après un choc de type COVID. En outre, nous étudions également l'efficacité de plusieurs politiques, notamment la très controversée "monnaie hélicoptère", pour éviter l'effondrement économique. Nous insistons donc sur l'importance des ABMs comme des "générateurs de scénarios" polyvalents, pour produire des résultats difficiles à prévoir en raison de la complexité intrinsèque de la dynamique macro-économique.

ACKNOWLEDGMENTS

To these I owe a debt past telling:
My several muses, harsh and kind;
My folks, who stood my sulks and yelling,
And (in the long run) did not mind;
Economists, dead or alive, whose orations
I've filched to mix my own potations;
Indeed, all those whose brains I've pressed,
Unmerciful, because obsessed;
My own dumb soul, which on a pittance
Survived to weave this doctoral spell;
And, gentle reader, you as well,
The fountainhead of all remittance.
Read me before good sense desists
You'll sprain neither your purse nor your wrists.

— Adapted from a poem by Vikram Seth, *A word of Thanks*

I would first of all like to thank my advisors Francesco and Jean-Philippe for their continued support over the course of my Ph.D. This Ph.D. would not have been possible without their encouragement and insights and my constant interactions with them helped me mature as a scientist. I also deeply appreciate the trust and confidence they placed in me to develop my projects and lent me the independence to discover and learn far beyond the scope of our projects.

I would also like to thank Marco for being a source of enthusiasm, encouragement, and joviality. Finally, I would like to thank Stan for all his help on the Covid project. I would also like to express my gratitude to Michael Benzaquen for welcoming me into the Econophysix chair which provided a valuable platform for discussion with other Ph.D. students and researchers in the field.

I would like to thank my office mates — Louise, Aldo and Emilio — for their friendship and for their patience, especially when I got frustrated with my simulations. I would like to thank Elisabeth, Ada, Valentina, Marco, Alessandro, Misaki, Harukuni, Eric, Antonio, Chen, and Francois for all their help and their advice. Even at the most difficult of times during my Ph.D., I looked forward to talking with you guys over lunch

or over multiple cups of coffee. A very special thanks to Elisabeth and Dmytro who made sure that I wasn't the only Marvel (or science fiction) fan in the lab.

As much as this experience was defined by the people within the lab, it was also defined by the discussions and camaraderie outside of it. I would like to thank Sudarshan for the uncountable number of beers shared after a hard day's work and for Nagarjuna, Manali, and Neeraj for the passionate discussions on cricket at the Bombardier. I would also like to thank the folks over at the Tech Help channel — Camille, Augustin, Geoff — for being a 24-7 tech support helpline. A very special shout out to "Escrime X2013" — Denis, Clara, Antoine (Pelletier, Saporta, and Berthier), Gaetan, Maxime, Greg — for putting up with my rants and my insistence on using Emacs for anything and everything under the sun. I am also immensely grateful to Pierre-Cyril and our countless discussions on matters relating to physics and economics. Finally, I am especially thankful to Jose who, for all our disagreements and heated debates, I am privileged to call a friend and colleague.

I would like to thank all my colleagues at "Cité des Sciences et de l'Industrie" who welcomed me into their small community of *passionés de science* and who gave me the opportunity to rediscover my love for science and research whenever my motivation was at its lowest. I would like to especially thank my colleagues — Graciela, Valerie, Marlene, Gilles, Benjamin, Agathe, Guillaume, Aurelie, JB, Nadege, Olivier, Julia, Soufiane, Alain, and most importantly Laetitia whose support and advice have helped me tremendously in becoming a better speaker, a better scientist, and in many ways a better person.

I want to thank my parents who pushed me towards excellence and set me on this path. Lastly, I would like to thank Lipsa without whom none of this would have been possible.

CONTENTS

LIST OF FIGURES

1

OUTLINE

This central theme of this thesis is the study of agent-based models for understanding emergent macroeconomic behavior. This thesis is made of two separate but interrelated parts. In the first part, we propose the paradigm of using constraint-satisfaction problems and integrate it within the behavioral rules of our agent-based model. In the second part, we study a well-established agent-based model (ABM) to simulate shocks and examine the evolution of the post-shock economy.

This thesis begins with a discussion on the current state of macroeconomic modeling and where agent-based modeling fits within the larger study of macroeconomics. In Chap 2 we discuss some of the shortcomings of mainstream models such as DSGE models and how the ABM paradigm provides a viable alternative. We also discuss how agent-based models incarnate the spirit of the microfoundations approach *à la Lucas*, and how this is related to statistical mechanics. We present a (short) history of the interactions between economists and physicist. We conclude by motivating the particular modeling point-of-view adopted in this thesis, inspired from recent advances in statistical physics, namely Constraint Satisfaction Problems or CSPs.

In Chap 3, we present Constraint Satisfaction Problems through the use of two examples - Boolean Satisfiability problems like K-SAT and the Perceptron. Through these examples, we develop the basic notions and tools from statistical physics that we shall employ in the rest of the thesis.

We then consider the particular example of the perceptron CSP and discuss its extension to a non-convex optimization problem in Chap 4. We then discuss a simple dynamical rule which leads to non-quenched dynamic disorder. We then describe the details of a dynamical phase transition that we have named "self-planting". Starting from a configuration in which the constraints are unsatisfied (or UNSAT), this transition separates two phases: one in which the dynamics finds a zero-energy (or SAT) configuration and another phase where the dynamics moves from one UNSAT configuration to another.

A simplified version of the perceptron then forms the basis of the agent-based model that we present in Chap 5. The perceptron CSP is used to model agents' budgetary constraints and we detail the multiple regimes that are found for this model as a function of allowed debt. Three phases are found called the Unstable (U), Stable (S), and Endogenous Crises (EC) phases. Through a simple model that couples the production output in the economy to agents' budgetary constraints, we are able to demonstrate the importance of debt for macroeconomic stability.

The beginning of the year 2020 saw the world reeling under the effects of the Coronavirus pandemic. Sanitary measures like lock-downs and partial to complete shutdown of economic activity were adopted to curb the spread of the virus. Such drastic measures led to a steep downturn in economic activity and the world economy was hit. As a means to understand the impact of the COVID-19 shock, we consider another ABM called "Mark-0" in Chap 6. We discuss the possible scenarios of economic recovery and the evolution of the economy with or without strong policy measures.

The work presented in Chapters 4, 5 and 6 are available in the following papers [1–3].

A reader interested in the economic applications of the thesis can read Chap 2 and Chap 3 till section 3.4. The reader may then skip Chap 4 entirely and move on to Chapters 5 and 6, where two different ABMs are discussed. On the other hand, the reader interested in the technical details of the perceptron may then read the thesis in order. For such readers, a technical appendix A is also provided.

We conclude with a discussion of perspectives for future work and some personal reflections on the usefulness of economic modeling in Chap 7. A recurrent concept in this thesis is that of *emergence* - phenomena that occur when the overall effect of individuals' actions is different from what the individuals themselves are doing. In such a case the behavior of the system differs from the individual actions of the agents that comprise the system. The thesis explores this idea and its applications to macroeconomics. While the work on the perceptron remains in the realm of scientific exploration and serves as a "proof-of-concept" for the CSP-inspired approach, the work on Covid-like shocks demonstrates the importance of agent-based models for understanding and applying the lessons learned to real-world economic situations.

2

INTRODUCTION

> The master-economist must possess a rare combination of gifts ... He must be mathematician, historian, statesman, philosopher - in some degree. He must understand symbols and speak in words. He must contemplate the particular in terms of the general and touch abstract and concrete in the same flight of thought. He must study the present in the light of the past for the purposes of the future. No part of man's nature or his institutions must lie entirely outside his regard.
>
> — John Maynard Keynes, *Keynes on Marshall [4]*

2.1 OVERVIEW

Macroeconomics is the study of the large-scale structure of the economy. Contrasted with microeconomics which seeks to understand how individual behave when they are trading with one another or make decisions based on prevailing economic conditions, macroeconomics interests itself in the study of the economic forces at play at the scale of large institutions like central banks or more commonly at the scale of nation states. Commonly used notions like GDP (Gross Domestic Product) or the Central Bank Interest rate are examined within the realm of macroeconomics. In short, it concerns itself with large-scale (or macroscopic) economic phenomena.

Whereas the study of economics itself dates back millennia with thinkers such as Kautilya in *Arthashastra* discussing the merits and demerits of high taxation, or Ibn-Khaldun in the medieval period, macroeconomics as practiced today is a fairly new development in the history of economic thought. One can safely say that modern macroeconomics was born with the publication of *The General Theory* by John Maynard Keynes in the 1930s. Keynes work formalized later by Hicks and Ramsey paved the way for a scientific approach towards understanding economic movements within nation states. *Keynesian* economics remained the central theory for macroeconomics for almost

half a century and was behind the rapid expansion of the industrialized west twice — once after the Great Depression of 1929 and then after the second war in 1945.

Macroeconomic models in the period between 1945 and the stagflation of the 1970s were empirical in nature. Modelers looked for *causal* explanations for the economic trends that the data revealed to them. An example can be found in Tinbergen's Nobel Prize lecture: the fluctuations in beef prices seen annually could be *fit* if one considered negative sign in the fodder prices [5]. While empirical models continue to be used within the profession, they have given way to models which take into account the economic *expectations* of economic actors. The most commonly used version of such models for studying macroeconomic fluctuations are the so-called Dynamic Stochastic General Equilibrium or DSGE models [6]. However, just as the classical ideas of Marshall, Jevons, and Edgeworth were called into question during the period of the great depression, the prevailing economic situation for over a decade have cast serious doubts on the use of DSGE models and especially their methodology for analyzing and predicting macroeconomic phenomena. A common criticism leveled at these models is their inability to predict the Great Financial Crisis (GFC) of 2008 [7]. Although no model can be expected to predict an event like the GFC, the models' inability to say something about the *probability* of a crisis occurring is indeed a glaring shortcoming [8]. Beyond this, these models or their variants were unable to warn us about the consequences of the crisis once it had struck.

This lack of predictive and prescriptive power is due to the simplifications that render DSGE models analytically tractable. This set of assumptions are based on a desire to construct macroeconomic models on sound "microfoundations".

2.2 THE MICROFOUNDATIONS CONUNDRUM

LUCAS CRITIQUE Keynesian ideas found their practical application in policy making through structural models of the economy which concentrated on the movement of aggregates[1] in the economy. These models were built through an empiricist viewpoint — adjustments were made to the models following what the data suggested. [9]. However, following the stagflation of the 1970s, macroeconomic modeling underwent a paradigm shift with such large-scale macroeconometric models giving way to models founded on principles of microeconomics. One of the criticisms leveled at the models of the former kind came from Lucas [10]. In an influential paper, he explained that a major shortcoming of these large-scale models was that they couldn't account for changes in

1 The word aggregate here refers to the macroscopic quantities or variables in the economy. These are quantities such as the GDP or the manufacturing index. Whenever we say aggregate in what follows, it can be used as a synonym for macroscopic.

people's behaviors after a shift in policy. This line of reasoning came to be known as the "Lucas critique". Indeed, models at the time were estimated on historical data and this estimated model was then used to produce predictions following a change in policy or other scenarios.

At its heart the Lucas critique makes the following observation: any model necessarily has to be estimated on historical data; however this model is estimated under a specific policy regime and it (or its estimated parameters) may not be the same had a different policy regime been in place. Since large-scale models were empirically founded, Lucas goes on to suggest that models must take in account changes in the behavior of agents once the underlying economic conditions have changed. It is reasonable to expect that robust models must explain how changes in policy will affect aggregate outcomes [11]. Thus, models should be constructed based on sound theoretical foundations and must allow for the formation of expectations of economic agents. These expectations must also be allowed *to change* as the policy regime changes.

As a means of circumventing the Lucas critique, it was suggested that macroeconomic models should be "microfounded" [12, 13]. In other words, macroeconomic models should be constructed by aggregating the individual properties and behaviors of the agents. It was hoped that since the agents' behavior was fully encoded, any change at the aggregate level due to policy could be modeled safely and lead to robust predictions. These models and their descendants have found widespread acceptance for their analytical elegance and their analytical tractability. Moreover, these models are able to reduce the complexities of the economic world into a small number of parameters.

These models still need to make assumptions about the behavior of individual agents. The assumption most often used is that of a purely self-interested individual who seeks to maximize its utility. It is further assumed that agents are infinitely lived and have perfect foresight. These set of assumptions go under the name of rational expectations. The DSGE paradigm takes it one step further by linearly aggregating multiple agents and reasoning in terms of a *representative agent*. The models also assume a general equilibrium structure with the dynamics driven by purely exogenous shocks[2]. Let us consider these assumptions in turn.

REPRESENTATIVE AGENT At the core of the many failings of the DSGE approach is the episteme of the rational representative agent. By positing a representative agent, DSGE models throw away the heterogeneity which is the key to much of the complexity seen in the economy. Even for policy discussion, models based on Representative agent

2 Shocks refer to abrupt, sharp changes in economic conditions that have systemic consequences. A simple mental picture for a shock could be a sharp knock with a hammer to a pendulum performing simple harmonic motion (SHM). It is clear that such a perturbation will surely take us away from the SHM regime, and economic shocks can be thought about in a similar way.

assumptions are in doubt since the aggregate response (i.e. of the macroeconomy) may be different from that of the hypothesized representative agent. In a prominent paper, Kirman demonstrated that in general aggregate behavior of a set of utility optimizing agents cannot be taken as the optimal decision by the representative agent [14]. As Kirman puts it, "The sum of the behavior of simple economically plausible individuals may generate complicated dynamics, whereas constructing one individual whose behavior has these dynamics may lead to that individual having very unnatural characteristics.". There is simply no direct relationship between individual and collective behavior.

RATIONAL EXPECTATIONS Beyond the simplifying assumption of the representative agent, this representative agent is then assumed to possess perfect rationality. This forms the *Rational Expectations* hypothesis and in its strongest form [15] endows our economic agent with a form of "Olympic" rationality [16]. Briefly, the rational expectations hypothesis assumes that an agent possesses complete information about the economy (the so-called "information set") to be able to predict the price of goods up to some error term. Further assumptions follow: the agent is assumed to know the behaviors of all other agents and the true values of exogenous parameters along with the underlying probability distributions of exogenous variables [11].

It is clear to see that such agents with perfect foresight are not be found in the real economy, and the kind of rationality attributed to individuals has its roots in the introspection of economists rather than in the empirical truth regarding how individuals actually behave [14]. The assumption of "Olympic" rationality is related to the aggregation issues with the representative agent hypothesis mentioned above — even if we were to assume that rationality exists at the level of individuals, this doesn't necessarily imply its existence at the aggregate level. Finally, economic actors live, react, and trade within a highly complex system. In such scenarios, there is mounting evidence that suggests that heuristics or rules-of-thumb are in fact more "rational". There is also evidence that suggests that the rational expectations hypothesis breaks down in multiple contexts, for instance in minority games [17, 18].

EQUILIBRIUM AND HISTORY DSGE models in their simplest forms have a unique equilibrium and assume that deviations from this situation are small and smooth. Equilibrium in neo-classical analysis results from economic actors solving maximization problems (usually utility maximization) with prices of goods as parameters. The notion of equilibrium actually has deep roots in physics. Lucas gives the game away thus, "The underlying idea seems to be taken from physics, as referring to as system *at rest*. In economics, I suppose such a static equilibrium corresponds to a prediction as to how an economy would behave should external shocks remain fixed over a long period, so that

households and firms would adjust facing the same set of prices over and over again and attune their behavior accordingly." [19]

Furthermore, we only consider situations of small exogenous shocks around a *stable* equilibrium: if the equilibrium were unstable, the smallest perturbation would disturb the system and the subsequent evolution of the system could not be captured as small oscillations around some steady state. This has a lot in common with mechanical notions of stability and equilibrium: think about the small oscillation assumption for simple pendulums. Consequently, one is limited to an exercise of "comparative statics" where two different equilibrium states are compared without discussing *how* any state is reached from some arbitrary starting point. This in turn seems in contradiction with the crux of the Lucas critique, that agents form expectations of the future and change behaviors. Thus policy changes will necessarily lead to agents changing their behavior and hence introduce non-stationarity into the system. This may indeed prevent the economy from reaching *any* equilibrium state. In simpler terms, the equilibrium assumption implies that history doesn't matter. While technical in nature, this does have deep consequences. As Joan Robinson reminds us, "An economy that has developed the technology for growing potatoes does not have the same spectrum of technical knowledge as one which only grows wheat.", and hence the two follow different *paths* to a (some) purported equilibrium states [20].

EXOGENOUS VS ENDOGENOUS SHOCKS Within DSGE models, the observed fluctuations are assumed to be caused by external shocks. Hence, almost by construction, the mechanisms creating these shocks are abstracted away and one can only study how these shocks propagate. The models generally consist of non-linear equations which don't readily admit closed-form solutions. Often modelers have to resort to log-linearization: this involves converting a non-linear equation into an equation which is linear in the log-deviations around the stationary values of the variables. Log-linearization then prevents the modeler from understanding the global dynamics driving the model and hence makes it difficult to properly take into account system-wide and large fluctuations [21]. Given that the relevant non-linear regime is out of reach, DSGE models are unable to guide us towards measures which could potentially reduce the risk of downturns. The analytical tractability of the model hence comes at the expense of providing useful policy advice when a crisis or downturn does occur. Linearization also leads to many second-order effects during calibration. DSGE models are generally calibrated by matching second moments such as variances and covariances. Financial crises and economic downturns are rare tail events and are hence not captured by the second moments [22]. In the words of Stiglitz, "macroeconometrics becomes little more than an exercise in curve fitting" [8].

In light of these various theoretical issues, how did DSGE models maintain their privileged position as the workhorse for macroeconomic modeling. It is indeed for their rigorously derived theoretical content that microfounded models attracted so much attention. DSGE models, for all their shortcomings, are still used for making forecasts. For instance, the Bank of England [23] and the US Federal Reserve have their own DSGE models for performing macroeconomic prediction [24]. [3]

Another important reason for the (academic) success of DSGE models are the prevailing economic conditions. The period when the DSGE saw the most success was also the period of the so-called "Great Moderation" when economic output was high with very little volatility. It led Lucas to remark that "the central problem of depression-prevention has been solved,"[4] and in fact there was little need for policy at all [26]. Paradoxically, it is precisely because policies were so successful that Lucas was led to believe in their irrelevance [27]. However, the crisis of 2008 laid bare the theoretical and methodological inconsistencies of DSGE type modeling. Interestingly enough, the response to the crisis relied not on DSGE type models, but the more traditional IS-LM Keynes-Hicksian approach [28]

One is then led to conclude that the limitations of the DSGE approach are not related to the modeling framework of "microfoundations". Rather, the limitations are due to the correctness of the assumptions involved. DSGE models have simply failed to incorporate details about how consumers, firms and other economic actors actually behave. They further fail to take into account research on behavioral economics and information economics [8, 29].

BAND-AID FOR DSGE The economics profession has recognized the limitations of the standard approach towards DSGE modeling and has sought to correct the errors of the past. Since the root cause for the crisis was in the financial sector, newer DSGE models now take into account non-Gaussian shocks and financial frictions [30]. On the theoretical side of things, DSGE models now take into account heterogeneous agents who possess *bounded rationality* rather than perfect rationality. However, these extensions of the DSGE paradigm have been likened to epicycles added to the Ptolemaic model of astronomy [21]. Stiglitz has claimed that these extensions are purely ad-hoc and Paul Romer has even called these models "post-real" [8, 31]. Methodologically, these extended models suffer from the same ailments as their ancestors — one can only reason about these models in a linear regime and when the shocks are small. Furthermore, since direct interactions between agents and institutions is missing, they are unable to

3 The French central bank, *Banque de France*, however uses an "old-school" structural model [25]

4 It should be noted though that this comment concerns mainly the US economy. In the years between 1987-2007, the world economy did undergo many other crises such as the Asian Debt crisis of 1997 and the Dot-com bubble in 2001 to name just a couple.

capture aspects of real-world dynamics. Furthermore, the set-up still remains that of a general equilibrium model and shocks are on the whole exogenous [16]. Thus once again, they have little to say about large system-wide shocks and, more importantly, little or nothing to say about the regions of parameter space where the linear approximation is stable and valid.

The author believes that the wrong lesson has been learned. Even though DSGE models suffered from theoretical deficiencies, the source of the crisis is not to be found within any kind of DSGE model. As recent histories of the crisis have suggested [32], it is far more important to understand the plumbing of the financial system rather than integrate it in an *ad-hoc* manner. We must then also understand how the internal plumbing of the financial system is connected to the larger global economic system — the collapse of a particular bank is a supposedly "local event". How a local event such as this could trigger massive economic strife is indeed the relevant question to ask. The devil as always is in the details, or rather the plumbing.

Hence, modifications to the DSGE model as discussed above acts as temporary band-aids — enough to cover the wound but not to fundamentally change the methodology. There have been calls to broaden the range of macroeconomic models and Blanchard has suggested a typology of *five* different classes of models[5] [33, 34]. The requirement for all these models is, however, for them to be microfounded; as Blanchard quips "Where else to start from?". It would be worthwhile to note that it is possible to find aggregate behavior which is extremely hard, even impossible to micro-found and it would be better to concentrate, as was done for "structural models", to focus on the aggregates [9].

2.3 AGENT-BASED MODELS - THE WAY AHEAD

> How selfish soever man may be supposed, there are evidently some principles in his nature which interest him in the fortunes of others, and render their happiness important to him, though he derives nothing from it except the pleasure of seeing it.
>
> — Adam Smith, *Theory of Moral Sentiments*

What is the amateur macro-modeler to do? Should we leave behind the condition of micro-founding our models? That would be throwing the baby out with the bathwater. The shortcoming of the DSGE approach is the set of assumptions about agent behavior, not microfoundations themselves. The Lucas critique doesn't constrain us to assume

5 Much to the author's chagrin, there is no place for ABMs, the subject of this work, within Blanchard's typology.

that agents are perfectly rational utility maximizing beings. Thus, our models can be microfounded as long as one is clear about the objectives. As has been noted, models should capture the essential characteristics of the behavior of firms and people, and not try to capture all relevant dynamics [34].

Agent-Based Models or ABMs are a step in this direction. These models take a truly "bottom-up" approach for modeling economic phenomena without aggregating linearly. ABMs seek to explain the evolution of the macroeconomy by simulating the behavior and interaction with the environment of each agent and then combining the micro-level details to produce a consistent macro-picture. Thus, they offer a complementary approach to macroeconomic modeling by incarnating the spirit of the microfoundations approach as proposed by Lucas. Finally, ABMs build upon the understanding that heterogeneities at the micro-level produce complex and emergent phenomena at the macroscopic level.

ABMs follow a different set of assumptions for modeling the behavior of economic agents, assumptions that tackle each one of the deficiencies of the DSGE paradigm as noted above. Firstly, at the micro-level, the heterogeneity among agents is placed front and center — agents can follow different preferences sets, have differing utility functions, or face different constraints. One is also not limited by simply modeling human economic actors — indeed the flexibility of ABMs allow us to model institutions such as banks as well. Finally, ABMs can be constructed to take into account not just individuals but also their interactions with their environment. For instance, a central bank may well be interested in how banks might behave under a change in the regulatory framework. In this example, the regulatory framework forms the environment for the banks.

ABMs also incorporate the lessons from behavioral economics allowing us to go beyond the rational expectations hypothesis, and model our agents as possessing only bounded rationality with limited foresight. A generally accepted set of microfoundations is based upon heuristics or "rules of thumbs" empirically observed in how consumers and firms make decisions [35, 36]. Adam Smith reminds us that economic actors don’t always engage in purely selfish or rational behavior all the time [37]. Our economic models should reflect these attitudes as well [8].

Another way in which ABMs differ from standard DSGE models is their modeling of the complete evolution of the economy rather than studying “equilibrium” states. These equilibrium states in more mainstream models are produced by imposing a set of restrictions, often market-clearing, which may not hold at all times. By simulating the trajectory of the individual agents, and then studying some well-chosen aggregates, we can uncover the path that the macroeconomy follows towards any, if at all they exist, equilibrium states.

Furthermore, given that interactions between agents (humans or institutions) is integrated within the model, ABMs can also, *endogenously*, generate business cycles.[6] Thus, rather than separating the study of macroeconomics into the short- and long-run, ABMs can account for short-term fluctuations along with long-term trends. However, given that the behavior of individual agents is taken into account, ABMs can also be useful tools for simulating exogenous shocks, as is done in Chap 6 in this thesis.

Overall, ABMs provide a complementary way to construct macroeconomic models based on sound microfoundations that integrate the assumptions of bounded rationality, agents with limited horizons, network effects, and multiple markets [11]. ABMs also allow for feedback mechanisms that can amplify small effects, such as system-wide panics that can cause crashes. Even though there is a strong case to be made for structural models that consider only aggregates, such "top-down" approaches will necessarily miss the actual local effects that lead to the emergence of large-scale phenomena. On the other hand, ABMs can prove to be supremely useful for examining the effects of "top-down" policy interventions since the behavior of all agents is being considered within the model.

CRITICISMS For the range of possibilities that ABMs open up for macroeconomic modeling, they have seen only slow acceptance within the larger community. One line of criticism relates to the use of simulations rather than fundamental theory. The study by simulation approach is however fairly standard when there are complex phenomena to be studied and the modeler is unaware of the key variables and parameters that drive the system [40]. Once a specific effect has been identified, one can then start to build a theoretical model.

Another line of argument against ABMs concerns the lack of a single "core" model as is the case for DSGE models. [33, 34]. As we have discussed, the core DSGE model is fraught with inconsistencies and such a model exists due to the strong restrictions imposed for the purposes of analytical tractability. ABMs offer a different and broader approach: rather than study a *single* "core" model, a framework is provided to garner insight into the complexities of the macroeconomy.

A related thread of criticism relates to ABMs' inability to precisely match historical data and produce accurate predictions. This however is misunderstanding the purpose of ABMs — by their construction, ABMs seek to provide evidence for "stylized" facts i.e. long-term regularities in historical data. Moreover, ABMs are indeed useful to produce accurate forecasts following a policy intervention or for identifying possible sources of systemic risk [40, 41]. Such an approach has much in common with the methodology of

6 Endogenous business cycles have been discussed well before the advent of modern DSGE models. See Kalecki [38]. The idea of business cycles within market economies goes all the way back to Marx [39]

epidemiologists who attempt to identify the conditions under which an outbreak might occur and mitigate those risks, rather than precisely predicting *when* an outbreak will occur [42]. There is however some progress being made towards an empirical validation and calibration of Agent-based models with recent work by Hommes where a complete ABM is estimated [43].

Finally, it is often claimed that ABMs have too many parameters. This however is a modeling choice and not something inherently problematic with the ABM paradigm. One must note that DSGE models themselves are not immune to such criticism since these models also require the estimation of a fairly large number of parameters [33]. A large parameter space can quickly become unwieldy. One can then seek to reduce the parameter space by carefully studying the relevance of certain parameters and the robustness of the effects seen under a change of these parameters. This is indeed the point of view adopted in the ABM studied in this thesis in Chap 5 and also in the Mark0 model [44] (see also Chap 6). A key test for any model should indeed be how parsimoniously it is able to reproduce or explain observed phenomena.

2.4 APPLICATIONS OF ABMS

Even though ABMs might have a slower path towards acceptance within the macroeconomic modeling community, they have found applications in many areas. We present here a non-exhaustive list of ABMs that have found applications in diverse fields, while a more comprehensive review can be found in [45].

One of the institutions that adopted the ABM paradigm quite early was the Bank of England. Due to their centrality within the larger financial system, one of the first ABMs studied there concerned the inter-bank network. In [46], it was shown how diversification at the individual bank level can generate systemic risk because of cyclic dependencies. Another ABM was developed at the BoE to understand the housing market within the UK. The ABM presented in [47] explored housing prices influenced larger macroeconomic trends. Within this model, different kinds of households are considered along with a banking sector and a central bank. Agents (households) in this model follow behavioral rules that are determined via a variety of economic factors such as expectations of real-estate appreciations and the prevailing interest rates which in turn influence mortgage costs. Given the nature of the housing market, this work presented a great example of how ABMs can tackle various levels of heterogeneity within the subset of actors (households are distinguished on the basis of age, income, wealth etc) and also heterogeneity between actors (individuals and lender banks). Another ABM, developed at the BoE and presented in [48], looked at possible feedback loops within the UK corporate bond market.

The role of heterogeneities was examined in [49] and it was shown, within an Agent-based framework, how inequalities in income and wealth distribution can seriously hamper long-term economic growth. On the subject of economic growth, Dosi and co-authors have developed the "K+S" (Keynes meets Schumpeter) series of models. In these models, the dynamics of aggregate demand is coupled to innovation, and it is shown that technological progress itself is endogenous to macroeconomic dynamics [50].

The dynamics of aggregates such as the GDP are emergent phenomena of the macroeconomy. The emergence of such aggregates and the influence of credit and capital is explored in the agent-based models studied by Assenza and co-authors [51, 52]. They also study the role of heterogeneities at the firm level and its importance for the transmission of shocks. Another model which focuses on emergent phenomena is the Mark-0 model [44, 53, 54], itself derived from the Mark-1 series of models developed by Delli-Gatti and co-authors [55, 56]. Finally, in the JAMEL series of models, the role of expectations for long-term macroeconomic stability is examined. In these models, hearkening to Keynes' animal spirits, agents' sentiment for the market, switching between pessimism and optimism, can generate endogenous business cycles [57, 58].

2.5 STATISTICAL PHYSICS AND ABMS

Physicists have been exchanging with economists since the Santa Fe conference of 1987. The conference organized by two physicists, David Pines and Phil Anderson, and the economist Kenneth Arrow led to rich cross-fertilization with the birth of many insightful ideas [59]. Continued interaction between physicists and practitioners from economics, sociology and other social sciences has led to enormous progress in the last 30 years [60]. The interaction between physics and economics goes much further back than the Santa Fe conference. The very first works by Walras and Jevons were inspired by the study of thermodynamics. Mirowski has demonstrated in some detail how the concept of "equilibrium" ubiquitous in economic modeling owes its introduction in economics to thermodynamics [61].

Statistical physicists in particular have been interacting with economists for some time.[7] Initial work from statistical physicists focused on understanding empirical relationships such as income distributions [62–65]. These were based on the application of theoretical models from statistical mechanics to explain and "derive" empirical distributions. Such approaches came under fair criticism from within the econophysics community [66] and also from the economics community [67]. The crux of the criticism has been on the lack of awareness of the available economics literature by physicists and

7 For the purposes of this discussion, we limit ourselves to macroeconomics. In finance however, statistical physicists have had much success

the "reality" of the statistical models being used. Claiming that the empirical income distribution comes about due to the same kinds of exchange processes in physics can only be termed tenuous. Later work has expanded beyond just finding "particle" like statistical models and focused more on reproducing stylized facts, with the foremost example being that of the Minority Game [17, 68]. More recent work has sought to understand an ensemble of complex economies via the use of techniques from statistical mechanics [69]. In this work, a complementary direction is taken — mathematical tools from statistical mechanics are borrowed freely, but a different methodological approach is also proposed. On the other hand economists too have realized the relevance of statistical physics for understanding socio-economic phenomena and have proposed using such tools [70].

As discussed above, ABMs seek to "sum-up" individual interacting agents and observe the macroscopic phenomena generated. This approach shares a lot of common features with statistical physics. Just as the microfoundations approach requires the aggregation of the behavior of individual agents to best model the macroeconomy, statistical physics seeks to understand the collective phenomena that emerge from the interaction of entities at the individual level. This identification between approaches allows analytical techniques to be applied for analyzing ABMs.

Beyond such a *rapprochement*, statistical mechanics has more to offer. Since ABMs aim for an understanding of a complex, interconnected system such as the economy, the parameter space of any agent-based model is bound to be high-dimensional. While simulations can be performed to understand the overall mechanics, due to the large number of parameters , an *intuitive* understanding of the model remains elusive. To overcome this barrier towards genuinely understanding our models, statistical physicists use the "phase diagram" as an important concept. While microscopically physical systems might be very complex to study, at the macroscopic level physical systems can show remarkable regularity. Importantly, the macroscopic properties are emergent and individual constituents don't necessarily possess these properties. For instance, water can be found in multiple states or *phases*; however, the humble water molecule itself can't be said to be liquid, or solid. It is remarkable that a solid phase even exists; indeed "we are so accustomed to the rigidity property ... that it is an *emergent property* not contained in the simple laws of physics, although it is a consequence of them" [71]. An important property of these macroscopic phases is their robustness to changes of parameters — within each phase, macroscopic properties show little fluctuations and are agnostic to microscopic details over a large range of parameters.

Things are not always so simple of course: the system might well sit at "phase boundaries" i.e. the region of parameter space delimiting two or more macroscopic phases. In such situations, the system is critical and small changes to parameters via external perturbations or even internal fluctuations can lead to the system switching from one

macroscopic phase to another macroscopic phase with very different properties. The transformation from liquid water to water vapor is an obvious example of such a phase shift or a "phase transition". Phase transitions appear in the macroeconomy as "dark corners" where the economy becomes strongly non-linear and the response to both small perturbations (exogenous or endogenous) can become catastrophically large [44].

For agent-based models, a critical first step should then be to identify such phases and reduce the parameter space to a few parameters. By their very nature, ABMs lend themselves to such an approach since the parameter space can be *scanned* through computer simulations. A reduction of parameter space and identification of the different phases or regimes with their corresponding phase boundaries provides a first step towards understanding the basic phenomenology of the model. This is indeed a crucial *methodological* concept missing in the examination of ABMs. In the present work, we will discuss phase diagrams in each chapter and this methodological point-of-view is adopted throughout this work. Such an understanding, necessarily qualitative in nature, comes at the expense of quantitative precision. It is however primordial to first understand well an *in silico* system whose rules are being set by us, before embarking upon the exercise of calibration. Reducing the parameter space in itself is an important step, even if one wants to undertake a quantitative study of the model.

In this thesis, this methodological viewpoint is developed through the study of an agent-based model, itself based on an important class of problems at the interface of optimization, machine learning and statistical physics. These are the so-called "Constraint Satisfaction problems" which we present briefly below, and in detail in Chap 3.

CONSTRAINT SATISFACTION PROBLEMS We will be discussing Constraint Satisfaction Problems at length in Chap 3. Nevertheless, we provide here a brief introduction to these problems. Constraint Satisfaction problems or CSPs are, as the name suggests, problems involving variables which must satisfy some constraints. These constraints can be equality constraints or inequality constraints or a mix of both. These problems are widely studied in the computer science literature to understand their computational complexity or to develop algorithms to most efficiently solve a given problem. Typically, one has a set of N variables, which have to satisfy a number M of constraints. For example, the variables could be the weights of a neural network, and each constraint imposes that the network satisfies the correct input-output relation on one of M training examples (e.g. distinguishing images of cats from dogs).

Statistical physics goes a step further and studies the typical case in the limit where the system size $N \to \infty$ and the number of constraints $M \to \infty$ (called the "thermodynamics limit" in physics). This dual limit is taken in such a way that the ratio $\alpha = \frac{M}{N}$ is a fixed constant. The thermodynamic limit is taken because, as is often the case with physical phenomena, the typical behavior of large systems is governed by only a

limited number of parameters. CSPs also commonly display phase transition between two phases named SAT and UNSAT. The SAT phase corresponds to a situation where all constraints can be satisfied and the UNSAT phase corresponds to the situation where at least one constraint is left unsatisfied. Generally, for CSPs the parameter which controls the passage from one phase to the other is α [72].

2.6 CONTRIBUTIONS OF THIS THESIS

This thesis then advances a methodology commonly used in statistical physics and applies it to the study of agent-based models. Two agent-based models are studied in this thesis — the first ABM presented here builds on the paradigm of constraint-satisfaction problems and integrates it within the behavioral rules of agent-based modeling, while the second ABM is used to understand the impact of shocks due to the ongoing coronavirus pandemic. The present thesis thus contributes to the field of agent-based models through three distinct but interrelated lines of study.

Firstly, an agent-based model is proposed which uses a budgetary constraint to determine agent behavior. Albeit simple, the model shows three distinct phases as a function of the debt level allowed to the agents. Two of these three phases, namely the stable(S) and endogenous crisis (EC) phases are accompanied with a spontaneous generation of structure within the economy, similar to the establishment of self-organized criticality in other complex systems [73]. The third phase named unstable(U) remains structureless. The model then underscores the importance of the level of debt for economic stability. This itself is related to the fact that we couple the debt to the eventual level of production within the economy. This work was previously published as [2].

The budgetary constraint used in [2] is related to a well-known constraint-satisfaction problem, namely the perceptron. An old model studied as a prototype for understanding the functioning of real neurons, it has reemerged as a useful model to understand the jamming of hard spheres [74]. The perceptron is here extended to a non-convex optimization problem with a rugged energy landscape [75]. We propose a new dynamical scheme "Remove & Replace" which produces planted states [76] for the perceptron optimization problem. We name this phenomenon "self-planting", since the planted state is not predefined but is found as a result of the co-evolution of the weights and patterns of the perceptron. This work was previously published as [1].

Turning our attention back to agent-based models, we study another model, Mark-0, previously studied to understand monetary policy transmissions and for inflation targeting [44, 53, 54]. This model produces endogenous effects such as tipping points [44] and synchronized default waves [77]. We extend Mark-0 to study exogenous shocks such as those caused by the loss in economic activity following the COVID-19 pan-

demic. Rather than predict possible trajectories, we present plausible scenarios as a function of the severity of the shock and its duration. We find that without substantial and swift government policy interventions, the economy risks entering a "bad-state" characterized in the commentariat as an "L-shaped" recovery [78]. A novel result is the possibility of a "W-shaped" recovery wherein the economy bounces back after the shock but suffers another recession even without a second wave of infections and lock-downs. This work was previously published as [3].

3

BACKGROUND

The first foray of statistical physicists in the domain of computer science was via the study of combinatorial optimization problems. These problems come in various flavors and settings with the most famous among them being the Traveling Salesman Problem (TSP) or the knapsack problem. Briefly, these problems seek to find a solution incurring the least "cost". For instance, in the TSP, we seek a route that goes through all the cities on the salesperson's itinerary but which requires the least amount of fuel. Another related set of problems occur in computer science and their examination has far reaching implications, especially in the context of computational complexity. These are generically called "Constraint-satisfaction problems".

3.1 CONSTRAINT SATISFACTION PROBLEMS

In its most general form, the aim of a constraint satisfaction problem is to find a configuration of a number N variables that satisfy a certain number M of constraints. Typical examples of these problems are the Boolean Satisfiability problems where the variables are boolean and we describe them briefly below. More specialized and exhaustive treatments (from a physicists' viewpoint) can be found in [79].

We consider here the SAT series of problems. We define a Boolean formula F written in the conjunctive normal form [80]

$$F = C_1 \wedge C_2 \wedge C_3 \cdots \wedge C_M \tag{3.1}$$

with the C_i taking the form

$$C_1 = (x_i \vee (\neg x_j) \vee \cdots), C_2 = \cdots, \tag{3.2}$$

Here $\{x_i\}$ are the set of boolean variables with $\neg x_i$ standing for the logical NOT operation. The C_i are the *clauses* constructed out of the variables that are the constraints, numbering M, that the variables $x_i \in \{0,1\}^N$ must satisfy.

Thus a general Boolean satisfiability (or SAT) problem can be stated as follows: given a formula F in (3.1), find an assignment (or configuration) of the set of variables $\{x_i\}$ such that F evaluates to true. If such an assignment can be found, F is then said to be "satisfied" and "unsatisfied" otherwise. K-SAT is the specific problem when each clause C_i in the formula F takes exactly K variables. We then talk of the K-SAT problem. Replacing the logical OR operations between the clauses in Eq (3.2) with the logical XOR (eXclusive-OR) operation, leads us to the K-XORSAT problem. These problems though fairly straightforward to state are however the centerpiece in the study of computational complexity. In particular, it has been shown that the K-SAT problem for $K \geq 3$ is in the so-called NP complexity class i.e. no algorithm (as yet) has been found that can find an assignment for the variables $\{x_i\}$ in polynomial time, but given an assignment we can verify in polynomial time whether the assignment is "satisfiable". It has been shown that K-SAT is in fact an NP-complete problem — other problems in the NP complexity class can be formally reduced to the K-SAT problem [81].

We can define a particular instance $\mathcal{K}$ of the problem as the choice of variables $x_a = (x_{i_a^1}, \cdots x_{i_a^K})$ occurring in the clause C_a for each clause $C_a \in \{C_1, \cdots C_M\}$. The choice of the instance hence fixes the formula F via Eq (3.1). The selection can be done probabilistically by defining a measure over the set of instances $\mathcal{K}$. One then talks of random-K-SAT. Two measures commonly used are:

1. *Uniform measure*: the measure is uniform i.e. the instance is chosen uniformly over the whole space of instances.

2. *Planted measure*: in this case, we choose a random assignment $\bar{x}$ of the variables, and then we choose the set of clauses C_a that satisfy the assignment. The clauses are chosen uniformly over the set of clauses. We will have more to say about the planted measure in the Chap 4.

Intuitively, we would expect that as the number of clauses increases, the likelihood of finding a configuration that satisfies all constraints for a randomly drawn instance becomes smaller. Indeed, while studying random instances of the K-SAT problem, a threshold phenomenon was discovered as a function of the ratio of clauses M to the number of variables N — beyond a critical value of $\alpha_c = \frac{M}{N}$, a randomly drawn instance of the problem becomes unsatisfiable with probability 1. Stated more precisely, if $P_K(N, \alpha)$ is the probability that a randomly drawn instance $\mathcal{K}$ is satisfiable in the set of all instances of random K-SAT, then it was found that:

$$\lim_{N \to \infty} P_K(N, \alpha) = \begin{cases} 1 & \text{if } \alpha < \alpha_c \\ 0 & \text{if } \alpha > \alpha_c. \end{cases} \tag{3.3}$$

This threshold phenomenon is called the SAT-UNSAT transition, with the region $\alpha < \alpha_c$ named the SAT (for Satisfiable) phase and the UNSAT (for Unsatisfiable) phase for $\alpha > \alpha_c$. To a physicist studying statistical mechanics, such threshold phenomena are ubiquitous in physical systems and go by the name of "phase transitions". In the next section, we establish some notions from statistical mechanics to make this idea more precise.

