Head Office:
Università degli Studi di Padova

Dipartment: Physics and Astronomy "Galileo Galilei"

Ph.D. COURSE IN PHYSICS

Thesis title:
Out of equilibrium dynamics: from an entropy of the growth to the growth of entropy production.

Supervisor:
Prof. Attilio L. Stella

Ph.D. student: Gianluca Teza

Series XXXII

Abstract

This thesis contains some of the main results obtained during my research activity as a PhD student in the course of Physics of the University of Padova, under the supervision of prof. Attilio L. Stella. The work presented in this thesis lies within the area of Statistical Mechanics and Complex Systems, with the concept of "entropy" acting as common denominator across most of the works presented.

Chapter 1 serves as a brief introduction that illustrates the lines along which my research developed during the course. It is kept brief because each chapter will begin with a specific introductory section.

In chapter 2 I will discuss the dynamical model of the exports in the period 1962-2000 aggregated at a global level [1] that inspired the main research topic of my PhD. This preliminary chapter is essential to properly understand the role played by each term of the system of stochastic differential equations that regulate the dynamics and also to illustrate results supporting the validity of the choices made in its construction.

In chapter 3 I will present results concerning the adaptation of the dynamical model of [1] to the case of the exports of a single country. The database used is more recent (1995-2015) and the analyses were carried out for 223 countries using more than 1000 products. Here we also introduce a new metrics inspired by Shannon's entropy to determine simultaneously a ranking for the best performing countries and the most important products in the global trade network.

In chapter 4 I will present an analysis of the G7 countries imports and exports in the time window 1962-2000, using around 200 categories of products for the description. The analysis focuses mainly on the network representing the correlations linking all the commodities composing the basket of traded goods of a country. Moreover, an analysis is carried out to explore the potential growth that might come by increasing the reallocation of resources in such network.

In chapter 5 a rigorous coarse-grainig approach of Markov jump models that allows the preservation of the whole distribution of produced entropy at stationarity is presented. This works allows to provide clear answers to open, longstanding questions concerning the formulation of fluctation theorems for coarse-grained systems.

In chapter 6 we explore Markov jump models characterized by the phenomenon of he negative dfferential mobility. This work provides insights in the mechanism that originate this behavior finding consistency with respect to experimental observations of such phenomenon.

In chapter A I'll present a side project concerning the phenomenon of segregation in networks characterized by discrete strong convictions that drive the dynamics of the rewiring of links. The choice of not including a spatial metrics (everyone can be connected with anyone) makes this model suitable to describe scenarios that are typically present in social networks.

The work presented in this thesis (with the exception of chapter 5) has been published in the following articles:

- Teza, Gianluca, Michele Caraglio, and Attilio L. Stella. "Growth dynamics and complexity of national economies in the global trade network." *Scientific reports* 8.1 (2018): 15230. [**2**]
- Teza, Gianluca, Michele Caraglio, and Attilio L. Stella. "Data Driven Approach to the Dynamics of Import and Export of G7 Countries." *Entropy* 20.10 (2018): 735. [**3**]
- Teza, Gianluca and Attilio L. Stella "Memory leaves entropy production fluctuations invariant under coarse-graining." *in preparation* (2019).
- Teza, Gianluca, et al. "Rate dependence of current and fluctuations in jump models with negative differential mobility." *arXiv preprint* arXiv:1904.05241 (2019). [**4**] (submitted to the peer-reviewed journal Physica A)
- Teza, Gianluca, et al. "Network model of conviction-driven social segregation." *Physical Review E* 99.3 (2019): 032310. [**5**]

Contents

CHAPTER 1

Introduction

"More is different" [6]. Three words that capture the essence of Statistical Mechanics and Complex Systems Physics. The whole reason of being of this field is based on the idea that, in many cases, we can provide accurate descriptions of the phenomenology of a macroscopic system only by taking a step back and looking at things from a wider perspective coarse-graining many details. This might sound counterintuitive at a first glance, but expecting to model phenomena that are observed at a macroscopic scale starting from the laws driving the evolution of microscopic components usually reveals to be both unfeasible and pointless.

Coarse-graining is indeed one of the fundamental pillars of Statistical Mechanics, which allowed physicists to model a wide variety of systems ranging across different scales, up to the point that they were able to address problems tangible (and understandable) by the whole society, regardless of their background. This suggested a redefinition of what had to be considered *fundamental* by the scientific community. Physics started to make contact with other natural sciences, making use of tools and methods ignored by these other disciplines, which allowed to address problems and provide answers supported by the formalism and the rigorousness that is proper of mathematics and statistics.

Interdisciplinarity led to the development of a whole new branch of Physics, namely the *Complex Systems Science*. The name of this discipline is meant to underline the fact this field deals with systems whose "degree of complexity" is so high that we are not able to model them with a fine-grained description. A peculiarity unique of the Complex Systems is that the shapes of the laws emerging in this field were found to be extremely common across completely unrelated contexts. The idea that such laws were characterized by some sort of universality was indeed one of the reasons that helped the flourishing of this field, making also the whole scientific community to recognize the value of this kind of research.

During my PhD course of studies I have been deepening my knowledge in statistical mechanics and applied the expertise gathered in this field in first place to the economic sciences. Starting from the project of my master thesis, I began to approach the relatively new field of Econophysics by looking at novel ways to tackle the problem of Economic Complexity, a field that was started in 2009 by R. Hausmann and C. Hidalgo with a paper [7] in which they tried to assess the level of development of a country by analyzing with an iterative algorithm how its basket of exports compares with those of other countries. In the following years other research groups started to provide contributions concerning this topic. A big step forward was made in 2012, when L. Pietronero's group published a work [8] that highlighted some issues concerning Haussmann's work, proposing an alternative scheme that solved the pathology of the original formulation.

Together with the research group in Padova, we started to tackle the issue from a different perspective. Inspired from a work [1] that addressed the dynamics that regulates the exports at a global level, we were able to build a dynamical model that managed to replicate the historical evolution of traded goods among

countries through a system of coupled stochastic differential equations [2]. The (few) parameters used in the construction of such model allowed to capture key features of every country, such as growth, fluctuations and resources reallocation rates. Then, to reconnect with the existing work in literature, we evaluated node-specific measures for both layers composing the bipartite network of countries-products, introducing an original scheme based on the Shannon's entropy function [9]. This scheme makes full use of the information provided by the data, while both Pietronero [8] and Haussman's [7] analysis relied on a binarization of the export matrix [10, 11] to determine which countries could be considered producers of which products. Such choice, apart from causing an enormous loss of information, introduced an arbitrariness in the choice of the thresholds that affects the whole analysis. One important result that emerged from our of analysis is that *diversification* is the fundamental trait that allows to capture the potential (in terms of future growth of an economy). Coupling this measure with the well known indicator Gross Domestic Product Per Capita ($\mathrm{GDP_{pc}}$) we extend the dimensions of the space that we use to monitor the economic evolution of the countries and we are therefore able distinguish among developed, underdeveloped, emerging and risky economies.

This work stimulated me to apply the knowledge I gathered during this project in other contexts and to start exploring new research areas. For instance, the entropic scheme that we constructed for studying the trade of exports in our work, is a measure that can, in principle, be applied to any bipartite (weighted or unweighted) network. One of the directions we are currently exploring for future work is the possibility of applying our results in fields like ecology (within mutualistic networks [12]) and microbiology (for characterizing richness of samples and prevalence of species [13]).

The extensive use of stochastic dynamics and entropy related concepts in the applications to Econophysics, naturally led me to become interested also in more fundamental problems in the general theory of dynamics out of equilibrium. Indeed, already when looking at exports aggregated at a global level we are considering a system that is far from being a closed one: the drifts regulating the deterministic evolution in the dynamical models force it to approach asymptotically a stationary steady state of non equilibrium [1]. The same goes of course in the more specific scenario of a single country, in which the imports add to the contributes coming from the country itself [2, 3].