3.2 FROM CSPS TO STATISTICAL MECHANICS

As introduced in Chap 2, the fundamental idea in statistical mechanics is the description of macroscopic phenomena as properties of the system that emerge statistically. Moreover, in statistical mechanics, we are not concerned with the particular microscopic configurations of the system — for instance the precise positions of a molecule in a gas; since essential, macroscopic properties like the temperature or pressure of the system do not depend on precise microscopic details. Hence the study of any physical system in statistical mechanics begins by defining a probability measure on the space of possible configurations $\overline{\sigma} = (\sigma_1, \sigma_2, \cdots, \sigma_3)$. The variables σ_i can take discrete (like spin variables) or continuous (particle positions in a liquid). To each configuration $\overline{\sigma}$ of the system, we associate an *energy* function $E(\overline{\sigma})$. The probability of finding the system in configuration $\overline{\sigma}$ at temperature T (or the inverse temperature $\beta = \frac{1}{T}$) is then given by the Gibbs measure,

$$\mu(\overline{\sigma}) = \frac{1}{Z} \exp\Big(-\frac{1}{T} E(\overline{\sigma}) \Big), \tag{3.4}$$

where Z is known as the partition function and defined as

$$Z = \sum_{\{\overline{\sigma}\}} \exp\Big(-\frac{1}{T} E(\overline{\sigma}) \Big). \tag{3.5}$$

Here $\{\overline{\sigma}\}$ stands for all the possible configurations $\overline{\sigma}$ that the system can be in. A central task when confronted with a statistical mechanical system is to compute the partition function Z of the system, since knowledge of the partition function allows us to compute all other statistical quantities of interest. For instance, we can compute the average energy of the system (with respect to the Gibbs measure $\mu(\overline{\sigma})$) via the partition function as

$$\langle E \rangle = \sum_{\{\sigma\}} \mu(\overline{\sigma}) E(\overline{\sigma}) = -\frac{\partial}{\partial \beta} \log Z. \tag{3.6}$$

A related quantity is the free-energy defined as $F = -\frac{1}{\beta}\log Z$. The free energy density $f = \frac{F}{N}$ is related to the average energy above as $\beta f = \beta e - s$ where s is the entropy which counts the (log of) number of configurations for a given value of the energy.

SIMPLE EXAMPLE- Let us now apply these notions to a simple system — a system of 3 spins all interacting with each other. Each spin $\sigma_i, i \in \{1, 2, 3\}$ can take two discrete values either +1 or −1. The spins σ_i correspond to magnetic orientations along some axis. The space of configurations for this system is 8-dimensional and for any given spin-configuration of the system, we can write the following energy function

$$E(\overline{\sigma}) = -\sum_{i,j,i\neq j}^{3} J_{ij}\sigma_i\sigma_j, \tag{3.7}$$

where J_{ij} are the "coupling constants". These modulate the strength of the interactions between any two pairs of spins in the system. In the simplest case, we take all the J_{ij} to be constant equal to J, and the sign of J determines the nature of the interaction: if $J > 0$ then the couplings are said to be ferromagnetic and the spins tend to align with each other and if $J < 0$ then the couplings are said to be anti-ferromagnetic and the spins tend to anti-align. We can then compute the partition function for the system which reads:

$$Z = 6\exp(-\beta J) + 2\exp(3\beta J) \tag{3.8}$$

This can be computed via simple enumeration: there are 2 configurations when all the spins are aligned (either all +1 or all -1) which have an energy of $-3J$. In the remaining 6 configurations, there is always one misaligned spin and hence all these configurations have an energy of J. The average energy of the system can then be computed as well via Eq (3.6) which gives

$$\langle E\rangle = -6J\Big(\exp(3\beta J) - \exp(-\beta J)\Big). \tag{3.9}$$

It is instructive at this point to consider two limiting cases:

1. *High Temperature limit or* $\beta \to 0$: In this case, the average energy is zero and the Gibbs measure is flat i.e. uniform over all the possible configurations of the system. The system is said to be in the paramagnetic state.

2. *Low temperature limit or* $\beta \to \infty$: In this case, the Gibbs measure is concentrated on two configurations of the system — either all spins are +1 or all spins are -1. Thus in both configurations, the spins are all aligned with each other. These two states are equivalent and correspond to the "ground state" configurations of the system. In this limit, the energy (Eq (3.7)) takes its minimal value and equals $-3J$.

The low-temperature limit motivates another interpretation of the energy as defined in Eq (3.7): it is the "cost" incurred whenever two spins are not aligned with each other and this "cost" is minimal whenever all the spins are aligned. This interpretation will prove to be crucial when trying to examine the K-SAT problem through the lens of statistical mechanics.

The energy function we wrote in Eq. (3.7) is fairly generic and can model many different systems. Our choice of taking the couplings $J_{ij} = J$ constant corresponds to the well-known Ising model. One can take the other extreme limit when J_{ij} are all themselves random variables. In such a case, the energy function itself becomes a random function and this choice corresponds to the Sherrington-Kirkpatrick or SK Model. The introduction of random couplings J_{ij} leads us naturally to the notion of "quenched disorder" — when computing physically relevant quantities such as the free-energy or the energy, we must also perform an average over the complete distribution over the couplings. However, a question remains — if the energy itself is a random function, then what is the physically relevant property of the energy distribution, insofar as we are interested in properties of the whole ensemble of systems and not just particular instances of it? It turns out that for quantities such as the free-energy f, the average over the disorder suffices to make claims about the properties of a typical sample. These quantities are then said to be self-averaging.

The actual computation of these quantities proceeds in two steps — first we compute the partition function $Z(J_{ij})$ keeping the J_{ij} fixed. Hence, we get $J-$ dependent quantities such as the free energy $f(J,\beta) = -\frac{1}{\beta N}\log Z(J_{ij},\beta)$. We then proceed to compute the average over all realizations of the couplings J_{ij}:

$$f(\beta) = \mathbb{E}_{J_{ij}}[f(J,\beta)] = \mathbb{E}_{J_{ij}}\Big[-\frac{1}{\beta N}\log Z(J_{ij},\beta)\Big], \tag{3.10}$$

where $\mathbb{E}[\cdot]$ stands for the expectation value taken over the distribution of the couplings J_{ij}. It is in this sense that J_{ij} constitutes quenched disorder — the randomness in the couplings J_{ij} model the possibility of impurities in the system which can possibly diffuse around the system. However, we make the assumption that the time-scale of this diffusion is larger than the time-scale for thermal fluctuations and we take the couplings J_{ij} to be constant hence the name quenched.

3.3 BACK TO K-SAT

We turn our attention back to K-SAT. Our brief foray into the central notions of statistical mechanics have taught us that a first step towards formulating a problem should be to identify an energy function or cost function whose minimum value corresponds to a ground-state configuration. For a boolean satisfiability problem such as K-SAT, this corresponds to a particular assignment of the variables when all the clauses evaluate to true (are satisfied). Thus, given an instance $\mathcal{K} = \bigwedge_{a=1}^{M} C_a$, and an assignment $\mathcal{B} : x_1 = b_1, x_2 = b_2, \cdots x_n = b_n$, an intuitive choice for the energy of the assignment $\mathcal{B}$ is:

$$E(\mathcal{B}) = \text{Number of clauses of } \mathcal{K} \text{ violated by } \mathcal{B}. \tag{3.11}$$

With this particular choice, minimizing the energy corresponds to minimizing the number of clauses which are violated: if the assignment $\mathcal{B}$ is satisfiable, then it will have zero energy. Furthermore, since we are interested in only the lowest energy solutions to the problem, we will naturally work in the $\beta \to \infty$ or $T = 0$ limit. We use spin variables to construct an appropriate energy function. Let σ_i be spin variables such that for $x_i \in \{0, 1\}$, we have $\sigma_i = (-1)^{x_i}$. Thus $\sigma_i = -1$ (respectively $+1$) when x_i is false (true). We define coupling variables J_{ai} in the following way: $J_{ai} = -1$ (respectively $+1$) if clause C_a contains x_i (respectively $\neg x_i$). Then for a given assignment S_i of the spin variables, the quantity $\mathbb{I}(S)$ defined as :

$$\mathbb{I}_a(S) = \prod_{i=1, i \in C_a}^{N} \frac{1 + J_{a,i} S_i}{2}, \tag{3.12}$$

where the product is only evaluated for those spin variables that occur in the clause C_a. Then $\mathbb{I}_a(S) = 1$ when all assignments S_i make the corresponding variables x_i false. As an example, consider the case for the clause $C_a = x_1 \vee x_2 \vee (\neg x_4)$ and the assignment $S_i : S_1 = -1, S_2 = -1, S_4 = +1$. For the clause C_a, the couplings take the form $J_{a,1} = -1, J_{a,2} = -1, J_{a,4} = +1$. We can then confirm by computing Eq (3.12) for this assignment that $\mathbb{I}_a(S) = 1$. The energy function for K-SAT is then simply:

$$E(S) = \sum_{a}^{M} \mathbb{I}_a(S) \tag{3.13}$$

We can then seek to minimize this cost function in the limit $\beta \to \infty$, which is equivalent to finding the ground state of the system. Remember that the ground state configurations are those that minimize the energy and for K-SAT these will correspond to

configurations which satisfy all clauses. Furthermore, at the ground state, the partition function simply counts the number of possible satisfying assignments and hence the Gibbs measure becomes uniform over the space of assignments. It is important to note that here the couplings J_{ia} play the role of quenched disorder and hence we must compute the free-energy through the Eq (3.10).

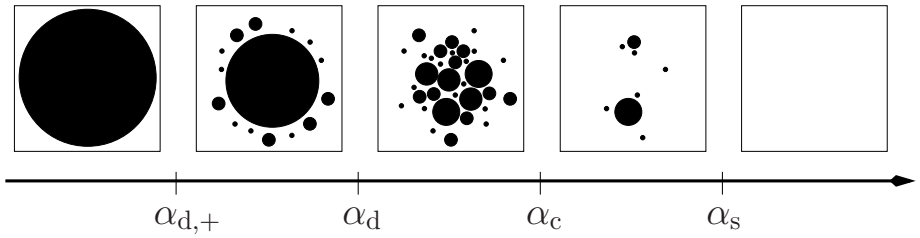


Figure 3.1: Schematic for the various transitions that a CSP can undergo as a function of the parameter $\alpha = \frac{M}{N}$ and a rich phenomenology is revealed. $\alpha_{d,+}$ is when clusters first appear, followed by the clustering transition at α_d. The exponential number of clusters then condense to a sub-exponential number of solutions following the condensation transition at α_c. Finally beyond α_S, the CSP admits no solutions. Figure from [82].

The computation of the quenched free-energy of CSPs such as K-SAT teaches us that these problems can undergo not just the SAT-UNSAT transition but multiple phase transitions as shown in Fig 3.1. For small values of α, the CSP is in the SAT phase — a random instance of the problem is satisfiable with high probability. For large values of $\alpha > \alpha_S$, the CSP is in the UNSAT phase — a random instance is almost surely unsatisfiable. Between these two extreme regimes, the set of solutions can split into small clusters at $\alpha_{d,+}$ and between $\alpha_{d,+}$ and α_d, these clusters constitute only a small (sub-exponential) fraction of solution space. At α_d, the CSP undergoes the "clustering" transition — the main cluster splits into an exponential number of smaller clusters. The CSP remains in such a "clustered" phase till α_c when the CSP undergoes the "condensation"phase transition. In this phase, the exponential number of clusters condense to a sub-exponential number of clusters. Finally, beyond α_S the CSP doesn't admit any solutions and undergoes the SAT-UNSAT transition. We hence observe that even within the SAT phase, when the CSP admits solutions with high probability, the *geometry* of the space of solutions is fairly intricate. Such a picture is revealed by a thorough analysis of the quenched free-energy of CSPs such as K-SAT. We will not provide the details of the computation of the quenched free-energy for a CSP described over a set of discrete (spin) variables such as K-SAT in this manuscript and we refer the reader to specialized texts [72, 79, 83, 84]. However, we can very easily generalize the

set of problems and the tools to study them to CSPs with continuous variables. In the rest of this chapter, we tackle the specific example of a CSP with continuous variables which forms the heart of this work — the perceptron.

3.4 CONTINUOUS CSPS — THE PERCEPTRON

Let us begin our discussion of continuous CSPs with a simple geometric problem. Consider M points $\vec{\xi}^1, \vec{\xi}^2, \ldots, \vec{\xi}^M$ living in the N-dimensional space $\mathbb{R}^N$. We denote the coordinates of the μ-th point $\vec{\xi}^\mu$ as $\vec{\xi}^\mu = \left(\xi_1^\mu, \xi_2^\mu, \ldots, \xi_N^\mu\right)$. We then ask the following question: does there exist a vector $\vec{p} \in \mathbb{R}^N$ such that its scalar-product with $\vec{\xi}^\mu$ is positive:

$$\vec{p} \cdot \vec{\xi}^\mu \equiv \sum_i p_i \xi_i^\mu \geq 0 \qquad \forall \mu \in 1 \ldots M \tag{3.14}$$

Stated differently, we want to know whether the M points $\vec{\xi}^\mu$ belong to the same half-space. For points chosen randomly on the N-dimensional hyper-sphere, we can ask a similar, probabilistic version of this question [85]. Just as in the case for discrete CSPs, we expect that as we increase the number of constraints, it becomes more difficult to find configurations that simultaneously satisfy all constraints. Indeed, we find that the probability of finding a set of M random points that belong to the same half-space depends on the ratio $\alpha = \frac{M}{N}$, and in the thermodynamic limit, a threshold phenomenon is retrieved. We find that:

$$\lim_{N\to\infty} P(N, \alpha) = \begin{cases} 1 & \text{if } \alpha < \alpha_S \\ 0 & \text{if } \alpha > \alpha_S. \end{cases} \tag{3.15}$$

.

The critical value is $\alpha_S = 2$. Below this critical value, there exists, almost surely, a vector $\vec{p}$ satisfying Eq. (3.14). Beyond the critical value, almost surely, no such vector can be found. Thus this geometric problem also undergoes a SAT-UNSAT transition at $\alpha_S = 2$.The problem of finding a vector $\vec{p}$ which has a positive dot product with the points $\vec{\xi}^\mu$ goes by the name of the perceptron problem.

We can of course generalize the constraints to have scalar products larger than some positive value, say σ. In the language of continuous CSPs above, the following reformulation is immediate. Let $\vec{X}$ be vectors on the N-dimensional sphere $\mathcal{S}_N$, such that $|\vec{X}|^2 = N$. Given M vectors $\vec{\xi}_\mu \in \mathcal{R}^N$, we seek for $\vec{X} \in \mathcal{S}_N$ such that:

$$h_\mu(\vec{X}) := \frac{1}{\sqrt{N}} \vec{\xi}_\mu \cdot \vec{X} - \sigma > 0 \qquad \forall \mu \in \{1 \ldots M\}\,. \tag{3.16}$$

The quantities h_μ will be called gaps in the following. We discuss how the constraint in Eq. (3.16) can encode a learning problem and also be used to model the sphere jamming problem.

3.5 THE PERCEPTRON AS A LEARNING PROBLEM

The perceptron was one of the first formal models of how neurons function. Starting with the pioneering work of Rosenblatt [86], it has continued to be both a paradigm and a cornerstone of machine learning. Statistical physicists became interested in the perceptron problem and significant progress was made in the late 1980's in a series of papers [87–91]. The perceptron is an element that gives an output σ^μ according to the following rule:

$$y^\mu = \mathrm{sgn}\Big(\sum_{j=1}^{N} p_j \xi_j^\mu\Big). \tag{3.17}$$

Here $\xi_1^\mu, \xi_2^\mu, \cdots, \xi_N^\mu$ are N inputs and $p_1, p_2, \cdots, p_N$ are called synaptic connections with μ running from $1 \cdots M$. Thus, whenever the input $\vec{p} \cdot \vec{\xi}^\mu$ exceeds 0, the perceptron (modeling a neuron) fires (y^μ = +1 and -1 otherwise). The corresponding learning problem consists of adjusting the weights $\vec{p} = [p_j]$ so that the inputs $[\xi_i^\mu]_{i=1,\cdots,N}$ (called patterns) are matched correctly to their corresponding output $y^\mu \in \{-1, +1\}$.

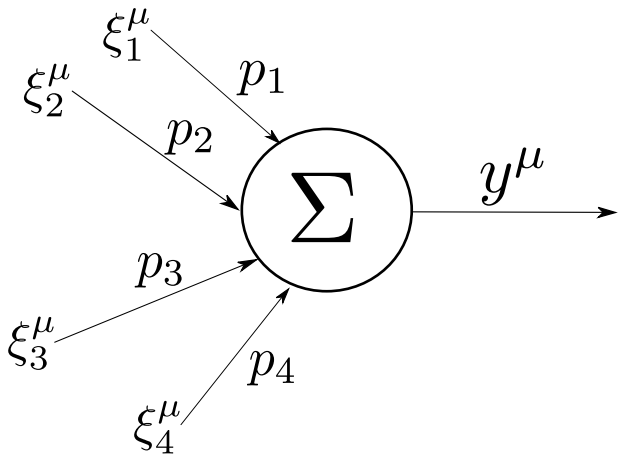


Figure 3.2: Schematic of a simple perceptron taking the inputs ξ_i^μ and generating an output y for the couplings $\vec{p}$.

In the limit of large N thus we normalize the input signal to get the form :

$$y^{\mu} = \mathrm{sgn}\left(\frac{1}{\sqrt{N}}\sum_{j} p_j \xi_j^{\mu}\right), \tag{3.18}$$

and where the couplings p_j are also normalized $\sum_{j=1}^{N} p_j^2 = N$. Eq (3.17) can also be recast as a classification problem, or more precisely a support vector machine (SVM) [92]. Consider the quantity

$$\Delta^{\mu} \equiv \frac{y^{\mu}}{\sqrt{N}}\sum_{j} p_j \xi_j^{\mu} > 0 \tag{3.19}$$

Then if all the constraints defined by Eq. (3.19) are satisfied, then there exists a hyperplane $\vec{\xi}\cdot\vec{W} = 0$ (with $\vec{W} = \vec{p}/\sqrt{N}$ and $|\vec{W}|^2 = 1$) which correctly classifies all inputs according to the rule $y_{\mu} = \mathrm{sgn}(\vec{W}\cdot\vec{\xi}_{\mu})$ (Eq (3.17)). More generally, we can have $\Delta^{\mu} > \sigma$ in which case we demand that the minimal distance Δ^{μ} of points $\vec{\xi}^{\mu}$ from the hyperplane found above be greater than σ (formally equivalent to the set of constraints in Eq (3.16)). σ then corresponds to the "margin" of the SVM.

One can also ask a dual question: given M random pairings of inputs and outputs $(\vec{\xi}_{\mu}, y_{\mu})$, what is the maximal number of such pairings that a perceptron can learn? Or formulated differently, in the absence of a relation $f(\vec{\xi}_{\mu})$ relating y_{μ} and $\vec{\xi}_{\mu}$, how many random patterns can the perceptron classify? This is known as the storage problem and the critical number M_c of patterns learned is noted as the storage capacity [87, 88].

3.6 PERCEPTRON AND SPHERE PACKING

It is known that the packing fraction ρ of a collection of hard spheres in 3 dimensions cannot exceed a maximum of 0.74. This is the well-known case of the close packed lattice such as the HCP (Hexagon Closed packing) lattice. However, if we begin with a loose collection of hard spheres with random initial conditions and compactify these spheres, then ρ_c can be less than 0.74. We say that the system *jams* during the compactification process before it crystallizes. Interestingly, $\rho_c < 0.64$, which is called the random closed packing limit.

There has been a lot of activity in the study of the jamming transition [93]. Recently, Franz and Parisi [74] consider the onset of jamming as a SAT-UNSAT transition, where the terms SAT and UNSAT carry the same meaning as in the previous section. They study a toy problem of jamming where M points ξ_i^{μ} are randomly distributed over the surface of a $N-$ dimensional sphere with radius $\sqrt{N}$. Here i runs from 1 to N and μ

runs from 1 to M. Suppose that we have a hard-core radius σ. Then if we were to add a new free particle with coordinates x_i, it would have to satisfy the condition

$$|\vec{\xi}^{\mu} - \vec{x}| > \sigma \tag{3.20}$$

The identification with the jamming problem is straightforward. The perceptron problem requires that the quantities

$$r_{\mu} = \frac{1}{\sqrt{N}} \sum_{i} \xi_{\mu}^{i} x_i - \kappa \tag{3.21}$$

must be all positive. By squaring 3.20, we observe immediately that the toy problem of jamming is actually a perceptron problem where we identify $\sigma^2 = 2N + \kappa$.

Franz and Parisi treat κ as a free parameter and show analytically that for $\kappa > 0$, the system is not jammed while for $\kappa < 0$ there is a jamming transition. In fact, the regime where $\kappa < 0$ is a non-convex optimization problem. It is in this regime that we encounter *replica symmetry breaking*, i.e. the breaking of the space of configuration into separate domains similar to the clustering transition of Fig. 3.1.

3.7 DISCUSSION

In the present chapter, we have looked at random constraint-satisfaction problems defined over discrete (K-SAT) and continuous variables (the perceptron). Although at first sight these problems may not be particularly relevant for a statistical physicist, their examination, especially in the thermodynamic limit, reveals many interesting features. The study of CSPs is a rapidly developing field and many other examples could have been discussed as well [79]. In the rest of this manuscript, we will focus on the perceptron, a CSP with continuous variables more directly relevant for an economic modeling point of view.

Two directions are hence proposed to the reader. In Chap 4, we embark on a more detailed and technical exploration of the Perceptron. We will examine some notions that were only briefly introduced in this chapter such as the planted ensemble and we will also consider the dynamical evolution of the disorder and its consequences for the perceptron CSP.

The more economically-minded reader can safely skip Chap 4 and can proceed directly to Chap 5, written so as to be independent from Chap 4. There we develop an agent-based model which builds upon the CSP paradigm to study emergent macroeconomic behavior. The vocabulary developed here will be used to reason about the

economic phenomena that emerge from the study of the proposed ABM. Chap 5 also constitutes one of two agent-based models studied in this manuscript, with the second one presented in Chap 6.

4

SELF-PLANTING IN THE PERCEPTRON CSP

In this chapter, we will define and study the dynamics of the perceptron CSP that we discussed in the previous chapter. Although this chapter can be read independently from the next chapter, let us nevertheless understand why we are interested in the perceptron and its dynamics for our economic ABM.

The primary objective of the macroeconomic ABM to be presented in Ch 5 is to demonstrate how the CSP paradigm can be used to encode real economic constraints and discover non-trivial macroscopic behavior. In such a setting, the perceptron constraint encodes agents' budgetary constraint. Agents' preferences for buying and selling goods are then encoded in the vector $\vec{\xi}_\mu$, which plays the role of quenched disorder in the perceptron CSP. Since in an economic setting agents' preferences are constantly evolving, we must consider the case of dynamic disorder for the perceptron CSP as well. Furthermore, as agents' go bankrupt (or in CSP parlance, unable to satisfy their constraints), they are removed from the economy and replaced with newer agents, who in turn have their *new* set of preferences. We are then led to consider a particular kind of dynamics called "Remove & Replace (R&R)" that we examine in this chapter. While the ABM to be presented in Ch 5 will have richer dynamics, it is instructive to thoroughly understand the simpler case first. The R&R dynamics deserves a closer look from a physical point of view as well, since it extends the discussion from the well-studied case quenched disorder for the perceptron CSP to the more interesting case of evolving, dynamic disorder.

4.1 A FEW PRELIMINARIES

Following up on the definition of CSPs in the previous chapter, we can define CSPs on continuous variables as follows: find an N dimensional vector $\vec{X} \in \mathbb{R}^N$ which satisfies the M constraints $h_\mu(\vec{X}) > 0 \ \forall \mu$ where μ runs from 1 to M. Here $h_\mu(\vec{X})$ is a function $h_\mu(\vec{X}) : \mathbb{R}^N \to \mathbb{R}$. In general, the functions h_μ can have quenched disorder.

To study such problems, we begin by defining a Hamiltonian $\mathrm{H}[\vec{X}]$, (the energy function) such that when all the constraints are satisfied, $\mathrm{H}[\vec{X}] = 0$ and non-zero if there is at least one unsatisfied constraint. The choice of the Hamiltonian is such that we incur a cost whenever a constraint is unsatisfied

$$\mathrm{H}(\vec{X}) = \sum_{\mu=1}^{M} W(h_\mu(\vec{X})) = \frac{1}{2}\sum_{\mu=1}^{M} h_\mu(\vec{X})^2\theta(-h_\mu(\vec{X})), \tag{4.1}$$

where $\theta(x)$ is the Heaviside function and $h_\mu(\vec{X})$ are the perceptron constraints:

$$h_\mu(\vec{X}) = \frac{1}{\sqrt{N}}\vec{\xi}_\mu \cdot \vec{X} - \sigma > 0 \qquad \forall \mu \in \{1 \dots M\}. \tag{4.2}$$

This cost itself is taken as the square of the distance from a satisfied configuration — if $h_\mu = 0$, then the constraint is satisfied, hence h_μ encodes the distance from this satisfied configuration. In what follows, we call h_μ the "gaps". We are interested in the limit $M \to \infty$, $N \to \infty$, with $\alpha = \frac{M}{N}$ kept constant. Given the Hamiltonian, we can compute the partition function

$$Z = \int [\mathrm{D}\vec{X}] e^{-\beta H} = \int [\mathrm{D}\vec{X}] e^{-\frac{\beta}{2}\sum_\mu h_\mu(\vec{X})^2\theta(-h_\mu)}, \tag{4.3}$$

where $\mathrm{D}\vec{X}$ is the integration measure, which might eventually contain certain conditions to be satisfied by $\vec{X}$, for instance normalization conditions. The variables ξ_i^μ play the role of quenched disorder. We once again study the zero-temperature limit of the partition function since in this limit, we can access the ground state solution of the problem. The free energy density is then given as usual:

$$f = -\frac{T}{N}\overline{\ln Z} = e - Ts, \tag{4.4}$$

where e is the energy density, s is the entropy densityand $\overline{Z}$ denotes a quenched average. The quenched average is performed over the distribution of the disorder generated by the patterns ξ_i^μ. We thus have to compute the quantity:

$$\overline{\log Z} = \int [\mathrm{d}\vec{\xi}_\mu] P(\vec{\xi}_\mu) \log\left(\int [\mathrm{D}\vec{X}] e^{-\frac{\beta}{2}\sum_\mu v_\mu(\vec{X})\theta(-h_\mu)} \right), \tag{4.5}$$

where $P(\vec{\xi}_\mu)$ is the probability distribution of the variables $\vec{\xi}_\mu$. Computing the quenched average is difficult and we employ the "replica trick" — rather than computing the quenched average of $\log Z$, we compute the quenched average of Z^n by using the following identity

$$\overline{\log Z} = \lim_{n\to 0} \frac{\overline{Z^n} - 1}{n} \tag{4.6}$$

The replica trick provides us an alternative way of computing the quenched average — we make n (with n some positive integer) copies of "replicas" of the partition function, compute the quenched average of these n copies and then take the limit $n \to 0$. As a mathematical identity, Eq (4.6) poses no mathematical problems *a priori*. However, the analytic continuation from n integral to n real, tending to 0 isn't perfectly rigorous. Indeed, for many years the replica trick remained just that, a trick until the rigorous proofs from Talagrand which transported Eq (4.6) from the physicists' "bag-o-tricks" to more solid ground [94, 95].

The introduction of replicas makes the computation of the quenched average simpler and the final partition function is then a function of the "overlap" between pairs of replicas. Briefly, the quantity

$$q_{ab} = \frac{1}{N}\vec{X}_a \cdot \vec{X}_b \qquad \forall a \neq b, \tag{4.7}$$

is the $n \times n$ replica overlap matrix which encodes the distance between the configurations $\vec{X}$ of two replicas a and b. For $a = b$, the overlap is of course 1 and it is customary to only look at off-diagonal terms. However, we do not possess the explicit values of the overlap matrix q_{ab} and moreover, since we introduced replicas for pure mathematical convenience, physical quantities such as the free-energy and the entropy should be independent of any particular value of the overlap matrix. This leads us to the "replica-symmetry" or RS solution — $q_{ab} = q_0$ for some value q_0 i.e. all replicas are equivalent and the free-energy has permutation symmetry in the replica indices $\{a, b\}$. We can then compute the final quenched free-energy via a saddle-point computation. The q_{ab} matrix under the RS ansatz has the following form, with all the off-diagonal entries equal to q_0:

$$q_{ab} = \begin{bmatrix} 0 & & & & \\ & 0 & & & q_0 \\ & & 0 & & \\ & q_0 & & 0 & \\ & & & & 0 \end{bmatrix} \tag{4.8}$$

Although the RS solution seems natural from a physical point of view, there are also situations when replica-symmetry is broken. We can show that the Replica-symmetric free-energy is not stable for all values of the temperature and can lead to unphysical results such as negative entropies (in the case of discrete models) [96, 97]. We then have to consider the possibility of violation of the symmetry between replicas — we are then led to the "Replica Symmetry Breaking" (RSB) solution. The system is then said to undergo the RS-RSB transition. First proposed by Parisi in a series of papers, the presence of RSB is a source of surprise but also great mathematical insight [96, 98, 99]. Let us briefly sketch some details of the RSB solution, and most importantly some of its physical implications.

The first-step towards breaking replica-symmetry consists of choosing an integer m_1 and then sub-dividing the matrix into $\frac{n}{m_1}$ blocks. Within each diagonal block, the replicas take the value q_1, with all off-diagonal blocks set to nq_0. The diagonal entries themselves are still zero. The overlap matrix under "1-step" replica-symmetry breaking then takes the form:

$$q_{ab} = \left[\begin{array}{ccc|ccc} 0 & q_1 & q_1 & & & \\ q_1 & 0 & q_1 & & q_0 & \\ q_1 & q_1 & 0 & & & \\ \hline & & & 0 & q_1 & q_1 \\ & q_0 & & q_1 & 0 & q_1 \\ & & & q_1 & q_1 & 0 \end{array}\right] \tag{4.9}$$

Further steps of replica-symmetry breaking are also possible. For instance in the second step, or "2-step" replica-symmetry breaking, the inner-most diagonal blocks are further subdivided into m_1/m_2 blocks by choosing an integer m_2. The off-diagonal entries of these blocks then are ascribed the value q_2, with the off-diagonal entries within these blocks taking the value q_1 as before. The remaining (larger) off-diagonal blocks remain with the value q_0. Under 2-step replica-symmetry breaking, the overlap matrix may take the following form:

$$q_{ab} = \left[\begin{array}{ccc|ccc|ccc|ccc} 0 & q_2 & q_2 & & & & & & & & & \\ q_2 & 0 & q_2 & & q_1 & & & & & & & \\ q_2 & q_2 & 0 & & & & & & & & & \\ \hline & & & 0 & q_2 & q_2 & & & & & & \\ & q_1 & & q_2 & 0 & q_2 & & & q_0 & & & \\ & & & q_2 & q_2 & 0 & & & & & & \\ \hline & & & & & & 0 & q_2 & q_2 & & & \\ & & & & & & q_2 & 0 & q_2 & & q_1 & \\ & & q_0 & & & & q_2 & q_2 & 0 & & & \\ \hline & & & & & & & & & 0 & q_2 & q_2 \\ & & & & & & & q_1 & & q_2 & 0 & q_2 \\ & & & & & & & & & q_2 & q_2 & 0 \end{array}\right] \tag{4.10}$$

In general, we can of course continue breaking replica-symmetry further as well — one then talks about K-step replica symmetry breaking. At each step of replica-symmetry breaking, we increase the number of variables that our free-energy depends on - we begin with q_0 under the RS-ansatz, then (q_0, q_1) under 1-RSB, (q_0, q_1, q_2) under 2-RSB and $(q_0, q_1, \cdots, q_K)$ under K-RSB. The most general case is the "full" replica-symmetry breaking ansatz which consists of taking the K itself to infinity. We then have a full "function's" worth of variables to take into account - q_{ab} is then a function $q(x)$ in the interval $[0, 1]$. The full-RSB ansatz is what we are confronted with when we study the perceptron CSP. A simple geometric problem does indeed hide a lot of complexity.

Our discussion on replica-symmetry and replica-symmetry breaking until now has been fairly abstract. Let us now consider some implications of these notions. The existence of Replica Symmetry breaking implies that the free-energy of the system becomes extremely complicated. This is best seen by contrasting it with the situation for the classic Ising model that we studied in Chap 3. The ground state of the Ising model, ones with lowest energy, are those where all the spins are pointing in the same direction. The total magnetization $m = (1/N)\sum_i S_i$ for these states ± 1. For all other states, the free-energy is higher and it can be shown that it varies smoothly as a function of the magnetization m. This is shown in the Fig 4.1 where we plot the free-energy as a function of some generic configuration of the system.

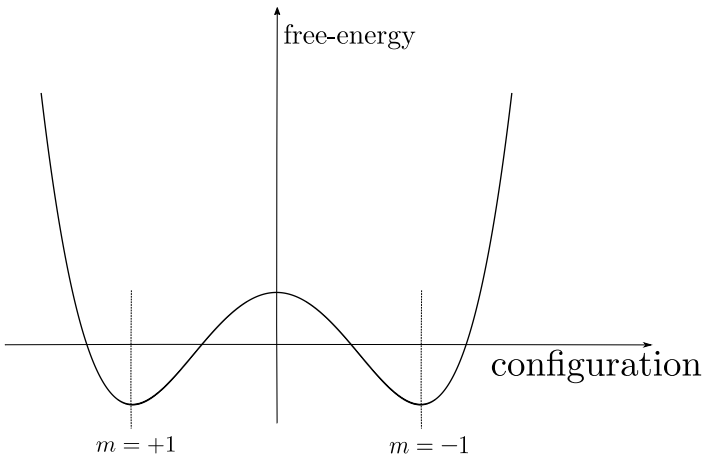


Figure 4.1: Schematic of the free-energy landscape for the Ising model drawn as function of the possible configuration. The two states $m = \pm 1$ correspond to the ground-state energies of the Ising model and all other states have energies higher than the ground-state energy (taken to be zero here).

Fig 4.1 is also known as the "energy landscape" of the Ising model: to each configuration of the spins, there is a specific value of the free-energy associated. The ground-states, corresponding to the two configuration $m = \pm 1$ are hence minima within this energy landscape. The energy landscape picture also informs us of the relative stability of the configurations — configurations with $m = \pm 1$ are stable since they are found at minima of the energy landscape. Furthermore, the two minima are symmetric and are separated by an energy barrier. Thus if we wish to switch the system for a $m = 1$ state to a $m = -1$ state, we must provide the system enough energy for it to surmount the barrier and spontaneous switches from one state to the next aren't allowed. For physical systems without disorder (standard Ising model) or for disordered systems in the Replica-symmetric phase, the energy landscape is fairly simple, akin to the one shown in Fig 4.1.

For systems that undergo replica-symmetry breaking however the situation is quite complex. The energy landscape is said to be "rough" and possesses an energy landscape structure shown schematically in Fig 4.2. The landscape has not a few but (exponentially) many minima which are separated by energy barriers which are of variable height. The system may then relax to one of the minima and if the energy barrier is sufficiently small to be overcome by thermal fluctuations, it may go over the barrier to find itself in a lower energy minima.

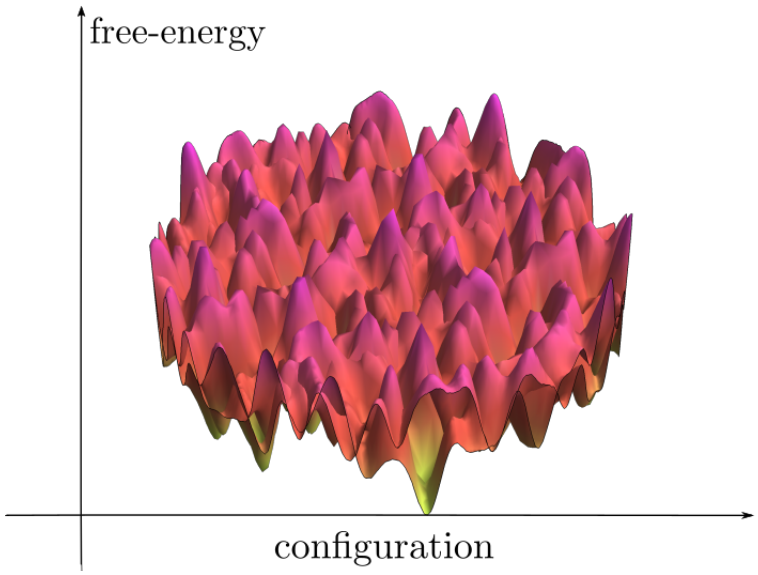


Figure 4.2: Schematic of the free-energy landscape of a system with Replica-symmetry breaking. The landscape has a large number of minimas separated by energy barriers with variable heights. The *configuration* axis represents a generic configuration, which itself is a higher-dimensional object.

It is important to recognize that the structure and geometry of the free-energy landscape depends on a particular realization of the disorder. Hence, if we choose a different realization of the disorder, by sampling a different $\vec{\xi}_\mu$ from $P(\vec{\xi}_\mu)$, the positions of the minima will change and so will the configurations $\vec{X}$ that corresponding to these minima. Fig 4.2 is then supposed to give a very general picture of the energy-landscape for a system that undergoes Replica-symmetry breaking. The existence of such an energy landscape is one of the most striking consequences of the possibility of Replica-symmetry broken solution. It will be found that the perceptron CSP also undergoes a RSB transition and hence the landscape picture developed here will prove to be very useful. The discovery of the RSB phase is fundamental for understanding the intricacies of various other disordered systems such as spin-glasses and structural glasses apart from their direct applications in the study of CSPs. These applications and other more specialized discussions about the RSB phase can be found in [100–104]

4.2 PHASE DIAGRAM OF THE PERCEPTRON CSP

Compared to the standard perceptron specification, the model we consider is slightly modified: apart from the spherical constraint on the X_i, we will also impose that the sum of the components of each of the vectors be positive: $\frac{1}{N}\sum_i X_i = m$, i.e., that the

magnetization is positive. In the economic ABM presented in Ch. 5, prices of goods will be represented by the vector $\vec{X}$. Hence we impose constant magnetization constraint here to model the positivity constraint on prices in our economic model. The Hamiltonian is hence modified to:

$$H_1(\vec{X}) = \frac{1}{2}\sum_{\mu=1}^{M} h_\mu^2 \theta(-h_\mu) - \Upsilon \sum_{i=1}^{N} X_i, \tag{4.11}$$

where Υ acts as a magnetic field. The field Υ and the magnetization m are related by a simple Legendre transform. One can then solve the model in the thermodynamic limit — $M, N \to \infty$ with $\alpha = \frac{M}{N}$ fixed — and retrieve the phase diagram similar to the standard perceptron. The addition of a magnetic field does not change the qualitative features of the phase diagram, which is illustrated in Fig. 4.3 for $m = 0.577$ and can be compared with the one for $m = 0$ reported in Ref. [75]. The SAT-UNSAT transition line separates the SAT and UNSAT regions, and each of these regions is separated in a replica symmetric (RS) region and a replica symmetry broken (RSB) one. The complete details of the computation are provided in Appendix A. Four phases are possible:

1. In the RS-UNSAT phase, instances are typically not satisfiable, and the Hamiltonian has a unique minimum at positive energy.

2. In the RSB-UNSAT phase, instances are again unsatisfiable, but the energy landscape is now rough, displaying many local minima, all at positive energy.

3. In the RS-SAT phase, instances are satisfiable, and the space of solutions is simply connected (i.e. the energy landscape has a flat connected bottom at zero energy).

4. In the RSB-SAT phase, instances are satisfiable, but the space of solutions is disconnected in many "clusters" of solutions.

The de Almeida-Thouless (AT) line [105] $\sigma_{\rm AT}(\alpha)$ which determines the RS-RSB boundary is independent of α (i.e., it is vertical in Fig. 4.3) in the UNSAT phase, and the corresponding value of $\sigma_{\rm AT}$ is the solution of

$$1 - m^2 = \frac{\int_{-\infty}^{\sigma_{\rm AT}} (h - \sigma_{\rm AT})^2 e^{-h^2/2} \mathrm{d}h}{\int_{-\infty}^{\sigma_{\rm AT}} e^{-h^2/2} \mathrm{d}h}. \tag{4.12}$$

Note that $\sigma_{\rm AT} = 0$ for $m = 0$ but $\sigma_{\rm AT} < 0$ for $m > 0$. The RSB-UNSAT phase corresponds to $\sigma < \sigma_{\rm AT}$.

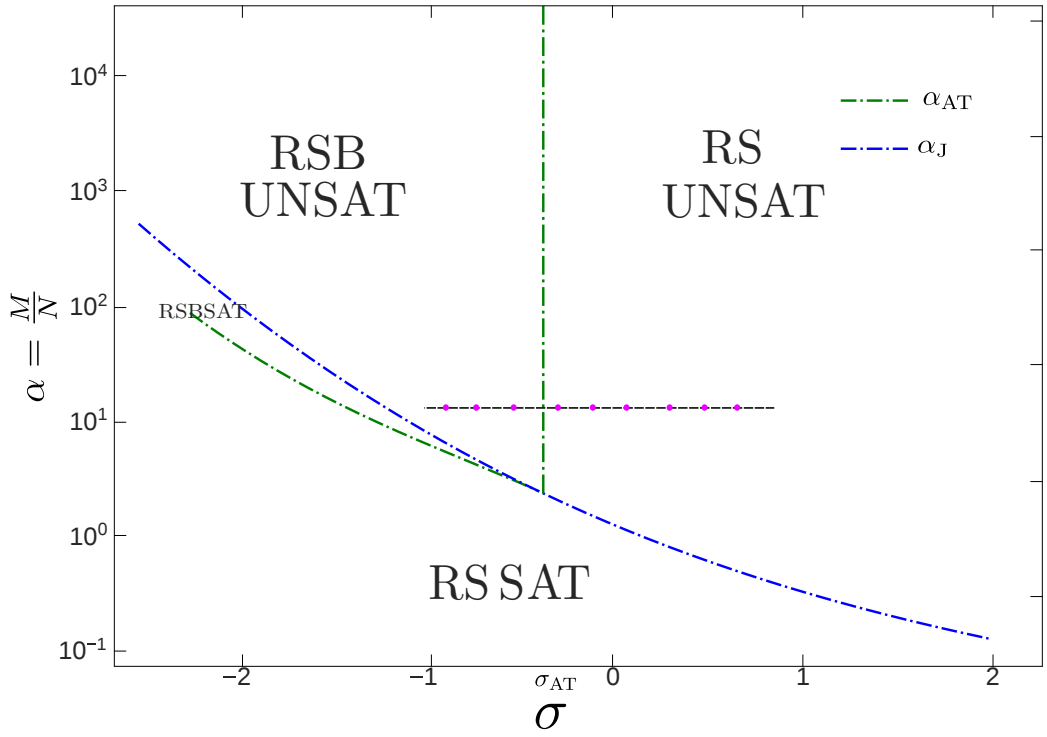


Figure 4.3: Thermodynamic zero-temperature phase diagram for the perceptron problem with $m = 0.577$. α_J is the SAT-UNSAT transition line, α_{AT} is the de Almeida-Thouless line. Note that the AT line in the UNSAT phase is shifted with respect to $m = 0$, and is located at a value of $\sigma_{AT} < 0$. There is also a region of the phase diagram at lower values of σ (outside the range of the figure) where the dAT line coincides with a continuous transition from a replica symmetric SAT phase to a stable 1RSB SAT phase. At even lower values of σ a RFOT type phenomenology is observed, with a dynamical and a Kauzmann transition to a 1RSB SAT phase [75]. Note that the SAT-UNSAT threshold is computed within the replica symmetric approximation, and hence is exact only for $\sigma > \sigma_{AT}$. The dotted black line corresponds to a cut at constant $\alpha = 20$: keeping α fixed we perform R&R dynamics at various values of σ.

4.3 PLANTING AND SELF-PLANTING

Till now our study of constraint satisfaction problems has relied on the examination of the properties of ensembles of random instances. These random instances can be created by choosing the couplings J_{ij} (for K-SAT) or the patterns $\vec{\xi}_\mu$ for the perceptron CSP. Another way of creating a random instance is to first assign a configuration to

all the variables, and then choose only those constraints that are compatible with the chosen configuration. This method by definition creates a satisfiable instance of the problem and goes by the name of *planting*. For the perceptron problem, this consists in choosing a random state $\vec{X}_{\text{plant}}$ and then sampling patterns $\vec{\xi}_\mu$ such that the constraint $h_\mu(\vec{X}_{\text{plant}}) > 0$ is satisfied. One can for instance extract $\vec{\xi}_\mu$ from the normal distribution $\mathcal{N}(0,1)$ and, if the constraint is violated, reject it and extract a new one, until the constraint is satisfied. Because the probability of satisfying a constraint is finite and equal to $\text{erfc}(\sigma/\sqrt{2})/2$, the process converges after a few steps.

The planting procedure for the perceptron also coincides with the "teacher-student" learning scenario. In this case, we first define a teacher perceptron who chooses the configuration $\vec{T} \in \mathbb{R}^N$ and generates output $y_\mu = \text{sgn}(\vec{T} \cdot \vec{\xi}_\mu)$ with the patterns $\vec{\xi}_\mu$ are sampled randomly. Another perceptron — the student — is provided the patterns $\vec{\xi}_\mu$ and the outputs y_μ. The task of the student perceptron is then to recover the weights $\vec{T}$. Formulated as a CSP (with $\sigma = 0$) with patterns $\Xi_\mu = y_\mu \vec{\xi}_\mu$ admits, by construction, the teacher configuration $\vec{T}$ as a solution. The patterns $\vec{\xi}_\mu$ then play the role of planted disorder from the point of view of the student perceptron.

One interesting question is whether the ensemble of such random planted problems is distinguishable from the completely random ensemble in which the $\vec{\xi}_\mu$ are i.i.d. normal variables[1]. The planted ensemble and the purely random ensemble have very different properties: since the disorder $\vec{\xi}_\mu$ depends on the planted configuration $\vec{X}^{\text{plant}}$, the disorder becomes correlated. [107]. Remarkably however it has been shown that in certain regions of parameter space, typical realizations of the disorder in the two ensembles have statistically identical properties — the planted ensemble coincides with the random one. This construction, named "quiet planting" [76, 107–109], is particularly useful as an algorithmic benchmark since we can then construct instances of constraint satisfaction problems whose solution is known while remaining representative of the random ensemble.

SELF-PLANTING: Within the RSB phase, the free-energy landscape is fixed by the choice of the disorder $\vec{\xi}_\mu$. In what follows, we extend our inquiry to dynamically changing landscapes. We ask how the structure of the minima of the perceptron CSP evolve as the energy landscape itself is modified by an external process. This leads to dynamic, non-quenched disorder. In this case the ξ_μ^i are dynamically updated, with

1 These two ensembles occur as specific cases of a more general teacher-student scenario in which the teacher (or reference) perceptron $\vec{T}$ is itself noisy [106]: tuning the level of noise of the teacher perceptron allows one to interpolate between the purely deterministic case, where $\vec{T}$ is exactly known, and the case of infinite noise, which corresponds to the storage capacity problem.

a rule that seeks to improve violated constraints. We call the procedure "Remove & Replace" (R&R in the following), as a generalization of the planting procedure — see section 4.4 for a more precise statement.

We detail how the R&R dynamical rule allows the system to reach a SAT configuration through the co-evolution of both $\vec{X}$ and the $\vec{\xi}_\mu$, only when $\sigma < \sigma_{\mathrm{RR}}$, where σ_{RR} is an algorithmic threshold. This scenario is similar to planting, but in our case, the planted solution $\vec{X}$ is not chosen *a priori* but is the result of our co-evolution rule — hence the name "self-planting". We also discuss how the landscape evolves as we cross the σ_{RR} threshold and compare the situation with that of the static case, i.e. quenched disorder.

4.4 "REMOVE AND REPLACE" DYNAMICS

The model is fully specified by the choice of parameters m, σ and α and the value of N. The dynamical process that we propose is then the following:

1. Start with randomly chosen vectors, $\vec{\xi}_\mu$ from the normal distribution $\mathcal{N}(0,1)$, and $\vec{X}$ uniformly on the sphere $|\vec{X}|^2 = N$.

2. Find $\vec{X}^{\mathrm{opt}}$ which (locally) minimizes the cost function $\mathcal{L}$, using a gradient descent algorithm, where $\mathcal{L}$ is defined as

$$\mathcal{L} = H_0(\vec{X}) + \frac{\lambda_1}{2}\left(\frac{1}{N}\sum_{i=1}^{N} X_i^2 - 1\right) + \frac{\lambda_2}{2}\left(\frac{1}{N}\sum_{i=1}^{N} X_i - m\right)^2. \tag{4.13}$$

   Here λ_1 and λ_2 are Lagrange multipliers which fix the spherical and linear constraints that $\vec{X}$ must satisfy.

3. **Remove** those vectors $\vec{\xi}_\mu$ such that $h_\mu(\vec{X}^{\mathrm{opt}}) < 0$. These are the constraints that are unsatisfied.

4. **Replace** the removed vectors $\vec{\xi}_\mu$ with new random vectors sampled from the normal distribution $\mathcal{N}(0,1)$.

5. Repeat steps 2-4. Stop if all the constraints are satisfied at step 3.

Because this dynamical scheme involves the removal and replacement of the vectors $\vec{\xi}_\mu$, we call this procedure Remove & Replace. Albeit quite simple, this dynamical scheme uncovers a rich picture which we elucidate below. Note that if $\vec{X}$ were not updated in step 2, this procedure would coincide with the standard planting procedure — so in that sense R&R should rather be called "Optimize, Remove & Replace".

In the following, we will simulate the R&R dynamics and focus on the following dynamical observables, which are all functions of the unsatisfied constraints or the "negative" gaps. These are the (intensive) energy,

$$e(t) = \frac{H_0(\vec{X}(t))}{N} = \frac{1}{2N}\sum_{\mu=1}^{M} h_\mu(t)^2\theta(-h_\mu(t))\,, \tag{4.14}$$

the average gap,

$$h(t) = -\frac{1}{M}\sum_{\mu=1}^{M} h_\mu(t)\theta(-h_\mu(t))\,, \tag{4.15}$$

and the average number of "contacts",

$$z(t) = \frac{1}{M}\sum_{\mu=1}^{M} \theta(-h_\mu(t))\,. \tag{4.16}$$

Note that we define time t as the number of the R&R steps that have been performed.

In this section, we present numerical simulation results for the R&R dynamics. The control parameters are α, σ, m, N and unless otherwise noted, we fix $\alpha = 20$ and $m = 0.577$. Note that if we start at low enough σ, all the constraints are satisfied with very high probability, going to one exponentially in N already in the initial state, and the R&R dynamics stops before any removal is made. Hence, we only consider large values of σ belonging to the thermodynamically UNSAT phase in the phase diagram of Fig. 4.3, in which instances are guaranteed to be UNSAT with probability one for $N \to \infty$.

4.4.1 *Long-time behavior of the energy*

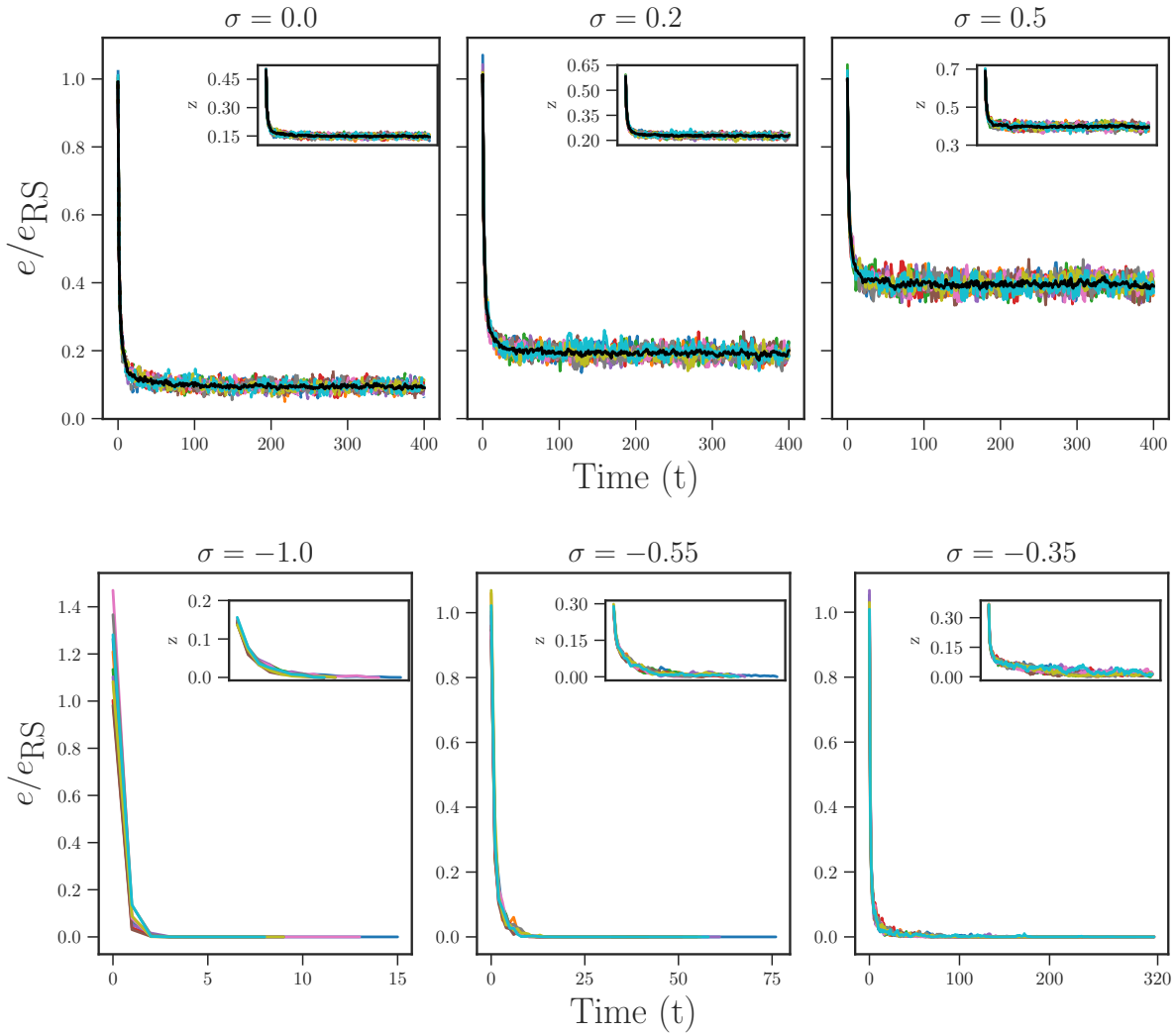


Figure 4.4: Variation of energy e as a function of time t in R&R dynamics, for $\alpha = 20$, $m = 0.577$, $N = 200$, and several values of σ. Individual curves for 10 runs are shown, together with their average (full black line). The inset shows the evolution of the fraction of unsatisfied constraints, z, as a function of time t. Top panel: R&R-UNSAT phase. Bottom panel: R&R-SAT phase.

To begin, we discuss the long-time behavior of the system under R&R dynamics. Fig. 4.4 shows the energy $e = H/N$ as a function of time for several values of σ. We rescale the energy by the thermodynamic value $e_{\mathrm{RS}} > 0$, which corresponds to a random choice

of the $\vec{\xi}_\mu$. In the RS phase this corresponds to the value of the energy at time $t = 0$. In the RSB phase, the value of $e_{\rm RS}$ is not accurate, but provides an approximation to the thermodynamic energy. Note that in the initial RS state, there is a unique configuration $\vec{X}^{\rm opt}$, which minimizes the energy for a given set of vectors $\vec{\xi}_\mu$, while in the RSB phase there are many local minima that can be found via gradient descent. This remains true, at least in the initial stages of the R&R dynamics, in which the energy landscape has not changed with respect to the one at $t = 0$. The evolution of the energy landscape itself is a subject of a later section below.

As the vectors $\vec{\xi}_\mu$ are removed and replaced, we observe that both $e/e_{\rm RS}$ and the fraction of unsatisfied constraints z are reduced, as expected. Interestingly, for large enough σ, R&R is unable to find a configuration such that the constraints are satisfied. After a brief transient, the energy fluctuates around some plateau value which we note e^∞. This long-time value $e^\infty/e_{\rm RS}$ decreases upon decreasing σ as shown in Fig. 4.4, top panel. For example, when $\sigma = 0$ the R&R algorithm finds configurations that are ten times "better" than random (i.e. $e^\infty/e_{\rm RS} \approx 0.1$). Upon further decreasing σ, we reach a region in which e^∞ is strictly zero — see Fig. 4.4, bottom panel. The presence of a zero-energy final state implies that we have been able to generate a satisfiable instance of the problem starting from an unsatisfiable one. Furthermore, this zero-energy state is often reached fairly quickly as we discuss below, within a few steps of the R&R dynamics, at least when σ is sufficiently small. We denote the final state as $\vec{X}^\infty$ and the corresponding configuration of disorder as $\vec{\xi}^\infty_\mu$. The pair $(\vec{\xi}^\infty_\mu, \vec{X}^\infty)$ is then our *self-planted* state.

4.4.2 *The R&R Transition*

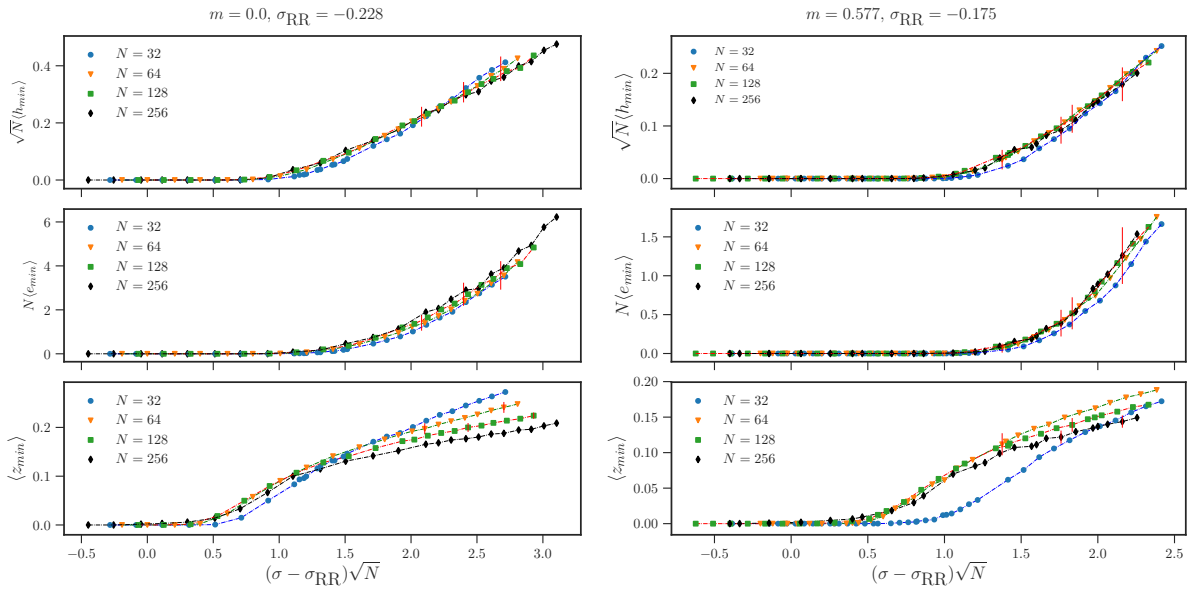
We expect that a sharp *algorithmic* phase transition separates two phases in the limit $N \to \infty$: an R&R-UNSAT phase at large (and positive) σ, in which the R&R dynamics is unable to find a self-planted state, and a R&R-SAT phase at low σ, in which the R&R dynamics converges to a self-planted state in a finite time. The situation for large but finite N can be summed up as follows:

- In the R&R-UNSAT phase, the dynamics typically converges to a stationary state with positive energy for $N \to \infty$. For finite N, fluctuations around the typical state are controlled by a large deviation function, $P(e) \sim \exp[-N\omega(e)]$. The function $\omega(e)$ attains a maximum at the typical value of e, and is finite for $e = 0$. This large-deviation function controls then predicts an exponentially small probability in N to reach zero energy through a fluctuation that produces a rare instance that satisfies all constraints. Because $e = 0$ is an absorbing state of the dynamics, the R&R procedure would stop at that point. We conclude that within this phase

a self-planted state can be found provided the simulation runs for a time $t \sim \exp[N\omega(e=0)]$ i.e. growing exponentially with N.

- On the other hand in the R&R SAT phase, the R&R dynamics typically reaches a self-planted state in a finite time, even when $N \to \infty$. In this case the function $\omega(e)$ is maximal at $e = 0$.

a scenario similar to the one investigated in [110] for the Walk-SAT algorithm for the standard Boolean K-satisfiability problem:



(a) Some label here

(b) Some label here

Figure 4.5: Finite size scaling around the algorithmic transition. For fixed $\alpha = 20$ and $m = 0$ (left) or $m = 0.577$ (right), we plot the minimal energy, gap and fraction of contacts observed during the first $T_{max} = 600$ iterations of R&R, averaged over several realizations. Data for different N are scaled to achieve collapse by a single parameter σ_{RR}. For better visualization, error bars are shown only for a few data points. Apart from the data for $N = 32$, the quality of the rescaling is acceptable.

We then seek to identify the transition by looking at the scaling with N of the time T_{sp} needed to reach a self-planted state; the time should be finite for $\sigma < \sigma_{RR}$ and should diverge exponentially in N for $\sigma > \sigma_{RR}$. This is a scenario similar to the Walk-SAT algorithm for the standard K-SAT problem - if the problem remains UNSAT, we choose a clause at random and then flip the variable (or flip the "spin"). Since K-SAT is defined

over boolean variables, flipping the spin is equivalent to R&R. [110]. Contrary to K-SAT however, the perceptron is a fully-connected model and numerical simulation of the model scales as N^2 compared to N for K-SAT. This in turn limits the study to mall systems, $N \leq 256$.

Nonetheless, a finite-size scaling analysis is possible for small system sizes $N \gtrsim 64$. We consider R&R runs of fixed time $T_{\max} = 600$ and we compute the minimal energy $e_{\min} = \min_{t \leq T_{\max}} e(t)$, the minimal average gap $h_{\min} \min_{t \leq T_{\max}} h(t)$, and the minimal fraction of contacts $z_{\min} = \min_{t \leq T_{\max}} z(t)$. Because $T_{\max}$ is finite, we expect these quantities to remain finite for $N \to \infty$ in the R&R-UNSAT phase, and to vanish in the R&R-SAT phase, except if we are close to the phase boundary. We consider the minimal value since it has smaller fluctuations as compared to the instantaneous value at $T_{\max}$ and also because it is agnostic to the initial transient.

In Fig. 4.5 we present the average of $e_{\min}$, $h_{\min}$ and $z_{\min}$ over samples in a finite size scaling plot. The scaling variable on the horizontal axis is, as in a standard jamming transition, $\sqrt{N}(\sigma - \sigma_{\mathrm{RR}})$ [111]. In the R&R-UNSAT phase, we expect $e_{\min} \propto (\sigma - \sigma_{\mathrm{RR}})^2$, $h_{\min} \propto (\sigma - \sigma_{\mathrm{RR}})$, $z_{\min} \propto (\sigma - \sigma_{\mathrm{RR}})^0$, which fixes the scaling of the vertical axis. All data can be collapsed by a single free parameter, i.e. the value of σ_{RR}, with deviations being observed only for the smallest size, $N = 32$. By this finite size scaling analysis, we can precisely estimate $\sigma_{\mathrm{RR}} = -0.228$ for $m = 0$ and $\sigma_{\mathrm{RR}} = -0.175$ for $m = 0.577$. For other values of m, we estimated $\sigma_{\mathrm{RR}}(m)$ by collapsing data for $N = 64$ and $N = 128$ only, and we report the result in Fig. 4.7.

4.4.3 *Interpreting the algorithmic transition*

We thus find that a sharp algorithmic transition at $\sigma_{\mathrm{RR}}(m)$ separates a R&R-SAT phase from a R&R-UNSAT phase for $N \to \infty$. Our interpretation of this transition is the following. If the value of $\vec{X}^{\mathrm{opt}}(t)$ does not change too much from one R&R step to the next, then R&R becomes essentially equivalent to standard planting and quickly converges to a zero-energy state. In order to confirm this picture, we compute the following overlaps:

$$r(t) := \frac{1}{N} \vec{X}^{\mathrm{opt}}(t+1) \cdot \vec{X}^{\mathrm{opt}}(t) \,, \tag{4.17}$$

$$s(t) := \frac{1}{N} \vec{X}^{\mathrm{opt}}(t) \cdot \vec{X}^{\mathrm{opt}}(0) \,. \tag{4.18}$$

$r(t)$ is the overlap between the optimum $\vec{X}^{\mathrm{opt}}$ at two consecutive time steps of the R&R dynamics and $s(t)$ is the overlap between the $\vec{X}^{\mathrm{opt}}(t)$ and the initial optimal solution $\vec{X}^{\mathrm{opt}}(0)$, corresponding to the initial choice of the disorder $\vec{\xi}_\mu$. The numerical results for $r(t)$ reported in Fig. 4.6 indeed confirm that in the R&R-SAT phase $r(t) \to 1$, indicating

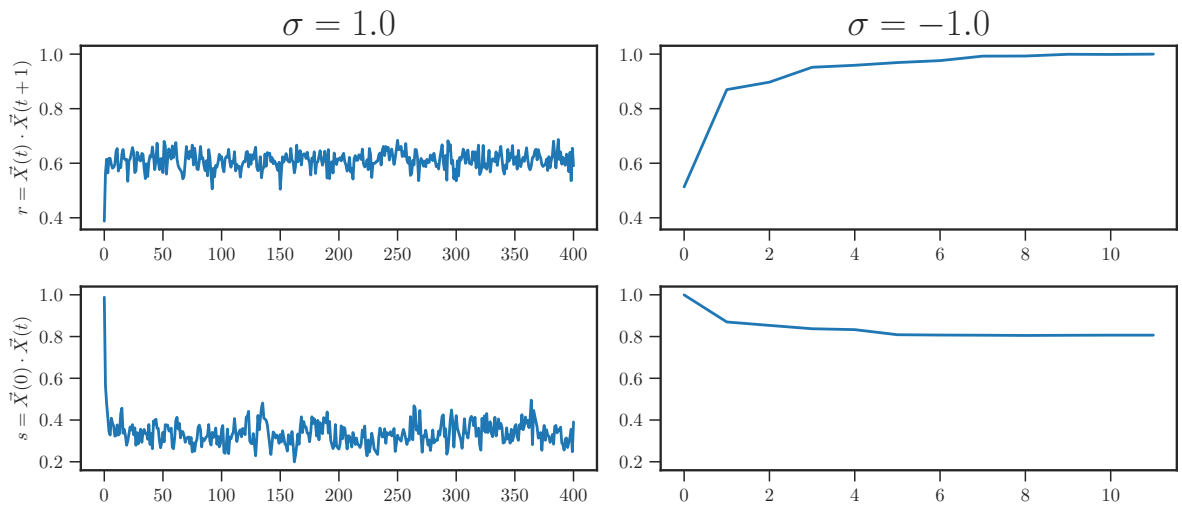


Figure 4.6: Overlaps $r(t)$ and $s(t)$ for $\alpha = 20$, $m = 0.577$, $N = 200$, and two values of $\sigma = 1.0$ and $\sigma = -1.0$ in the R&R-UNSAT and R&R-SAT phases respectively.

that $\vec{X}^{\text{opt}}(t)$ stabilizes to a constant value. On the contrary, in the R&R-UNSAT during the R&R dynamics, at each step we are sampling significantly different realizations of the $\vec{\xi}_\mu$ and therefore new optimal solutions $\vec{X}^{\text{opt}}$. The dynamical process never converges and $\vec{X}^{\text{opt}}$ keeps evolving randomly on the sphere $\mathcal{S}_N$, with $r(t) < 1$.

There are at least two mechanisms that can stabilize $\vec{X}^{\text{opt}}$ and thus ensure that self-planting is achieved. First of all, if m is large, the space of allowed $\vec{X}$ is very small, hence $\vec{X}$ cannot change too much from one step to the next. In the limiting case $m = 1$, $\vec{X} \equiv (1, \cdots, 1)$ is constant, and R&R is equivalent to planting. We expect that this mechanism dominates for m close enough to one, and indeed $\sigma_{\text{RR}}(m) \to \infty$ when $m \to 1$, and R&R (like planting) is achieved at all values of σ.

Another mechanism that can stabilize $\vec{X}^{\text{opt}}$ is the roughness of the energy landscape. Indeed, if the Hamiltonian has a single energy minimum (as in the thermodynamic RS-UNSAT phase), then we expect the location of this minimum to depend sensitively on the disorder. Conversely, if the energy landscape is rough, we expect that a local energy minimum is present close enough to each configuration $\vec{X}$. Therefore, in the RSB phase we would expect R&R to converge to zero energy more easily.

In order to confirm this picture, in Fig. 4.7 we compare the results for $\sigma_{\text{AT}}(m)$ and $\sigma_{\text{RR}}(m)$. At small m, we find that the two values are indeed quite close. R&R converges if started from the thermodynamic RSB phase, except very close to the de Almeida-Thouless transition. At larger m, R&R becomes more and more efficient, and it converges

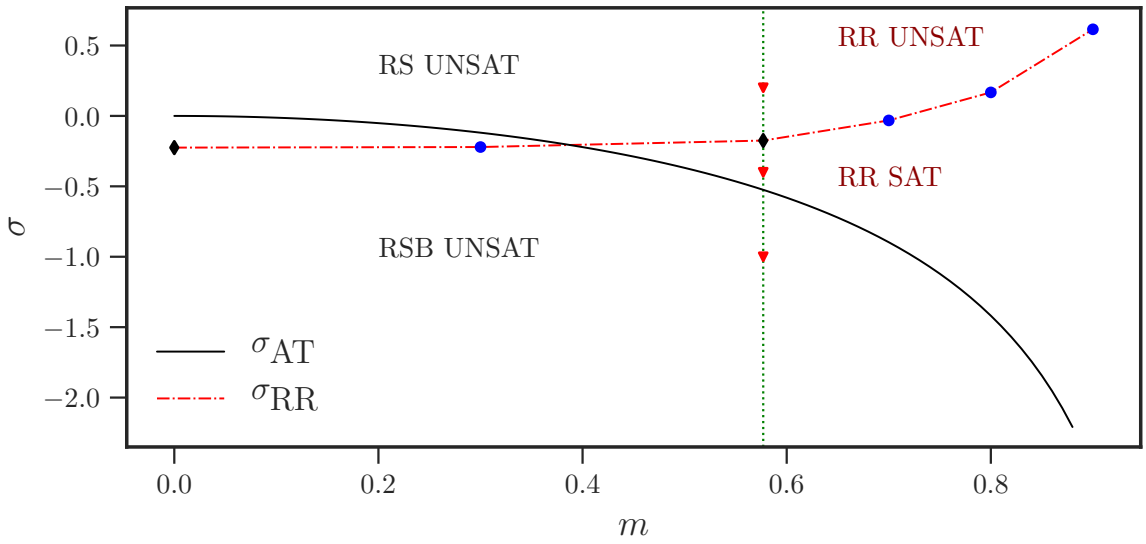


Figure 4.7: The full line is the de Almeida-Thouless transition σ_{AT}, below which replica symmetry is broken in the UNSAT phase (RSB UNSAT), obtained analytically for $N \to \infty$. Dots indicate the R&R algorithmic transition σ_{RR}, below which R&R dynamics converges to a self-planted state (R&R SAT), obtained numerically from the finite size scaling (diamonds, with four values of N as reported in Fig. 4.5; circles, with two values of N, not shown), for $\alpha = 20$ and as a function of m. The vertical dashed line and triangles indicate the points that are further investigated in Fig. 4.8.

even when initialized in the thermodynamic RS phase, presumably because the phase space is restricted by the magnetization constraint.

Finally, one can ask whether the disorder $\vec{\xi}_\mu^\infty$ acquires a special structure in the self-planted state. A way to investigate this is to study the spectrum of the covariance matrix $C_{\mu\nu} = \vec{\xi}_\mu^\infty \vec{\xi}_\nu^\infty$. In the absence of any structure, we expect that the spectrum is of the Marcenko-Pastur type [112], whereas non-trivial structure would lead to some outliers [113]. We find that when $\sigma > 0$, the constraint $\vec{X} \cdot \vec{\xi}_\mu^\infty > \sigma\sqrt{N}$ is strong enough to induce a rank one correlation between the $\vec{\xi}_\mu^\infty$. When $\sigma < 0$, on the other hand, the most probable configurations are such that $\vec{X} \cdot \vec{\xi}_\mu^\infty \approx 0$ and no correlation is generated. In other words, no non-trivial correlations are found beyond those imposed by geometrical constraints.

4.4.4 *Energy landscape dynamics*

We discussed the similarity between the phenomenology of the Walk-SAT algorithm [110] and the R&R dynamics in Sec 4.4.2. A significant way in which the current dynamics differs from the Walk-SAT algorithm is in the evolution of the landscape itself. During R&R dynamics, we observe that the energy landscape itself evolves and in what follows we discuss the structure that emerges from the dynamics. To do so, we use the following procedure. At a given time t, the R&R dynamics produces a pair $\{\vec{\xi}_\mu(t), \vec{X}^{\rm opt}(t)\}$. The disorder $\vec{\xi}_\mu(t)$ encodes the energy landscape as defined by the Hamiltonian in Eq. (4.2), while $\vec{X}^{\rm opt}(t)$ is a *local* minimum of the Hamiltonian, selected by the dynamical history. In order to characterize the energy landscape, for a fixed $\vec{\xi}_\mu(t)$, we construct a series of random configurations $\vec{X}_{\rm new}$ on the sphere $|\vec{X}|^2 = N$, with constant magnetization $m = \sum_i X_i/N$, and with given overlap $q_{\rm ini}$ with $\vec{X}^{\rm opt}(t)$, i.e., such that

$$q_{\rm ini} = \frac{1}{N}\vec{X}^{\rm opt}(t)\cdot\vec{X}_{\rm new}\,. \tag{4.19}$$

Such new configurations are then evolved under standard gradient-descent (again, with fixed disorder). We then measure the overlap of each final configuration $\vec{X}_{\rm fin}$ with the reference state,

$$q_{\rm fin} = \frac{1}{N}\vec{X}^{\rm opt}(t)\cdot\vec{X}_{\rm fin}\,. \tag{4.20}$$

A scatter plot of $q_{\rm fin}$ versus $q_{\rm in}$ (called "overlap plot" in the following) gives an idea of the structure of minima in the landscape.

For a simple RS landscape, there is a unique minimum, and all configurations fall in the unique minimum leading to $q_{\rm fin} = 1$, irrespective of the initial distance $q_{\rm in}$. The overlap plot is therefore a single horizontal line. If, on the contrary, the landscape is rough (as, for example, in a full-RSB phase), then the gradient descent procedure can terminate in one of many possible nearby minima. In this case, we expect that far away initial configurations, with $q_{\rm in} \ll 1$, will fall in far away minima, $q_{\rm fin} < 1$, and we expect a monotonically increasing overlap plot.

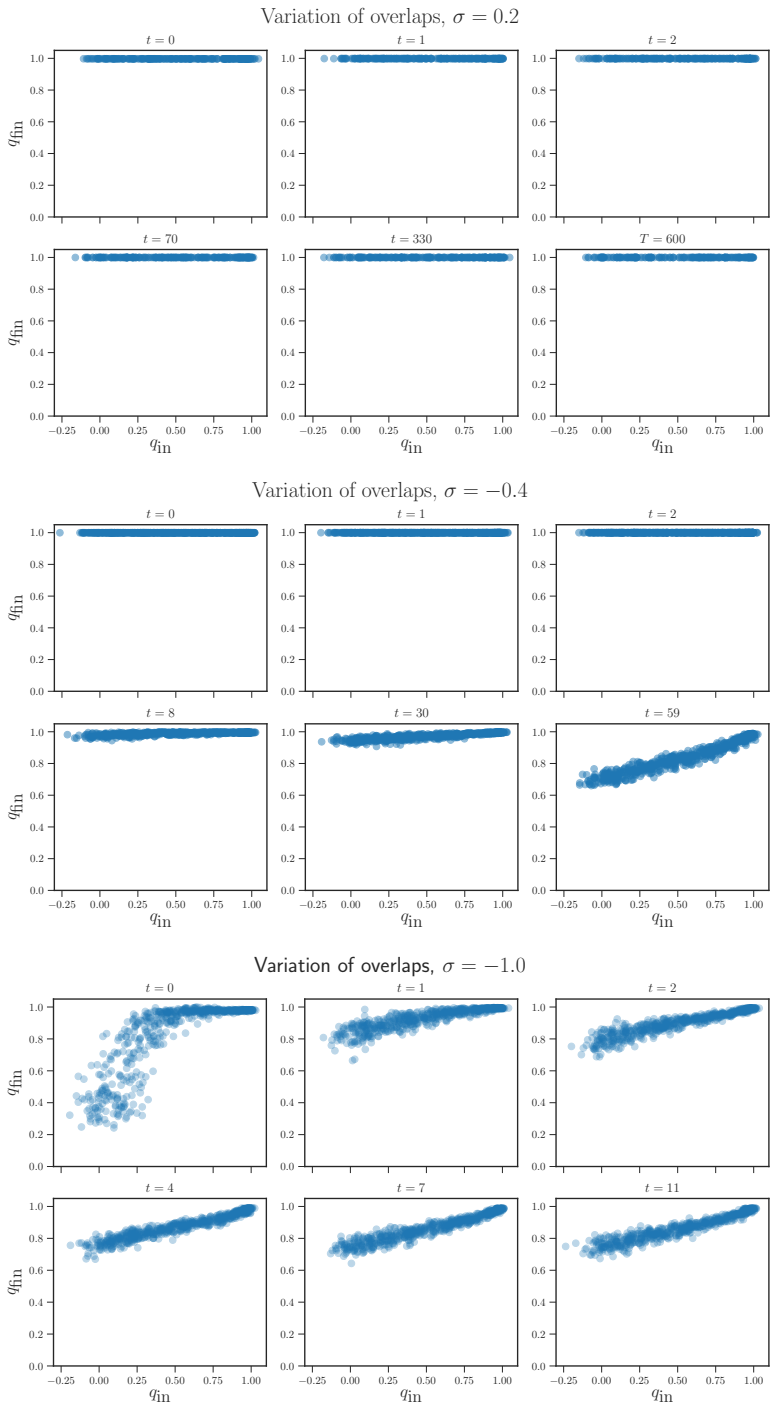


Figure 4.8: Variation of the final overlaps q_{fin} as a function of the initial overlap q_{in}, for $\alpha = 20$, $N = 200$, $m = 0.577$. Overlap plots are shown for three distinct values of σ corresponding to different R&R and thermodynamic phases, as shown in Fig 4.7. $\sigma = 0.2$ is in the R&R-UNSAT (thermodynamically RS) phase whereas $\sigma = -1.0$ is in the R&R-SAT (thermodynamically RSB) phase. $\sigma = -0.4$ is an intermediate case: it is in the R&R-SAT, but thermodynamically RS, phase.

In Fig. 4.8, we show numerical results for the overlap plot, at selected values of σ and t in the different phases, along the vertical dashed line of Fig. 4.7 (red triangles):

- For $\sigma = 0.2$, the initial state ($t = 0$) is in a thermodynamic RS UNSAT phase. Indeed, we observe that the overlap plot is flat with $q_{\text{fin}} = 1$, indicating a trivial landscape with a single energy minimum. Under R&R dynamics, we know that the energy remains finite and self-planting is not achieved. We expect, as discussed above, that this is due to the fact that the landscape remains simple, and its unique minimum changes strongly at each R&R step, as indicated by $r(t) < 1$ (Fig. 4.6). Consistently, we observe that the overlap plot is unchanged at positive times $t > 0$.

- For $\sigma = -0.4$, the initial state is still thermodynamically RS and UNSAT, but now R&R achieves self-planting, and asymptotically the energy goes to zero and $r(t)$ to one. In this case, we observe that the overlap plot is initially flat, but develops structure as time increases, indicating that the self-planting process is accompanied by the development of a rough landscape, at least in the vicinity of the self-planted state.

- Finally, for $\sigma = -1.0$ the initial state is thermodynamically RSB, and the initial overlap plot at $t = 0$ is indeed monotonically increasing, indicating a rough landscape. For subsequent times, we observe that the landscape remains rough. However, for configurations with $q_{\text{in}} \sim 0$, final configurations are closer to the self-planted one. This indicates the presence of many low-energy minima, close to the self-planted one.

4.5 DISCUSSION & CONCLUSION

In this chapter we have discussed the existence of a new sharp algorithmic transition (in the thermodynamic limit) for the R&R dynamics applied to the perceptron model, which separates a R&R-UNSAT phase at large σ (in which the system is unable to produce a SAT, or self-planted configuration on finite time-scales) from a R&R-SAT phase at low σ (in which a zero-energy configuration is found in a finite time). To precisely pin-down the transition, we used a finite-size scaling analysis of the numerical data. We have shown that the transition to the SAT phase is generally associated to the fact that the optimal configuration does not evolve during the R&R dynamics, so that it reduces to simple planting. This can be driven by (at least) two different mechanisms: in the limit of large m the space of allowed configurations is very small, and $\vec{X}^{\text{opt}}$ cannot change for trivial geometrical reasons. On the other hand at small m the satisfiability of the R&R dynamics can be attributed to the roughness of the energy landscape. This rough landscape then produces a large number of local minima and

quasi-optimal configurations close to any initial condition. And even if the disorder changes significantly in a single R&R step (by changing the $\vec{\xi}_\mu$), a new minimum close to the old one can be found. More generally, in this chapter we have sought to answer the question: are there situations for which the RSB structure (the rough energy landscape) persists when the disorder itself evolves. The answer is in the affirmative: the R&R procedure proposed here is one such example of a simple dynamics in which the roughness of the landscape is preserved under evolving disorder, leading to the self-planting transition.

Self-planting is thus accompanied by a modification of the energy landscape. An exact analytical treatment of the phenomenon is difficult since the R&R dynamics couples the gradient descent (or zero-temperature) dynamics of $\vec{X}$ under fixed disorder with an evolution of the disorder itself.

It would be very interesting to see whether the self-planting phenomenon also occurs in other models characterized by a large connectivity of the constraints, such as the Hopfield model [114] or poly-disperse soft spheres in high dimensions [115]. An equivalent to R&R dynamics in the latter case would then be the removal and replacement of overlapping spheres after performing a gradient descent starting from a random initial condition. Finally, we can use the budgetary constraint encoded by the perceptron as the building block for a CSP based agent-based model. This is discussed in Chap 5 where we have found how the macroeconomic state of the economy changes dramatically as we tune σ.