More specifically, we recently obtained very promising results in the attempt to develop approximate and rigorous coarse-graining procedures for stochastic systems. We treated both jump processes governed by Master equations for discrete states, and Langevin dynamics in the continuum, establishing how such descriptions are transformed under a change in the spatial resolving power of the measuring apparatus.

The reason we started exploring this path was also encouraged by the fact that understanding how fluctuation relations apply to the dynamics of coarse-grained systems was an open, longstanding question. Such theorems [14, 15, 16, 17] are fundamental for the whole field of stochastic thermodynamics, and in the past decades numerous experiments carried out at different levels of space resolution showed to be in agreement with such theorems. Therefore it is a crucial issue to understand if and how such theorems should change when the system they're describing undergoes a coarse-graining procedure. Moreover, very recently, some interesting work [18, 19] carried out on active matter at the micro-scale relied heavily on the coarse-graining of experimental data for getting evidence of the entropy production and then to understand if it was inside or outside equilibrium.

Looking into the literature, the contributes we could find concerning coarse-graining of Markov jump processes went back to almost a decade ago [20, 21], and

exploited approximate methods (numerical or perturbative) to address the problem, to a limited set of system, in which enormous gaps in the time scale of the rates allowed for natural coarse-grainings that grouped together sites connected by fast transition rates. In the end this resulted in the authors asking explicitly if the observation of deviations from the expected fluctuation relations could be a hint of the presence of hidden degrees of freedom, and also if it was possible to define a "coarse-grained entropy" that instead satisfied exactly such relations.

Our contribute in this field turned out to be more substantial and far reacing than its inclusion in an already ample and articulated project could let us suspect at the start: indeed we were able to understand how to perform coarse-graining procedures of Markovian systems in an analytic and exact way that allows to preserve the whole distribution of the entropy production of the original system. This implies naturally that for what concerns the fluctuation theorem, since the original system was satisfying exactly such relation then the same applies to the coarse-grained system.

The price to pay in order to preserve the whole statistics through a coarse-graining procedure is however to extend the class of models beyond markovianity, because the analytic procedure introduces terms into the equation of the evolution of the probability which introduce memory effects. However, an enormous advantage that comes from such procedure is that we are able to identify exactly the terms responsible for such memory effects. This allowed us to build an algorithm that, at each step, eliminates one or multiple sites of the network and "markovianizes" the system by neglecting properly such terms and at the same time preserving the average rate of entropy production.

The main result emerging from this analysis is the general rule that the whole distribution of entropy production should be preserved when performing coarse-graining operations. Entropy should be considered an invariant that allows to assess the validity of this kind of procedures. Our intentions for the future are indeed to understand if the same rules can apply to different modelizations in stochastic thermodynamics.

Following the lines of this research, we investigated the mechanisms that lead to the phenomenon of the Negative Differential Mobility (NDM), which is the phenomenon where increasing a force leads to a decrease of mobility due to, for example, trapping or crowding. It is often studied theoretically in driven lattice gas models [**22, 23, 24, 25, 26, 27, 28, 29**] and has been observed in several experiments on electronic properties of materials [**30, 31, 32, 33**]. It can also occur in gel electrophoresis of polymers [**34, 35, 36, 37**] and chemical reaction networks [**38**]. Markov jump models reveal to be very useful when trying to build a system that is characterized by NDM, which however are characterized by a very high level of abstraction and imply considerable coarse-graining. So we wanted to understand up to what extent these jump models can be considered a faithful representation of a real physical process. This example shows that an investigation of how a microscopic description can lead under coarse-graining to a jump process description is in principle needed to assess the precise choice of transition rates. However, without undertaking such an ambitious program, it is important to explore the physical implications that different choices of rates can have by analyzing specific prototype models. The results obtained in this work highlight the need of a deeper understanding of how Markov jump models and in general how coarse grained description relate with the real world, providing ulterior support to the research we carried out concerning fluctuation relations of entropy production.

As already anticipated at the beginning of the introduction, one peculiarity of the physics of complex systems is the possibility of exploiting the same tools to perform analyses on completely unrelated topics. In my case, methods related to statistical mechanics and network dynamics revealed to be very effective in the study of a phenomenon of the social science, namely the problem of segregation [**39, 40**]. In order to predict and, possibly, prevent such phenomenon, it is necessary to understand the factors that cause it. While in most available descriptions [**41, 42, 43, 44**] space plays an essential role, one outstanding question is whether and how this phenomenon is possible in a well-mixed social network.

To create a model that helped us in the understanding of such phenomenon we defined and solved a simple model of segregation on networks based on discrete convictions [**5**]. In our model, space does not play a role, and individuals never change their conviction, but they may choose to connect socially to other individuals based on two criteria: sharing the same conviction, and individual popularity (regardless of conviction). The trade- off between these two moves defines a parameter, analogous to the "tolerance" parameter in classical models of spatial segregation. Numerical simulations and analytic calculations show that this parameter determines a true phase transition (somewhat reminiscent of phase separation in a binary mixture [**45**]) between a well-mixed and a segregated state. Additionally, we were able to show that minority convictions segregate faster and inter- specific aversion alone may lead to a segregation threshold with similar properties. Together, our results highlight the general principle that a segregation transition is possible in absence of spatial degrees of freedom, provided that conviction-based rewiring occurs on the same time scale of popularity rewirings.

CHAPTER 2

Export dynamics in the network of global economy.

1. Introduction

The work presented in this introductory chapter was fundamental in the definition of my main PhD research topic. It covers the results of the analysis of the exports aggregated at a global level published in [1], finalized by the research group led by professor A. Stella when I joined them during my master thesis.

In this work is presented an analysis of export data aggregated at world global level of 219 classes of products over a period of 39 years (1962-2000). The main result is the construction of a dynamical model that identifies and quantifies the mechanisms by which the evolutions of the various exports affect each other. A stochastic differential description is used to build the model, which outlines a complex network of transfer rates describing how resources are shifted between different product classes, determining also how casual favorable conditions for one export can spread to the other ones.

Here, in the first sections, I cover the process that lead to the realization of this work, following in detail the construction of the stochastic model that defines the dynamics, both to justify adequately the choices that lead to the final model and also to allow the reader to gain a proper insight on the roles played by each term.

Then I'll illustrate the calibration procedure that allows to fit the four free parameters of the model to the data, which in some points differs from the one originally used in [1]. Such differences allow to better understand also the work we did on the model single country analysis, presented in the following chapters.

Finally, some main results will be presented, which will highlight the validity of the model and the potential introduced by the possibility of performing synthetic simulations through with the dynamics. We perform *counterfactual analyses* in which we explore alternative hypothetical scenarios that a country could have faced in its history by choosing different investment policies. Moreover, a long-time extrapolation highlights how the optimal growth conditions change with respect to a 39 year forecast.

2. Data and model presentation

2.1. NBER-UN database and commodities classification. The data used to build the countries-products network have been extracted from the NBER-UN database (*National Bureau of Economic Research - United Nations*) [**46**], which in turn is built upon the UN-Comtrade dataset (*United Nations Commodity Trade Statistics Database*) [**47**]. It contains yearly records of roughly 95% of all the goods traded all over the World. The NBER-UN database covers a time window between the years 1962 and 2000, and it is the result of a cleaning procedure of the original UN-Comtrade dataset in which the amount of traded goods declared by the exporting country were made consistent with those declared by the importing ones.

Commodities are identified and grouped following the Standard International Trade Classification (2nd revision) [48], in which commodities are organized in roughly 1300 representative categories. Each class is identified by a 4-digit code in which the leftmost digit indicates the most general group to which the product belongs to, the second one is more specific and so on. Below we show an illustrative example that exemplifies the structure of the database:

- 0000: Food and live animals;
  - .....
- 1000: Beverages and tobacco;
  - 1100: Beverages;
    - 1110: Non-alchoolic beverages, n.e.s.;
    - 1120: Alcoholic beverages;
      - 1121: Wine of fresh grapes;
      - 1123: Beer made from malt;
      - .....
  - 1200: Tobacco;
    - 1210: Tobacco unmanufactured, tobacco refuse;
      - 1211: Tobacco, not stripped;
      - .....
- 3000: Mineral fuels, lubricants and related materials;
  - .....

Following this notation, the database reports the amount (in thousands of current US dollars) of products exported from a certain country to another one for every year. Unfortunately, for some countries, the data available do not cover the whole time window of 1962-2000. This problem does not show up when looking at data aggregated at a global level, however, as we'll see in chapter 4, it might lead to some issues in the study of the exports relative to a single country. For a more accurate and detailed description of the criteria used in the cleaning procedure of the database look at [49].

Finally, before performing our analysis, we performed a coarse-graining of the NBER-UN database aggregating all the products sharing the first 3 digits of the SITC code, reducing the classification of products to roughly 200 classes (226 in the case of the global aggregation).

3. Dynamical model

A system of stochastic differential equations (SDEs) is used to model the time evolution of the exported products. The system of equations is inspired by a model which was first introduced in the year 2000 by Bouchaud and collaborators in a work [50] in which the authors were trying to address the problem of the distribution of richness in a society. In this scenario, the nodes of the network are identified by the variable $Z_p(t)$ and represent the richness of a "being" (individual or company) p at a certain time t. Its evolution in time is defined by a deterministic term taking into account all the interaction with the other p' nodes through a matrix $J_{pp'}$, while a Gaussian time correlated noise $\eta_p(t)$ is responsible for the stochastic contribution:

$$\partial_t Z_p(t) = \sum_{p' \neq p} \left(J_{pp'} Z_{p'}(t) - J_{p'p} Z_p(t) \right) + \eta_p(t) Z_p(t) \tag{1}$$

$$\langle \eta_p(t_1) \eta_{p'}(t_2) \rangle = \delta_{pp'} \frac{\sigma^2}{\tau} e^{-\frac{|t_2 - t_1|}{\tau}} \tag{2}$$

In a work published in 2014 [51], Bouchaud applies this same model to study the problem of growth in the framework of simple toy-model networks lacking a

deterministic drift. The issue he wanted to address had an important economic footprint but could be easily extended to other contexts [**52, 53, 54, 55**]: he wanted to understand up to which extent it's more convenient to invest a capital in the production of a certain product rather than transfering resources towards new and unexplored directions. Given the absence of a deterministic drift, the growth will be affected only by the initial conditions and the organization of the reallocation of resources among the network.

We can easily see that Eq. 1 (in the case of a regular d-dimensional lattice in which $J_{ij} \equiv J$) is a discretized version of the stochastic heat equation

$$\partial_t Z(\vec{x},t) = J\nabla^2 Z(\vec{x},t) + \eta(\vec{x},t)Z(\vec{x},t) \tag{3}$$

which, after a Cole-Hopf transformation, becomes the famous Kardar-Parisi-Zhang equation [**56**] used to model the growth of surfaces:

$$Z(\vec{x},t) = e^{h(\vec{x},t)} \Longrightarrow \partial_t h(\vec{x},t) = J\nabla^2 h(\vec{x},t) + J\left(\vec{\nabla} h(\vec{x},t)\right)^2 + \eta(\vec{x},t) \tag{4}$$

where $h(\vec{x},t)$ is the growth of the observed surface.

Our model (see [**1**]) is directly inspired by Eq. 1. We write:

$$\frac{\partial Z_{p,t}}{\partial t} = \sum_{p' \neq p} \left(J_{pp'} Z_{p',t} - J_{p'p} Z_{p,t}\right) + \left(\eta_{p,t} + \mu_t\right) Z_{p,t} \tag{5}$$

where Z_p represents the value of the exports of a product p. We also introduced an additional deterministic drift term $\mu(t)$ that takes into account all the average export growth (comprehensive of the inflation).

We will also need to consider a spatial correlation among the nodes, because every Z_p will react in its own peculiar way to favorable (or unfavorable) fluctuations. We will therefore need to modify the noise 2 replacing the Kronecker delta with a correlation matrix $d_{pp'}$ that we will define later on.

$$\langle \eta_{p,t_1} \eta_{p',t_2} \rangle = d_{pp'} \frac{\sigma^2}{\tau} e^{-\frac{|t_2 - t_1|}{\tau}} \tag{6}$$

In the case studied in [**51**] the diagonal form of the correlation was sufficient because different Z_p's reacted to fluctuations in an uncorrelated way.

The SDE that we are introducing can't be integrated analytically, therefore we'll rely on numerical integration schemes to find approximate solutions of the systems, which will take the following form:

$$\mathrm{d}\vec{Z}_t = (\mathbf{J} - \mathbf{J}^T)\vec{Z}_t \,\mathrm{d}t + (\bar{\bar{\eta}}_t + \mu_t)\,\vec{Z}_t \,\mathrm{d}t \tag{7}$$

where $\bar{\bar{\eta}}_t$ is a diagonal matrix whose elements are the η_p's.

In the sections that follows we will provide an in-depth analysis of all the terms composing the equation together with a detailed derivation.

3.1. Ranking. The Z_p's of each commodity seem to provide a ranking that is relatively stable throughout all the years observed, especially in the final ones (see Fig. 1). We formalize this ranking by evaluating the fraction of the value of product p over the total export averaged over the last 10 years (in the time period 1991-2000). We therefore define the ranking of each product as:

$$z_p = \frac{1}{10} \sum_{n=29}^{38} \frac{Z_{p,n}}{\sum_{k=1}^{219} Z_{k,n}} \tag{8}$$

Commodities whose value have dominated the network in the final years will be on the top of the ranking induced by the z_p's. A first result that we observe is that by associating a color the each product, whose wavelength is directly proportional to such ranking, leads to a *rainbow effect* persistent throughout all years. Some

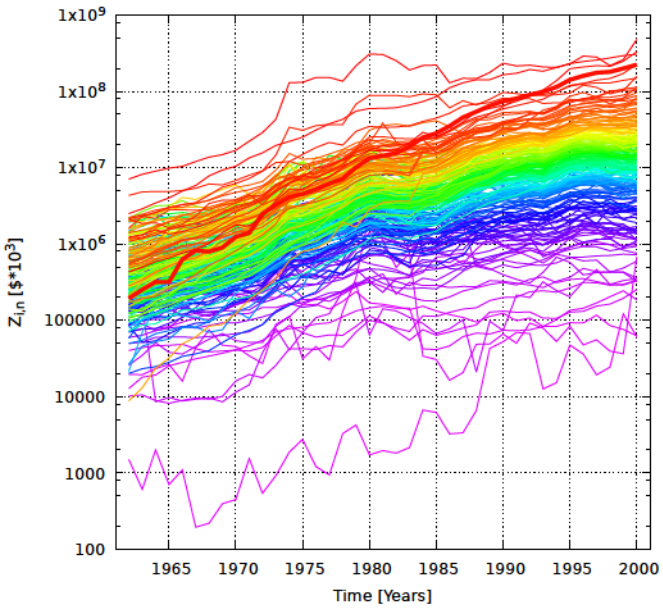


FIGURE 1. Yearly time series of the exports aggregated at a global level between the year 1962 and 2000. Coloring is related to the ranking induced by the z_p's. The thick line highlights the commodities representing electronic goods.

exceptions are of course present: we highlight the case of the electronic goods, who underwent an anomalously exceptional growth in the second half of the twentieth century.