5

A CSP BASED AGENT-BASED MODEL

5.1 OVERVIEW

In the previous chapter, we learned that general constraint satisfaction problems undergo phase transitions from a SAT phase (where all constraints can be satisfied) to an UNSAT phase (where at least one constraint remains unsatisfied). The transitions between these two regimes become sharp in the "thermodynamic" limit — $M, N \to \infty$ with the ratio $\alpha = \frac{M}{N}$ fixed. These transitions are also accompanied with modifications in the landscape of solutions: solutions become rarer and become clustered within disconnected "regions" of variable space [72, 116]. As M further increases, these clusters may disappear, forcing the system to re-adapt to a completely different configuration.

Economic problems too involve several types of constraints, such as budget constraints, leverage constraints, temporal constraints, and non-substitutability effects. It is our belief that CSP paradigm can also lead to important modeling insights in economic situations as well. As argued in the introduction, agent-based models provide a perfect "sandbox" for testing exactly such an idea — we can build a toy economy and introduce such constraints which would in turn determine how agents would behave. The ABM approach that we propose requires careful study of the aggregation of the behavior of individual agents to observe the macroscopic (or macroeconomic more precisely) phenomena that emerges. Hence, ABMs also provide a bridge between the CSP paradigm and macroeconomic modeling.

In this chapter, we propose an agent-based model in a simple setting — a market of goods with producers and consumers. These economic agents then have to satisfy budget and production constraints, and they adjust their strategies optimally all the while satisfying these constraints. Formally, our model is identical to the so-called "perceptron model", another classical and well studied CSP which has a rich phenomenology and in particular it can show multiple equilibria in certain regions of parameter space [1, 75]. A complete and more technical discussion of the full "perceptron model" is presented in Chapters 3 and 4 and we restrict ourselves in this chapter to an examination of

the economic implications of a simpler version of the model. We consider a regime where there is always a unique solution (a single equilibrium) to the perceptron problem. Nevertheless, we discover many interesting phenomena within our CSP-inspired macroeconomic ABM. We observe the model has three distinct regimes — the economy transitions from one devoid of any structure to one where we observe a spontaneous speciation of goods accompanied with endogenous business cycles.

5.2 THE ECONOMY AS A CONSTRAINT SATISFACTION PROBLEM

In our model, we consider M agents and N products with prices $p_1 \dots p_N$. Each agent, labeled by $\mu = 1 \cdots M$, wants to buy or sell a certain quantity ξ_μ^i of the product i (positive for selling the product and negative for buying it). We normalize the prices such that at all times $\frac{1}{N}\sum_i p_i = 1$ and that all prices are bounded from below $p_i \geq x_m$, where x_m is a non negative real number, possibly zero. The evolution of prices is given by a "market" that attempts to enforce simple budgetary constraints for all agents, while the evolution of preferences is decided by the agents according to a simple behavioral rule, as we now describe.

5.2.1 *Budget constraint and formation of prices*

The quantity $\pi_\mu = \sum_i \xi_\mu^i p_i$ is the total money that the agent μ is willing to spend (or earn) in the market in a given round. This quantity is subject to a budget constraint, namely that $\pi_\mu \geq \sigma$. Here, $\sigma < 0$ if the agent is allowed to borrow to cover losses and $\sigma > 0$ if the agent is required to make a profit. The products prices p_i thus depend dynamically on agents' preferences. Given the matrix of preferences $\vec{\xi}_\mu$, we assume that prices are determined by the market, in such a way that the least number of agents have unsatisfied budgetary constraints. We define a variable called the "gap" as follows:

$$h_\mu(\vec{p}) := \vec{\xi}_\mu \cdot \vec{p} - \sigma > 0 \qquad \forall \mu \in \{1 \dots M\}\,. \tag{5.1}$$

The gap variable hence encodes the distance from the configuration where the budgetary constraint for agent μ is on the verge of being unsatisfied ($h_\mu = 0$). The price vector $\vec{p}$ is then determined by minimizing the number of agents whose constraints are unsatisfied. This can be achieved by minimizing the following cost function:

$$\mathrm{H}(\vec{p}) = \frac{1}{2}\sum_{\mu=1}^{M} h_\mu^2 \Theta(-h_\mu), \tag{5.2}$$

under the constraints that $\frac{1}{N}\sum_i p_i = 1$ and $p_i \geq x_m$, where x_m is a small positive number thus ensuring that all prices are positive. The Heaviside Θ function is equal to 1 when

its argument is positive, and zero otherwise. This function (or Hamiltonian) takes its minimal value (i.e. zero) when all agents are satisfied and is positive if at least one agent is unsatisfied. In what follows, we use the formulation "unsatisfied agent" to mean that the agent's budgetary constraint is presently unsatisfied.

Note that the constraint on prices $\frac{1}{N}\sum_i p_i = 1$ sets the price units. Since it is a linear constraint, the above optimisation problem is convex and the solution space remains connected.

5.2.2 *Preferences update: supply and demand*

Agents' preferences $\vec{\xi}_\mu$ are allowed to evolve in reaction to supply/demand imbalances and prices, according to a simple behavioral rule similar to that used in the "MarkI" and "Mark0" models [44, 55]. Contrary to the infinite horizon, profit maximizing framework which is *de rigueur* in standard microeconomic models, we posit reasonable, heuristic rules to model the behavior of agents. These rules take the following form:

- Supply side. If the agent is a supplier for product i ($\xi^i_\mu > 0$), then it adapts as a function of the mismatch between the supply S_i and demand D_i of product i, defined as:

$$S_i = \sum_\mu \xi^i_\mu \Theta(\xi^i_\mu), \qquad D_i = \sum_\mu \xi^i_\mu \Theta(-\xi^i_\mu). \tag{5.3}$$

  The supplier updates their preference as:

$$\begin{aligned} S_i > D_i &\implies \xi^i_\mu(t+1) = \xi^i_\mu(t)(1-\epsilon_D \mathrm{u}), \\ S_i < D_i &\implies \xi^i_\mu(t+1) = \xi^i_\mu(t)(1+\epsilon_D \mathrm{u}), \end{aligned} \tag{5.4}$$

  where u is a random number sampled independently from the uniform distribution on $[0,1]$, and ϵ_D is the speed of adjustment to supply-demand pressure.

- Demand side. If the agent is a buyer for product i ($\xi^i_\mu < 0$), it adapts its preferences looking at the relative price level for the product as follows:

$$\begin{aligned} p_i > 1 &\implies \xi^i_\mu(t+1) = \xi^i_\mu(t)(1-\epsilon_P \mathrm{u}), \\ p_i < 1 &\implies \xi^i_\mu(t+1) = \xi^i_\mu(t)(1+\epsilon_P \mathrm{u}), \end{aligned} \tag{5.5}$$

  where once again u is sampled independently from the uniform distribution and ϵ_P denotes the speed of adjustment to price pressure.

5.2.3 *Transactions, production costs and redistribution*

Agents then perform transactions, which determines the actual amounts of products sold and bought. The update rules depend on the ratio of supply and demand for product i, noted $\zeta_i = \frac{S_i}{D_i}$. If $\zeta_i > 1$, then the suppliers in the market are unable to dispose off all their inventory and sell only a fraction of it. Conversely if $\zeta_i < 1$, then the buyers are unable to satisfy their demands for product i. These transactions are encoded in the matrix of realized supply/demand $\bar{\xi}^i_\mu$, which takes the following form:

$$\begin{aligned} \zeta_i > 1, \xi^i_\mu > 0 &\implies \bar{\xi}^i_\mu = \frac{\xi^i_\mu}{\zeta_i}, \\ \zeta_i < 1, \xi^i_\mu < 0 &\implies \bar{\xi}^i_\mu = \zeta_i \xi^i_\mu. \end{aligned} \tag{5.6}$$

The matrix $\bar{\xi}^i_\mu$ hence enters the computation of the true money exchanged by the agents.

We posit that each product has a production cost γ_i associated with it, paid by all the suppliers of product i and sampled from a uniform distribution $[0, \gamma]$. The production cost of all goods in the economy is redistributed to the agents as wages. We define the total production cost as $W(t) = \sum_{\mu,i} \gamma_i \xi^\mu_i(t) \Theta(\xi^i_\mu(t))$. This is then uniformly distributed to all agents with each agent getting $w = \frac{W}{M}$ back as "wage".

The "full" profit (or money exchanged) at time t by agent μ thus reads:

$$\bar{\pi}_\mu(t) = \sum_i \bar{\xi}^i_\mu p_i(t) - \sum_i \gamma_i \xi^\mu_i \Theta(\xi^i_\mu) + w(t), \tag{5.7}$$

where the first term corresponds to transactions, the second to production costs, and the third to the wage. Note that because of the "market clearing" condition, $\sum_\mu \bar{\xi}^i_\mu = 0$, we have $\sum_\mu \bar{\pi}_\mu(t) = 0$, so that the total amount of circulating money is conserved. Note also that the production cost and wage terms, on average, compensate each other. This is the reason why we do not take them into account in the price formation process described in section 5.2.1.

5.2.4 *Removal and Replacement of agents*

At the end of these steps, it is possible that there exist agents whose budget constraint is unsatisfied. The economic interpretation of an unsatisfied agent is straightforward: the agent cannot participate in the economy and should in principle go bankrupt and be removed. However, instead of looking at the instantaneous value of the constraint (which could be sensitive to random local fluctuations), we impose that the budget is satisfied on average over some time window.

More precisely, we take an exponentially moving average of the money that agents exchange and compare it to their budget. We define $\pi_\mu^{\text{ema}}(t)$ which is the averaged money exchanged over multiple time steps (over a duration of the order of ω^{-1}):

$$\pi_\mu^{\text{ema}}(t) = \omega \bar{\pi}_\mu(t) + (1-\omega)\pi_\mu^{\text{ema}}(t-1). \tag{5.8}$$

The agent μ is then removed if $\pi_\mu^{\text{ema}}(t) < \sigma$ and replaced by a new agent. The new agents' preferences $\vec{\xi}^\mu$ are sampled independently from the Normal distribution $\mathcal{N}(0,1)$. We also initialize π_μ^{ema} for the new agent with the average $\bar{\pi}_\mu$ of the remaining surviving agents. Note that while the average profit over all agents is zero, as discussed above, the average profit of surviving agents can be non-zero and typically positive, because the distribution of profits of surviving agents is biased towards the positive values. Note that when $\omega = 1$, agents violating their budget constraint are immediately removed.

5.2.5 *Summary of the parameters*

With the dynamical rules above, our model thus has the following parameters:

1. $\alpha = \frac{M}{N}$: the ratio of the number of agents to the number of products
2. x_m: which fixes the lower bound of the prices,
3. σ: which is the budgetary constraint
4. ϵ_D, ϵ_P: susceptibilities of the agents to demand and price pressures
5. γ: parameter of the uniform distribution which fixes the production cost of the products
6. ω^{-1}: timescale over which the agents' average profit is computed.

The model is completely specified once these parameters are provided along with the initial distributions for the price vector $\vec{p}$ and agents' preferences $\vec{\xi}_\mu$.

RELATION TO THE PERCEPTRON MODEL The Hamiltonian in Eq. (5.2) is formally identical to that of the perceptron model, as presented in chapters 3. There are some crucial differences however — for our agent-based model, agents' preferences evolve as price changes and hence we are in the realm of time-dependent (annealed) disorder and not quenched disorder. Agents are also removed from the economy, if they don't satisfy their (time-averaged) budgetary constraints. This echoes the "Remove & Replace" dynamics studied in Chap 4 where vectors $\vec{\xi}_\mu$ were removed as soon as the constraint h_μ

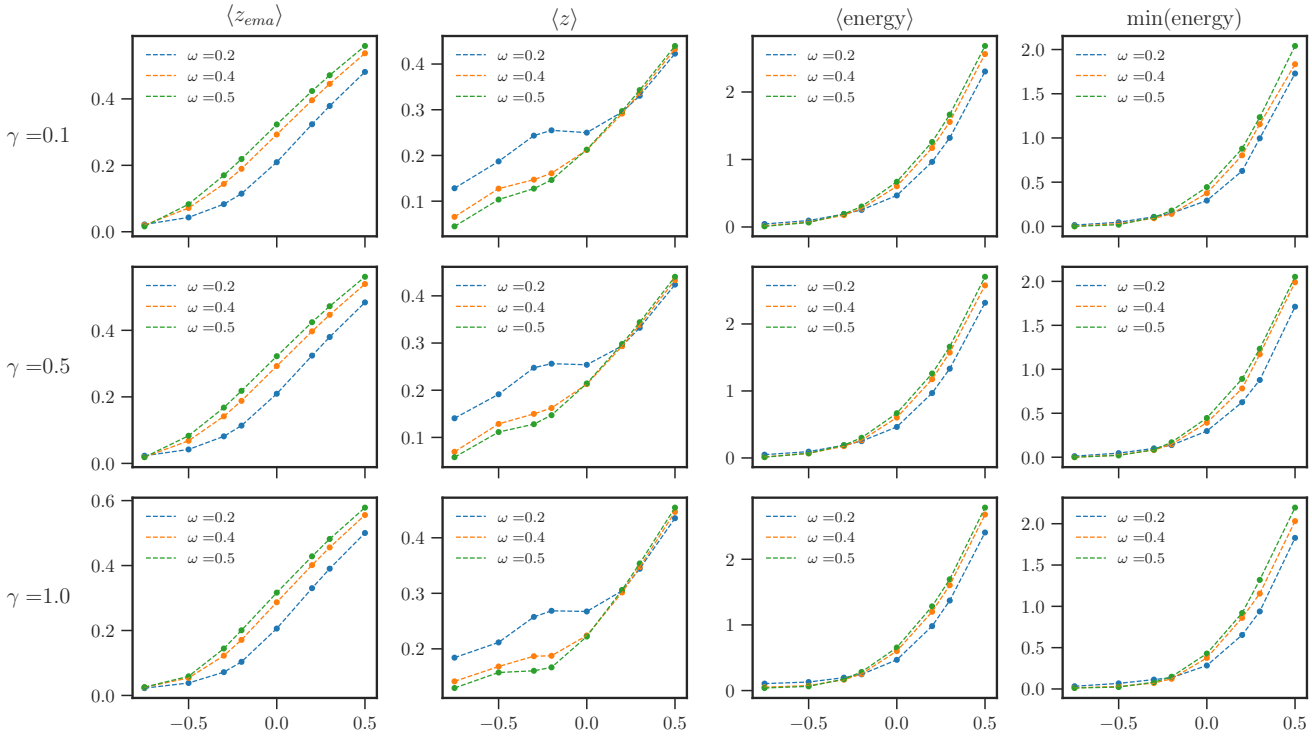
was violated. There we found an algorithmic transition, as a function of the threshold σ, between a phase where the dynamics quickly finds a zero-energy minima — one where all the constraints are satisified or SAT in the language of Ch 3) — and one where the dynamics keeps moving from a UNSAT configuration to another. We would like to achieve a similar scenario within our CSP-based ABM too i.e. find a configuration of prices and agents' preferences such that everyone's budgetary constraints are satisfied. In what follows, we will discover that σ again plays a crucial role in determining the macroscopic state of the system. However a crucial difference between the model here and the one studied in Chap 4 is the presence of a linear and positivity constraint on the price vector (and no spherical constraint). This makes the optimisation problem for finding the minimum of the Hamiltonian convex.

5.3 REDUCING THE SPACE OF PARAMETERS

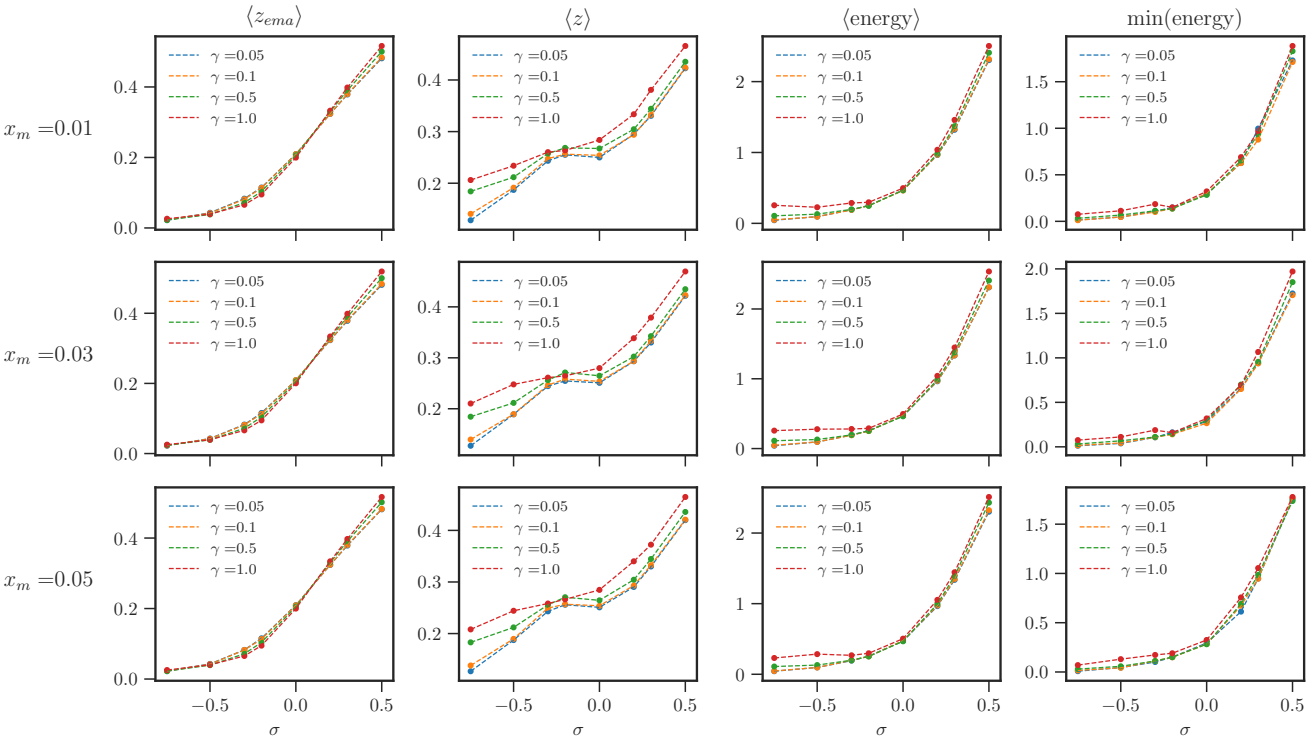
We now turn our attention to the results obtained through numerical simulation of the dynamics of the model. Our model has seven free parameters and as a first step towards understanding its behavior, we must ascertain how the macroscopic state of the economy depends on the choice of these parameters. This follows the methodology proposed in Chap 2 and the objective is to determine the various phases or regimes in the economy as a function of a reduced set of parameters. To this end, we choose dynamical observables all of which depend on the "gap" variable h_μ (Eq. (5.1)) :

1. z_{ema} is defined as the fraction of agents who do not satisfy the time averaged constraint: $\pi_\mu^{\text{ema}}(t) < \sigma$;
2. z is defined as the fraction of agents that do not satisfy the instantaneous constraint, Eq. (5.1), at time t.
3. energy is defined as the value of the cost function $\mathrm{H}[\vec{X}, \vec{\xi}_\mu]$ (5.2).
4. min(energy) is the minimum value of the energy attained over the course of a single run. Using the minimum value removes the dependence on the transient, especially during the start of the simulation.

In what follows and unless otherwise noted, we set $N = 100$ products, $\alpha = 10$ (i.e. $M = 1000$ agents), $\epsilon_P = 0.05$ and $\epsilon_D = 0.05$. Agents' preferences $\vec{\xi}_\mu$ are sampled from a standard normal distribution. Fixing these parameters still leaves us with four parameters. We choose σ to be our control parameter and we examine in turn how the dynamical variables defined above vary within the parameter space of $\{\gamma, x_m, \omega\}$.



(a) Variation of dynamical variables as a function of σ (as control parameter), γ and ω for a fixed value of $x_m = 0.01$.



(b) Variation of dynamical variables as a function of σ (as control parameter), x_m and γ for a fixed value of $\omega = 0.2$.

Figure 5.1: Variation of dynamical variables as a function of the control parameter σ for a fixed value of $x_m = 0.01$ (*Top*) and $\omega = 0.2$ (*Bottom*). *Top*: Each row corresponds to a fixed value of γ and for each value of γ, we vary the production cost parameter ω. *Bottom*: Similarly, each row corresponds to a particular value of x_m and for each value of x_m, we vary γ. We observe that the variation of all the thermodynamic quantities in question shows very little variation as we change x_m and γ.

We begin first by setting $x_m = 0.01$ and then vary $\gamma \in \{0.1, 0.5, 1.0\}$ and $\omega \in \{0.2, 0.4, 0.5\}$. For each value of γ and ω, we vary σ in the range $[-0.75, 0.5]$ i.e. from a region of high debt to a region of low debt. For each choice of these parameters, we present the average value (averaged over 10 independent runs) attained by each of the dynamical variables. We observe that as σ is reduced from larger to smaller values, all the dynamical variables tend to zero on average as shown in Fig. 5.1. In a similar fashion, we also examine how the macroscopic state of the economy varies as we vary x_m. To this end, we vary $x_m \in \{0.01, 0.03, 0.05\}$ for a fixed value of $\omega = 0.2$. Once again, σ is chosen as the control parameter and we vary $\gamma \in \{0.05, 0.1, 0.5, 1.0\}$. As before, we observe in Fig. 5.1 that apart from the dependence on the budgetary constraint σ, the other parameters have only a small effect on the dynamical observables. We thus observe that the qualitative behavior of the model remains unaffected by the other parameters of the model and it is only σ that determines the emergent behavior of the macroeconomy.

5.4 ROLE OF THE DEBT LIMIT: MACRO-LEVEL

Having established that the sole parameter which influences the emergent macroscopic behavior of the economy is the budgetary constraint parameter σ, we proceed with our analysis by fixing $x_m = 0.01$, $\omega = 0.2$ and $\gamma = 0.1$. We also focus on two dynamical variables z_{ema} and z. In Fig 5.2, we plot these two variables along with their minimum and maximum values obtained during the simulations and we observe that as σ is reduced the fraction of unsatisfied agents goes to zero, on average. This points to the interpretation that, as expected, an increased debt level lends flexibility to agents and permits them to live another day. However, beyond a certain negative value of σ (≈ -0.75), we observe that the maximum value attained by z_{ema} during the whole time series actually starts increasing, see Figure 5.2-b.

To better understand why the maximum of z_{ema} increases beyond a certain value of σ, we plot in Figure 5.3 the time series for z_{ema} for three values of σ. Three distinct behaviors are observed: for positive, or mildly negative values of σ, the default rate is large. For $\sigma \sim 0.2$, at every time step close to 30% of the agents are being removed at every step. As σ is reduced, this default rate goes down and for $\sigma = -0.75$, a small fraction of the agents are removed at any time, about 3%. On reducing the value of σ further (~ -1.6), we observe a clear periodicity in the time series of z_{ema}, with the appearance of regular spikes (which persist forever in the simulation). At the peaks, close to 10% of the agents are removed from the economy. These oscillations in the rate of bankruptcies is reminiscent of a similar effect discovered in the Mark-0 model [44, 53, 54] (described in detail here in Chap 6). The mechanism for these oscillations will be described below.

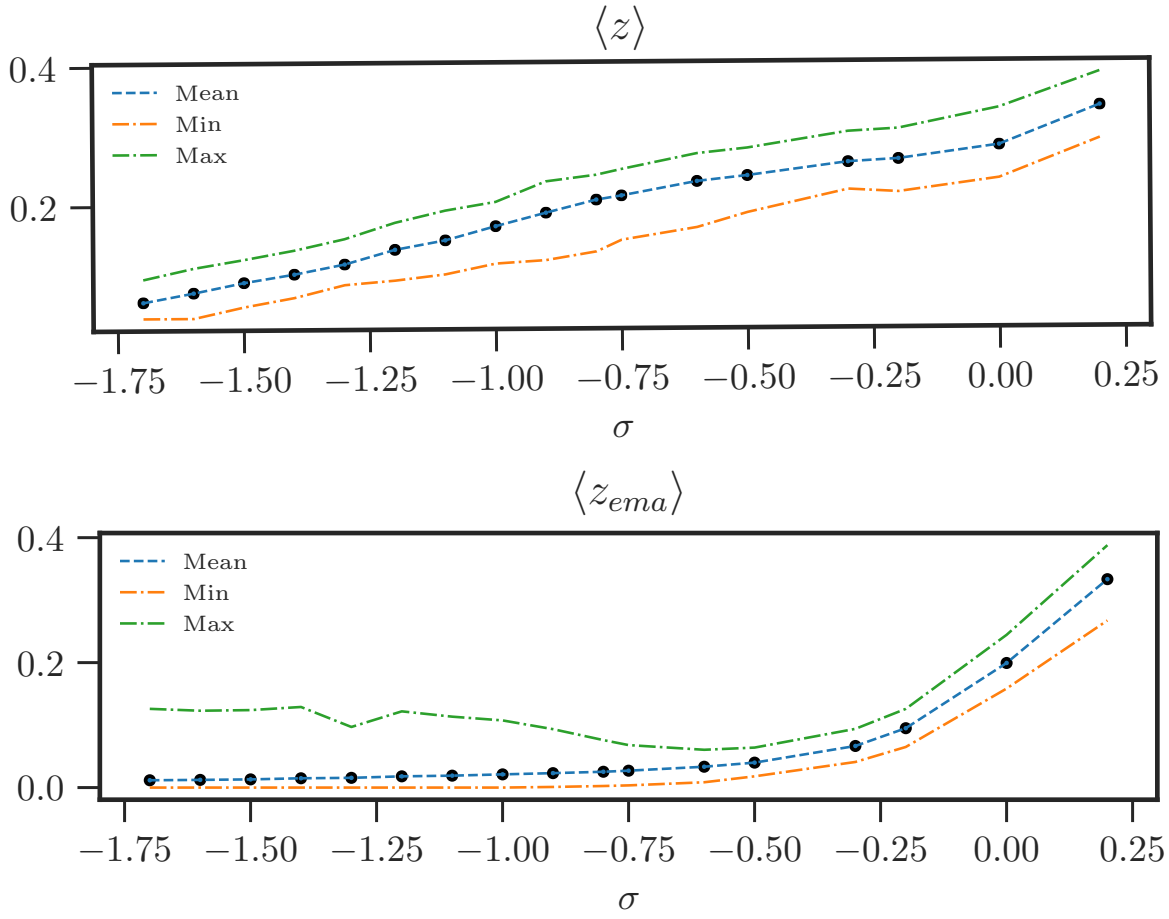


Figure 5.2: Behavior of thermodynamic variables as a function of σ: $\langle z \rangle$ and $\langle z_{\text{ema}} \rangle$. Also represented are the maximum and minimum values attained during the simulation.

This suggests the existence of three regimes:

1. Endogenous Crises (EC): this occurs for small values of σ (high levels of allowed debt), $\sigma \lesssim -0.75$, where periodic spikes of bankruptcies are observed;

2. Stable Phase (S): this occurs for an intermediate range of debt, $-0.75 \lesssim \sigma \lesssim -0.25$, where the economy reaches a stationary state characterized by a few bankruptcies;

3. Unstable phase (U): this occurs for positive values of minimal profits σ, where the economy features a high rate of bankruptcies.

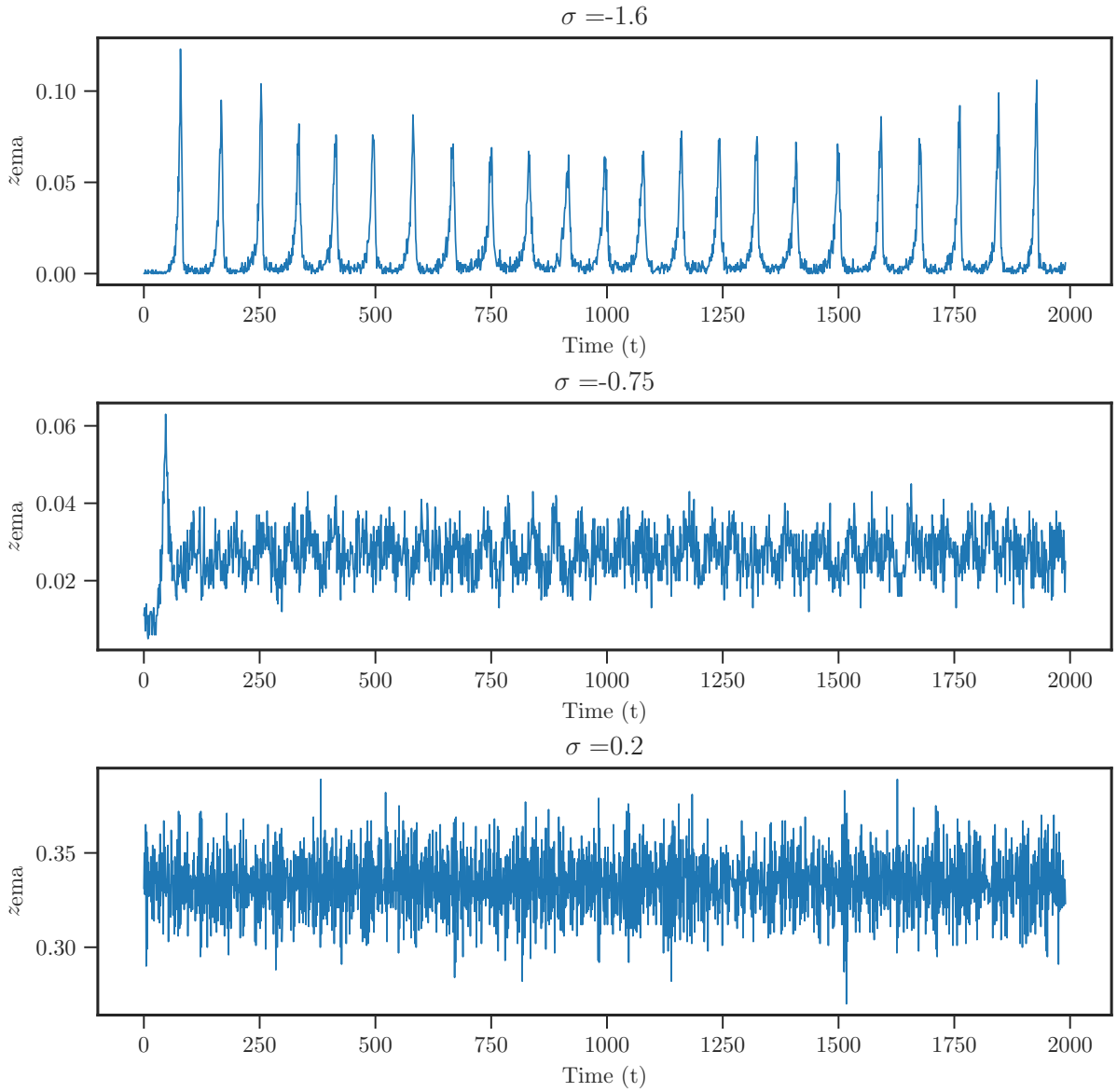


Figure 5.3: Variation of the fraction of unsatisfied agents z_{ema} as a function of time for three values of σ, in the EC (top), S (middle) and U (bottom) regimes. For $\sigma = -1.6$, the agents are removed and replaced periodically whereas for $\sigma = -0.75$, periodicity is lost. In the bottom panel, the $\sigma = 0.2$ case corresponding to high-profit is shown.

The behavior for intermediate (S) and low (EC) levels of debt is reminiscent of the self-planting phenomenon discovered in [1] and discussed at length in Chapter 4. In that case, below some critical value σ_c the dynamics was able to find a configuration where all constraints (agents) were satisfied, while above σ_c the dynamics reached

a steady-state value for the fraction of unsatisfied constraints. Differently from [1], however, in the present context the dynamics is not halted since agents can continue to participate in transactions and update their preferences according to the movement of the prices.

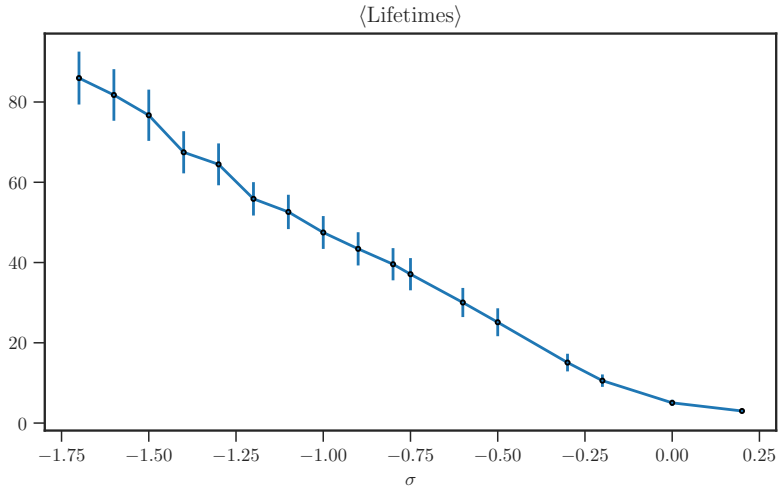


Figure 5.4: Dependence of agent lifetimes as a function of σ.

We also measure the average number of time-steps during which an agent participates in the economy. The results are shown in Figure 5.4, where we show how agent lifetimes change and the average number of times an agent is removed as a function of σ. It is natural to expect that the lifetime of agents is higher when σ is small, since it corresponds to a situation where z_{ema} is low on average. Fewer agents are removed in each time-step in this regime (EC or S) and hence agents participate and live longer. The opposite situation is produced when there is no possibility of debt leading to the frequent removal of a large fraction of agents which in turn leads to smaller agent lifetimes. The U regime is characterized by the fact that agents live for not more than 3-4 time-steps, i.e. the timescale of the exponential moving average $\sim 1/\omega = 5$. On the contrary, the lifetimes of agents in the EC or S regimes are much larger than the averaging timescale. In the EC regime, the agent lifetime is comparable to the period of the "business cycle" (i.e. the period in z_{ema} oscillations), which can be of the order of hundreds of time-steps. If a "time-step" is 3 months, this corresponds to an economic activity lifetime of 25 years.

Having considered how macro indicators behave as a function of σ, we now study how σ influences levels of supply and demand. The EC and S regimes are characterized by agents persisting for long times with only a few agents being replaced at any given time step (except during crises in the EC regime). Longer living agents influence the

distributions of the goods being exchanged in the economy. We show how the demand distributions vary as we tune σ in Figure 5.5. We observe that for intermediate values of $\sigma = -0.75$ in the S regime, the distributions is bimodal. The bi-modality persists for a lower value of σ in the EC regime, with the second peak presenting a heavy right tail. For positive σ in the U regime, on the other hand, the distribution remains Gaussian, since agents' preferences are initially sampled from a normal distribution, and the selection process that biases these preferences does not have time to operate before agents are replaced.

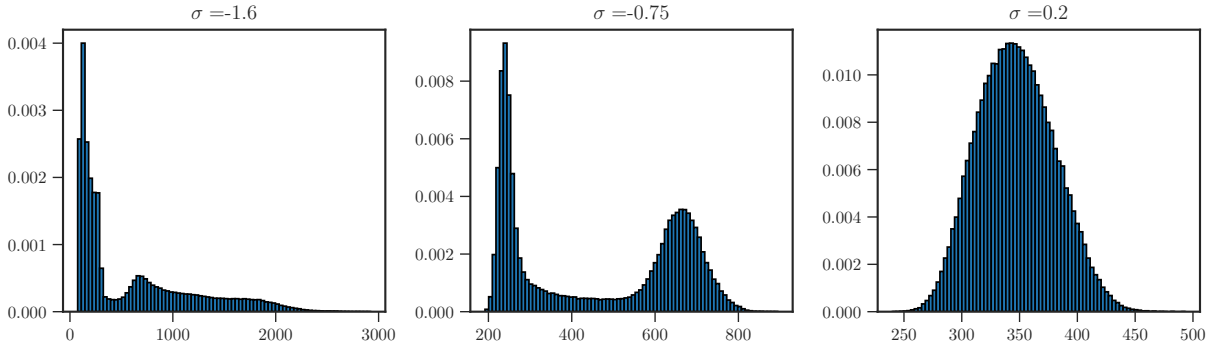


Figure 5.5: The demand distributions for different values of σ. We note that in the S regime (middle), the distribution is bimodal, while in the U regime (right) the distributions remain quasi-Gaussian. In the EC regime (left), the distribution has a long tail. These distributions are obtained by taking the distribution of the complete demand time series over all goods.

The rules governing agents' behavior also couple the level of supply and demand to the price of goods. In general, we expect that cheaper goods will be in higher demand (and hence supply) and vice-versa for more expensive goods. We test this hypothesis in Figure 5.6, where we show scatter plots of supplies and demands versus prices. We find that, as expected, both demand and supply of goods are anti-correlated with their price in the EC and S regimes. In the U regime, instead, such behavior is not observed: the process of price formations is completely random because agents fail before being able to adapt their preferences. Figure 5.6 also shows the signature of the bi-modality of the demand and supply distributions in the EC and S regimes. This is confirmed by the observation that the price distribution is bimodal as well, as shown in Figure 5.7.

Hence, our model generates three well distinct regimes. The U regime is somehow pathological: agents are replaced immediately after their introduction, hence their preferences are completely random and unable to adapt, and prices are also random, as a consequence. On the contrary, in the EC and S regimes, agents remain in the economy long enough to be able to adapt their prices. In the EC regime, we observe periodic

spikes of bankruptcies during which all agents are replaced; the agents' lifetime is then comparable to the periodicity of crises. In the S regimes, agents are replaced at a constant but very low rate. In both cases, one observes what could be called a "speciation" of goods. While all goods are *a priori* equivalent, the system self-organizes in such a way to create two categories of goods: cheap goods in high demand on the one hand, and expensive goods in low demand on the other. Note however that since this speciation is endogenous to the dynamics, it is also temporary: goods switch from one group to the other with time. The details of this dynamical process is what we examine next.

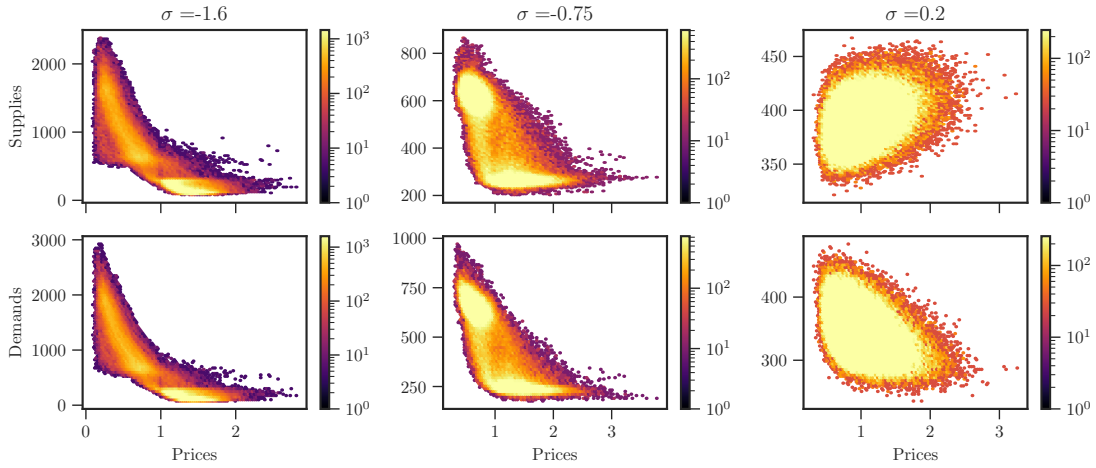


Figure 5.6: Scatter plot of supply and demand versus the prices of goods. In the first column, we have $\sigma = -1.6$ with supplies and demands concentrated at the extremes, corroborating the bimodal nature of their distributions. The intermediate case with $\sigma = -0.75$ is similar with some goods' supplies occurring between the two extremes. The last column shows the case for $\sigma = 0.2$ where the supplies and demands are uncorrelated with the prices of the goods. The color map corresponds to the number of points in a given region.

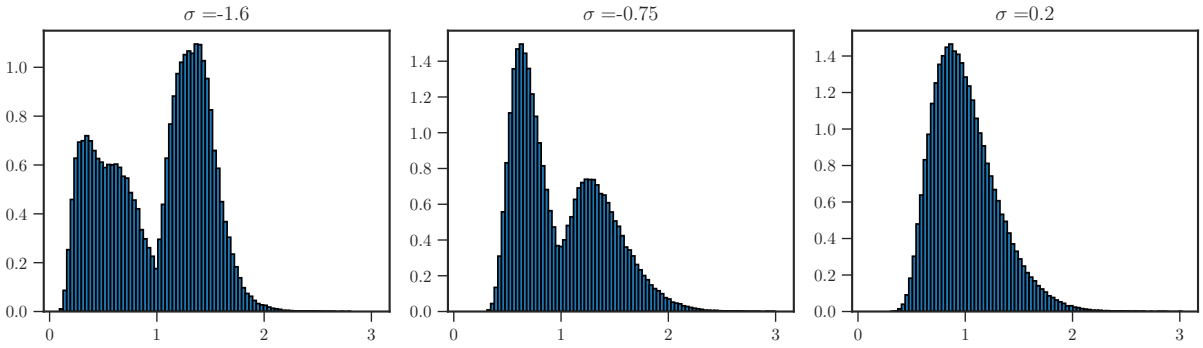


Figure 5.7: Price distributions for different values of σ, as in Figure 5.5. These distributions are computed over the complete time-series of the prices of goods.

5.5 ROLE OF THE DEBT LIMIT: DYNAMICS

The study of macroscopic observables has demonstrated that σ is a key control parameter which drives the system through three distinct regimes. We now turn to an analysis of the influence of σ on the dynamics of our toy economy.

We begin by observing how the demand of individual goods varies with time in these three regimes. In Figure 5.8 we show time-series of the demand of three randomly chosen goods for three values of σ. In the EC regime, we observe that the demand level of the goods is also periodic. This is a consequence of the existence of periodic crashes: as a large number of agents are periodically removed, the corresponding demand for goods also undergoes periodic swings. Interestingly, goods with low demand are out of phase with goods with high demand. In the U regime ($\sigma = 0.2$), the demand for goods fluctuates around an average value, within $\sim \pm 10\%$. This follows the behavior observed in the z_{ema} time-series as well: a significant proportion of agents is removed at each time step and hence the level of demand remains stable around its average.