3.2. Correlation matrix. Barabasi et al. [11] made the first attempt to build a network that captured the similarities connecting all products through the concept of *proximity*, which relied on the correlations introduced by analogies in terms of methods and structures necessary for the production of the commodities. For instance, apples and pears are strictly connected because they need similar production facilities and technologies, whereas apples and oil are not that closely related from such perspective. In our model we decided to substitute the measure of proximity with one related to the concept of investment, which allows to highlight connections undetectable by the proximity. For instance, both apples and pears rely deeply on oil for transportation and we cannot neglect these kind of correlations, which are necessary to describe more adequately growth phenomena.

The building block of such correlation matrix is the *logarithmic return*[1]

$$R_{p,n} = \log\left(Z_{p,n}/Z_{p,n-1}\right) \tag{9}$$

which is normalized with respect to time

$$r_{p,n} = \frac{R_{p,n} - \langle R_{p,n}\rangle_n}{\sqrt{\mathrm{Var}(R_{p,n})_n}} \tag{10}$$

[1]Logarithmic returns are widely used in finance in place of a simple return for two main reasons: first, given the value of a logarithmic return, the value corresponding to the associated price is always positive, moreover in finance usually prices are distributed log-normally in time, therefore such returns will follow a Gaussian distribution.

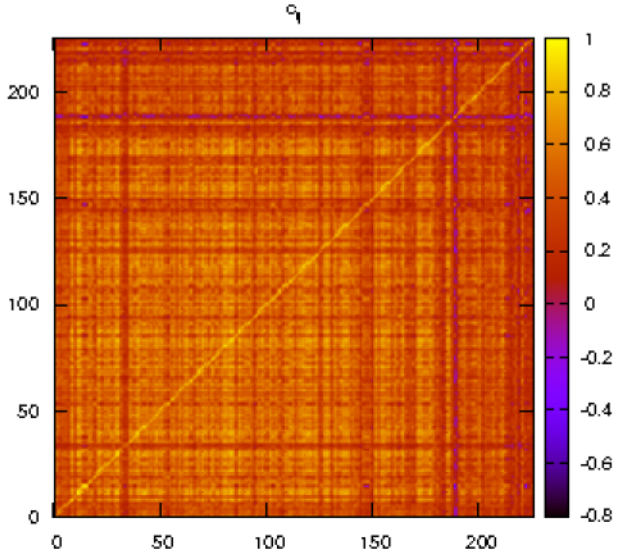


FIGURE 2. Correlation matrix c_{ij} for the 226 products composing the global trade network.

where average and variance are evaluated throughout the years as follows

$$\langle R_{p,n} \rangle_n = \frac{1}{38} \sum_{n=1}^{38} R_{p,n} \tag{11}$$

$$\mathrm{Var}(R_{p,n})_n = \frac{1}{38} \sum_{n=1}^{38} R_{p,n}^2 - \left(\frac{1}{38} \sum_{n=1}^{38} R_{p,n} \right)^2 \tag{12}$$

This standardization of the returns was performed because we are interested in the relative displacement with respect to the average return, since each R_p has its own distribution with different means and variances.

We therefore define an empirical correlation between two products as:

$$c_{pp'} = \frac{1}{38} \sum_{n=1}^{38} r_{p,n} r_{p',n} \tag{13}$$

In the following paragraphs we'll see how this matrix is exploited both in the definition of the transfer matrix $J_{pp'}$ and in the definition of the correlated noise $\eta_{p,n}$.

3.3. Matrix $J_{pp'}$. In many different applications (see [51] and references therein) the transfer matrix has a very low correlation length and reflects the regularities of the lattice on which it is defined, so that all the Z_p's are equivalent. This is not our case, as it is clearly visible from the rainbow effect depicted in Fig. 1. The $J_{pp'}$ matrix therefore must reflect this feature, and will have to play the important role of establishing persistently the correct ranking of the Z_p's in time.

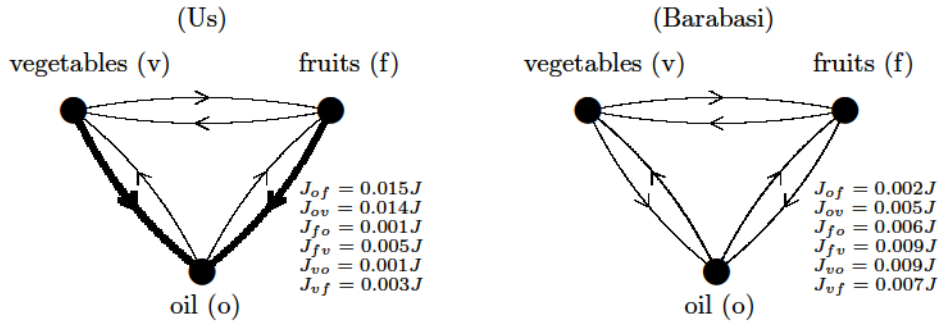


FIGURE 3. Transfer matrix $J_{pp'}$ between oil (o), vegetables (v) and fruit (f) obtained with our and Barabasi's model [11]. The width of the lines is proportional to the value of the $J_{pp'}$'s. From this comparison we see how our model is able to correlate products products that in Barabasi network result far with one another. In our model is also clearly visible the asymmetry that characterizes the transfer matrix.

In order to fulfill this task, we take a look at the Eq. 5 in the absence of noise:

$$\frac{\partial Z_{p,t}}{\partial t} = \sum_{p' \neq p} (J_{pp'} Z_{p',t} - J_{p'p} Z_{p,t}) + \mu_t Z_{p,t} = \sum_{p'} A_{pp'} Z_{p',t} + \mu_t Z_{p,t} \tag{14}$$

where the matrix $\mathbf{A}$ has entries $A_{pp'} = J_{pp'}$ for $p \neq p'$ and $A_{pp} = -\sum_{p' \neq p} J_{p'p}$ as diagonal terms[2]. The solutions of such equation tend to a long-time attractor whose ranking is defined by the kernel of the matrix $\mathbf{A}$. We can exploit such fact by choosing

$$J_{pp'} \propto z_p \tag{15}$$

With such choice, on average, the ranking of a solution to Eq.r̃efeq:dynamics-equation converges to the ranking of the real data. An important feature that follows this choice is that an asymmetry of the matrix $J_{pp'}$ (see Fig. 3). In this way we are able to address problems deriving by the fact that, given a pair of products, the two will be sensible differently to the fluctuations of one another. For instance, if the export of apples increases, also the amount of oil used as fuel from apple producers will increase, while they opposite is also true but at a smaller magnitude.

Eq. 5 with the choice 15 is also consistent (on average) with the *gravitational model* for what concerns the mass-related terms. Such model is used in economics to predict a positive correlation in international trades between countries, where both the size of a country (measured as Gross Domestic Product) as well as the distance separating them, play an important role.

To build a matrix fully consistent with this model we need to introduce a quantity that acts as distance, which, in our case, means that it should provide a measure that quantifies how much a product can influence the value of the export of another one. One possibility, is to use the correlation matrix $c_{pp'}$ as an inverse distance: the higher the correlation, the closer the two products will be. The choice

[2]This matrix is analogous to the transition matrices in continuous Markov chains, in fact is column-stochastic (i.e. summing over the entries of each columns yields zero) [57].