For the intermediate S regime ($\sigma = -0.75$), the situation is quite complex. The demand for individual goods shows large variations: one good might start with high demand before being replaced by another good, within rather short time intervals. This suggests that the goods that are present in the two peaks of the bimodal distribution of demands (Figure 5.5) are not the same but keep changing in time. We indeed observe that goods keep switching from the right peak (high demand) to the left peak (low demand) dynamically, while maintaining the bimodal character of the distributions globally intact. A similar scenario with endogenous switches is found in [117]: firms compete among themselves to gain market share with boundedly rational consumers choosing firms to maximize their utility. This interaction leads to firms dynamically exchanging positions of monopoly with each other.

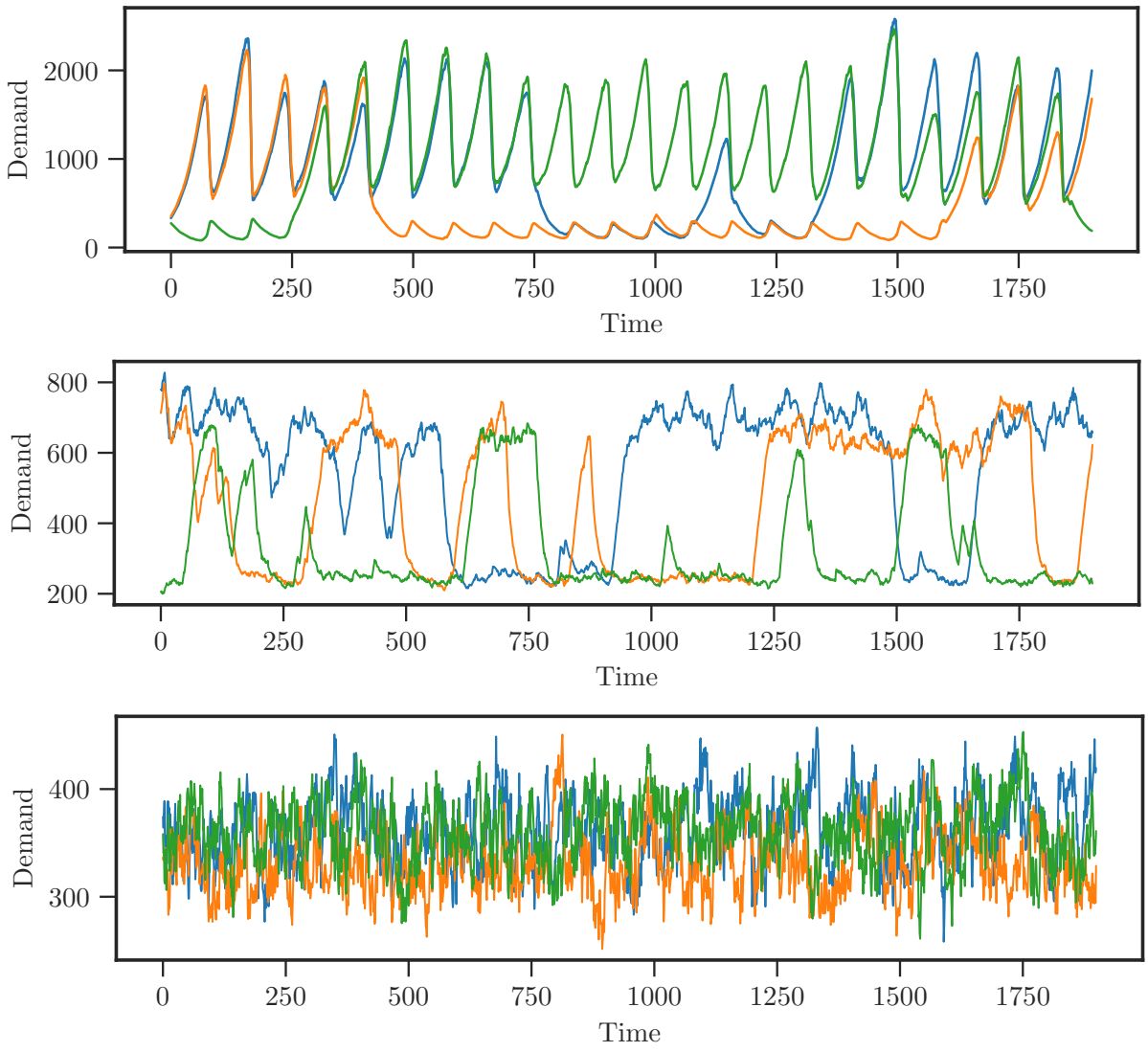


Figure 5.8: Time series of the demand of three random goods for three values of σ. Top: $\sigma = -1.6$ (EC regime) where we have cyclical rise and fall in demand. Middle: $\sigma = -0.75$ (S regime) where goods can switch from having high demand to having low demand (and vice-versa). As one good demand falls, another low-demand good takes over. Bottom: $\sigma = 0.2$ (U regime) where no coherent trend is found.

As a means of understanding the dynamical switching between goods in the S regime, we consider the case of a particular good whose demand (and hence supply) falls. A preliminary observation we make is that the behavioral rules (Eqs (5.4), (5.5)) for the agents imply that as soon as the price of a good becomes high (greater than 1 which is the average price), then agents will reduce their demand. Hence, we expect that as

soon as the price of a good reaches 1, we will observe a fall in demand, which will lead to a fall in supply. The fall in supply in turn will produce a further rise in prices since sellers will seek to maintain their previous profit levels at reduced demand. This increase in price feeds back into suppressed demand and this feedback loop leads to a rapid collapse. This situation is indeed borne out in the data as shown in Figure 5.9. In the left column, we show how agents reduce their demand and how supply follows in lockstep. At the aggregate level, we observe the rapid fall of the supply and demand of the good as soon as $p_i > 1$ (right column top).

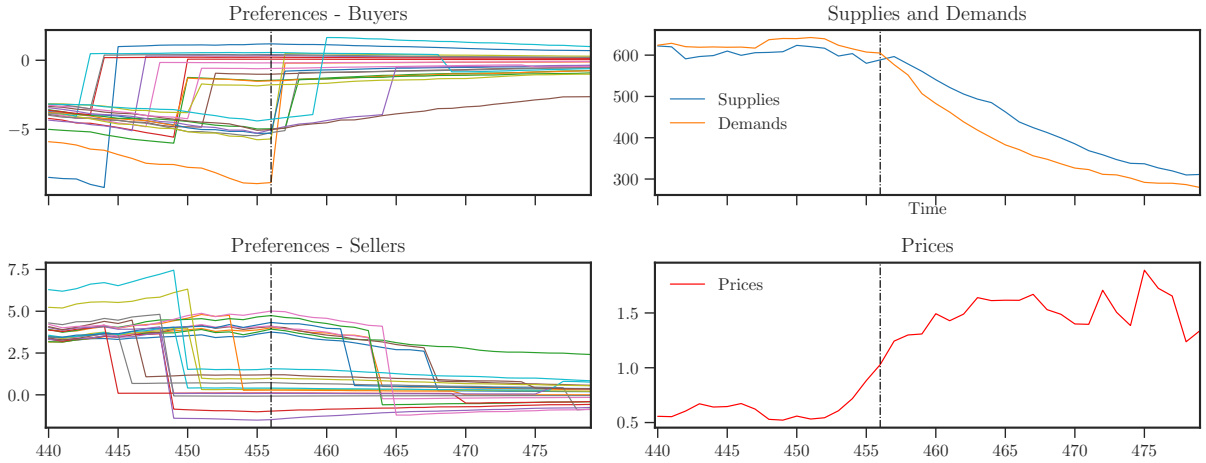


Figure 5.9: Understanding the switch from high to low demand of a good. *Left column*: The individual trajectories of the preferences of agents who are buyers (top) and sellers (bottom) is shown. The removal of an agent is shown by an abrupt jump or fall in the preference. *Right column*: The total supply and demand for the good (top) and its price (bottom) are shown. The black dotted line is the point when $p_i > 1$.

Still, two questions remain unanswered. The first question is: what factors lead to the initial price increase of the good itself? We observe in Figure 5.9 that the price continues to increase for a certain number of time-steps before reaching $p_i = 1$. This increase in price can be understood by the failure of big buyers of the good. Producers of this good then face reduced demand and hence must lower production according to Eq. (5.4). The prices then have to adjust upwards since producers have to satisfy their budget constraint at a lower scale of production. The second question is: what precipitates the

failure of a big buyer or big seller? To understand this better, we define a quantity called the f-index:

$$
\begin{aligned}
f_i^{\text{sellers}}(t) &= \sum_{\mu} \xi_\mu^i(t)\,\Theta(\xi_\mu^i(t))\,\Theta\Big(\sigma - \pi_\mu^{\text{ema}}(t+1)\Big), \\
f_i^{\text{buyers}}(t) &= \sum_{\mu} |\xi_\mu^i(t)|\,\Theta(-\xi_\mu^i(t))\,\Theta\Big(\sigma - \pi_\mu^{\text{ema}}(t+1)\Big).
\end{aligned}
\tag{5.9}
$$

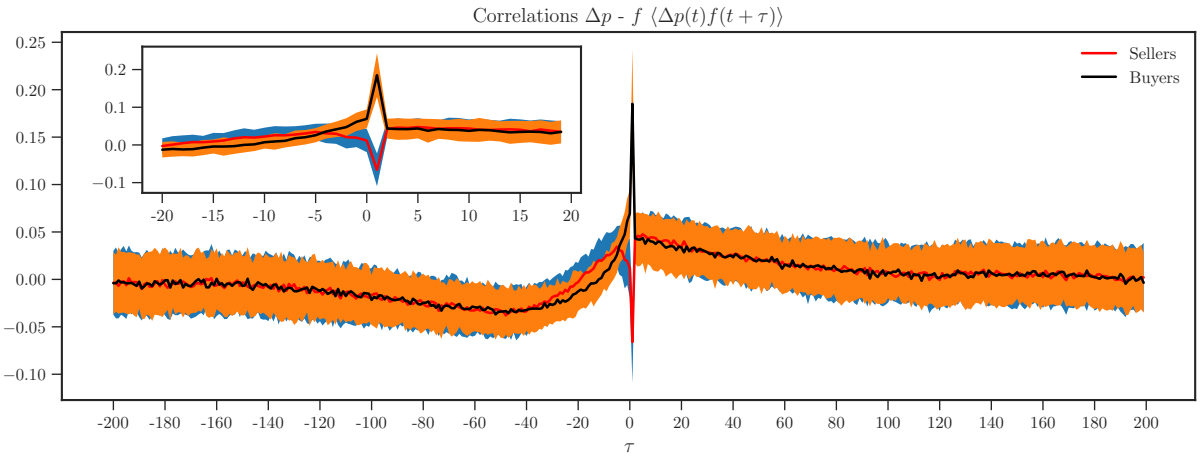


Figure 5.10: Correlation between changes in prices Δp with the f-index. The correlation shows a sharp peak at $\tau = 1$. Inset shows a zoom near the peak of the correlation function. The correlation here is $\langle \Delta p(t) f(t+\tau) \rangle$ averaged over all the goods.

Thus the f-index computes the decrease in supply or demand for a good due to the agents who will go bankrupt in the next step. We measure the correlation of the f-index with the change in prices. The average of this correlation over all goods is shown in Figure 5.10. We observe that for both buyers and sellers the correlations are peaked at $\tau = 1$. The positive correlation observed for buyers suggests that an increase in the price of goods is accompanied with the removal of buyers in the next time step. On the other hand, the correlation for sellers is negative implying that a reduction in the prices corresponds to the sellers being removed.

We thus conclude that purely random fluctuations in the price of a good can engender, through a feedback loop, the failure of sellers and buyers and produce the peculiar dynamics observed in the middle panel of Figure 5.8. One might argue that this is due to the myopic nature of our agents and the strategy they use. This might well be; on the other hand, it is difficult to bet the stability of the economy on the purported rationality of agents.

We now move to the dynamics of individual agents. For sufficiently small values of σ, in the EC regime, we have observed that z_{ema} is cyclical: a significant fraction of the agents are removed during crashes, which cover about 10-20 time-steps separated by periods with low removal rates. In the top panel of Figure 5.11 we show the π_μ^{ema} time-series for 20 randomly chosen agents. We observe that those agents who are well above their threshold and with disparate budgets are nevertheless removed together during a crash. Furthermore, the lower panel of Figure 5.11 shows that over one period of the cycle almost all the agents are removed and replaced.

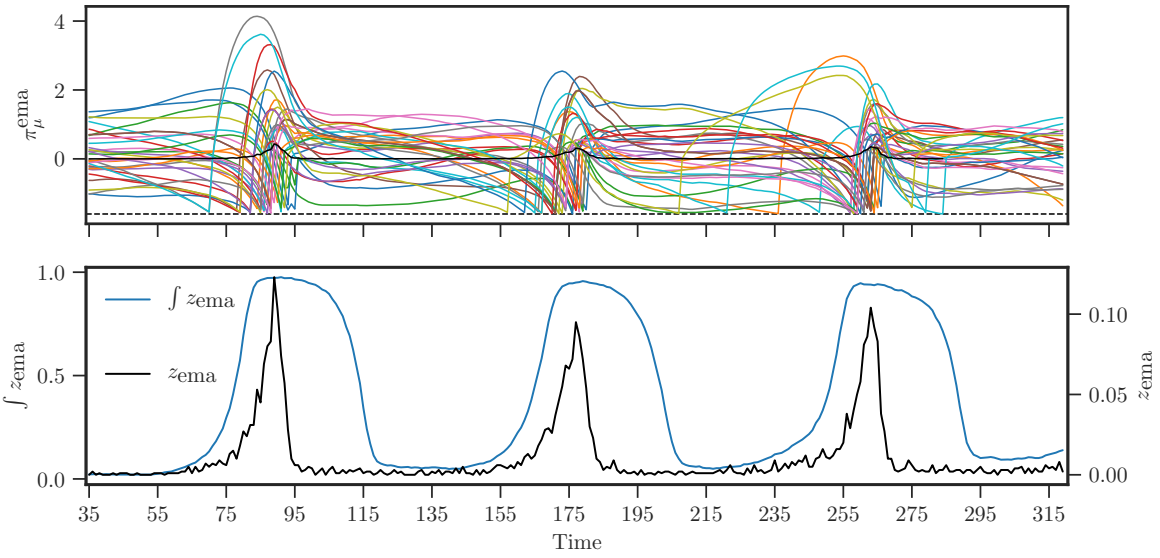


Figure 5.11: *Top*: π_μ^{ema} time-series for $\sigma = -1.6$ is shown for 20 agents. The dark black line is the average value of π_μ at which agents are re-injected. *Bottom*: z_{ema} (in black) with a moving integral of z_{ema} showing that during the crash almost all agents are removed and replaced.

The mechanism leading to the cyclical behavior can be understood via the feedback that exists between the failure of agents and the profit trajectories of the agents. As discussed in [77], a biased random walk along with an absorbing boundary condition and re-injections can under certain conditions lead to synchronized crisis waves. In our case, we observe that the removal of an agent produces a small but systematic reduction in π_μ^{ema} for all the other surviving agents — i.e. the failure of one agent fragilises the rest of the community. This is exactly the mechanism at the origin of synchronization discussed in [77], leading to periodic crises of fully endogenous nature.

In order to provide support for this interpretation, in Figure 5.12, we plot the change in profits, for all surviving agents, $\Delta\pi_{\text{ema}}^{\mu} = \pi_{\text{ema}}^{\mu}(t+\tau) - \pi_{\text{ema}}^{\mu}(t)$ against the fraction

of agents removed z_{ema} at each time-step. This shows that the average change (over surviving agents) in profits is nearly always negative for the surviving agents and this is true irrespective of the number of agents removed, even when this number remains small (see Figure 5.12, right).

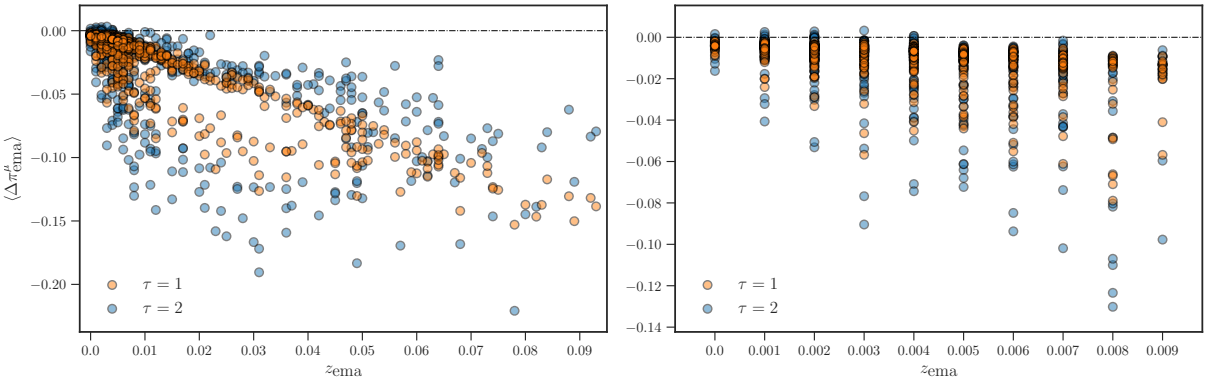


Figure 5.12: Variation of the change in the profits π^{ema}_{μ} averaged over all remaining agents following the removal of an agent plotted against z_{ema}. Two distinct cases are shown, $\tau = 1$ and $\tau = 2$ corresponding to changes in the profit of surviving agents one time-step or two time-steps after the removal of agents. *Left*: $\Delta\pi^{\mu}_{\text{ema}}$ against z_{ema} computed over the complete evolution of the economy. *Right*: Zoom of $\Delta\pi^{\mu}_{\text{ema}}$ for small values of z_{ema}.

5.6 DISCUSSION & CONCLUSION

In this chapter we studied a prototype agent-based model with a budgetary constraint. Within our model, agents' preferences, which in turn determine the supply and demand for goods, cannot adapt immediately. Then the prices for the goods are set by the market so as to satisfy every agent's budgetary constraints. These budgetary constraints lead to the interpretation of the economy as a constraint satisfaction problem (CSP) — given agents' preferences and their budgets, find the set of prices that satisfies each agent.

The model exhibits three regimes as a function of the amount of debt that each agent can accumulate before defaulting. At high debt (large negative σ), endogenous cycles appear with the economy going from a stable state with few bankruptcies to a state with a high bankruptcy rate. The emergence of waves of defaults in the endogenous crisis regime is triggered by the same mechanism as the one proposed in [77]: surviving agents profit margins reduce as one agent goes bankrupt. Lowering the allowed debt (increasing σ) takes the model to a regime where the economy is stable with very low bankruptcy rate and no aggregate-level crises. In this regime, the system self-organizes to create two categories of goods: cheap goods in high demand on the one hand, and

expensive goods in low demand on the other. Goods switch from one category to the other with time with switches triggered by random fluctuations of the prices. Finally, there is a low debt or high required profit regime (positive σ) where the bankruptcy rate is extremely high with a large fraction of the agents going bankrupt at every time step. In other words, in this regime agents never have time to adapt and the economy remains structure-less.

The model confirms the central role that debt levels play in the stability of the economy: too high a debt level and we have periodic crises, too low a debt level and agents cannot sustain themselves long enough and the economy remains structure-less with agents going bankrupt all the time. Understanding the coupling between debt-levels and economic stability is paramount to understanding macroeconomic dynamics and has become even more important following the Great Financial Crisis of 2008. There is now increasing evidence that financial crises are preceded by protracted periods of credit expansion. Empirically, it has been observed that higher the private debt buildup, deeper the downturn [118]. Both ABM and standard DSGE approaches have sought to understand the details of the leverage cycle: Ref. [119] studies the leverage of financial entities inter-mediating between households and industry. They conclude that increased leverage does lead to increased output but at the cost of systemic risk. Aymanns and co-authors build an agent-based model to uncover the dynamics of the leverage cycle based on the risk perception of actors like banks [120, 121]. They propose a detailed study underlying the importance of managing leverage systematically to dampen endogenously created debt-driven boom-bust cycles. Importantly, they find sustainable levels of leverage exceeding which the dynamics becomes unstable and chaotic.

The work presented in this chapter hence was also a direction towards understanding how the "real" economy interacts with the financial side through access and levels of debt. Our work does present a break from previous studies on the leverage cycle by coupling the production output and trading within the economy with agents' budgetary constraints. However, agents in our model get credit for free and hence an interesting direction for future exploration is the introduction of a bank (or banking sector) to set a price for borrowing through the interest rate.

Finally, the approach presented in this chapter is related to the literature on the so-called "occasionally-binding constraint" (OBC) type models studied in macroeconomic literature [122, 123]. Within such models, we study the effect of a strong constraint on macroeconomic dynamics when it becomes binding. These constraints can be seen as modeling two disparate regimes within a single model: under one regime, the constraint is non-binding (or slack) while in the other the constraints are binding or strict. A simple example is the Zero-Lower bound (ZLB) for short term rates set by central banks: in normal times these rates are well above the zero-bound and the constraint is non-

binding. However, as crisis hits and central banks cut rates, these rates can become close to zero and then the ZLB becomes a binding constraint. Contrary to such OBC models, the framework presented here focuses on understanding the consequences of making individual constraints interact with each other through a price-setting mechanism. It is indeed this interaction through the price-setting mechanism that leads to the complex and rich behavior that we have seen in this chapter.

We restricted our analysis in the present chapter to the simplest case where the global constraint on prices is linear, fixing the average price to unity. This implies that the CSP is, at each time step, convex and its solution unique. This is no longer the case for other types of constraints, for example $\frac{1}{N}\sum_i (p_i - p_i^0)^2 = 1$ where p_i^0 are some reference prices or prices prevailing at the previous step, as considered in [1], and described in detail in Chap 4. Such non-linear constraints, or adding linear costs to the price update step, might generate an even richer phenomenology due to the existence of a large number of equilibria [1, 75].

6

SIMULATING COVID-LIKE SHOCKS TO AN ABM

In the previous chapter, we have studied an ABM to understand the importance of debt for economic stability. There the dynamics generated endogenous business cycles along with spontaneous speciation of goods. In this chapter, we turn our attention to study another ABM with the aim of understanding the impact of exogenous macroeconomic shocks.

During the early months of 2020, the world saw the unexpected arrival of the COVID-19 virus and very quickly a local epidemic was transformed into a worldwide pandemic. Taking a serious human toll, it was necessary to implement emergency measures such as complete lock-downs and shutting down all but the most necessary economic activity and as a consequence the world economy suffered a massive shock.

Just as soon as the lock-downs were imposed, questions concerning the economic impact of such measures began to be asked. Initial concerns about the economic recovery and the longer term impact of the economic were the subject of almost daily discussion. Following the shock, the commentariat has discussed the various scenarios of the "post-covid" evolution of the economy: A V-shaped recovery was expected initially wherein after a sharp downturn, the economy would quickly pick up the pace as sanitary measures are lifted; this expectation was gave way to concerns of a "U-Shaped" recovery with the drop deeper than expected as lock-downs were extended, but giving way nonetheless to a quick recovery; less hopeful commentators feared an "L-shaped" crisis with a permanent loss of output. Other scenarios like the "swoosh" (rapid drop giving way to a extremely slow recovery) or the "W-shaped" recovery with one economic downturn followed by another economic drop due either to a second outbreak or premature lifting of emergency economic measures. Indeed, a veritable zoology of shapes [78]. Within the academic economics community, a lot of work has been done. While some have coupled classical economic models with SIR-like epidemic models, with the underlying assumption that the economy is somehow slaved to the dynamics of COVID [124], others have reasoned in terms of traditional economic models. There has been analytical support for both quick (V or U shaped) recoveries [125] and prolonged

(L-shaped) crisis due to a stagnation trap (poor economic forecasts leading to lower consumption leading to lower investment) [126]. Given how different sectors of the economy are affected disproportionately (some completely shut, some are not), there are fears of a deep recession due to a Keynesian supply shock — a deep demand shock greater in magnitude to the supply shock that caused it [127].

Further investigation into the economic consequences of the pandemic have taken the direction of modeling the dynamics of the pandemic to propose optimal lockdown policies [128–130]. This has been done with the point of view of balancing a trade-off between the health of the economy and the health of the citizenry. However the evolution of the pandemic remains uncertain before vaccines become widely available, and computations based on short-term trade-offs may cause harm since policy-makers might err on the side of protecting the economy and lifting the lockdowns too early [131].

In this chapter, our aim is to explore how the economy can recover from a COVID-shock which has led to a rapid drop of both supply and demand, even assuming quick return to normal in terms of sanitary measures. To this end, we perform numerical simulations using the Mark-0 ABM which has been previously studied for understanding the transmission of monetary policy and for optimal inflation targeting. The model is able to reproduce the various scenarios that have been proposed — from rapid V-shaped recoveries to prolonged L-shaped crises. The intensity and the rapidity of the recovery is determined by the particular policy measures that are adopted. The lesson learned from the experiments presented here is to "do whatever it takes" [132]: either by providing prolonged support to firms to weather a difficult economic situation or by boosting household consumption through lower taxes or increased savings.

In what follows, a conscious decision has been taken to not model the dynamics of the pandemic along with the dynamics of the economy. The evolution of the pandemic is contingent not just on an optimal policy, but depends crucially on its timing, on society's compliance with that policy, on the availability of an efficient vaccine, etc. We focus on modeling the evolution of the economy itself, with the pandemic effectively modelled as one or several supply and demand shocks of different amplitudes and duration. Importantly, lifting of lockdowns in itself is not sufficient for the economy to recover since there is still widespread fear of the disease [133].

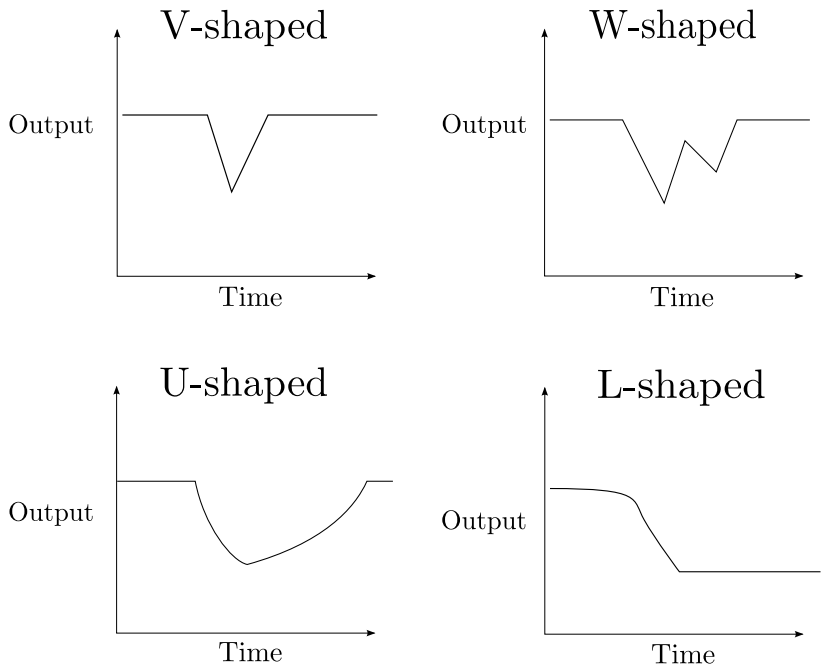


Figure 6.1: Various recovery "shapes" discussed in the commentariat [78].

6.1 DESCRIPTION OF MARK-0

The Mark-0 model with a Central Bank (CB) and interest rates has been described in full details in [44, 53, 54], where pseudo-codes are also provided. It was originally devised as a simplification of the Mark family of ABMs, developed in [55, 56]. Since the details of the model are spread across multiple papers, for purposes of completeness we summarize the fundamental aspects of the model below.

First, we establish some basic notation. The model is defined in discrete time, where the unit time between t and $t+1$ will be chosen to be ~ 1 month. Each firm i at time t produces a quantity $Y_i(t)$ of perishable goods that it attempts to sell at price $p_i(t)$, and pays a wage $W_i(t)$ to its employees. The demand $D_i(t)$ for good i depends on the global consumption budget of households $C_B(t)$, itself determined as an inflation rate-dependent fraction of the household savings. To update their production, price and wage policy, firms use reasonable "rules of thumb" [44] that also depend on the inflation rate through their level of debt (see below). For example, production is decreased and employees are made redundant whenever $Y_i > D_i$, and vice-versa. As a consequence of these adaptive adjustments, the economy is on average always 'close' to the global market clearing condition one would posit in a fully representative agent framework. However, small fluctuations persists in the limit of large system sizes

giving rise to a rich phenomenology [44], including business cycles. The model is fully "stock-flow consistent" (i.e. all the stocks and flows within the toy economy are properly accounted for). In particular, there is no uncontrolled money creation or destruction in the dynamics, except when explicitly stated.

In Mark-0 we assume a linear production function with a constant productivity, which means that output Y_i and labor N_i coincide, up to a multiplicative factor ζ: $Y_i = \zeta N_i$. The unemployment rate u is defined as:

$$u(t) = 1 - \frac{\sum_i N_i(t)}{N}, \tag{6.1}$$

where N is the number of agents. Note that firms cannot hire more workers than available, so that $u(t) \geq 0$ at all times — see Eq. (6.11) below. The instantaneous inflation rate, noted $\pi(t)$ is defined as:

$$\pi(t) = \frac{\overline{p}(t) - \overline{p}(t-1)}{\overline{p}(t-1)}, \tag{6.2}$$

where $\overline{p}(t)$ is the production-weighted average price $\overline{p} = \frac{\sum_i p_i Y_i}{\sum_i Y_i}$. We define $\overline{w}$ as the production-weighted wage as well: $\overline{w} = \frac{\sum_i w_i Y_i}{\sum_i Y_i}$. We further assume that households and firms adapt their behavior not to the instantaneous value of the inflation $\pi(t)$, but rather to a smoothed averaged value. We name this the "realized inflation" and it is defined as:

$$\pi^{\text{ema}} = \omega\pi(t) - (1-\omega)\pi^{\text{ema}}(t-1) \tag{6.3}$$

Households and firms form expectations for future inflation by observing the realised inflation and the target inflation set by the Central Bank. This is modeled as follows:

$$\widehat{\pi}(t) = \tau^R \pi^{\text{ema}} + \tau^T \pi^* \tag{6.4}$$

The parameters τ^R and τ^T model the importance of past inflation for forming future expectations and the confidence that economic actors have in the central bank's ability to achieve its target inflation.

The Mark-0 economy is made of firms producing goods, households consuming these goods, a banking sector and a central bank. Households and the banking sector are described at the aggregate level by a single "representative household" and "representative bank" respectively. In what follows, we describe how they interact within our toy economy.

6.1.1 *Households*

We assume that the total consumption budget of households $C_B(t)$ is given by:

$$C_B(t) = c\Big[S(t) + W(t) + \rho^d(t)S(t)\Big]\,, \tag{6.5}$$

where $S(t)$ is the savings, $W(t) = \sum_i W_i(t)N_i(t)$ the total wages, and $\rho^d(t)$ is the interest rate on deposits, and $c(t)$ is the "consumption propensity" of households. $c(t)$ is chosen such that it increases with increasing inflation:

$$c(t) = c_0[1 + \alpha_c(\pi^{\text{ema}}(t) - \rho^d(t)] \tag{6.6}$$

Here α_c is a parameter that modulates the sensitivity of households to the real interest (inflation adjusted) rate on their savings $\rho^d(t)$. With this choice of the consumption propensity, Eq. (6.5) describes an inflation-mediated feedback on consumption similar to the standard Euler equation in DSGE models [6].

The total household savings then evolve according to:

$$S(t+1) = S(t) + W(t) + \rho^d(t)S(t) - C(t) + \Delta(t), \tag{6.7}$$

where $\Delta(t)$ are the dividends received from firms' profits (see below). $C(t) \leq C_B(t)$ is the actual consumption of households, determined by the matching of production and demand and computed as :

$$C(t) = \sum_i^{N_F} p_i \min\{Y_i, D_i\} \leq C_B(t) = \sum_i^{N_F} p_i(t)D_i(t) \tag{6.8}$$

The demand for goods D_i itself is modeled via an intensity of choice model with a parameter β which defines the dependence of the demand on the price:

$$D_i(t) = \frac{C_B(t)}{p_i(t)} \frac{\exp^{-\beta p_i(t)}}{\sum_i \exp^{-\beta p_i(t)}} \tag{6.9}$$

6.1.2 *Firms*

The model contains N_F firms (we chose $N_F = N$ for simplicity [44]), each firm being characterized by its workforce N_i and production $Y_i = \zeta N_i$, demand for its goods D_i,

price p_i, wage W_i and its cash balance $\mathcal{E}_i$ which, when negative, is the debt of the firm. We characterize the *financial fragility* of the firm through the debt-to-payroll ratio

$$\Phi_i = -\frac{\mathcal{E}_i}{W_i N_i}. \tag{6.10}$$

Negative Φ's describe healthy firms with positive cash balance, while indebted firms have a positive Φ. If $\Phi_i < \Theta$, i.e. when the flux of credit needed from the bank is not too high compared to the size of the company (measured as the total payroll), the firm i is allowed to continue its activity. If on the other hand $\Phi_i \geq \Theta$, the firm i defaults and the corresponding default cost is absorbed by the banking sector, which adjusts the loan and deposit rates ρ^ℓ and ρ^d accordingly (see below). The defaulted firm is replaced by a new one at rate φ, initialized at random. The parameter Θ controls the maximum leverage in the economy, and models the risk-control policy of the banking sector.

PRODUCTION UPDATE If the firm is allowed to continue its business, it adapts its price, wages and production according to reasonable (but of course debatable) "rules of thumb" — see [44, 54]. In particular, the production update is chosen as follows:

$$\begin{aligned} &\text{If } Y_i(t) < D_i(t) \quad \Rightarrow \quad Y_i(t+1) = Y_i(t) + \min\{\eta_i^+ (D_i(t) - Y_i(t)), \zeta u_i^\star(t)\} \\ &\text{If } Y_i(t) > D_i(t) \quad \Rightarrow \quad Y_i(t+1) = Y_i(t) - \eta_i^- [Y_i(t) - D_i(t)] \end{aligned} \tag{6.11}$$

where $u_i^\star(t)$ is the maximum number of unemployed workers available to the firm i at time t, which depends on the wage the firm pays:

$$u^* = \frac{\exp^{\beta W_i(t)/\overline{w}(t)}}{\sum_i \exp^{\beta W_i(t)/\overline{w}(t)}}, \tag{6.12}$$

where $\overline{w}$ is the production-weighted average wage. Firms hire and fire workers according to their level of financial fragility Ψ_i: firms that are close to bankruptcy are arguably faster to fire and slower to hire, and vice-versa for healthy firms. The coefficients $\eta^\pm \in [0,1]$ express the sensitivity of the firm's target production to excess demand/supply. We posit that the coefficients $\eta_i^\pm$ for firm i (belonging to $[0,1]$) are given by:

$$\eta_i^- = [[\eta_0^- (1 + \Gamma \Phi_i(t))]] \tag{6.13}$$

$$\eta_i^+ = [[\eta_0^+ (1 - \Gamma \Phi_i(t))]], \tag{6.14}$$

where $\eta_0^\pm$ are fixed coefficients, identical for all firms, and $[[x]] = 1$ when $x \geq 1$ and $[[x]] = 0$ when $x \leq 0$. The factor $\Gamma > 0$ measures how the financial fragility of firms

influences their hiring/firing policy, since a larger value of Φ_i then leads to a faster downward adjustment of the workforce when the firm is over-producing, and a slower (more cautious) upward adjustment when the firm is under-producing. Γ itself depends on the inflation-adjusted interest rate and takes the following form:

$$\Gamma = \max\left[\alpha_\Gamma(\rho^l(t) - \pi^{\text{ema}}(t), \Gamma_0\right], \tag{6.15}$$

where α_Γ is a free parameter, similar to α_c that captures the influence of the real interest rate.

PRICE UPDATE Following the initial specification of the Mark series of models [55], prices are updated through a random multiplicative process, which takes into account the production-demand gap experienced in the previous time step and if the price offered is competitive (with respect to the average price). The update rule for prices reads:

$$\begin{aligned}
\text{If } Y_i(t) < D_i(t) \quad \Rightarrow \quad &\begin{cases} \text{If } p_i(t) < \overline{p}(t) \quad \Rightarrow \quad p_i(t+1) = p_i(t)(1+\gamma\xi_i(t))(1+\hat{\pi}(t)) \\ \text{If } p_i(t) \geq \overline{p}(t) \quad \Rightarrow \quad p_i(t+1) = p_i(t) \end{cases} \\
\text{If } Y_i(t) > D_i(t) \quad \Rightarrow \quad &\begin{cases} \text{If } p_i(t) > \overline{p}(t) \quad \Rightarrow \quad p_i(t+1) = p_i(t)(1-\gamma\xi_i(t))(1+\hat{\pi}(t)) \\ \text{If } p_i(t) \leq \overline{p}(t) \quad \Rightarrow \quad p_i(t+1) = p_i(t) \end{cases}
\end{aligned} \tag{6.16}$$

where $\xi_i(t)$ are independent uniform $U[0,1]$ random variables and γ is a parameter setting the relative magnitude of the price adjustment. The factor $(1+\widehat{\pi}(t))$ models firms' anticipated inflation $\widehat{\pi}(t)$ when they set their prices and wages.

WAGE UPDATE The wage update rule follows the choices made for price and production. Similarly to workforce adjustments, we posit that at each time step firm i updates the wage paid to its employees as:

$$\begin{aligned}
W_i^T(t+1) = W_i(t)[1+\gamma(1-\Gamma\Phi_i)(1-u(t))\xi_i'(t)][1+g\hat{\pi}(t)] \quad \text{if} \quad &\begin{cases} Y_i(t) < D_i(t) \\ \mathcal{P}_i(t) > 0 \end{cases} \\
W_i(t+1) = W_i(t)[1-\gamma(1+\Gamma\Phi_i)u(t)\xi_i'(t)][1+g\hat{\pi}(t)] \quad \text{if} \quad &\begin{cases} Y_i(t) > D_i(t) \\ \mathcal{P}_i(t) < 0 \end{cases}
\end{aligned} \tag{6.17}$$

where $\mathcal{P}_i$ is the profit of the firm at time t, $\xi_i'(t)$ an independent $U[0,1]$ random variable and g modulates how wages are indexed to the firms' inflation expectations. If $W_i^T(t+1)$ is such that the profit of firm i at time t with this amount of wages would have been negative, $W_i(t+1)$ is chosen to be exactly at the equilibrium point where $\mathcal{P}_i(t) = 0$; otherwise $W_i(t+1) = W_i^T(t+1)$. Here, Γ is the same parameter introduced in Eq. (6.13).

Note that within the current model the productivity of workers is not related to their wages. The only channel through which wages impact production is that the quantity $u_i^\star(t)$ that appears in Eq. (6.11), which represents the share of unemployed workers accessible to firm i, is an increasing function of W_i. Hence, firms that want to produce more (hence hire more) do so by increasing W_i, as to attract more applicants.

The above rules are meant to capture the fact that deeply indebted firms seek to reduce wages more aggressively, whereas flourishing firms tend to increase wages more rapidly:

- If a firm makes a profit and it has a large demand for its good, it will increase the pay of its workers. The pay rise is expected to be large if the firm is financially healthy and/or if unemployment is low because pressure on salaries is high.
- Conversely, if the firm makes a loss and has a low demand for its good, it will attempt to reduce the wages. This reduction is more drastic if the company is close to bankruptcy, and/or if unemployment is high, because pressure on salaries is then low.
- In all other cases, wages are not updated.

PROFITS AND DIVIDENDS Finally, the profits of the firm $\mathcal{P}_i$ are computed as the sales minus the wages paid with the addition of the interest earned on their deposits and the interest paid on their loans:

$$\mathcal{P}_i = p_i(t)\min\{Y_i(t), D_i(t)\} - W_i(t)Y_i(t) + \rho^d \max\{\mathcal{E}_i(t), 0\} - \rho^l \min\{\mathcal{E}_i(t), 0\}. \tag{6.18}$$

If the firm's profits are positive and they possess a positive cash balance, they pay dividends as a fraction δ of their cash balance $\mathcal{E}_i$:

$$\Delta(t) = \delta \sum_i \mathcal{E}_i(t)\theta\Big(\mathcal{P}_i(t)\Big)\theta\Big(\mathcal{E}_i(t)\Big), \tag{6.19}$$

where θ is the Dirac delta-function. These dividends are then reduced from the firms cash balance and added to the households savings in Eq. (6.7).

6.1.3 *Banking sector*

The banking sector in Mark-0 consists of one "representative bank" and a central bank which sets baseline interest rates. The central bank also has an inflation targeting mandate. The central bank sets the base interest rate ρ_0 via a Taylor-like rule:

$$\rho_0(t) = \rho^* + \phi_\pi[\pi^{\text{ema}} - \pi^*]. \tag{6.20}$$

Here, ϕ_π modulates the intensity of the central bank policy, ρ^* is the baseline interest rate and π^* is the inflation target for the central bank. The banking sector then sets interest rates for deposits ρ^d (for households) and loans ρ^l (for borrowing by firms). Defining $\mathcal{E}^+ = \sum_i \max(\mathcal{E}_i, 0)$ (equivalent to firms' positive cash balance) and $\mathcal{E}^- = -\sum_i \min(\mathcal{E}_i, 0)$ (equivalent to firms' total debt), the interest rates are set as:

$$\rho^l(t) = \rho_0(t) + f\frac{\mathcal{D}(t)}{\mathcal{E}^-(t)} \tag{6.21}$$

$$\rho^d(t) = \frac{\rho_0\mathcal{E}^-(t) - (1-f)\mathcal{D}(t)}{S + \mathcal{E}^+(t)}, \tag{6.22}$$

where $\mathcal{D}(t)$ are the total costs accrued due to defaulting firms. The parameter f then determines how the impact of these defaults fall upon lenders and depositors — f interpolates between these costs being borne completely by the borrowers ($f = 1$) or fully by the depositors ($f = 0$). The total amount of (central-bank) money M in circulation is kept constant and the balance sheet of the banking sector reads:

$$M = S(t) + \mathcal{E}^+(t) - \mathcal{E}^-(t) \tag{6.23}$$

This is simply a restatement of the fact that the sum of households savings and the deposits and debts of firms is equal to the total amount of money in circulation.