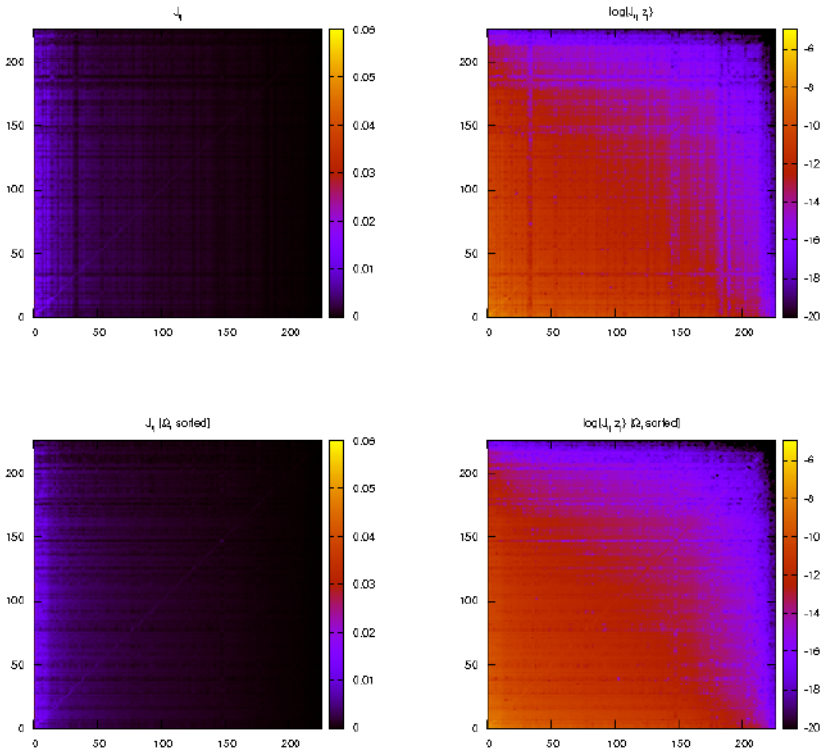


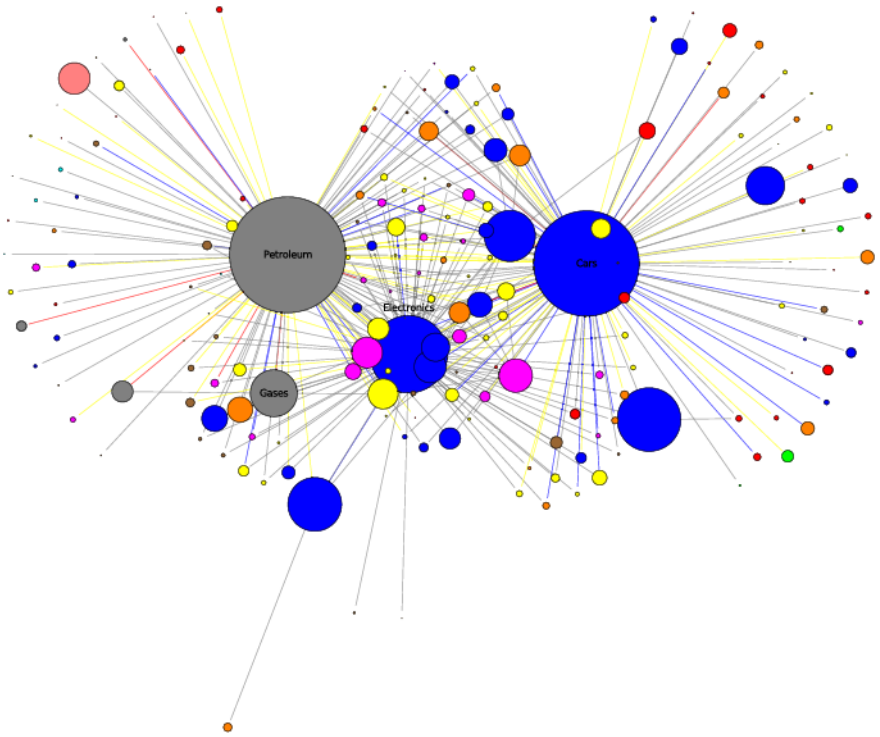
FIGURE 4. Transfer matrices $J_{pp'}$ (on the left) and $\log [J_{pp'} z_{p'}]$ (on the right) with $G = 1$. In the bottom figures the matrices have been sorted by the connectivity Ω_p.

for defining the correlation matrix will therefore be:

$$J_{pp'} = G z_p \, |c_{pp'}| \tag{16}$$

where G is a coupling constant, which will be determined in the calibration procedure of the model. The validity of such choice will be shown later on with a comparison with real data. In Fig. 4 we see the matrices $J_{pp'}$ and $\log [J_{pp'} z_{p'}]$ in the case $G = 1$. In the two plots on the bottom the entries have been sorted with respect to the connectivity of every product defined as $\Omega_p = \sum_{q \neq p} J_{pq}$.

Finally, in Fig. 5 we show a graphical representation of the network with adjacency matrix $L_{pp'} = (J_{pp'} + J_{p'p})/2$ (G is set to be equal to 1). Oil and cars are clearly the dominating nodes, which are connected with almost every other commodity.



Food and live animals	Beverages and tobacco
Crude materials, inedible, except fuels	Mineral fuels, lubricants and related materials
Animal and vegetable oils, fats and waxes	Chemicals and related products, n.e.s.
Manufactured goods classified chiefly by material	Machinery and transport equipment
Miscellaneous manufactured articles	Commodities and transactions not classified elsewhere in the SITC

FIGURE 5. Graphical representation of the global products network. We plotted only edges with a weight $L_{pp'} = (J_{pp'} + J_{p'p})/2 > 0.01$. The size of the node is proportional to the ranking of each product, while the associated color is representative of the category they belong to, according to the SITC4 classification shown in the legend. The links are colored depending on the weight $L_{pp'}$: gray if $L_{pp'} < 0.015$, yellow if $0.015 < L_{pp'} < 0.0175$, blue if $0.0175 < L_{pp'} < 0.02$ and red if $L > 0.02$.

3.4. Noise. Growth models [**51, 55**] usually do not include correlation among nodes, and assume only a time correlation as in 2. Using this kind of noise, we find that we are unable to reproduce the correct amount of traded goods with simulations. This justifies the need to introduce a correlation matrix $d_{pp'}$ in place of a Kronecker delta in the definition of the noise:

$$\langle \eta_{p,t_1} \eta_{p',t_2} \rangle = d_{pp'} \frac{\sigma^2}{\tau} e^{-\frac{|t_2 - t_1|}{\tau}} \tag{17}$$

The parameter τ quantifies the characteristic duration of either a crisis or a favorable fluctuation, while σ quantifies the weight of the stochastic contribution to the dynamics. Both these parameters are determined in the calibration procedure.

Concerning the node-to-node correlation, our choice falls on $d_{pp'} = c_{pp'}$, the same matrix used in the definition of the transfer matrix. The validity of such choice is tested in the simulations of synthetic histories after the calibration procedure has been performed. Below we illustrate the procedure used to create a noise with such correlation feature, in order to understand also how we will be able to simulate it when integrating the equation numerically.

First we'll take care of the spatial correlation. We start with a zero-mean Gaussian noise $\xi_p(t)$ correlated accordingly with the matrix $c_{pp'}$ and with a Dirac delta in time:

$$\langle \xi_p(t) \rangle = 0 \tag{18}$$

$$\langle \xi_p(t_1) \xi_{p'}(t_2) \rangle = c_{pp'} \delta(t_1 - t_2) \tag{19}$$

To produce this kind of noise we start with an uncorrelated white Gaussian noise $\left\langle \tilde{\xi}_p(t_1) \tilde{\xi}_{p'}(t_2) \right\rangle = \delta_{pp'} \delta(t_1 - t_2)$ and we decompose the matrix $\mathbf{C}$ following a variation of the Cholesky decomposition, called the *LDLT decomposition*. Such procedure allows to express the original matrix as a product of a lower-triangular matrix ($\mathbf{L}$), a diagonal one ($\mathbf{D}$) and the transposed of the first one ($\mathbf{L}^T$)

$$\mathbf{C} = \mathbf{L}\mathbf{D}\mathbf{L}^T \tag{20}$$

Therefore by choosing $\xi(t) = \mathbf{L}\mathbf{D}^{1/2}\tilde{\xi}(t)$ we end up with a noise that satisfies the spatial correlation requirements.