6.2 SUMMARY

The model, as presented above, has a rich phase diagram and even though the model possesses many free parameters, it has been found that only R (Hiring/firing ratio) and Θ (bankruptcy threshold for firms) are important [44, 54]. The model hence present four phases in the $\Theta - R$ plane:

- Full Employment (FE) phase: characterized by positive average inflation and close to full employment,

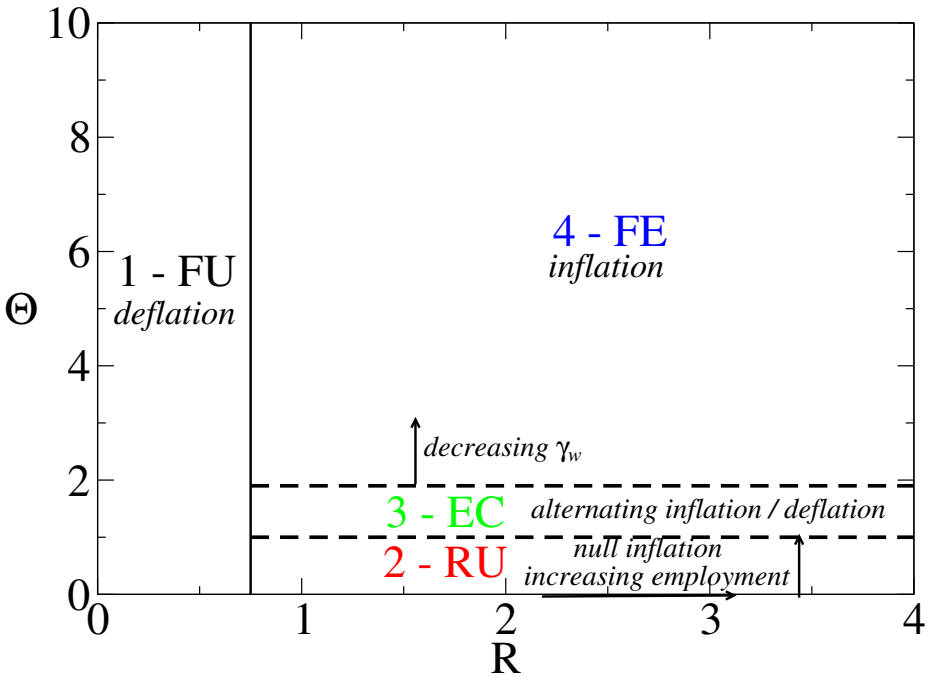


Figure 6.2: The phase diagram of the Mark-0 model is shown. Four distinct phases are identified. The Full Unemployment(FU) and Full Employment (FE) are accompanied by deflation and inflation respectively. The Residual unemployment (RU) phase has zero inflation but some unemployment remains. The Endogenous crises (EC) phase is marked by high employment (and inflation) on average punctuated by deflationary periods of high unemployment. Source [54].

- Full Unemployment (FU) phase: characterized by high unemployment and deflation,
- Residual Unemployment (RU) phase: characterized by zero inflation but with some residual unemployment, a
- Endogenous crises (EC) phase: characterized by low unemployment and inflation on average, but with the intermittent spikes of "endogenous crises" accompanied by high unemployment and deflation.

These phases are shown in Fig.6.2. In what follows, we fix a certain number of parameters which have been summarized in Table 6.1. The baseline values of these parameters leads to an economic situation with a low-level of unemployment, near maximal output for the given productivity $\zeta = 1$. The inflation level is $\approx 1.3\%$/year and the average financial fragility $\langle \Phi \rangle$ of the firms is ≈ 1 — a debt equal to one month of wages — far from the baseline bankruptcy threshold $\Theta = 3$. This corresponds to the Full

Number of firms	N_F	10000
Consumption propensity	c_0	0.5
Intensity of choice parameter	β	2
Price adjustment parameter	γ	0.01
Firing propensity	η_-^0	0.2
Hiring propensity	η_+^0	$R\eta_-^0$
Hiring/firing ratio	R	2
Fraction of dividends	δ	0.02
Bankruptcy threshold	Θ	3
Rate of firm revival	φ	0.1
Productivity factor	ζ	1
Financial fragility sensitivity	Γ_0	0
Exponentially moving average (ema) parameter	ω	0.2

Table 6.1: Parameters of the Mark-0 model relevant for this chapter, together with their symbol and baseline values.

Employment (FE) phase in Fig 6.2. In our examination of the Covid-induced shocks to our toy economy, certain simplifying assumptions have been made:

1. The inflation expectations are set to zero i.e. in Eq (6.4), τ^R and τ^T are both set to zero.

2. We neglect the feedback of inflation on consumption i.e. $\alpha_c = 0$ in Eq (6.6). This implies that the consumption propensity c is no longer a function of time and is constant, equal to c_0.

3. Firms' financial fragility doesn't affect their hiring/firing rates i.e. α_Γ and Γ_0 are set to zero in Eq (6.15).

We also make the strong assumption of not using the traditional monetary policy channel, i.e. the central bank is switched off — the base interest rate $\rho_0 = 0$ and the central bank performs no inflation targeting, $\pi^* = 0$.

6.3 SIMULATING SHOCKS TO THE ECONOMY

The specificity of the ongoing Covid crisis is that it induced both a supply and a demand shock [134]. We model this effect by a sudden drop in the productivity of firms,

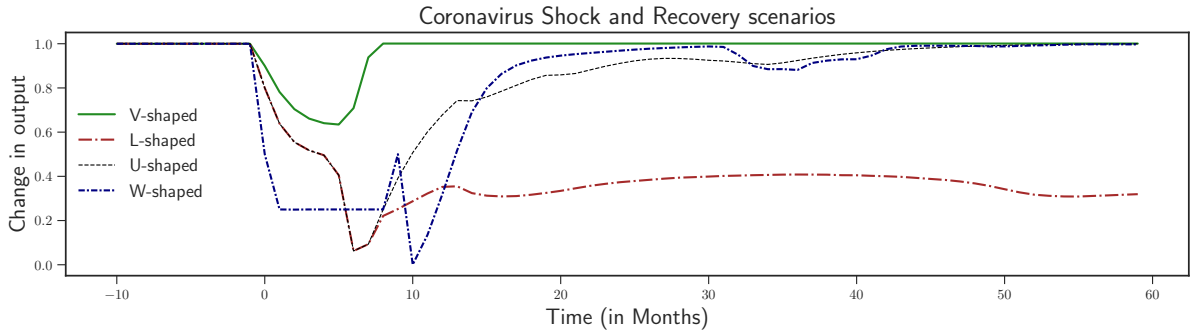


Figure 6.3: Some recovery patterns following the coronavirus shock, which starts at time $t = 0$. We show, as a function of time, the fall in output relative to the no-shock scenario. For Mild shocks ($\Delta c/c = 0.3$, $\Delta\zeta/\zeta = 0.1$, lasting six months), the economy contracts but quickly recovers (V-Shaped recovery). For more severe shocks over the same time period ($\Delta c/c = 0.3$, $\Delta\zeta/\zeta = 0.2$), the economy contracts permanently and never recovers (at least on the time scale of the simulation). This is the dreaded L-shaped scenario, in the absence of any government policy. An increase in consumption propensity to $c = 0.7$ (0.2 above pre-crisis levels) and a helicopter money drop at the end of the shock leads to a faster recovery (U-shaped recovery). Finally, a W-shaped scenario can also be found for stronger shocks ($\Delta c/c = 0.3$, $\Delta\zeta/\zeta = 0.5$, lasting nine months) with strong government policy and a helicopter drop.

$\zeta \to \zeta - \Delta\zeta$, and in the consumption propensity of households, $c \to c - \Delta c$. This is meant to model the effect of a lock-down on the economy, that leads both to a drop in supply (employees must stay at home and either not work at all or work remotely with reduced productivity) and a drop in demand (people are prohibited from stepping out or are simply too afraid). Given that the infection continues to spread, it is uncertain how long the effects of the crisis would last. Hence, an important parameter describing the shock is its duration T. In what follows, we choose three values for the duration of the shock: 3 months, 6 months and 9 months. Lockdown measures can be partially lifted which may lead to an increased value for both ζ and c during the shock period. We also note that although lockdowns might not last for as long as 9 months, the effects of the pandemic on consumer behavior might persist for longer. As has already been noted, the economic impact on people's lives may last well beyond the lifting of the lockdowns with consumer spending lower than pre-pandemic levels [133]. A drop of consumption propensity could reflect a long lasting loss of confidence, for example. These values for T are thus meant to represent an effective length of the shock, accounting for the fact that lock-down measures can be partially lifted, which leads to an increased value of both ζ and c during the shock period and hence a shorter effective shock duration.

PHENOMENOLOGY: We begin with an exploration of the various scenarios possible for the evolution of the economy following the shock. In Fig. 6.3, we show several typical crisis and recovery shapes, depending on the strength of the shock and the policy used to alleviate the severity of the crisis. For small enough Δc and/or $\Delta\zeta$, there is no drop of output at all. For larger shock amplitude or duration, one observes a V-shape recovery, as expected when the shock is mild enough not to dent the financial health of the firms. Stronger shocks can however lead to a permanent dysfunctional state (L-shape), with high unemployment, falling wages and savings, and a high level of financial fragility and bankruptcies. An L-shaped scenario can however be prevented if after the shock, consumer demand rises i.e. by increasing c. To facilitate and boost consumption, a one-time policy of helicopter money can move the economy towards a path of recovery over the scale of a few years (U-shape). Interestingly, for shocks lasting for nine months, a one-time helicopter drop of money without any other interventions can lead to a W-shaped recovery.

PHASE DIAGRAMS To better understand the influence of these shocks, we once again call upon our trusted tool — the phase diagram and plot it in the plane $\Delta c/c, \Delta\zeta/\zeta$, for $T = 3, 6$ and 9 months in Fig. 6.4. The phase diagrams are all constructed in the absence of policy since we want to first understand which regions of phase-space are ones where policy would indeed be necessary. Within each phase diagram, we present in order (a) the probability of a "dire" (L-shaped) crisis, (b) the peak value of unemployment during the shock and (c) the peak value of unemployment after the shock. Black regions indicate that the economy recovers well — no dire crisis, or short crises with low levels of unemployment. This occurs, as expected, for small shocks, small $\Delta c/c, \Delta\zeta/\zeta$, in the lower left corner of the graphs. These regions shrink as T increases implying that even mild shocks but sustained over a longer period of time can lead to dire crises. We also note that mild shocks lasting only a short time ($T = 3$ months) can also cause lasting damage. Indeed, we observe that low rates of unemployment during the shock are not representative of future evolution. Interestingly, there is an abrupt, first order transition line (a "tipping point" in the language of [44]) beyond which crises have a very large probability to become permanent, with high levels of unemployment. In reality, the precise location of such tipping points is hard to estimate. We are led to conclude that governments and other institutions should err on the side of caution with their policy making.

We also focus on the $\Delta\zeta = 0$ line of these two-dimensional plots, corresponding to a consumption shock without a simultaneous productivity shock. We show in Fig. 6.5 the same three quantities as in Fig. 6.4. An abrupt transition between no dire crises and dire crises can be seen for $\Delta c/c \sim 0.4$ when $T = 9$ months.

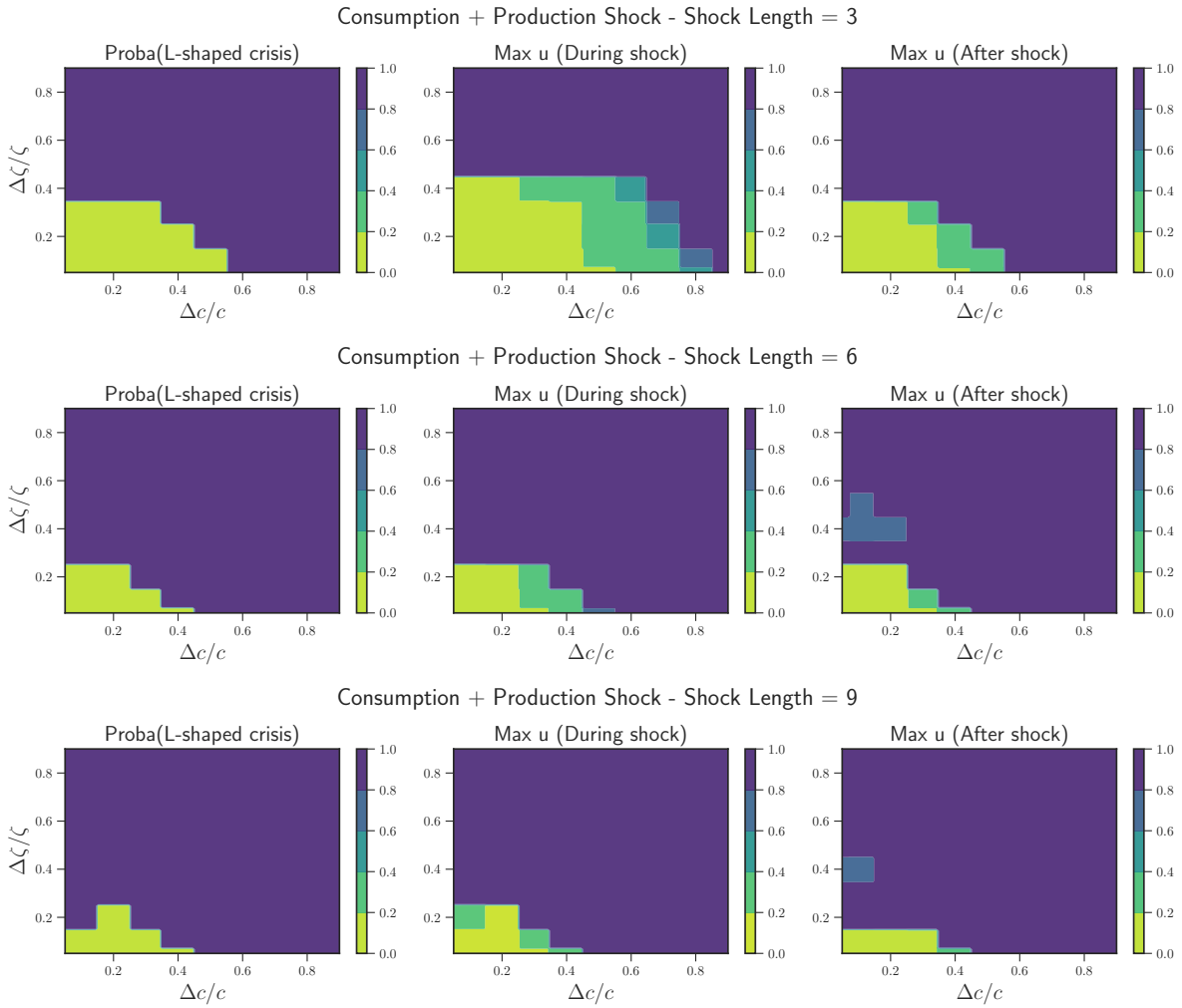


Figure 6.4: Phase diagrams in the $\Delta c/c$ - $\Delta\zeta/\zeta$ plane for different shock lengths. *Top Row:* Shock lasting for 3 months. The region of parameter space with no L-shaped crisis is quite large allowing for strong consumption shocks ($\Delta c/c \lesssim 0.5$) and mild productivity shocks ($\Delta\zeta/\zeta \lesssim 0.3$). Note that for such a short shock, the effects on unemployment are seen after the shock has ended. *Middle Row:* Shock lasts for 6 months. A decrease in the region of no-crisis is observed. Mild shocks ($\Delta c/c \sim 0.4$) can also lead to prolonged crises. During the shock itself, extremely high rates of unemployment can be seen. *Bottom Row:* Shock lasts for 9 months. Only for extremely mild shocks does the economy not undergo a prolonged crisis.

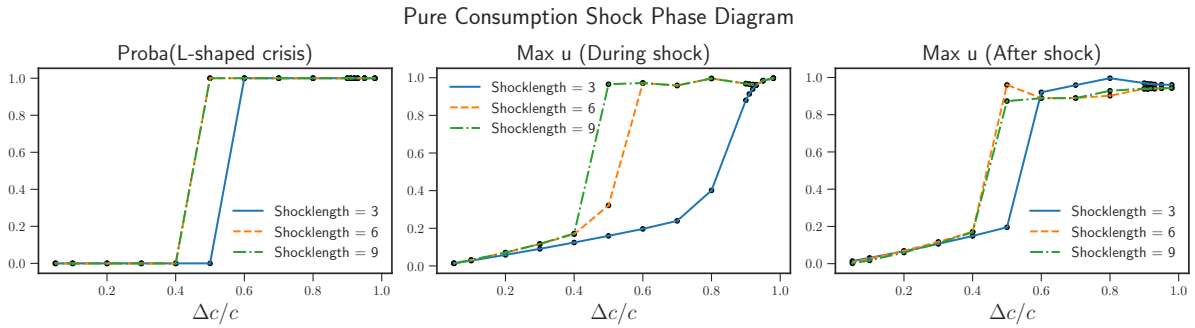


Figure 6.5: Phase diagram for a pure consumption shock ($\Delta\zeta/\zeta = 0$). We observe an abrupt transition to an L-shaped crisis for consumption shocks beyond $\Delta c/c = 0.5$ for $T = 3$ months, and beyond $\Delta c/c = 0.4$ for $T = 6$ or 9 months. Below these shock amplitudes, there is no prolonged crises but short-lived crises are observed. This can be seen by observing the maximum unemployment rate during the shocks: for $\Delta c/c = 0.4$, we reach about 15% unemployment.

6.4 POLICY PROPOSALS FOR A QUICK RECOVERY

> Within our mandate, the ECB is ready to do "whatever it takes" to preserve the euro. And believe me, it will be enough.
>
> — Mario Draghi, *Speech at the Global Investment Conference, London, 26 July 2012*

The toolbox developed in the aftermath of the Global Financial Crisis (GFC) of 2008 puts monetary policy at the center of economic crisis management. This takes the form of either direct interest rate cuts or, as was seen recently for the GFC interventions, even stronger measures such as quantitative easing. Given that the interest rates are already very low, the interest-rate channel itself might not be effective, and might lead to a stagnation trap and a L-shaped recovery [126]. Hence, we disregard the interest-rate channel as a possible policy tool, since it cannot be used as an emergency measure in the face of a collapsing supply sector. Consequently, we discuss two other channels: easy credit for firms, and "helicopter money" for households.

POLICY-1: EASY CREDIT ACCESS TO FIRMS A way to loosen the stranglehold on struggling firms is to give them easy access to credit lines, independently of their financial situation. Within our model, this is equivalent to an increase of the bankruptcy threshold Θ. The policy we investigate is then the following: during the whole duration of the shock, we set $\Theta = \infty$ — all firms are allowed to continue their business as usual and accumulate debt. When the shock is over, the value of Θ is reduced to its pre-crisis

levels i.e. $\Theta = 3$. Tuning down the credit access line can be done in multiple ways. One extreme possibility, termed "naive" below, is to set Θ to its pre-shock value as soon as the shock is over. Intuitively, when the shock is short enough, allowing endangered firms to survive might be enough. For longer shocks, such a naive policy might not be helpful since firms that have pulled through the crisis have become more fragile (accumulated too much debt) by the end of the shock. In this scenario, many firms fail when credit is tightened, and the economy plunges into recession as if no policy were applied.

Another possibility, that we call "adaptive", is to reduce Θ progressively, in a way that is adapted to firms' average fragility: the government measures the instantaneous value of $\langle\Phi\rangle$ averaged over firms still in activity weighted by production, $\langle\Phi\rangle = \sum_i \Phi_i Y_i / \sum_i Y_i$, and sets Θ as:

$$\Theta = \max(\theta\langle\Phi\rangle, 3), \qquad (t > T), \tag{6.24}$$

where θ is some offset that we chose to be $\theta = 1.25$. This means that only the most indebted firms, whose fragility exceeds the average value by more than 25%, will go bankrupt as the effective threshold Θ is progressively reduced.

A primary reason for focusing on firms' credit limits and not the banking sector directly is that the major risks from the Covid shock are not going to be transmitted via a systemic crisis in the financial sector. It is the private non-banking sector which is facing the brunt of the health measures such as lock-downs. As argued in [135], 2020 is not 2008 — while 2008 was an endogenous crisis transmitted through the interlinked nature of modern finance, the Covid shock is well and truly exogenous. Furthermore, given the experience of the 2008 crisis, central banks and other monetary institutions are well-equipped to prevent any systemic risks within the financial sector — they have indeed "flattened the curve" of financial panic [136]. Hence we focus on a policy which allows firms' to stay afloat and weather the economic shock.

POLICY-2: "HELICOPTER MONEY" Another possibility for mitigating losses due to the pandemic is for the governments to take a more active role. Governments can undertake financing for emergency requirements by issuing debt. However, this mechanism leaves future governments vulnerable to interest rate changes [137]. Thus an alternative is to inject cash in the economy to boost consumption and facilitate recovery [138]. This is often nicknamed "helicopter money". This involves the expansion of the money supply by the central bank with the newly created money lent directly to the government. The central bank then immediately writes off this "loan". The government can then spend it on emergency healthcare requirement or other infrastructure projects.

The distribution of the newly-created reserves has to be intermediated via the banking sector. To overcome this, following Friedman's original proposal [139], central banks can

provide the money directly to households. This is the version of a "helicopter money" drop examined here. This policy has been considered radical due to the fear that an expansion in money supply might lead to runaway inflation. In normal times, there might be support for such a view, but it has been shown that a helicopter drop may not always be inflationary [140]. Given the enormity of the crisis, there have been calls from all corners for central banks to break "taboos" [132, 138, 141, 142] and do what is necessary.

In this work, we implement a helicopter-money drop by assuming that the government distributes money to households multiplying their savings by a certain factor $\kappa > 1$: $S \to \kappa S$. The distribution takes place at the end of the shock, and we study here how the "naive policy" (for which Θ goes back to its baseline value immediately after the shock) can be improved by some helicopter money. In what follows $\kappa = 1.5$ i.e. households' savings are increased by 50% via a one-time money injection.

Of course, the efficacy of helicopter money relies on the consumption propensity c of households. Here, we assume that c recovers its pre-shock value immediately, but a loss of confidence may lead to persistently low values of c, weakening the effect of the policy. The effect of a slow recovery of confidence, through a coupling between the value of c and unemployment, could easily be included in the model, following [53, 54].

IMPLEMENTATION OF POLICIES We now implement these governmental policies. In what follows, we choose $\Delta c/c = 0.3$ and $\Delta\zeta/\zeta = 0.5$ as reasonable values to represent the severity of the Covid shock [143, 144], and T takes values 3 and 9 months. In Figs 6.6,6.7 and 6.8, we present a dashboard of the state of the economy: output and unemployment, financial fragility and the bankruptcy rate, inflation and wages, savings and interest rates. In each case, we present 4 scenarios: a baseline scenario without any policy in the first column ("No policy"), the scenario with the naive policy in the second column ("Naive policy"), with a "helicopter drop" in the third column ("Naive policy + Helicopter") and the situation with the adaptive policy in the fourth column ("Adaptive policy"). In all cases, in the absence of any active policy, the economy collapses into a deep recession, with an output reduced by 2/3 compared to pre-shock levels.

It is instructive to remark how the duration of the shock influences which policy is successful. For a shock that lasts for 3 months, we observe that the economy contracts and remains in a depressed state. A significant loss in wages is observed accompanied by a high unemployment rate. We also observe that a large number of firms go bankrupt and households savings are reduced permanently. Finally, with bankruptcy rate quite high, the interest rate on loans ρ^l also sees a dramatic increase. This L-shaped scenario can be improved by extending the credit limits of the firms for the length of the crisis (naive policy). This improves the situation of the economy even though a temporary contraction in output can not be avoided (V-shaped recovery). The other two policies —

"helicopter money" and adaptive policy — perform similar to the naive policy. These scenarios are presented in Fig. 6.6. Note that at the end of the shock, when c returns to its original value, households start to over-spend with respect to the pre-crisis level, because their savings increase during the shock (mirroring the increase of firms' debt) and they want to spend a fixed fraction of them, see Eq. (6.5). However, this over-spending can be insufficient to drive back the economy to its pre-crisis state.

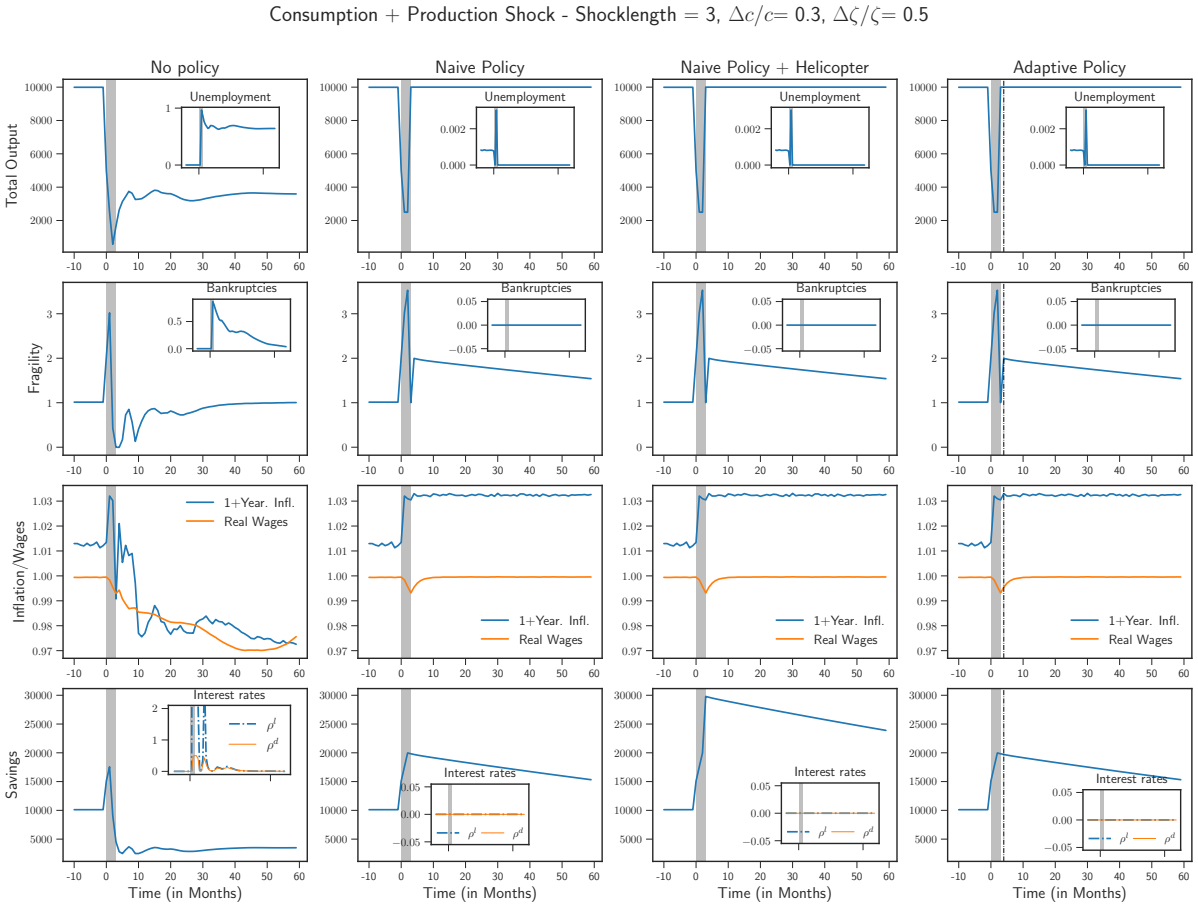


Figure 6.6: Scenarios for shock length of 3 months with and without policy is shown. *First column:* Without any policy, the economy suffers a deep contraction. *Second column:* The introduction of the naive policy improves the situation of the economy and a V-shaped recovery is observed. The results for the other two policies are equivalent to those with the naive policy

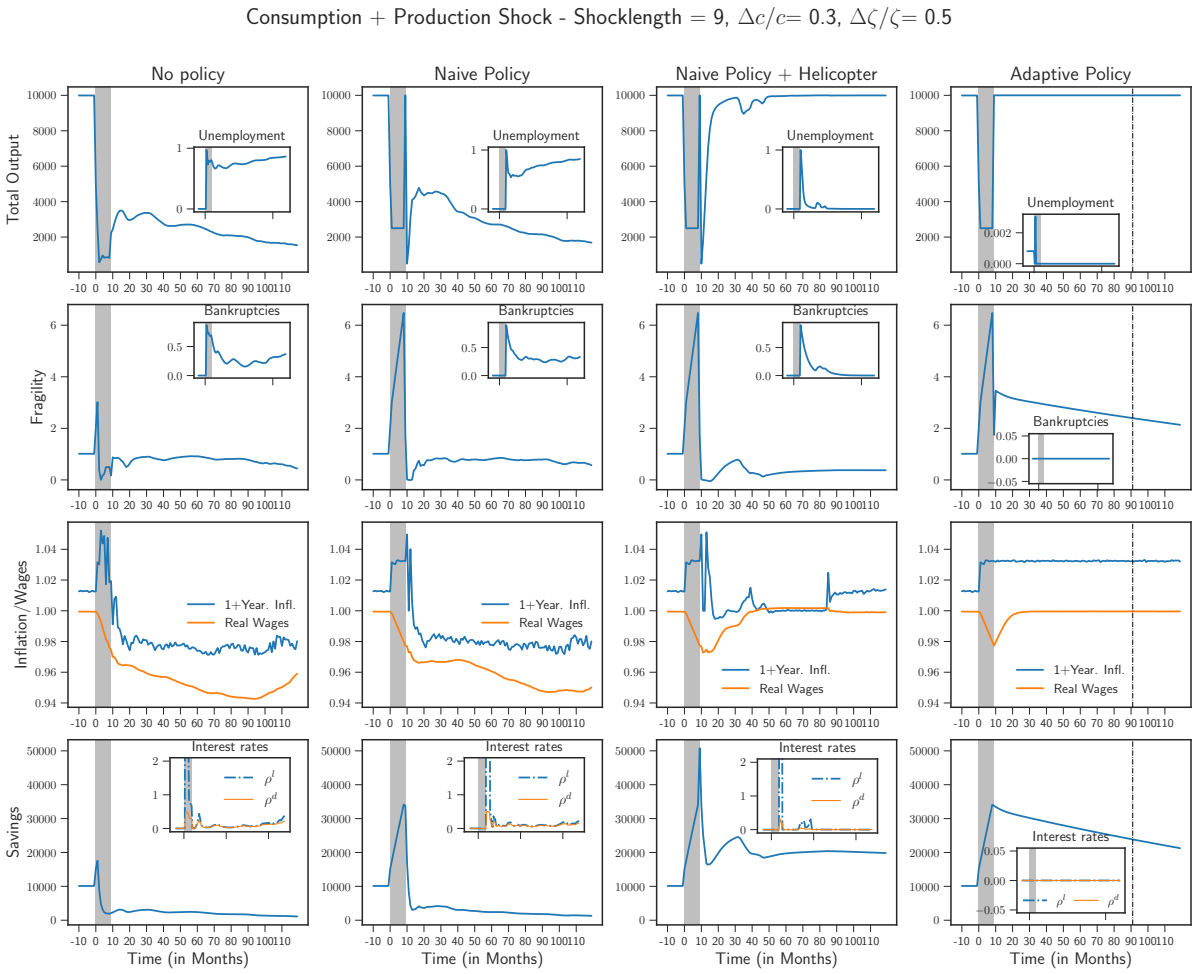


Figure 6.7: Scenarios for shock length of 9 months (marked in gray) with and without policy.*First column:* Similar to Fig. 6.6, the economy undergoes a severe and prolonged contraction. However, given the length of the shock, there is a deeper fall in the level of real wages with firms continuing to go bankrupt far after the shock has occurred. The naive policy (*Second column*) alone isn't successful but the introduction of helicopter money (*Third column*) is able to rescue the economy at the expense of a second milder shock later. Finally, the adaptive policy (*Fourth column*) is able to achieve a smooth recovery with little to no bankruptcies and low unemployment.

The situation for a longer shock duration — 9 months— is markedly different. As before, without any policy intervention, the economy suffers a severe and prolonged contraction, similar to the case for a shock duration of 3 months. However, given the length of the shock, there is a deeper fall in the level of real wages with firms continuing to go bankrupt far after the shock has ended. The introduction of the naive policy

in this case is unable to rescue the economy — extending credit limits of the firms during the crisis prevents them from going bankrupt during the crisis but as soon as the policy is removed, these indebted firms fail and unemployment shoots up drastically, with further downward pressure on the wages. The introduction of helicopter money improves upon the naive policy.

Nonetheless an interesting effect appears, in the form of a W-shape, or relapse of the economy, *even in the absence of a second lock-down period*. This "echo" of the initial shock is due to financially fragile firms that eventually have to file for bankruptcy when credit lines have been tightened post-shock. This second downturn is however temporary and the economy manages to recover fairly quickly. This experiment shows the importance of boosting consumption when the shock is over. A similar effect would be obtained if instead of the savings S, the consumption propensity c was increased post-lock-down. A combination of the two might indeed lead the economy to a faster recovery as shown in the U-shape recovery in Fig. 6.3.

Finally as shown in Fig 6.7 (fourth column) the adaptive policy is indeed the successful one. A contraction during the crisis is inevitable but the economy recovers 100% of its pre-shock output at the end of the shock. As can be seen from the plot of the average fragility, this comes at the price of $\langle \Phi \rangle$ reaching very high values for a while — for example $\langle \Phi \rangle \approx 6$ — and, much higher post-crisis inflation ($\approx 3\%$). But the slow removal of the easy credit policy allows the economy to revert smoothly to its pre-shock state, with bankruptcies kept low. It is important to note that this policy needs to be implemented for a rather long period of time after the shock (almost 7 years in our simulations).

We also studied the case of a pure, rather severe consumption shock $\Delta c/c = 0.7$ lasting $T = 9$ months in Fig. 6.8. We observe that a prolonged drop in consumption, without any loss in production, can still lead to long-lived crisis. The "naive"policy in this case is not enough to hasten the recovery. Direct cash transfer to households via helicopter money drop helps the economy recover faster but leads to a slow, W-shape recovery. Finally, the "adaptive" policy again works best in keeping unemployment low and ensures a rapid recovery.

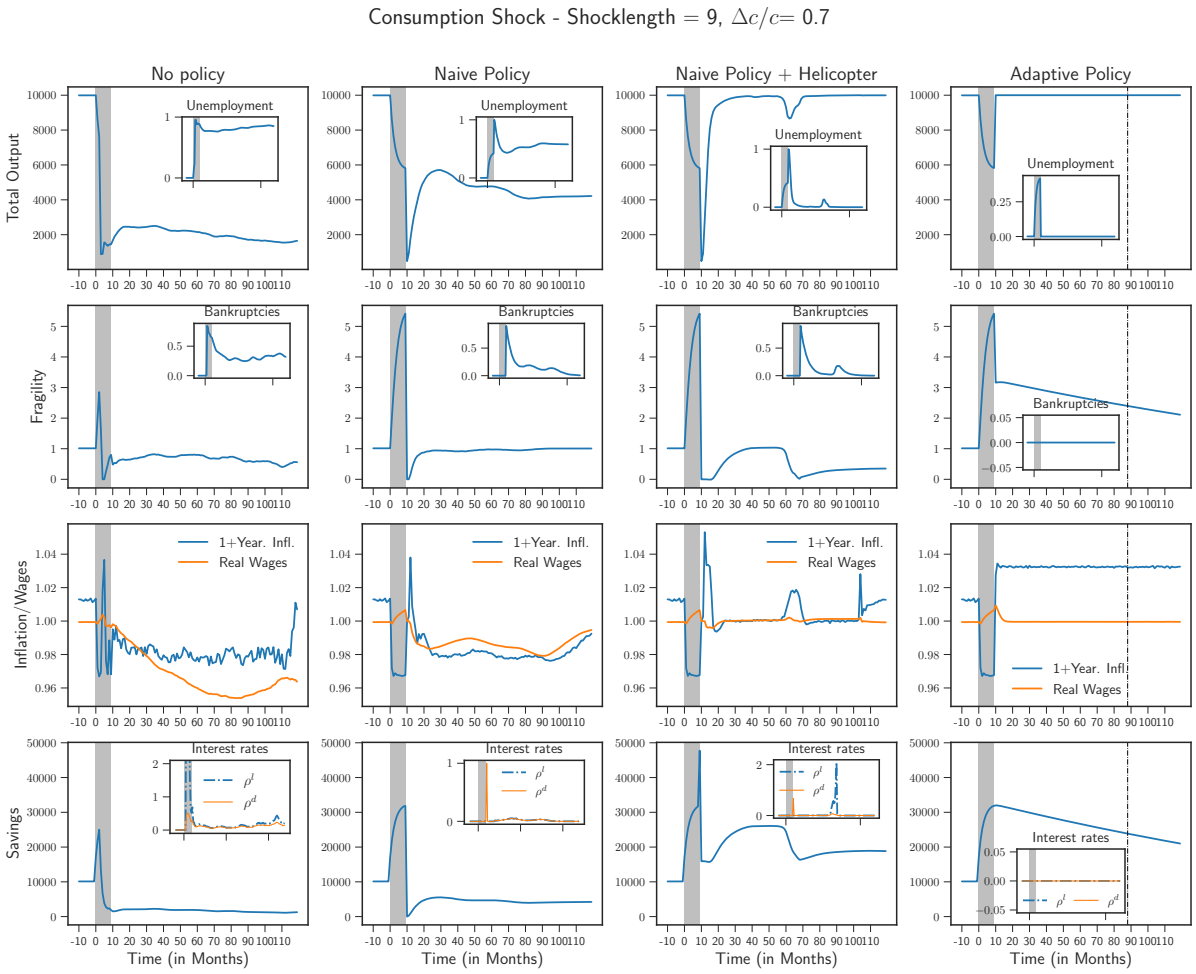


Figure 6.8: Scenarios for a severe consumption shock $\Delta c/c = 0.7$ of length of 9 months (marked in Gray) with and without policy for a pure consumption shock. Similar to the situation shown in Fig. 6.7, a prolonged contraction is observed with only the adaptive policy being the successful one to prevent a permanent contraction.

MULTIPLE LOCKDOWNS While we have discussed the possibility of an endogenous W-shaped scenario in Fig. 6.7, we now discuss how multiple lockdowns affect economic outcomes. Indeed, in many countries, a second series of lockdowns has become unavoidable as the number of Covid-19 cases spiralled out of control in the fall of 2020.

In what follows, we briefly discuss a simple scenario with two lockdowns, each lasting 3 months, with the lockdown being lifted for 4 months between them. We make the further assumption that after the first lockdown, consumption patterns for households and firm productivity do not instantaneously go back to their pre-lockdown values,

but recover in a linear fashion. Once again, we test the four policies discussed above and compare the outcome with the case of a single lockdown. Results are presented in Fig. 6.9.

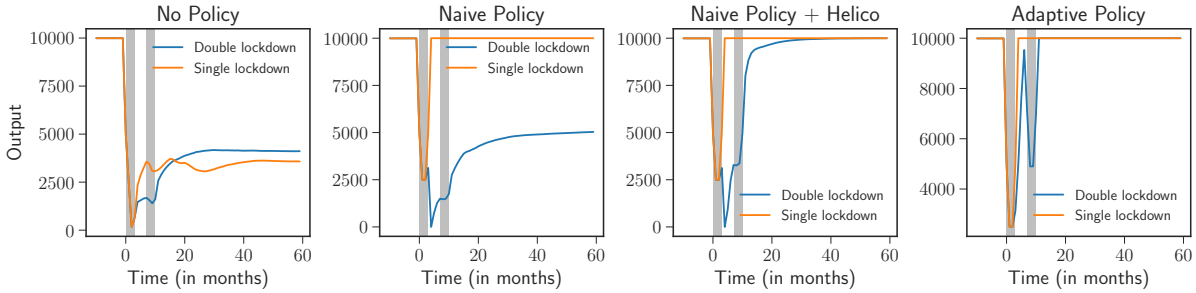


Figure 6.9: Comparison between scenarios following a single lockdown of 3 months and two lockdowns of 3 months with a 4 month lockdown free period. The naive policy, which is able to stabilize the economy in the former case, is unable to do so in the latter. Directly boosting households' savings via helicopter drops improves upon the naive policy. The adaptive policy does best and we observe an exogenous W-shaped scenario.

With two lockdowns, the situation without any policy intervention remains dire and the economy is found in a permanently depressed state. There is a small uptick in output during the lockdown free months, but at the end of the second series of lockdowns, the economy is pushed into a "L-shaped" scenario. The naive policy, when put in place for just the first 3 months, is not sufficient for a quick recovery, even though such a policy was enough for a rapid recovery in the single lockdown case (Fig. 6.6). With consumption habits and productivity coming back only slowly, the naive policy is no longer enough for a quick recovery. Boosting households' savings via helicopter drops helps the economy recover and improves upon the naive policy.

Finally, the best policy is still the adaptive one, which extends supports to firms in guaranteeing that the economy does not tip into a permanently bad state. Interestingly, we observe a W-shaped recovery which is now due to exogenous reasons, namely a second wave of the disease.

6.5 DISCUSSION & CONCLUSION

In this chapter, we have discussed the impact of a Covid-like shock on the toy economy described by the Mark-0 Agent-Based Model. We have shown that, depending on the amplitude and duration of the shock, the model can describe different kinds of recoveries (V-, U-, W-shaped), or even the absence of full recovery (L-shape). Indeed,

as discussed in [44], the non-linearities and heterogeneities of Mark-0 allow for the presence of "tipping points" (or phase transitions in the language of physics), for which infinitesimal changes of parameters can induce macroscopic changes of the economy. The model displays a self-sustained "bad" phase of the economy, characterized by absence of savings, mass unemployment, and deflation. A large enough shock can bring the model from a flourishing economy to such a bad state, which can then persist for long times, corresponding to decades in our time units.

We have also examined how government policies can prevent an economic collapse. We considered two policies: helicopter money for households and easy credit for firms. We find that some kind of easy credit is needed to avoid a wave of bankruptcies, and mixing both policies is effective, provided policy is strong enough. We also highlight that, for strong enough shocks, some flexibility on firm fragility might be needed for long time (a few years) after the shock to prevent a second wave of bankruptcies [145]. Too weak a policy intervention is not effective and can result in a "swoosh" recovery or no recovery at all. Again, a threshold effect is at play, with potentially sharp changes in outcome upon small changes of policy strength.

We note that a key factor in permitting future recovery is to prevent any permanent losses in productive capacity. It is already feared that some firms and sectors facing difficulties may never recover, thus preventing the economy from going back to 100% capacity [133, 146]. Our results then suggest that governments should try to be on the safe side and do "whatever it takes" to prevent the economy to fall in a bad state, and stimulate a rapid recovery. However, we find that when policy is successful, inflation post-crisis is significantly increased compared to the pre-crisis period. This however shouldn't be a cause for concern, the emphasis should be on making a fast recovery. In a recent paper, it has been shown that if central banks base their monetary policy decisions on future expectations of inflation during uncertain times, it can in fact lead to mistakes leading to even more future volatility [147].

Our model hence captures the basic phenomenology of a Covid-like shock and is flexible enough to be used as an efficient "telescope of the mind" [148]. Other extensions and avenues of investigation are possible. For instance, one could impose a different shock on the economy by reducing the value of R (the ratio of firms' hiring and saving propensity), either by increasing the speed at which firms fire their employees, or by reducing the hiring rate.A low enough value of R indeed drives the Mark-0 economy to a bad state (see Fig.6.2). A different policy, which we did not consider here, is for the government to pay the wages directly, allowing firms to maintain their financial health unscathed. This can also easily implemented in Mark-0, but should be roughly equivalent to an increase of Θ. Hence introducing another element like the government or the "public" sector would be a useful extension of the current model.

As discussed above, we have not investigated in this study the standard monetary policy tool, namely interest rate cuts. Whereas such cuts are not expected to play a major role in the short term management of the crisis, their effect on the long-term fate of the economy (in particular when the recovery is L-shaped) can be important and should be examined as well.

Finally, the most needed extension for the current model is the introduction of inequalities — of firm sizes and of household wealth and wages. One of the peculiarities of the Covid shock has been the asymmetry in the way the crisis has affected households, with lower spending by high-income households compounding the situation for low-income households [149]. A disparity in the outcomes following economic shock is already being seen with the prospect of a K-shaped recovery, with one section of the citizenry recovering quickly with the other struggling to recover [150]. Taking into account such heterogeneities in income and effects of the shock thus becomes paramount and will indeed be an important step towards making Mark-0 a better representation of the real-world.

7

SUMMARY AND CONCLUSIONS

> Macroeconomics is about general equilibrium
>
> — Olivier Blanchard, *On the future of macroeconomic models [33]*

> I accuse the classical economic theory of being itself one of these pretty, polite techniques which tries to deal with present by abstracting from the fact that we know very little about the future.
>
> — John Maynard Keynes,*The General Theory of Employment [151]*

Neoclassical economics has held an iron-grip over economic departments all over the world to the detriment of other, more nuanced approaches towards answering questions of economic importance. Central to this world-view is the insistence upon considering economic actors as perfectly rational, utility maximizing agents. For macroeconomics in particular, this world-view results in representative-agent models, where the heterogeneities and bounded rationality are sacrificed for analytical tractability. Most importantly, an insistence on general equilibrium effects misses understanding of phenomena that can emerge beyond equilibrium. In fact, the very dynamism that characterizes modern economies is *de facto* abstracted away. This thesis seeks to develop and apply a complementary approach to macroeconomic modeling through the use of Agent-based models or ABMs.

7.1 SUMMARY OF THE MAJOR RESULTS

In this work two macroeconomic ABMs have been presented. The first of these ABMs was described in Ch 5. Within this model, we described an economy of buyers and producers exchanging goods with the price of goods being set so as to "satisfy" the

budgetary constraints of a maximum number of agents at any given point of time. This model then built upon the paradigm of Constraint Satisfaction Problems (CSPs), a class of models used previously to study complex systems. A numerical examination of this model revealed that our model economy could settle into any one of three phases as a function of allowed debt: the Unstable (U) phase, the Stable (S) phase, and a phase characterized by boom-bust endogenous cycles (EC) [2]. While the U phase remained structure-less due to strong budgetary constraints with no possibility of indebtedness, the S phase presented a remarkable set of properties. Within this phase, the economy self-organizes with goods separating themselves into cheap goods in high demand and expensive goods in low demand. Furthermore, we also observed dynamical switching between these two sets of goods: abundant and cheap goods can lose their luster, due to price fluctuations, and become expensive and rare goods. This "speciation" and switching of goods is also found in the EC phase. However, this phase with high debt levels suffers endogenous cycles with periods of relative calm followed by a wave of bankruptcies.

The CSP which formed the basis for agents' budgetary constraint in this model is related to the perceptron problem. A well-studied problem in machine learning and statistical physics, it was extended to a non-convex optimization problem to study the jamming of hard spheres [74]. The non-convex regime of the perceptron is the interesting one since it allows for a natural introduction of debt within our model economy. Within the ABM presented in Ch 5, agents' preferences, serving as the source of disorder in the perceptron CSP, evolve continuously as a function of supply-demand mismatches and also changes in prices. Hence, in Ch 4, we proposed a dynamical scheme, "Remove & Replace" (R&R), to determine whether a rugged landscape (an RSB phase) persists with dynamic, annealed disorder. We discovered that such a phase can exist and in fact R&R can find zero-energy minima within this rugged landscape. These zero-energy minima correspond to configurations where all constraints in a CSP are satisfied. We dubbed this phenomenon "self-planting" [1].

A second ABM, studied in Ch 6, extended the Mark-0 model [44, 53, 54] to study the impact of economic shutdown following the worldwide coronavirus pandemic. While other approaches towards understanding the economic impact due to COVID-19 design a model around a plausible scenario, the minimal Mark-0model is able to generate multiple plausible scenarios following the Covid-19 shock: we discovered 4 possibilities characterized by a letter of the alphabet: V (rapid recovery), U (sharp downturn followed by slow recovery), W(sharp downturn and recovery followed by a second recession) and L (permanent loss in output). Multiple policies, including the much debated "helicopter money" drop, were studied to mitigate the effects of the Covid shock [3]

The two models hence serve to demonstrate the versatility of the agent-based approach. On the one hand, we were able to elucidate endogenous cycles while on the other hand, we could simulate exogenous shocks and glean useful insights, inaccessible otherwise. Furthermore, by not shying away from using computer simulations to assess and understand the economy, the ABM approach moves beyond the realm of Keynes' "pretty, polite" techniques and seeks to model the economy as closely as possible, warts and all.

Nevertheless, agent-based models still have some way to go in terms of matching real-world data, even though they have had much success in reproducing qualitatively many of the stylized facts observed in the economy. Once again given their flexibility, agent-based models are hard to calibrate. Substantial progress has however been made in this direction with the development of newer machine learning methods [152] and more classical Bayesian methods [153, 154]. Calibration of the models presented in this thesis remains difficult due to their simplistic nature.

While calibrating the models to real-world data remains a possibility, a conscious decision has been taken in this thesis to embark on an alternative route. This thesis proposes the use of ideas of statistical mechanics not just as a source of analytical techniques but for building a new methodology. The proposed methodology involves the reduction of the often high-dimensional parameter-space of the model to a few important parameters. Underlying this methodology is the concept of the "phase diagram". We suggest that a crucial first step towards the understanding of *any* ABM should be the description of its phase diagram. This allows the modeler to quickly gauge the different regimes of the model while reducing the complexity of the model. Furthermore, an examination of the phase diagram also provides a rapid overview of the more interesting regimes of the model, namely the phase boundaries or "dark corners".

7.2 PERSPECTIVES

While ABMs are quite flexible, this flexibility is also limiting. Over the course of the thesis while multiple phenomena were explored, many other paths were left unexplored. A natural first extension of the CSP-based ABM is the introduction of a banking institution. In this model, credit is available for "free"— there is no cost for accruing debt. Integrating a banking system via a banking network or via a representative bank to set a price on debt i.e. interest rates is hence important. With the introduction of interest rates, we will be able to model the effect of monetary policy to the real economy, since the budgetary constraints in the model couple the level of debt with production output.

For the same model the perceptron CSP formed the basis for the budgetary constraint. However, the imposition of positive prices via a linear constraint, though simple, left the problem within the realm of convex optimization with a unique minima for a given set of agent preferences. This model's non-convex version presented in Ch 4 proved unconvincing as an economic model since the spherical constraint led to negative prices for the goods. Hence, an important extension of the present model would be to consider constraints that could capture economic reality but at the same time lead to regimes where multiple (macroscopic) minima are possible. We could then consider models which have multiple *macroscopic* equilibria and consider different questions: the conditions under which these equilibria are stable, the passage from one equilibria to another, equivalence (if any) between different macroeconomic equilibria.

While the first ABM presented here is a prototype, the second ABM, Mark-0, is a more complete model. Nonetheless, an important extension to the model would be the introduction of a government sector. Through the government sector, one could then model possible fiscal policy transmissions and even more interestingly study the interplay between fiscal and monetary policy. A more longer term goal would be to transform Mark-0 to a truly representative model of the economy by integrating heterogeneities in the form of inequalities between firms, and between households. One could then study the impact of economic shocks, exogenous or endogenous, across the income distribution. It is indeed paramount that ABMs begin to provide answers to the problems of inequality which pose risks to the long-term health of the advanced economies.

CONCLUSION While the methodology proposed is built on a technical idea from statistical mechanics, it can have real-world consequences. One (of the many) criticisms leveled against DSGE models is that they make for bad communication devices [33]. So while the DSGE paradigm has all the characteristics of modern science, its use as tool for *explaining* its results to the casual reader remains a lot to be desired.

Such opacity persists not just within the realm of economic modeling but is also a hallmark of regulatory frameworks too - while the first set of Basel regulations was merely 30 pages long, subsequent iterations have tended to added more and more rules with Basel III ending up being more than 1000 pages long [155]. This raises important concerns about the robustness and validity of the framework itself since models integrating such regulatory frameworks in turn become extremely large and liable to over-fitting.

One can however raise the very legitimate question: we live in a highly complex, constantly evolving world; surely our models must be complex as well. Hence is it any wonder that results from a DSGE model can't be readily explained to the general public? It is true that our world is massively interconnected and it is indeed very difficult

to capture its complex, adaptive nature. Nonetheless, everyday billions of people do navigate this complicated world, go to work, earn a living, and (at least in the advanced economies) lead comfortable lives. So how do we do it?

The answer was provided more than half a century ago by Herbert Simon. He concluded that under complex environment, the rational thing to do was to follow *simple*, not complex rules. "Viewed as behaving systems, human beings are actually simple", Simon remarks [156]. Thus, our models should indeed aim for capturing the complexity of the world, but they must be built upon simpler "heuristics". This in essence is the agent-based modeling approach: constructing "bottom-up" models of the world by integrating (as much as possible) how humans truly behave.

7.3 BROADER PURPOSE

> This demands a different science. The "microfoundations" of current economics are precisely what is standing in the way of this. Any new, viable science will either have to draw on the accumulated knowledge of feminism, behavioral economics, psychology, and even anthropology to come up with theories based on how people actually behave, or once again embrace the notion of emergent levels of complexity—or, most likely, both.
>
> — David Graeber, *Against economics [157]*

In any enterprise undertaken over a substantial period of time, it is natural to consider the implications of the work. This is indeed true for doctoral work, and even more so if the work touches upon a subject like economics. Let me then conclude this thesis by sharing some personal reflections on the importance of economic modeling for society at large and also place within this context the work on ABMs presented here.

Economics as a discipline influences every aspect of modern life, and has been the cornerstone of human relation since the dawn of humanity. Economists on the other hand haven't always been as influential. While economists have been providing economic advice since before Adam Smith, their out-sized influence is a more recent phenomenon. Economic modeling itself in its complete mathematical avatar only dates back to within the last 50 years or so. Consequently, as the importance of economists within the public sphere has grown, so has the importance of economic reasoning through economic models.

This however wasn't always the case. As Binyamin Appelbaum in his book explains [158], before the early 1970's, academic economists were barely even present within the halls of power. He goes on to provide a complete historical account starting from the Nixon years through the Volcker era to the current day. Applebaum's book

makes for interesting reading since it details how supposedly bipartisan, apolitical advice can in fact lead to long-term societal consequences[1]. The conclusion is that economic ideas, however innocuous, when linked to policy-making carry enormous power.[2]

TWO EXAMPLES Economic modeling thus also possesses the capacity to influence the real-world. A primary example is the so-called "output gap" which measures the gap between realized GDP and a hypothetical potential GDP. As Adam Tooze in a scathing column titled "Output Gap Nonsense" remarks [160], the statistics of the output gap generated by the IMF and other institutions are misleading at best and obscure the genuine policy debates that are required. The output gap forms a key component in determining fiscal policy within the EU. Since the computation of the output gap is predicated on past data, it can lead to truly perverse effects: for instance after 2008, as the Eurozone went into a recession, the estimates of the output gap were updated downwards. This led to the absurd conclusion in the case of Italy whose economy was declared to be overheating even as unemployment touched 10%. Given that according to output gap estimates Italy was already close to full output, it wasn't allowed to undertake fiscal measures that would have helped it mitigate the consequences of the recession.

Another example of some of the real consequences of the assumptions in economic models once again concerns the Eurozone. In an influential paper, Alesina and Sardagna proposed the intellectual foundation for austerity or large-scale budget cuts [161]. These cuts, they claimed via a rational expectations type argument, when perceived as permanent, would lead consumers to anticipate a reduction in their tax burden and a permanent increase in their lifetime disposable income. The argument relies upon the assumption that consumers are able to compute their lifetime consumption based on the signal sent by the government that spending cuts are imminent, and they are then able to integrate this information to accurately assess their future spending decisions. And for the final nail in the coffin, they suggest that budgets should indeed be cut at times of "fiscal stress" since the signal for future cuts is even more credible. Thus, we find ourselves with massive contraction in government spending when it would be most required.

1 Consider the first example of Nixon ending the military draft in 1973 which he argues has led to a more aggressive foreign policy in the US

2 In India, the country of the author's birth, important topics such as development economics or questions of distribution have given way to utilitarian ideas as well. Insofar as development economics exists, it is mostly studied within a very limited framework of Randomized Control Trials (see [159] for a detailed discussion).

Why do we evoke these two examples? In part to show some of the suspect assumptions within economic models, but in part to also demonstrate the real consequences of these models. Statistics of the output gap and ensuing austerity have had very negative effects for Europe's political economy - "bad estimates of the output gaps ... make for terrible politics" [160]. These two examples also demonstrate that analyses of economic policies are anything but dispassionate or apolitical. Finally, they also serve to show that much of economic thought has drifted away from the purported objectives of the field — to achieve prosperity for all — and has been replaced by Keynes' pretty little games.

Where does that leave us? Circling back to the remark about the communication problems with DSGE models, one can imagine that Blanchard's "reader" is in fact a fellow economist and not say a layperson. This is, in the author's opinion, economics' true communication problem: while the field has advanced, its results which directly affect the livelihoods of people remains opaque.

One might object that technical details of an advanced subject will always necessarily remain inaccessible to the larger public. However, the lack of an attempt to render things accessible to the larger public is a severe hindrance. Barring a few public economists such as Adam Tooze, Thomas Piketty, Daniela Gabor, Mark Blyth, and Yanis Varoufakis, there are vast swaths of economics research which remain away from the public eye.

The expectation that economists engage with the public is not a completely new one either. Keynes once again presents himself as an exemplar of the public economist. Starting from his *Economic Consequences of the Peace* to the *General Theory* to his lesser known works such as *Economic Possibilities for our grandchildren*, he always kept the fundamental objective of economics at the center of his thought: that once the problem of economics is solved, "people will be able to live wisely, agreeably and well".

Where do Agent-based models fit within this discussion? The author believes that ABMs suffer less from DSGE's communication problem. At their core, ABMs integrate human behavior through rules that mimic heuristics. The complexity is indeed emergent but a careful reduction of the parameter space can make the complexity tractable. Furthermore, what ABMs lack in empirical precision, they make up for as scenario generators, narrative builders, and economic story-telling devices. Finally, with computing resources becoming ubiquitous, one can hope that economic ABMs can become accessible to a wider public[3]

We began our thesis with a statement by Keynes where he described a master economist as someone who is not only an economist but also an historian, a mathematician, and a philosopher. The world has transformed since Keynes' day. Our study of agent-based models suggests that we must also integrate insights from the realm of

3 This would also be in the spirit of open and reproducible science.

human behavior through psychology and also the history of human society. We must then update Keynes:

Today's master economists must possess a rare combination of interests. They must be interested in history, philosophy, and also psychology and anthropology. They must be at ease with mathematics, computer programming, and ideas from statistical physics. In brief, no part of man's nature and no part of scientific progress must lie entirely outside their regard.

Only then can an economist said to possess, as Marshall put it, as someone "with a cool head but warm heart".

A

DERIVATION OF THE PERCEPTRON PHASE DIAGRAM

In this appendix, we present the details of the computation of the phase-diagram for the modified perceptron model studied in Chapter 4.

A.1 SETTING UP THE PROBLEM

Let us recall the variables at play in the problem. Let $h_\mu(X)$ stand for the $\mu = 1...M$ constraints. The constraints have the form

$$h_\mu(\vec{X}) = \frac{1}{\sqrt{N}}\vec{X}\cdot\vec{\xi}_\mu - b_\mu - \sigma, \tag{A.1}$$

where $\vec{X}$ is an N -dimensional vector, $\vec{\xi}_\mu$ is the price vector for each agent μ, b_μ is the budget for the agent and σ is some overall constant. We take the ξ^i_μ to be normally distributed with variance 1. The budget for each agent is also taken from a gaussian distribution with mean 0 and variance σ.

The constraint satisfaction problem (CSP) is the determination of all $\vec{X}$ such that for each μ, we have

$$h_\mu(\vec{X}) > 0. \tag{A.2}$$

with the added constraint

$$\frac{1}{N}\sum_i^N X_i \quad = \quad m \tag{A.3}$$

To this end, we introduce a term corresponding to the addition of a magnetic field which forces the $\vec{X}$ to point in a particular direction. Thus, we have the following Hamiltonian,

$$\mathrm{H}[\vec{X}] = \frac{1}{2}\sum_{i=1}^{M} h_\mu(\vec{X})^2\theta(-h_\mu) + \Upsilon\sum_{i}^{N} X_i, \tag{A.4}$$

where Υ is the *global* magnetic field. To simplify the notations in what follows, we write $v(h) = \frac{1}{2}h^2\theta(-h)$.

A.2 CALCULATION OF THE PARTITION FUNCTION

We want to compute the partition function

$$Z = \int\prod_i \mathrm{d}X_i e^{-\beta \mathrm{H}(X_i)} = \int\prod_i \mathrm{d}X_i \exp(-\beta\sum_{i=1}^{M} v(h_\mu) - \beta\Upsilon\sum_{i=1}^{N} X_i). \tag{A.5}$$

Note that we are omitting factors of 2π and other normalization factors. We introduce the variable r_μ and rewrite the partition function as

$$Z = \int\prod_i \mathrm{d}X_i \prod_{\mu=1}^{M}\Big(\int \mathrm{d}r_\mu e^{-\beta v(r_\mu-\sigma)}\delta(r_\mu - \frac{1}{\sqrt{N}}\vec{X}\cdot\vec{\xi}_\mu + b_\mu)\Big)\prod_i e^{-\beta\Upsilon N X_i} \tag{A.6}$$

Integrating over the variable r_μ sets the value of r_μ to h_μ. To compute the free energy, we must compute $\overline{\ln Z}$. The disordered average of $\ln Z$ is a difficult quantity to compute directly. We thus employ the "replica trick" [96]. This involves the introduction of n copies of the system and using them the following identity to compute $\overline{\ln Z}$:

$$\overline{\ln Z} = \lim_{n\to 0}\frac{\overline{Z^n}}{n} \tag{A.7}$$

We thus compute $\overline{Z^n}$ for integer n and write the free-energy as a saddle-point over the $n\times n$ replica overlap matrix

$$q_{ab} = \frac{1}{N}\vec{X}_a\cdot\vec{X}_b \tag{A.8}$$

Introducing n replicas to perform the disorder average and compute the quenched free energy, we have

$$Z^n = \int \Big(\prod \mathrm{d}\vec{X}_a\Big)\Big(\prod_{a\mu} \mathrm{d}r_a^\mu \mathrm{d}\hat{r}_a^\mu\Big)\exp\Big(\sum_{a\mu}\Big[i\hat{r}_a^\mu r_a^\mu - \frac{1}{\sqrt{N}}\vec{X}_a\cdot\vec{\xi}^\mu + b_\mu) - \beta v(r_a^\mu - \sigma)\Big] + \sum_{ai} -\beta\Upsilon X_a^i\Big) \tag{A.9}$$

Here we have used the integral representation of the delta function via the new auxiliary variable $\hat{r}$. We now proceed to perform the disorder average over the variables ξ and b. This gives, for ξ,

$$\overline{\exp\Big(-\frac{1}{\sqrt{N}}\sum_{a\mu} i\hat{r}_a^\mu \vec{X}_a\cdot\vec{\xi}^\mu\Big)} = \exp\Big(-\frac{1}{2N}\sum_{ab\mu}\hat{r}_a^\mu\hat{r}_b^\mu\vec{X}_a\cdot\vec{X}_b\Big), \tag{A.10}$$

and for b, we get

$$\overline{\exp\Big(\sum_{a\mu} i\hat{r}_a^\mu b_\mu\Big)} = \exp\Big(-\frac{\eta}{2}\sum_{ab\mu}\hat{r}_a^\mu\hat{r}_b^\mu\Big) = \exp\Big(-\frac{\eta}{2}\sum_{ab\mu}\hat{r}_a^\mu\hat{r}_b^\mu\Big) \tag{A.11}$$

Putting it all back together, we get

$$\overline{Z^n} = \int\Big(\prod_{ia}\mathrm{d}X_a^i\Big)\Big(\prod_{a\mu}\mathrm{d}\hat{r}_a^\mu\mathrm{d}r_a^\mu\Big)\exp\Big[-\frac{1}{2}\sum_{ab\mu}\hat{r}_a^\mu\hat{r}_b^\mu\Big(\frac{\vec{X}_a\cdot\vec{X}_b}{N}+\eta\Big) + \sum_{a\mu} i\hat{r}_a^\mu r_a^\mu - \beta\sum_{a\mu} v(r_a^\mu-\sigma) - \beta\Upsilon\sum_{ai}X_a^i\Big]. \tag{A.12}$$

We introduce the variable $q_{ab} = \frac{1}{N}\vec{X}_a\cdot\vec{X}_b$, which determines the overlap between two replicas. These variables are computed variationally. Thus this gives us

$$\begin{aligned}\overline{Z^n} = &\int\prod_{ab}\mathrm{d}\hat{q}_{ab}\mathrm{d}q_{ab}\exp\Big(i\frac{N}{2}\sum_{ab}q_{ab}\hat{q}_{ab}\Big)\\ &\times\prod_i\Big[\int\prod_a\mathrm{d}\vec{X}_a\exp\Big(-\beta\Upsilon\sum_a X_a - \frac{i}{2}\sum_{ab}\hat{q}_{ab}\vec{X}_a\cdot\vec{X}_b\Big)\Big]\\ &\times\prod_\mu\Big[\int\prod_a\mathrm{d}\hat{r}_a^\mu\mathrm{d}r_a^\mu\exp\Big(-\frac{1}{2}\sum_{ab}\hat{r}_a^\mu\hat{r}_b^\mu(q_{ab}+\eta) + i\sum_a\hat{r}_a^\mu r_a^\mu - \beta\sum_a v(r_a-\sigma)\Big)\Big].\end{aligned} \tag{A.13}$$

Following [75], the third and last term in the equation above can be reformulated as, we note it as A_3:

$$A_3 = \alpha N \log\Big(\exp\Big[\frac{1}{2}\sum_{ab}(q_{ab}+\eta)\frac{\partial^2}{\partial h_a \partial h_b}\prod_c e^{-\beta v(h_c)}\Big|_{h_c=-\sigma}\Big]\Big), \tag{A.14}$$

with $\alpha = \frac{M}{N}$. The second term in A.13, contains a gaussian integral. We write it as

$$A_2' = \int \prod_a \mathrm{d}X_a \exp\Big(-\beta\Upsilon\sum_a X_a - \frac{i}{2}\sum_{ab} q_{ab}X_aX_b\Big), \tag{A.15}$$

and hence the second term can be written as $A_2 = e^{N\log A_2'}$. The integral A_2 is a standard integral and we get,

$$A_2' = \frac{(\sqrt{2\pi})^n}{\sqrt{\det(-\hat{q})}} e^{\sum_{ab}\frac{(\beta\Upsilon)^2}{2}(\hat{q}^{-1})_{ab}}, \tag{A.16}$$

which gives for A_2,

$$A_2 = \frac{n}{2}\log(2\pi) - \frac{1}{2}\log(\det(-\hat{q})) + \frac{(\beta\Upsilon)^2}{2}\sum_{ab}\hat{q}^{-1}_{ab}. \tag{A.17}$$

Writing all these integrals together, we get

$$\overline{Z^n} = \int \prod_{ab} \mathrm{d}\hat{q}_{ab}\mathrm{d}q_{ab} \exp\Big[\sum_{ab}\Big(\frac{iN}{2}\hat{q}_{ab}q_{ab} + \frac{(\beta\Upsilon)^2}{2}\hat{q}^{-1}_{ab}\Big) - \frac{1}{2}\log(\det(-\hat{q})) + A_3\Big] \tag{A.18}$$

Following [162], we can compute the integral over $\hat{q}_{ab}$, by performing a saddle point approximation where we have neglected terms of order n^2 in the calculation. These terms only supply $\frac{1}{N}$ corrections to the eventual replicated partition function. We get finally

$$\overline{Z^n} = \int \prod_{ab} \mathrm{d}q_{ab} e^{NS(q,\Upsilon,\beta)} \tag{A.19}$$

$$\begin{aligned} S(q,\Upsilon,\beta) &= \frac{1}{2}\log\det(q) + \frac{(\beta\Upsilon)^2}{2}\sum_{ab} q_{ab} \\ &\quad + \alpha\log\Big(\exp\Big[\frac{1}{2}\sum_{ab}(q_{ab}+\eta)\frac{\partial^2}{\partial h_a\partial h_b}\prod_c e^{-\beta v(h_c)}\Big|_{h_c=-\sigma}\Big]\Big), \end{aligned} \tag{A.20}$$

Here once again, we have neglected the prefactors of $\log(2\pi)$ etc.

A.3 HIERARCHICAL ANSATZ FOR q_{ab}

A.3.1 *Some preliminary details*

The replica overlap matrix q_{ab} can be written as a function on the interval $[0,1]$[96]. From the Parisi algebra, we know that

$$\lim_{n\to 0}\frac{1}{n}\log\det(q) = \log(1-q_M) + \frac{q_M}{\lambda(0)} - \int_0^1 \mathrm{d}x\frac{\dot{q}(x)}{\lambda(x)}, \tag{A.21}$$

with

$$\lambda(x) = 1 - xq(x) - \int_x^1 \mathrm{d}q(y), \tag{A.22}$$

from which we note that, $\dot{\lambda}(x) = -x\dot{q}(x)$. For the limiting cases, we have for $\lambda(x)$,

$$\lambda(0) = \lambda(x_m) = 1 - \int_0^1 \mathrm{d}y q(y) \tag{A.23}$$

$$\lambda(1) = \lambda(x_M) = 1 - q_M \tag{A.24}$$

Following [163], we have that,

$$\lim_{n\to 0}\frac{1}{n}\log\left(\exp\left[\frac{1}{2}\sum_{ab}(q_{ab}+\eta)\frac{\partial^2}{\partial h_a \partial h_b}\prod_c e^{-\beta v(h_c)}\Big|_{h_c=-\sigma}\right]\right) = \gamma_{q_m+\eta} * f(0,h)|_{h=-\sigma}, \tag{A.25}$$

where we have defined $*$ as the convolution operation

$$\gamma_q * g(h) = \int \frac{\mathrm{d}q}{\sqrt{2\pi q}} e^{-\frac{y^2}{2q}} g(h-y) \tag{A.26}$$

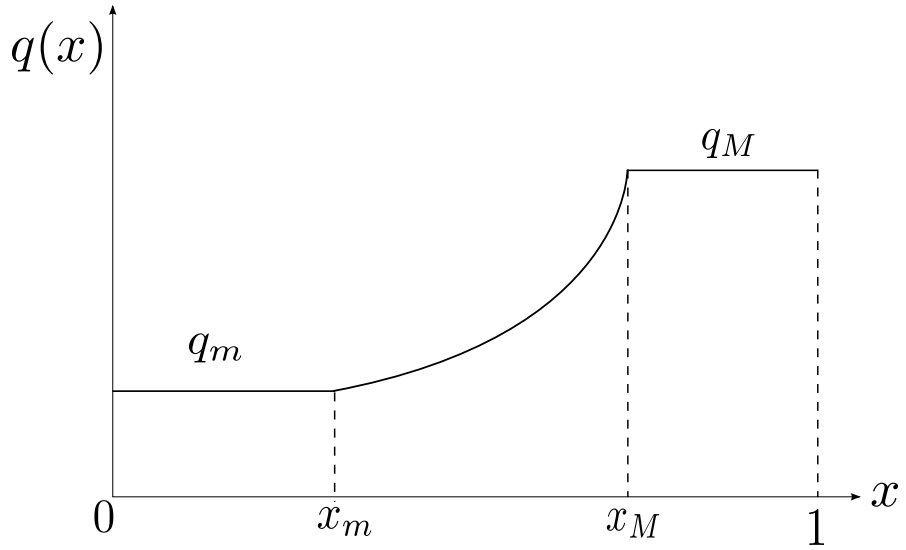


Figure A.1: The form of the function $q(x)$ under the Parisi ansatz in the Replica Symmetry breaking case.

Here, $f(x,h)$ satisfies Parisi's equation

$$\frac{\partial f}{\partial x} = -\frac{1}{2}\frac{\mathrm{d}q}{\mathrm{d}x}\left[\frac{\partial^2 f}{\partial h^2} + x\left(\frac{\partial f}{\partial h}\right)^2\right], \quad x_m < x < x_M \tag{A.27}$$

For $x \notin [x_m, x_M]$, $\dot{q}(x) = 0$ and hence $\frac{\partial f}{\partial x} = 0$, that is $f(x)$ is independent of x. For the boundary condition, we write $f(1,h) = \log g(1,h)$, where $g(m_K,h)$ is the value of the function g restricted to the largest sub-block in the Parisi hierarchical ansatz for the matrix q_{ab} [96, 163]. Thus, we have

$$g(m_K,h) = \exp\left(\frac{1}{2}(q_{K+1} - q_K)\frac{\partial^2}{\partial h^2}\prod_{a=1}^{n} e^{-\beta v(h_a)}\Big|_{h_a=-\sigma}\right) \tag{A.28}$$

This can then be written in the limit $m_K \to 1$,

$$f(1,h) = \log \gamma_{1-q_M} * e^{-\beta v(h)} \tag{A.29}$$

The equations for f are written in terms of the variable x whilst the parameter in the equation is q. We introduce $x(q)$ that is the inverse of $q(x)$ in the interval $[q_m, q_M]$. We then recast the Parisi equation and the boundary conditions for $f(q,h) = f(x(q),h)$ as follows

$$\begin{aligned} \dot{f}(q) = & \quad -\frac{1}{2}\Big[f''(q,h) + x(q)f'(q,h)^2\Big], \qquad q_m < x < q_M \\ f(1,h) = & \qquad f(q_M,h) = \log \gamma_{1-q_M} * e^{-\beta v(h)} \end{aligned} \tag{A.30}$$

The last step is to recast $\lambda(x)$ as $\lambda(q)$. We finally have

$$\begin{aligned} -\beta F = & \quad \frac{1}{2}\Big[\log(1-q_M) + \frac{q_m}{\lambda(q_m)} - \int_{q_m}^{q_M} \frac{\mathrm{d}q}{\lambda(q)}\Big] + \alpha \gamma_{q_m} * f(q_m,h)|_{h=-\sigma} \\ \lambda(q) = & \qquad\qquad\qquad\qquad\qquad\qquad 1 - q_M + \int_q^{q_M} x(p)\mathrm{d}p \end{aligned} \tag{A.31}$$

THE ADDITIONAL MAGNETIC FIELD TERM The introduction of the magnetic field Υ in the system introduced the term $\frac{(\beta\Upsilon)^2}{2}\sum_{ab} q_{ab}$ to the function replicated free energy. In the continuous hierarchical ansatz for the matrix q_{ab}, we find that the sum $\sum_{ab} q_{ab}$ is written as :

$$\lim_{n\to 0} \frac{1}{n} \sum_{ab} q_{ab} = 1 - \int_0^1 \mathrm{d}q(x). \tag{A.32}$$

We express this in terms of the inverse function $x(q)$, to obtain that

$$\frac{(\beta\Upsilon)^2}{2} \sum_{ab} q_{ab} = \frac{(\beta\Upsilon)^2}{2}\Big[1 - q_M - \int_{q_m}^{q_M} x(p)\mathrm{d}p\Big] \tag{A.33}$$

Hence, the replicated free energy with the hierarchical ansatz for q_{ab}, reads

$$\begin{aligned} -\beta F[x(q)] = & \frac{1}{2}\Big[\log(1 - q_M + \frac{q_m}{\lambda(q_m)} - \int_{q_m}^{q_M} \frac{\mathrm{d}q}{\lambda(q)}\Big] + \alpha \gamma_{q_m+\eta} * f(q_m,h)|_{h=-\sigma} \\ & + \frac{(\beta\Upsilon)^2}{2}\Big[1 - q_M - \int_{q_m}^{q_M} x(p)\mathrm{d}p\Big] \end{aligned} \tag{A.34}$$

A.4 REFORMULATION IN TERMS OF THE MAGNETIZATION M

We perform a simple Legendre transform on the complete free energy A.34. We write the free energy as

$$\begin{aligned} -\beta F[x(q)] &= K + \frac{(\beta\Upsilon)^2}{2}\Big[1 - q_M - \int_{q_m}^{q_M} x(p)\mathrm{d}p\Big] \\ &= K + \frac{(\beta\Upsilon)^2}{2}\lambda(q_m), \end{aligned} \tag{A.35}$$

where

$$K = \frac{1}{2}\Big[\log(1-q_M) + \frac{q_m}{\lambda(q_m)} - \int_{q_m}^{q_M}\frac{\mathrm{d}q}{\lambda(q)}\Big] + \alpha\gamma_{q_m+\eta} * f(q_m, h) \tag{A.36}$$

Let $m = \frac{\partial F}{\partial \Upsilon}$. Then $m = -\beta\Upsilon\lambda(q_m)$. Hence, we obtain

$$\Upsilon = -\frac{m}{\beta\lambda(q_m)} \tag{A.37}$$

Hence, the Legendre transformed free energy which we write as $F'[x(q), m] = F[x(q)] - m\Upsilon$. Performing the computation we get:

$$-\beta F[x(q), m] = \frac{1}{2}\Big[\log(1-q_M) + \frac{q_m}{\lambda(q_m)} - \int_{q_m}^{q_M}\frac{\mathrm{d}q}{\lambda(q)}\Big] + \alpha\gamma_{q_m+\eta} * f(q_m, h) - \frac{m^2}{2\lambda(q_m)} \tag{A.38}$$

A.5 REPLICA SYMMETRIC SOLUTION

In the replica symmetric case, we have $q_{ab} = q_M$ for all $a \neq b$. Thus in this case, the free energy reduces to

$$-\beta F_{RS}(q_M) = \frac{1}{2}\Big[\log(1-q_M) + \frac{q_M}{1-q_M}\Big] + \alpha\gamma_{q_M+\eta} * f(q_M, h)\Big|_{h=-\sigma} + \frac{(\beta\Upsilon)^2}{2}(1-q_M) \tag{A.39}$$

We are interested in the zero temperature limit. In this limit, we can write the function f as

$$\begin{aligned} f(q_M, h) &= \lim_{\beta\to\infty} \log \gamma_{1-q_M} * e^{-\beta v(h)} \\ &= \log \int_{-\infty}^{h} \frac{\mathrm{d}z}{\sqrt{2\pi(1-q_M)}} e^{-\frac{z^2}{2(1-q_M)}} \end{aligned} \tag{A.40}$$

The integral above can be recast in terms of the error function. We introduce $\Theta(x) = \frac{1}{2}(1 + \mathrm{Erf}(x))$. We thus obtain for $f(q_M, h)$

$$f(q_M, h) = \log \Theta\Big(\frac{h}{\sqrt{2(1-q_M)}}\Big) \tag{A.41}$$

A.5.1 *SAT Phase*

In the SAT phase, $-\beta F[x(q), m]$ has a finite limit. The saddle point equation for q_M in this case gives:

$$\frac{q_M}{(1-q_M)^2} = \alpha \int_{-\infty}^{\sigma} \frac{\mathrm{d}h}{\sqrt{2\pi(q_M+\eta)}} e^{-\frac{h^2}{2(q_M+\eta)}} \frac{(h-\sigma)^2}{(1-q_M)^2} - \frac{m^2}{2(1-q_M)^2} \tag{A.42}$$

A.5.2 *UNSAT Phase*

In the UNSAT phase, we have $q_M = 1 - \chi T$ and hence, the function f is in the low-temperature limit:

$$f(1 - \frac{\chi}{\beta}, h) \simeq -\frac{\beta h^2}{2(1+\chi)} \theta(-h) \tag{A.43}$$

and hence the energy is

$$E_{RS} = -\frac{1-m^2}{2\chi} + \frac{\alpha}{2(1+\chi)} \Big[\int_{-\infty}^{\sigma} \frac{\mathrm{d}h}{\sqrt{2\pi(1+\eta)}} e^{-\frac{h^2}{2(1+\eta)}} (h-\sigma)^2 \Big] \tag{A.44}$$

We name the integral in the equation above I. I can then be expressed as:

$$I = \frac{1-m^2}{\alpha_J} \tag{A.45}$$

Hence, we get E_{RS},

$$E_{RS} = -\frac{1-m^2}{2} \Big[\frac{\alpha}{\alpha_J} - \frac{1}{\chi} \Big] \tag{A.46}$$

We finally express χ in terms of the other parameters, by setting $\frac{dE_{RS}}{d\delta} = 0$. We get hence

$$\left(1 + \frac{1}{\chi}\right)^2 = \frac{\alpha}{\alpha_J(\eta, \sigma, m)} \tag{A.47}$$

Finally, we get for E_{RS}

$$E_{RS} = \frac{1 - m^2}{2}\left(\sqrt{\frac{\alpha}{\alpha_J}} - 1\right)^2 \tag{A.48}$$

A.5.3 *Jamming Limit*

The jamming limit demarcates the boundary between the SAT and UNSAT phases. It is attained in the limit $q_M \to 1$. Using the properties of the error function Erf(x), we obtain

$$\alpha_J(\eta, m, \sigma) = (1 - m^2)\left[\int_{-\infty}^{\sigma} \frac{dh}{\sqrt{2\pi(1+\eta)}} e^{-\frac{h^2}{2(1+\eta)}(h-\sigma)^2}\right]^{-1} \tag{A.49}$$

A.5.4 *Stability Analysis*

We now proceed to study the stability of the replica-symmetric phase and how replica-symmetry is broken. Replica-symmetry breaking (RSB) phase is characterized by a continuous d'Almeida-Thouless instability (dAT) [105]. To compute the dAT continuous instability, we investigate the eigenvalues of the Hessian matrix,

$$H_{ab;cd} = \frac{d^2(-\beta F[q])}{dq_{ab}dq_{cd}}\Big|_{q_{a\neq b}=q_M} \tag{A.50}$$

where q_M satisfies the Replica-symmetric saddle point equation (A.42). The instability occurs when an eigenvalue of this Hessian is zero.