When implementing this decomposition numerically we need however to keep into account the possibility of introducing approximation errors that might cause the diagonal terms to be equal to zero or negative. In our case, we are dealing $\mathbf{C}$ is a 219×219 matrix, and this problems arises starting from the 39th term (the products were sorted depending on the ranking, therefore the $d_{pp'}$ terms decrease as the index p grows). Since these numbers become increasingly small, we opted for setting them equal to zero and proceeding with the decomposition procedure.

In order to obtain the desired time-correlation we apply the following transformation to the $\xi(t)$ we obtained previously:

$$\begin{cases} \eta_p(t_0) = & \frac{\sigma}{\sqrt{\tau}} \xi(t_0) \\ \eta_p(t) = & \rho \eta_p(t - \mathrm{d}t) + \sqrt{1 - \rho^2} \frac{\sigma}{\tau} \xi_p(t) \end{cases} \tag{21}$$

where t_0 represents the lower bound of the time interval on which we will integrate the equation and we set for convenience $\rho = e^{-\mathrm{d}t/\tau}$.

Now let's check that the noise defined in this way satisfies 17. Setting $t_1 = t$ and $t_2 = 0$ without loss of generality (correlation is invariant with respect to time translations), we get:

$$\eta_p(0) = \frac{\sigma}{\sqrt{\tau}} \xi_p(0) \tag{22}$$

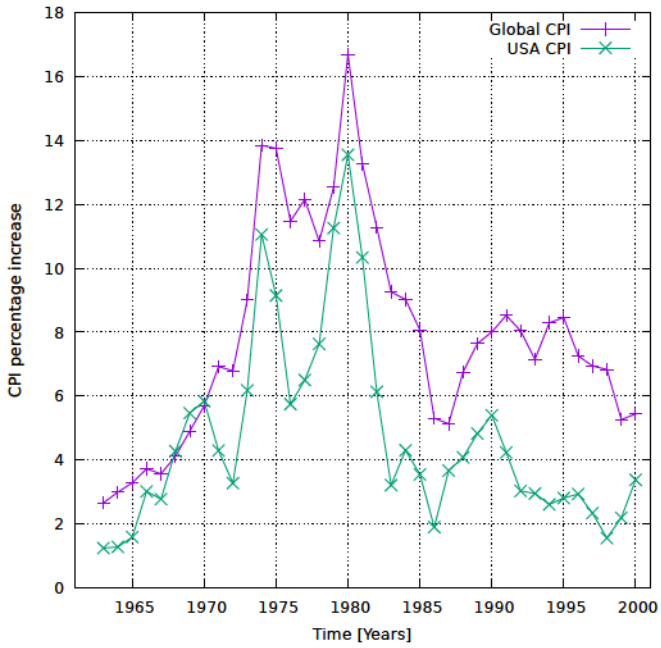


FIGURE 6. Global vs. USA inflation in the time period 1962-2000.

Anno	I_t	Anno	I_t	Anno	I_t	Anno	I_t
1963	2.66	1973	9.01	1983	9.25	1993	7.15
1964	2.99	1974	13.85	1984	9.03	1994	8.29
1965	3.28	1975	13.75	1985	8.07	1995	8.46
1966	3.72	1976	11.47	1986	5.30	1996	7.26
1967	3.56	1977	12.15	1987	5.13	1997	6.94
1968	4.11	1978	10.85	1988	6.75	1998	6.83
1969	4.90	1979	12.53	1989	7.64	1999	5.26
1970	5.69	1980	16.70	1990	8.00	2000	5.44
1971	6.92	1981	13.27	1991	8.52		
1972	6.79	1982	11.28	1992	8.04		

TABLE 1. Global inflation rate in the period 1963-2000.

$$\eta_p(t) = \rho\eta_p(t-t) + \sqrt{1-\rho^2}\frac{\sigma}{\tau}\xi_p(t) = \tag{23}$$
$$= \rho\frac{\sigma}{\sqrt{\tau}}\xi_p(0) + \sqrt{1-\rho^2}\frac{\sigma}{\tau}\xi_p(t)$$

where in this case $\rho = e^{-t/\tau}$. Evaluating their correlation yields

$$\langle\eta_p(t)\eta_{p'}(0)\rangle = \rho\frac{\sigma^2}{\tau}\langle\xi_p(0)\xi_{p'}(0)\rangle + \sqrt{1-\rho^2}\frac{\sigma^2}{\tau}\langle\xi_p(t)\xi_{p'}(0)\rangle = \tag{24}$$
$$= \rho\frac{\sigma^2}{\tau}c_{pp'} + \sqrt{1-\rho^2}\frac{\sigma^2}{\tau}c_{pp'}\delta(t)$$

which coincides with 17.

3.5. Drift and inflation. The term μ_t represents the deterministic part contributing to the growth rate, which is common to every single product independently on the

existence of the network. One would expect every commodity to have its own time-dependent growth rate, however we chose to consider an average growth rate $\bar{\mu}$ shared among all products in order to contain the number of parameters of the model. The time dependence is instead captured by the inflation component I_t

$$\mu_t = \bar{\mu} + I_t \tag{25}$$

This assumption was suggested by the average overall trend observed throughout the years in Fig. 1. In the end the model has only one parameter that captures growth contributions coming from external inputs, since inflation is simply a step function in time whose values are given.

However inflation is a country-specific quantity, while here we are dealing with a global aggregated network. One choice could be to use the inflation of the country that lead the economic panorama in the years observed (the USA) as in [1]. However, taking inspiration from a work found in an economics journal [58], we defined a *global inflation* as a weighted average of the inflation of each country in which the Gross Domestic Product (GDP) acted as weight. The resulting inflation is reported in Fig. 6 while the numerical values are shown in Table 1.

$$I_t^{global} = \frac{\sum_{c\in\{countries\}} I_t^c \cdot GDP_t^c}{\sum_{c\in\{countries\}} GDP_t^c} \tag{26}$$

The data used for the inflation rates were taken from the OECD database [59], while those of the GDP from a database redacted by the World Bank Group [60].

4. Integration and dynamics

In order to solve Eq. 5 we need to proceed with a numerical integration of the following system of SDEs:

$$\mathrm{d}Z_{p,t} = \sum_{p'\neq p} \left(J_{pp'}Z_{p',t} - J_{p'p}Z_{p,t}\right)\mathrm{d}t + \left(\bar{\mu} + I_t + \eta_{p,t}\right)Z_{p,t}\,\mathrm{d}t \tag{27}$$

with the quantities previously introduced.

In a regime in which $\mathrm{d}t \ll \tau$ different integration schemes lead to identical results. We opted to follow Itô's prescription, which is the most widely used in finance and economics.