Rather than computing the full Hessian, we follow a different procedure. First off, we add lagrange parameters $P(q_M, h)$ and $P(q, h)$ to set the differential equation satisfied by the function f as well as the boundary conditions (A.30). Putting everything together we get:

$$
\begin{aligned}
-\beta F[x(q),m] = \frac{1}{2}\Big[& \log(1-q_M) + \frac{q_m}{\lambda(q_m)} + \int_{q_m}^{q_M} \frac{\mathrm{d}q}{\lambda(q)} \Big] \\
& + \alpha \gamma_{q_m+\eta} * f(q_m,h)|_{h=-\sigma} \\
& - \alpha \int \mathrm{d}h P(q_M,h) \Big[f(q_M,h) - \log \gamma_{1-q_M} * e^{-\beta v(h)} \Big] \\
& + \alpha \int \mathrm{d}h \int_{q_m}^{q_M} P(q,h) \Big[\dot{f}(q,h) + \frac{1}{2}\Big(f''(q,h) + x(q) f'(q,h)^2 \Big) \Big] \\
& - \frac{m^2}{2\lambda(q_m)}
\end{aligned}
\tag{A.51}
$$

The variational of the free energy with respect to $x(q)$ gives

$$
\frac{q_m - m^2}{\lambda(q_m)} + \int_{q_m}^{q} \frac{\mathrm{d}p}{\lambda(p)^2} = \alpha \int \mathrm{d}h P(q,h) f'(q,h)^2 \tag{A.52}
$$

Taking a further derivative with respect to q gives:

$$
\frac{1}{\lambda(q)^2} = \alpha \int \mathrm{d}h P(q,h) f''(q,h)^2 \tag{A.53}
$$

This last equation will serve us to understand where the limit between the convex and non-convex phases lies. The line of instability can be retrieved by approaching the instability point from the RS phase. We remark that as we approach the RSB phase from the Replica-symmetric phase, the function $q(x)$ becomes non-constant but this change in behavior of the function happens continuously. Hence close to the instability point, the function $q(x)$ is close to a constant and hence A.53 reduces to its expression as computed on the RS solution to give,

$$
\frac{1}{(1-q_M)^2} = \alpha \int_{-\infty}^{\infty} \mathrm{d}h \gamma_{q_M+\eta}(h+\sigma) \Big[\frac{\mathrm{d}^2}{\mathrm{d}h^2} \log \gamma_{1-q_M} * e^{-\beta v(h)} \Big]^2 \tag{A.54}
$$

We must then compute this expression on the value of q_M that satisfies the (A.42) which in turn defines the instability line $\alpha_{\mathrm{dAT}}(\sigma)$.

1. **SAT Phase** : In the SAT Phase, the equation A.54 above becomes:

$$
\frac{1}{(1-q_M)^2} = \alpha \int_{-\infty}^{\infty} \mathrm{d}h \gamma_{q_M+\eta}(h+\sigma) \Big[\frac{\mathrm{d}^2}{\mathrm{d}h^2} \log \Theta\Big(\frac{h}{\sqrt{2(1-q_M)}} \Big) \Big]^2 \tag{A.55}
$$

At first sight, the magnetization m doesn't seem to effect the stability. However, q_M depends implicitly on the value of m via the saddle-point equation (A.42). Thus, to find the dAT line within the SAT phase, the equation above is computed with the value of q_M satisfying the saddle point equation for a fixed value of m.

2. **UNSAT Phase** : In the UNSAT phase, A.54 reduces to

$$\left(1+\frac{1}{\chi}\right)^2 = \alpha \int_{-\infty}^{\sigma} \frac{\mathrm{d}z}{\sqrt{2\pi(1+\eta)}} e^{-\frac{z^2}{2(1+\eta)}} \tag{A.56}$$

Injecting the expression for δ, we get

$$1-m^2 = \frac{\int_{-\infty}^{\sigma} \mathrm{D}h_\eta (h-\sigma)^2}{\int_{-\infty}^{\sigma} \mathrm{D}h_\eta} \tag{A.57}$$

where $\mathrm{D}h_\eta$ stands for the gaussian measure.

$$\mathrm{D}h_\eta = \frac{\mathrm{d}h}{\sqrt{2\pi(1+\eta)}} e^{-\frac{h^2}{2(1+\eta)}} \tag{A.58}$$

Equation A.57 couples m to the parameter σ. That is, for a fixed value of $m = m_0$, we look for the value of σ^* which satisfies the equation A.57. Then $\forall \alpha > \alpha_J(m_0, \eta, \sigma^*)$, we are in the non-convex regime in the UNSAT phase.

A.5.5 *Phase diagram*

We observe that even for large values of m (~ 0.9), replica symmetry breaking occurs in the UNSAT phase. In figure A.2, we present the the complete phase diagram with the dAT instabilities in the SAT and UNSAT phases for $m = 0.5$. From the stability analysis, we remark that the jamming line is exact only for the case where $\sigma > 0$ or $\sigma > \sigma^*$. σ^* is the solution of the equation A.57. Beyond this point we have the RSB phase or the non-convex regime. Similarly, in the SAT phase the $\alpha_{dAT}(\sigma)$ line signals the onset of the RSB phase. Physically, we observe that we can approach the jamming line from the SAT and UNSAT phases. If we approach the jamming line $\alpha_J(\sigma)$ from the SAT phase, then the volume of the space of solutions that solve the constraints shrinks continuously to zero. Furthermore, the overlap in a cluster of solutions in this phase, characterized by

q_M, tends to 1. On approaching the jamming line from the UNSAT phase, the energy of the system E_{RS} goes to zero continuously.

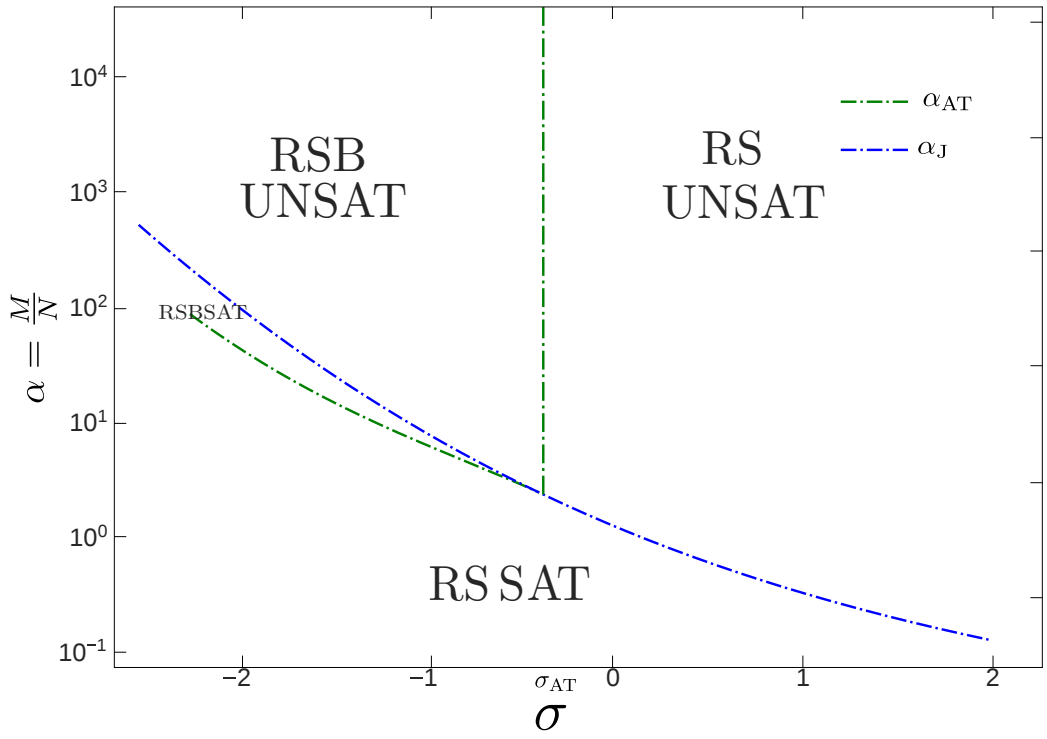


Figure A.2: Complete phase diagram of the perceptron model with the added magnetic field. Here $m = 0.577$

Independently of jamming, in the SAT phase, the volume of the space of solutions is finite. In this situation, the zero-temperature solutions of the system form clusters. Within each of these clusters, the overlap between any two solutions has an overlap q_M. If two solutions are not found within the same cluster, they then have an overlap $q < q_M$. The distribution of the overlap between such solutions is then given by the inverse function $x(q)$.

REFERENCES

[1] D. Sharma, J. P. Bouchaud, M. Tarzia, and F. Zamponi, "Self-planting: digging holes in rough landscapes," Journal of Statistical Mechanics: Theory and Experiment, 1–17 (2019) (cit. on pp. 2, 16, 53, 62, 63, 73, 100).

[2] D. Sharma, J.-P. Bouchaud, M. Tarzia, and F. Zamponi, "A Constraint-Satisfaction Agent-Based Model for the Macro-economy," SSRN Electronic Journal, 1–14 (2020) (cit. on pp. 2, 16, 100).

[3] D. Sharma, J.-P. Bouchaud, S. Gualdi, M. Tarzia, and F. Zamponi, "V -, U -, L - or W-shaped recovery after COVID? Insights from an Agent Based Model," SSRN Electronic Journal, 1–15 (2020) (cit. on pp. 2, 17, 100).

[4] J. M. Keynes, "Alfred Marshall, 1842-1924," The Economic Journal **34**, 311 (1924) (cit. on p. 3).

[5] J. Tinbergen, "The Use of Models: Experience and Prospects," Nobel Prize Lecture (1969) (cit. on p. 4).

[6] F. Smets and R. Wouters, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," Journal of the European Economic Association **1**, 1123–1175 (2003) (cit. on pp. 4, 79).

[7] R. C. Fair, "Has macro progressed?" Journal of Macroeconomics **34**, 2–10 (2012) (cit. on p. 4).

[8] J. E. Stiglitz, "Where modern macroeconomics went wrong," Oxford Review of Economic Policy **34**, 70–106 (2018) (cit. on pp. 4, 7, 8, 10).

[9] S. Wren-Lewis, "Ending the microfoundations hegemony," Oxford Review of Economic Policy **34**, 55–69 (2018) (cit. on pp. 4, 9).

[10] R. E. Lucas, "Expectations and the neutrality of money," Journal of Economic Theory **4**, 103–124 (1972) (cit. on p. 4).

[11] A. G. Haldane and A. E. Turrell, "An interdisciplinary model for macroeconomics," Oxford Review of Economic Policy **34**, 219–251 (2018) (cit. on pp. 5, 6, 11).

[12] R. E. Lucas, "Econometric policy evaluation: A critique," in Carnegie-rochester confer. series on public policy, Vol. 1, C (1976), pp. 19–46 (cit. on p. 5).

[13] F. E. Kydland and E. C. Prescott, "Time to build and aggregate fluctuations.," Econometrica **50**, 1345–1370 (1982) (cit. on p. 5).

[14] A. P. Kirman, "Whom or What Does the Representative Individual Represent?" Journal of Economic Perspectives **6**, 117–136 (1992) (cit. on p. 6).

[15] J. F. Muth, "Rational Expectations and the Theory of Price Movements," Econometrica **29**, 315 (1961) (cit. on p. 6).

[16] G. Dosi and A. Roventini, "More is different.. and complex! the case for agent-based macroeconomics," Journal of Evolutionary Economics, 1–37 (2019) (cit. on pp. 6, 9).

[17] D. Challet, M. Marsili, and Y. C. Zhang, "Minority games and stylized facts," Physica A: Statistical Mechanics and its Applications **299**, 228–233 (2001) (cit. on pp. 6, 14).

[18] D. Challet, M. Marsili, and R. Zecchina, "Statistical mechanics of systems with heterogeneous agents: Minority games," Physical Review Letters **84**, 1824–1827 (2000) (cit. on p. 6).

[19] R. E. Lucas, "Methods and Problems in Business Cycle Theory," Journal of Money, Credit and Banking **12**, 696 (1980) (cit. on p. 7).

[20] J. Robinson, "History versus Equilibrium," Indian Economic Journal **21**, 12 (1974) (cit. on p. 7).

[21] G. Fagiolo and A. Roventini, "Macroeconomic policy in DSGE and agent-based models redux: New developments and challenges ahead," Journal of Artificial Socieities and Social Simulation **20**, 10.18564/jasss.3280 (2017) (cit. on pp. 7, 8).

[22] A. Korinek, "Thoughts on DSGE Macroeconomics: Matching the Moment, But Missing the Point?" SSRN Electronic Journal, 1–12 (2017) (cit. on p. 7).

[23] S. Burgess et al., "The Bank of England's Forecasting Platform: COMPASS, MAPS, EASE and the Suite of Models," SSRN Electronic Journal, 10.2139/ssrn.2266506 (2013) (cit. on p. 8).

[24] F. Brayton, T. Laubach, and D. Reifschneider, "The FRB/US Model : A Tool for Macroeconomic Policy Analysis," FEDS Notes, FEDS Notes **2014**, 10.17016/2380-7172.0012 (2014) (cit. on p. 8).

[25] M. Lemoine, H. Turunen, A. Lepetit, A. Zhutova, P. Clerc, and J.-P. Laffargue, "The FR-BDF Model and an Assessment of Monetary Policy Transmission in France," Working PaperSeries no. 736, Banque de France. (2019) (cit. on p. 8).

[26] R. E. Lucas, "Macroeconomic Priorities," American Economic Review **93**, 1–14 (2003) (cit. on p. 8).

[27] D. Vines and S. Wills, "The rebuilding macroeconomic theory project: An analytical assessment," Oxford Review of Economic Policy **34**, 1–42 (2018) (cit. on p. 8).

[28] P. Krugman, "Fighting off Depression," New York Times (2009) (cit. on p. 8).

[29] C. F. Camerer, G. Loewenstein, and M. Rabin, eds., *Advances in Behavioral Economics* (Princeton University Press, 2011) (cit. on p. 8).

[30] J. Lindé, F. Smets, and R. Wouters, "Challenges for Central Banks' Macro Models," in *Handbook of macroeconomics*, Vol. 2, 1st ed. (Elsevier B.V., 2016), pp. 2185–2262 (cit. on p. 8).

[31] P. Romer, "The Trouble with macroeconomics," The American Economist, 1–20 (2016) (cit. on p. 8).

[32] A. Tooze, *Crashed - How a Decade of Financial Crises Changed the World* (Penguin Books, 2018), p. 720 (cit. on p. 9).

[33] O. Blanchard, "On the future of macroeconomic models," Oxford Review of Economic Policy **34**, 43–54 (2018) (cit. on pp. 9, 11, 12, 99, 102).

[34] O. Blanchard, "On the Need for (At Least) Five Classes of Macro Models," PIIE Real-time economics watch blog (2017) (cit. on pp. 9, 10, 11).

[35] C. H. Hommes, "Heterogeneous Agent Models in Economics and Finance," Amsterdam and Rotterdam, 2005 (cit. on p. 10).

[36] T. Assenza, P. Heemeijer, C. H. Hommes, and D. Massaro, "Managing self-organization of expectations through monetary policy: A macro experiment," Journal of Monetary Economics, 10 . 1016 / j . jmoneco . 2019 . 12 . 005 (2020) (cit. on p. 10).

[37] A. Smith, *The Theory of Moral Sentiments* (J. Richardson, 1822) (cit. on p. 10).

[38] M Kalecki, *Essays in the Theory of Economic Fluctuations* (Routledge, London, 2013) (cit. on p. 11).

[39] K. Marx, "Chapter 17," in *Theories of surplus value* (1861) (cit. on p. 11).

[40] A. Turrell, "Agent-based models: understanding the economy from the bottom up," Quarterly Bulletin of the Bank of England **Q4**, 173–188 (2016) (cit. on p. 11).

[41] A. G. Haldane and A. E. Turrell, "An interdisciplinary model for macroeconomics," Oxford Review of Economic Policy **34**, 219–251 (2018) (cit. on p. 11).

[42] N. M. Ferguson, D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke, "Strategies for mitigating an influenza pandemic," Nature **442**, 448–452 (2006) (cit. on p. 12).

[43] S. Poledna, M. Miess, and S. Thurner, "Economic Forecasting with an Agent-based Model Agent-based Model for the Austrian Economy," SSRN Electronic Journal (2019) (cit. on p. 12).

[44] S. Gualdi, M. Tarzia, F. Zamponi, and J. P. Bouchaud, "Tipping points in macroeconomic agent-based models," Journal of Economic Dynamics and Control **50**, 29–61 (2015) (cit. on pp. 12, 13, 15, 16, 55, 60, 77, 78, 79, 80, 83, 87, 97, 100).

[45] H. Dawid and D. Delli Gatti, "Agent-Based Macroeconomics," in *Handbook of computational economics*, Vol. 4 (2018) (cit. on p. 12).

[46] M. Bardoscia, S. Battiston, F. Caccioli, and G. Caldarelli, "Pathways towards instability in financial networks," Nature Communications **8**, 14416 (2017) (cit. on p. 12).

[47] R. Baptista, M. Hinterschweiger, K. Low, and A. Uluc, "Macroprudential Policy in an Agent-Based Model of the UK Housing Market," 2016 (cit. on p. 12).

[48] K. Braun-Munzinger, Z. Liu, and A. Turrell, "An Agent-Based Model of Dynamics in Corporate Bond Trading," London, 2018 (cit. on p. 12).

[49] A. Caiani, A. Russo, and M. Gallegati, *Does inequality hamper innovation and growth? An AB-SFC analysis*, Vol. 29, 1 (Journal of Evolutionary Economics, 2019), pp. 177–228 (cit. on p. 13).

[50] G. Dosi, G. Fagiolo, and A. Roventini, "Schumpeter meeting Keynes: A policy-friendly model of endogenous growth and business cycles," Journal of Economic Dynamics and Control **34**, 1748–1767 (2010) (cit. on p. 13).

[51] T. Assenza and D. Delli Gatti, *The financial transmission of shocks in a simple hybrid macroeconomic agent based model*, Vol. 29, 1 (2019), pp. 265–297 (cit. on p. 13).

[52] T. Assenza, D. Delli Gatti, and J. Grazzini, "Emergent dynamics of a macroeconomic agent based model with capital and credit," Journal of Economic Dynamics and Control **50**, 5–28 (2015) (cit. on p. 13).

[53] J.-P. Bouchaud, S. Gualdi, M. Tarzia, and F. Zamponi, "Optimal inflation target: insights from an agent-based model," Economics: The Open-Access, Open-Assessment E-Journal, 1–19 (2018) (cit. on pp. 13, 16, 60, 77, 91, 100).

[54] S. Gualdi, M. Tarzia, F. Zamponi, and J.-p. Bouchaud, "Monetary policy and dark corners in a stylized agent-based model," Journal of Economic Interaction and Coordination **12**, 507–537 (2017) (cit. on pp. 13, 16, 60, 77, 80, 83, 84, 91, 100).

[55] E. Gaffeo, D. Delli Gatti, S. Desiderio, and M. Gallegati, "Adaptive Microfoundations for Emergent Macroeconomics," Eastern Economic Journal **34**, 441–463 (2008) (cit. on pp. 13, 55, 77, 81).

[56] D. Delli Gatti, S. Desiderio, E. Gaffeo, P. Cirillo, and M. Gallegati, *Macroeconomics from the Bottom-up*, Vol. 1, New Economic Windows (Springer Milan, Milano, 2011) (cit. on pp. 13, 77).

[57] P. Seppecher, "Flexibility of wages and macroeconomic instability in an agent-based computational model with endogenous money," Macroeconomic Dynamics **16**, 284–297 (2012) (cit. on p. 13).

[58] P. Seppecher and I. Salle, "Deleveraging crises and deep recessions: a behavioural approach," Applied Economics **47**, 3771–3790 (2015) (cit. on p. 13).

[59] "The Economy as an Evolving Complex System," in Santa fe institute studies in the sciences of complexity. Edited by P. W. Anderson, K. J. Arrow, and D. Pines (1987) (cit. on p. 13).

[60] J.-P. Bouchaud and J.-P. Nadal, "From Statistical Physics to social sciences," Compte Rendus Physique **20**, 241–386 (2019) (cit. on p. 13).

[61] P. Mirowski, "The Rise and Fall of the Concept of Equilibrium in Economic Analysis," Louvain Economic Review **55**, 447–468 (1989) (cit. on p. 13).

[62] A. Dragulescu and V. M. Yakovenko, "Evidence for the exponential distribution of income in the USA," European Physical Journal B **20**, 4 (2000) (cit. on p. 13).

[63] A. Drăgulescu and V. M. Yakovenko, "Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States," in Physica a: statistical mechanics and its applications, Vol. 299, 1-2 (2001), pp. 213–221 (cit. on p. 13).

[64] A. A. Dragulescu and V. M. Yakovenko, "Statistical Mechanics of Money, Income, and Wealth: A Short Survey," 10.1063/1.1571309 (2002) (cit. on p. 13).

[65] A. Dragulescu and V. Yakovenko, "Statistical mechanics of money," The European Physical Journal B **378**, 373–378 (2000) (cit. on p. 13).

[66] M. Gallegati, S. Keen, T. Lux, and P. Ormerod, "Worrying trends in econophysics," Physica A: Statistical Mechanics and its Applications **370**, 1–6 (2006) (cit. on p. 13).

[67] S. G. Reddy, "What is an Explanation? Statistical Physics and Economics," SSRN Electronic Journal, 1–18 (2020) (cit. on p. 13).

[68] D. Challet, M. Marsili, and Y.-C. Zhang, *Minority Games* (Oxford University Press, New York, 2005) (cit. on p. 14).

[69] M. Bardoscia, G. Livan, and M. Marsili, "Statistical mechanics of complex economies," Journal of Statistical Mechanics: Theory and Experiment **2017**, 43401–43401 (2017) (cit. on p. 14).

[70] S. N. Durlauf, "How can statistical mechanics contribute to social science?" Proceedings of the National Academy of Sciences **96**, 10582–10584 (1999) (cit. on p. 14).

[71] P. W. Anderson, *Basic Notions of Condensed Matter Physics*, edited by P. W. Anderson (CRC Press, 2018) (cit. on p. 14).

[72] F. Altarelli, R. Monasson, G. Semerjian, and F. Zamponi, "A review of the Statistical Mechanics approach to Random Optimization Problems," in *Handbook of satisfiability* (IOS Press Van Diemenstraat 94 1013 CN AmsterdamNetherlands, 2009) (cit. on pp. 16, 25, 53).

[73] P. Bak, *How Nature Works: the science of self-organized criticality* (Springer Science and Business Media, 2013) (cit. on p. 16).

[74] S. Franz and G. Parisi, "The simplest model of jamming," J. Phys. A: Math. Theor **49**, 0 (2016) (cit. on pp. 16, 28, 100).

[75] S. Franz, G. Parisi, M. Sevelev, P. Urbani, and F. Zamponi, "Universality of the SAT-UNSAT (jamming) threshold in non-convex continuous constraint satisfaction problems," SciPost Phys. **019**, 1–37 (2017) (cit. on pp. 16, 38, 39, 53, 73, 110).

[76] W. Barthel et al., "Hiding solutions in random satisfiability problems: A statistical mechanics approach," Physical Review Letters **88**, 1887011–1887014 (2002) (cit. on pp. 16, 40).

[77] S. Gualdi, J. P. Bouchaud, G. Cencetti, M. Tarzia, and F. Zamponi, "Endogenous crisis waves: Stochastic model with synchronized collective Behavior," Physical Review Letters **114**, 1–5 (2015) (cit. on pp. 16, 70, 71).

[78] P. Hannon and S. Chaudhuri, *Why the Economic Recovery Will Be More of a 'Swoosh' Than V-Shaped*, 2020 (cit. on pp. 17, 75, 77).

[79] M. Mézard and A. Montanari, *Information, Physics, and Computation*, Vol. 9780198570 (Oxford University Press, 2009), pp. 1–584 (cit. on pp. 19, 25, 29).

[80] *Conjunctive Normal Form* (cit. on p. 19).

[81] S. A. Cook, "The complexity of theorem-proving procedures," in Proceedings of the third annual acm symposium on theory of computing - stoc '71 (1971), pp. 151–158 (cit. on p. 20).

[82] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborov{\'a}, "Gibbs states and the set of solutions of random constraint satisfaction problems.," Proceedings of the National Academy of Sciences of the United States of America **104**, 10318–23 (2007) (cit. on p. 25).

[83] M. Mezard and R. Zecchina, "The random K-satisfiability problem: from an analytic solution to an efficient algorithm," 1–38 (2002) (cit. on p. 25).

[84] O. C. Martin, R. Monasson, and R. Zecchina, *Statistical mechanics methods and phase transitions in optimization problems*, 2001 (cit. on p. 25).

[85] T. Cover, "Geometrical and Statistical properties of systems of linear inequalities with applications in pattern recognition," IEEE Transactions on Electronic Computers **EC-14**, 326–334 (1965) (cit. on p. 26).

[86] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," Psychological Review **65**, 386 (1958) (cit. on p. 27).

[87] E Gardner, "The space of interactions in neural network models," Journal of Physics A: Mathematical and General **21**, 257–270 (1988) (cit. on pp. 27, 28).

[88] E Gardner and B. Derrida, "Optimal storage properties of neural network models," Journal of Physics A: Mathematical and General **21**, 271–284 (1988) (cit. on pp. 27, 28).

[89] N. Brunel, J. P. Nadal, and G Toulouse, "Information Capacity of a Perceptron Au Brunel, N," J. Phys. A **25**, 5017–5037 (1992) (cit. on p. 27).

[90] W. Krauth, M. Mezard, and J. P. Nadal, "Basins of attraction in a perceptron-like Neural network," Complex Systems **2**, 387–408 (1988) (cit. on p. 27).

[91] W. Krauth and M. Mézard, "Storage capacity of memory networks with binary couplings," Journal de Physique **50**, 3057–3066 (1989) (cit. on p. 27).

[92] J Friedman, T Hastie, and R Tibshirani, *The Elements of Statistical Learning*, Vol. 1, 10 (2001) (cit. on p. 28).

[93] A. J. Liu and S. R. Nagel, "The jamming transition and the marginally jammed solid," Annual Review of Condensed Matter Physics **1**, 347–369 (2010) (cit. on p. 28).

[94] M. Talagrand, *Spin glasses: a challenge for mathematicians:* (Springer Science and Business Media, Heidelberg, 2003) (cit. on p. 33).

[95] M. Talagrand, "The Parisi formula," Annals of Mathematics **163**, 221–263 (2006) (cit. on p. 33).

[96] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific Publishing Company, 1987) (cit. on pp. 34, 108, 111, 112).

[97] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing An Introduction* (2001), p. 252 (cit. on p. 34).

[98] M. Mezard, G. Parisi, N. Sourlas, G. Toulouse, and M. Virasoro, "Replica Symmetry Breaking and the Nature of the Spin Glass Phase.," Journal de physique Paris **45**, 843–854 (1984) (cit. on p. 34).

[99] G. Parisi, "Infinite number of order parameters for spin-glasses," Physical Review Letters **43**, 1754–1756 (1979) (cit. on p. 34).

[100] B. Seoane and F. Zamponi, "Spin-glass-like aging in colloidal and granular glasses," soft Matter **14**, 5222–5234 (2018) (cit. on p. 37).

[101] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, "Fractal free energy landscapes in structural glasses," Nature Communications **5**, 1–6 (2014) (cit. on p. 37).

[102] G. Parisi, P. Urbani, and F. Zamponi, *Theory of Simple Glasses* (Cambridge University Press, 2020) (cit. on p. 37).

[103] G. Parisi and F. Zamponi, "Mean-field theory of hard sphere glasses and jamming," Reviews of Modern Physics **82**, 789–845 (2010) (cit. on p. 37).

[104] H. Yoshino and F. Zamponi, "Shear modulus of glasses: Results from the full replica-symmetry-breaking solution," Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **90**, 10.1103/PhysRevE.90.022302 (2014) (cit. on p. 37).

[105] J. D. Almeida and D. Thouless, "Stability of the Sherrington-Kirkpatrick solution of a spin glass model," Journal of Physics A: Mathematical and General **11**, 983–990 (1978) (cit. on pp. 38, 116).

[106] G. Györgyi, "Inference of a rule by a neural network with thermal noise," Physical Review Letters **64**, 2957–2960 (1990) (cit. on p. 40).

[107] L. Zdeborová and F. Krzakala, "Statistical physics of inference: thresholds and algorithms," Advances in Physics **65**, 453–552 (2016) (cit. on p. 40).

[108] F. Krzakala and L. Zdeborová, "Hiding quiet solutions in random constraint satisfaction problems," Physical Review Letters **102** (2009) (cit. on p. 40).

[109] F. Krzakala, M. Mézard, and L. Zdeborová, "Reweighted belief propagation and quiet planting for random K-SAT," Journal on Satisfiability (2012) (cit. on p. 40).

[110] G. Semerjian and R. Monasson, "Relaxation and metastability in a local search procedure for the random satisfiability problem," Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics **67**, 18 (2003) (cit. on pp. 45, 46, 49).

[111] M. Baity-Jesi, C. P. Goodrich, A. J. Liu, S. R. Nagel, and J. P. Sethna, "Emergent SO(3) Symmetry of the Frictionless Shear Jamming Transition," Journal of Statistical Physics **167**, 735–748 (2017) (cit. on p. 46).

[112] S. Franz, G. Parisi, P. Urbani, and F. Zamponi, "Universal spectrum of normal modes in low- temperature glasses," Proceedings of the National Academy of Sciences **112**, 14539–14544 (2015) (cit. on p. 48).

[113] J. Bun, J.-P. Bouchaud, and M. Potters, "Cleaning large correlation matrices: Tools from Random Matrix Theory," Physics Reports **666**, 1–109 (2017) (cit. on p. 48).

[114] J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," Proceedings of the National Academy of Sciences **79**, 2554–2558 (1982) (cit. on p. 52).

[115] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, "Glass and Jamming Transitions: From Exact Results to Finite-Dimensional Descriptions," Annual Review of Condensed Matter Physics **8**, 265–288 (2017) (cit. on p. 52).

[116] F. Antenucci, S. Franz, P. Urbani, and L. Zdeborová, "On the glassy nature of the hard phase in inference problems," 1–9 (2018) (cit. on p. 53).

[117] T. Yanagita and T. Onozaki, "Dynamics of market structure driven by the degree of consumer's rationality," Physica A: Statistical Mechanics and its Applications **389**, 1041–1054 (2010) (cit. on p. 66).

[118] O. Jorda, M. Schularick, and A. M. Taylor, "When Credit Bites Back: Leverage, Business cycles, and Crises," 2011 (cit. on p. 72).

[119] T. Adrian and N. Boyarchenko, "Intermediary Leverage Cycles and Financial Stability," SSRN Electronic Journal, 10 . 2139 / ssrn . 2133385 (2012) (cit. on p. 72).

[120] C. Aymanns and J. D. Farmer, "The dynamics of the leverage cycle," Journal of Economic Dynamics and Control **50**, 155–179 (2015) (cit. on p. 72).

[121] C. Aymanns, F. Caccioli, J. D. Farmer, and V. W. Tan, "Taming the Basel leverage cycle," Journal of Financial Stability **27**, 263–277 (2016) (cit. on p. 72).

[122] K. Bluwstein, M. Brzoza-Brzezina, P. Gelain, and M. Kolasa, "Multiperiod Loans, Occasionally Binding Constraints, and Monetary Policy: A Quantitative Evaluation," Journal of Money, Credit and Banking **52**, 1691–1718 (2020) (cit. on p. 72).

[123] M. Werner, "Occasionally Binding Liquidity Constraints and Macroeconomic Dynamics," SSRN Electronic Journal, 10.2139/ssrn.3635692 (2020) (cit. on p. 72).

[124] M. Eichenbaum, S. Rebelo, and M. Trabandt, *The Macroeconomics of Epidemics*, tech. rep. (National Bureau of Economic Research, Cambridge, MA, 2020) (cit. on p. 75).

[125] W. J. McKibbin and R. Fernando, "The Global Macroeconomic Impacts of COVID-19: Seven Scenarios," SSRN Electronic Journal, 1–43 (2020) (cit. on p. 75).

[126] L. Fornaro and M. Wolf, "Covid-19 Coronavirus and Macroeconomic Policy," 2020 (cit. on pp. 76, 89).

[127] V. Guerrieri, G. Lorenzoni, L. Straub, and I. Werning, *Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?* Tech. rep. 9 (National Bureau of Economic Research, Cambridge, MA, 2020), pp. 1689–1699 (cit. on p. 76).

[128] T. Assenza et al., "The Hammer and the Dance: Equilibrium and Optimal Policy during a Pandemic Crisis," Toulouse, 2020 (cit. on p. 76).

[129] F. Alvarez, D. Argente, and F. Lippi, "A Simple Planning Problem for COVID-19 Lockdown," Cambridge, MA, 2020 (cit. on p. 76).

[130] M. Farboodi, G. Jarosch, and R. Shimer, "Internal and External Effects of Social Distancing in a Pandemic," Cambridge, MA, 2020 (cit. on p. 76).

[131] M. Wolf, *The risks of lifting lockdowns prematurely are very large*, 2020 (cit. on p. 76).

[132] R. Baldwin and B. W. di Mauro, *Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes* (2020), pp. 1–227 (cit. on pp. 76, 91).

[133] *The 90% Economy: Life after Lockdowns*, 2020 (cit. on pp. 76, 86, 97).

[134] R. M. del Rio-Chanona, P. Mealy, A. Pichler, F. Lafond, and D. Farmer, "Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective," Covid Economics (2020) (cit. on p. 85).

[135] J. Danielson, R. Macrae, D. Vayanos, and J.-P. Zigrand, *The coronavirus crisis is no 2008*, 2020 (cit. on p. 90).

[136] A. Tooze, *How coronavirus almost brought down the global financial system*, 2020 (cit. on p. 90).

[137] C. Pacitti, R. Hughes, J. Leslie, C. McCurdy, J. Smith, and D. Tomlinson, *The fiscal costs of lockdown: Three scenarios for the UK*, 2020 (cit. on p. 90).

[138] J. Gali, "Helicopter Money: the time is now," Vox CEPR Policy Portal (2020) (cit. on pp. 90, 91).

[139] M. Friedman, "A monetary and fiscal framework for economic stability," The American Economic Review **38**, 245–264 (1948) (cit. on p. 90).

[140] W. H. Buiter, "The Simple Analytics of Helicopter Money: Why It Works — Always," Economics: The Open-Access, Open-Assessment E-Journal **8**, 10.5018/economics-ejournal.ja.2014-28 (2014) (cit. on p. 91).

[141] S. Kapoor and W. H. Buiter, *To fight the COVID pandemic , policymakers must move fast and break taboos*, 2020 (cit. on p. 91).

[142] O. Blanchard and J. Pisani-Ferry, *Monetisation: Do not panic*, 2020 (cit. on p. 91).

[143] M. Carney and B. Broadbent, *Bank of England Monetary Policy Report May 2020*, tech. rep. May (Bank of England, 2020) (cit. on p. 91).

[144] P. Aldama et al., *Macro-economic projections 2000-2020*, tech. rep. (Banque de France, 2010), pp. 1–13 (cit. on p. 91).

[145] M. W. Walsh, *A Tidal Wave of Bankruptcies Is Coming*, 2020 (cit. on p. 97).

[146] D. Rees, "What comes next?" 2020 (cit. on p. 97).

[147] P. de Grauwe and Y. Ji, *Central bankers should not be forward-looking in times of crisis*, 2020 (cit. on p. 97).

[148] M. Buchanan, *This Economy does not compute*, 2008 (cit. on p. 97).

[149] E. Badger and A. Parlapiano, *The Rich Cut Their Spending. That Has Hurt All the Workers Who Count on It.* 2020 (cit. on p. 98).

[150] E. Morath, T. Francis, and J. Baer, *The Covid Economy Carves Deep Divide Between Haves and Have-Nots*, 2020 (cit. on p. 98).

[151] J. M. Keynes, "The General Theory of Employment," The Quarterly Journal of Economics **51**, 209 (1937) (cit. on p. 99).

[152] F. Lamperti, A. Roventini, and A. Sani, "Agent-based model calibration using machine learning surrogates," Journal of Economic Dynamics and Control **90**, 366–389 (2018) (cit. on p. 101).

[153] J. Grazzini, M. G. Richiardi, and M. Tsionas, "Bayesian estimation of agent-based models," Journal of Economic Dynamics and Control **77**, 26–47 (2017) (cit. on p. 101).

[154] D. D. Gatti and J. Grazzini, "Rising to the Challenge : Bayesian Estimation and Forecasting Techniques for Macroeconomic Agent-Based Models," 2019 (cit. on p. 101).

[155] Andrew G Haldane, "Andrew G Haldane: The dog and the frisbee," Financial Stability, Bank of England, 1–34 (2012) (cit. on p. 102).

[156] H. A. Simon, "Rational choice and the structure of the environment.," Psychological Review **63**, 129–138 (1956) (cit. on p. 103).

[157] D. Graeber, *Against Economics*, 2019 (cit. on p. 103).

[158] B. Appelbaum, *The Economists' Hour: False Prophets, Free Markets, and the Fracture of Society* (Little, Brown and Company, 2019), p. 448 (cit. on p. 103).

[159] S. G. Reddy, "Randomise This ! On Poor Economics," Review of Agrarian Studies **2**, 60–73 (2013) (cit. on p. 104).

[160] A. Tooze, *Output Gap nonsense*, 2019 (cit. on pp. 104, 105).

[161] A. Alesina and S. Ardagna, "Tales of fiscal adjustment," Economic Policy **13**, 487–545 (1998) (cit. on p. 104).

[162] A. Crisanti and H. J. Sommers, "The spherical p-spin interaction spin glass model: the statics," Zeitschrift für Physik B Condensed Matter **87**, 341–354 (1992) (cit. on p. 110).

[163] B Duplantier, "Comment on Parisi's equation for the SK model for spin glasses," Journal of Physics A: Mathematical and General **14**, 283–285 (1981) (cit. on pp. 111, 112).

ABSTRACT

Agent-based models (ABMs) have emerged as a complementary paradigm for modeling macroeconomic phenomena. Compared to other, more established models such as DSGE (Dynamic Stochastic General Equilibrium) models, ABMs provide a flexible framework for understanding the complexity of the macroeconomy while at the same time taking into account the heterogeneous nature of economic actors, institutions and markets without making overly restrictive assumptions. ABMs take a "bottom-up" approach towards macroeconomic modeling by simulating the behavior of each individual agent in the economy and then aggregating to reveal emergent phenomena such as endogenous business cycles or flash crashes.

The object of this thesis is to advance a methodology commonly used in statistical physics and apply it to the study of two macroeconomic agent-based models. In both models studied here, we first determine the "phase-diagram" of the model to identify the relevant macroscopic regimes to develop an intuitive understanding of the macro-dynamics using a small subset of parameters.

The first ABM presented here builds upon the paradigm of constraint satisfaction problems (CSPs) and integrates it within the model's behavioral rules via agents' budgetary constraints. These constraints, similar to the well-studied perceptron CSP, reveal the existence of three regimes and underscore the importance of debt for macroeconomic stability: at low-levels of debt, the economy remains structure-less with frequent bankruptcies while high debt leads to *endogenous* business cycles. Between these two extremes, an intermediate regime of relative stability is found with low levels of bankruptcies for all times.

Within this ABM, agents' preferences, serving as the source of disorder in the CSP, evolve continuously in time. We thus study a simple dynamical scheme for the perceptron and discover that a rugged landscape can indeed exist with dynamic, annealed disorder.

Finally, we extend the Mark-0 ABM to simulate exogenous consumption and productivity shocks due to the Covid pandemic. Whereas standard approaches design a model to understand a particular outcome, this model can generate a variety of scenarios after a Covid-like shock. Furthermore, we also investigate the efficacy of several policies, including the much-debated "helicopter money" drop, in avoiding economic collapse. We thus highlight the importance of ABMs as multi-purpose "scenario generators", for producing outcomes that are difficult to foresee due to the intrinsic complexity of macro-economic dynamics.

keywords: macroeconomics; statistical physics; agent-based models; complex systems; disordered systems.

RÉSUMÉ

Les modèles à agents (Agent-Based Models ou ABMs) sont apparus comme un paradigme complémentaire pour la modélisation des phénomènes macro-économiques. Par rapport à d'autres modèles plus établis, tels que les modelés DSGE (Dynamic Stochastic General Equilibrium), les ABMs offrent un cadre flexible pour comprendre la complexité de la macroéconomie tout en prenant en compte la nature hétérogène des acteurs économiques, des institutions et des marchés, sans faire d'hypothèses trop restrictives. Ces modelés adoptent une approche "bottom-up" de la modélisation macro-économique en simulant le comportement de chaque agent individuel dans l'économie puis en s'agrégeant pour révéler des phénomènes émergents tels que les cycles économiques endogènes ou les crashs soudains.

L'objet de cette thèse est de faire progresser une méthodologie communément utilisée en physique statistique et de l'appliquer à l'étude de deux modèles macro-économiques. Dans les deux modèles étudiés ici, nous déterminons d'abord le "diagramme de phase" du modèle pour identifier les régimes macroscopiques pertinents afin de développer une compréhension intuitive de la macro-dynamique en n'utilisant qu'un petit sous-ensemble de paramètres.

Le premier modèle présenté ici s'appuie sur le paradigme des problèmes de satisfaction des contraintes (de l'anglais Constraint Satisfaction Problems, CSPs) et l'intègre dans le cadre des règles de comportement du modèle via les contraintes budgétaires des agents. Ces contraintes, similaires à celles du perceptron, un CSP bien-etudié, révèlent l'existence de trois régimes et soulignent l'importance de la dette pour la stabilité macro-économique : à un faible niveau d'endettement, l'économie reste sans structure et les faillites sont fréquentes, alors qu'à un niveau élevé la dette conduit à des cycles économiques *endogènes*. Entre ces deux extrêmes, l'on trouve un régime intermédiaire de stabilité relative avec de faibles niveaux des faillites tout le temps.

Dans ce modèle, les préférences des agents, qui sont à l'origine du désordre dans le CSP, évoluent continuellement dans le temps. Nous étudions donc un schéma dynamique simple pour le perceptron et découvrons qu'un paysage rugueux peut en effet exister avec un désordre dynamique.

Enfin, nous généralisons l'ABM Mark-0 pour simuler les chocs exogènes de consommation et de productivité dus à la pandémie de COVID. Alors que les approches standards élaborent un modèle pour comprendre un résultat particulier, ce modèle peut générer une variété de scénarios après un choc de type COVID. En outre, nous étudions également l'efficacité de plusieurs politiques, notamment la très controversée "monnaie hélicoptère", pour éviter l'effondrement économique. Nous insistons donc sur l'importance des ABMs comme des "générateurs de scénarios" polyvalents, pour produire des résultats difficiles à prévoir en raison de la complexité intrinsèque de la dynamique macro-économique.

mots-clés: macroeconomie; physique statistique; modèles à agents; systèmes complexes; systèmes désordonnées.