We use the correlated Gaussian noise we previously defined to build a Brownian motion:

$$\mathrm{d}W_{p,t} = \sqrt{\mathrm{d}t}\xi_{p,t} \tag{28}$$

so that we can rewrite $\eta_{p,t}$ as

$$\eta_{p,t} = \rho\eta_{p,t-\mathrm{d}t} + \sqrt{1-\rho^2}\frac{\sigma}{\sqrt{\tau\,\mathrm{d}t}}\mathrm{d}W_{p,t} \tag{29}$$

Substituting this term in the SDE and separating the deterministic and stochastic parts we find

$$\mathrm{d}Z_{p,t} = \left[\sum_{p'\neq p}\left(J_{pp'}\frac{Z_{p',t}}{Z_{p,t}} - J_{p'p}\right) + \rho\eta_{p,t-\mathrm{d}t} + \bar{\mu} + I_t\right]Z_{p,t}\,\mathrm{d}t + \sqrt{1-\rho^2}\frac{\sigma}{\sqrt{\tau}}Z_{p,t}\sqrt{\mathrm{d}t}\,\mathrm{d}W_{p,t} \tag{30}$$

In finance one usually prefers to use integration schemes characterized by a high weak convergence, since the interest is directed towards the ability to reproduce on average the exact solution rather than reproducing correctly one single evolution. The choice fell on a second order Runge-Kutta scheme adapted to SDE systems characterized by both weak and strong convergences of order 1.

Following such method, we define the coefficients $\vec{a}(t, \vec{Z}_t)$ for the deterministic part and $\vec{b}(t, \vec{Z}_t)$ for the stochastic one:

$$\vec{a}(t, \vec{Z}_t) = \left[\sum_{p' \neq p} \left(J_{pp'} \frac{Z_{p',t}}{Z_{p,t}} - J_{p'p} \right) + \rho \eta_{p,t-\mathrm{d}t} + \bar{\mu} + I_t \right] \vec{Z}_t \tag{31}$$

$$\vec{b}(t, \vec{Z}_t) = \sqrt{1-\rho^2} \frac{\sigma}{\sqrt{\tau}} \sqrt{\Delta t} \vec{Z}_t \tag{32}$$

where Δt is the analogous to the differential $\mathrm{d}t$ and represents the integration step.

With $\{\vec{w}_n\}_{n=0,N}$ being the approximation of the solution evaluated in the points of the integration mesh $[0, T]$, the integration scheme is

$$\begin{cases} \vec{w}_0 = \vec{X}_0 \\ \vec{w}_{n+1} = \vec{w}_n + \frac{\vec{K}_{1,n} + \vec{K}_{2,n}}{2} \\ \vec{K}_{1,n} = \vec{a}(t_n, \vec{w}_n)\Delta t + \vec{b}(t_n, \vec{w}_n) \left(\Delta W_n - S_n \sqrt{\Delta t} \right) \\ \vec{K}_{2,n} = \vec{a}(t_{n+1}, \vec{w}_n + \vec{K}_{1,n})\Delta t + \vec{b}(t_{n+1}, \vec{w}_n + \vec{K}_{1,n}) \left(\Delta W_n + S_n \sqrt{\Delta t} \right) \end{cases} \tag{33}$$

where S_n is a random variable which can assume values ± 1 each with probability $1/2$.

Methods with higher order of convergence are known, but since the relatively high (226) number of SDEs composing the system the time required for the computation becomes a matter that needs to be taken into account. Setting an integration step of $\Delta t = 1/1000$, takes around 8 seconds to simulate the whole dynamics using a single-thread processor with 3.1 GHz clock frequency.

5. Calibration

The dynamical model we developed in the previous sections is characterized by a set of four free parameters: G, $\bar{\mu}$, σ and τ. Therefore we need to find values for such parameters that allow to reproduce faithfully (quantitatively and qualitatively) the historical data.

We divide Eq. 30 by $Z_{p,t}$ without expanding iteratively the noise $\eta_{p,t}$:

$$\frac{\mathrm{d}Z_{p,t}}{Z_{p,t}} = \left[\sum_{p'\neq p}\left(J_{pp'}\frac{Z_{p',t}}{Z_{p,t}} - J_{p'p}\right) + \eta_{p,t} + \bar{\mu} + I_t\right]\mathrm{d}t \tag{34}$$

and we integrate in the time interval $[n_1, n_2]$ obtaining

$$\log\frac{Z_{p,n_1}}{Z_{p,n_2}} = \int_{n_1}^{n_2}\sum_{p'\neq p}\left(J_{pp'}\frac{Z_{p',t}}{Z_{p,t}} - J_{p'p}\right)\mathrm{d}t + \int_{n_1}^{n_2}(\eta_{p,t} + \bar{\mu} + I_t)\,\mathrm{d}t \tag{35}$$

For the sake of readability, it is convenient to define the following functions

$$f_p(n_1, n_2) = \frac{1}{n_2 - n_1}\left(\log\frac{Z_{p,n_1}}{Z_{p,n_2}} - \int_{n_1}^{n_2} I_t\,\mathrm{d}t\right) \tag{36}$$

$$g_p(n_1, n_2) = \frac{1}{n_2 - n_1}\int_{n_1}^{n_2}\sum_{p'\neq p}\left(z_p\,|c_{pp'}|\,\frac{Z_{p',t}}{Z_{p,t}} - z_{p'}\,|c_{pp'}|\right)\mathrm{d}t \tag{37}$$

where we expressed $J_{pp'} = G z_p\,|c_{pp'}|$. Eq. 35 takes the form

$$f_p(n_1, n_2) = G g_p(n_1, n_2) + \bar{\mu} + \frac{1}{n_2 - n_1}\int_{n_1}^{n_2}\eta_{p,t}\,\mathrm{d}t \tag{38}$$

For the evaluation of g_p we need to approximate the integral with a summation since the Z_p's are not continuous in time but are reported yearly. We therefore write

$$g_p(n_1, n_2) = \frac{1}{n_2 - n_1}\sum_{p'\neq p}\sum_{n=n_1+1}^{n_2}\left[z_p\,|c_{pp'}|\left(\frac{Z_{p',n-1}}{Z_{p,n-1}} + \frac{Z_{p',n}}{Z_{p,n}}\right) - z_{p'}\,|c_{pp'}|\right] \tag{39}$$

Now we get into the details of the calibration procedure, which will provide us the values for the free parameters G,σ, τ and $\bar{\mu}$ (in this order).

5.1. Calibration of G. The parameter G is easily found to be the slope of Eq. 38. We therefore evaluate all the pairs $f_p(n_1, n_2)$ and $g_p(n_1, n_2)$ making the indexes vary in the ranges $p = 1, ..., 219$, $n_1 = 0, ..., 37$ and $n_2 = n_1 + 1, ..., 38$ and perform a linear regression.

One problem arises in this procedure: when $|g_p|$ is too small, the stochastic component of the noise does not allow to correctly estimate G. After synthetic simulations of a 39 years period starting always from the same initial conditions but with arbitrarily chosen values for G, we verified that using only the 10th of the points with higher $|g_p|$ we were able to estimate the correct value of G with a confidence level of 10% (see Fig. 7). Given the high number of pairs composing the scatter plot we exploited the *quicksort algorithm* to perform the sorting procedure.

The parameters obtained with this regression are:

$$G = 0.042 \pm 0.001\ \mathrm{y}^{-1} \tag{40}$$

$$k = 0.044 \pm 0.001\ \mathrm{y}^{-1} \tag{41}$$

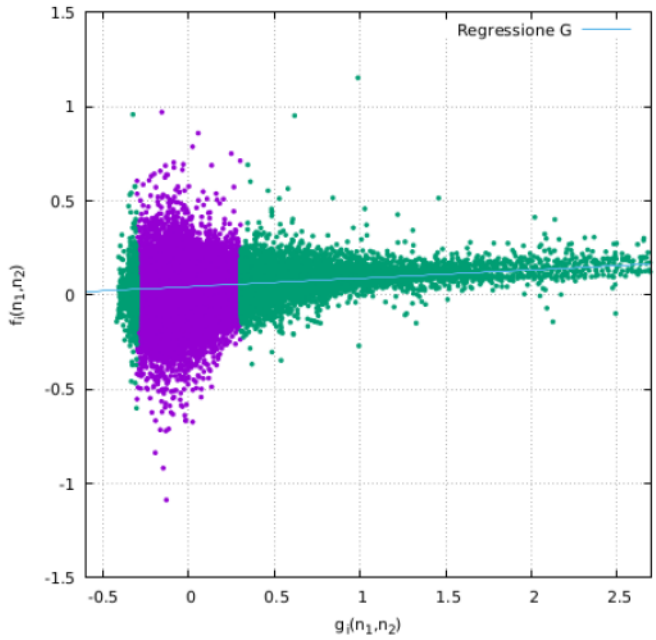


FIGURE 7. Calibration of the parameter G. Scatter plot of $f_p(n_1, n_2)$ vs. $g_p(n_1, n_2)$, with $p = 1, ..., 226$, $n_1 = 0, ..., 37$ e $n_2 = n_1 + 1, ..., 38$. In green we see the 10% of the points with grater $|g_p|$ while in purple we see the remaining 90%. The thick blue line represents the linear regression $y = Gx + k$.

5.2. σ and τ calibration. These two parameters are calibrated simultaneously in the same procedure. We exploit the fact that the variance of the integral of the noise is a function of the parameters σ and τ. Isolating the noise in Eq. 38 evaluated in the interval $[0, n]$ we find

$$\int_0^n \eta_{p,t}\, \mathrm{d}t = n\left[f_p(0,n) - G g_p(0,n) - \bar{\mu}\right] \tag{42}$$

and by evaluating the variance we get rid of the constant term $\bar{\mu}$

$$\mathrm{Var}\left[\int_0^n \eta_{i,t}\, \mathrm{d}t\right] = n^2 \mathrm{Var}\left[f_i(0,n) - G g_i(0,n)\right] \tag{43}$$

The right hand side of the equation is evaluated on the real historical data, while the left one can be evaluated in an explicit way. First we rewrite the variance as

$$\mathrm{Var}\left[\int_0^n \eta_{p,t}\, \mathrm{d}t\right] = \mathbb{E}\left[\left(\int_0^n \eta_{p,t}\, \mathrm{d}t\right)^2\right] - \mathbb{E}\left[\int_0^n \eta_{p,t}\, \mathrm{d}t\right]^2 \tag{44}$$

so we are left with the evaluation of the two averages on the right hand of the above equation. Exploiting the recursive form of the noise (Eq. 21) we find that its

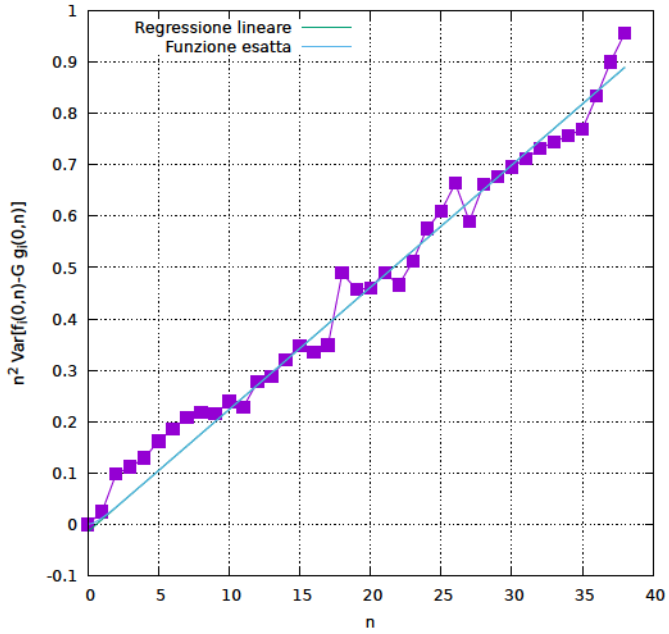


FIGURE 8. Calibration of the parameters σ and τ. It is clear how already starting from the first year we are not able to distinguish the exact function from the linear regression.

expectation is

$$\begin{aligned}(45)\qquad \mathbb{E}\left[\int_{n_1}^{n_2}\eta_{p,t}\,\mathrm{d}t\right] &= \mathbb{E}\left[\int_{n_1}^{n_2}\left[\rho\eta_{p,t-\mathrm{d}t}-\sqrt{1-\rho^2}\frac{\sigma}{\sqrt{\tau}}\xi_{p,t}\right]\mathrm{d}t\right]\\ &= \rho\mathbb{E}\left[\int_{n_1}^{n_2}\eta_{p,t-\mathrm{d}t}\,\mathrm{d}t\right]-\sqrt{1-\rho^2}\frac{\sigma}{\sqrt{\tau}}\int_{n_1}^{n_2}\underbrace{\mathbb{E}\left[\xi_{p,t}\right]}_{=0}\mathrm{d}t\\ &= \rho\int_{n_1}^{n_2}\mathbb{E}\left[\eta_{p,t-\mathrm{d}t}\right]\mathrm{d}t\end{aligned}$$

where we used the fact that $\xi_{p,t}\sim\mathcal{N}(0,1)$ and therefore $\mathbb{E}\left[\xi_{p,t}\right]=0$. Iterating such process we finally get to the evaluation of the term $\eta_{p,0}=\xi_{p,0}\rho\sigma/\sqrt{\tau}$ which is identically zero for the same reason. In the end we were able to prove that

$$(46)\qquad \mathbb{E}\left[\int_{n_1}^{n_2}\eta_{p,t}\,\mathrm{d}t\right]=0$$

Now we address the second term by expanding the squared power:

$$\begin{aligned}(47)\qquad \mathbb{E}\left[\left(\int_0^n\eta_{p,t}\,\mathrm{d}t\right)^2\right] &= \mathbb{E}\left[\int_0^n\int_0^n\eta_{p,t}\eta_{p,t'}\,\mathrm{d}t\,\mathrm{d}t'\right]\\ &= \int_0^n\int_0^n\mathbb{E}\left[\eta_{p,t}\eta_{p,t'}\right]\mathrm{d}t\,\mathrm{d}t'\end{aligned}$$

exploiting the noise correlation of Eq. 17 with $c_{pp}\equiv 1$ we get

$$(48)\qquad \mathbb{E}\left[\left(\int_0^n\eta_{p,t}\,\mathrm{d}t\right)^2\right]=\frac{\sigma^2}{\tau}\int_0^n\int_0^n e^{-\frac{|t-t'|}{\tau}}\,\mathrm{d}t\,\mathrm{d}t'$$

Now we perform the variable change $t - t' = t''$ in the integral and we decompose the absolute value in its positive and negative parts:

$$\begin{aligned}(49)\qquad \mathbb{E}\left[\left(\int_0^n \eta_{p,t}\,\mathrm{d}t\right)^2\right] &= \frac{\sigma^2}{\tau}\int_0^n \mathrm{d}t' \int_{-t'}^{n-t'} e^{-\frac{|t''|}{\tau}}\,\mathrm{d}t'' \\ &= \frac{\sigma^2}{\tau}\int_0^n \mathrm{d}t' \left(\int_{-t'}^{0} e^{\frac{t''}{\tau}}\,\mathrm{d}t'' + \int_0^{n-t'} e^{-\frac{t''}{\tau}}\,\mathrm{d}t''\right) \\ &= \sigma^2 \int_0^n \mathrm{d}t' \left(2 - e^{-\frac{t'}{\tau}} - e^{-\frac{n-t'}{\tau}}\right) \\ &= 2\sigma^2\left[n + \tau\left(e^{-\frac{n}{\tau}} - 1\right)\right]\end{aligned}$$

which gives us finally the following relation

$$(50)\qquad n^2 \mathrm{Var}\left[f_p(0,n) - G g_p(0,n)\right] = 2\sigma^2\left[n + \tau\left(e^{-\frac{n}{\tau}} - 1\right)\right]$$

Performing a regression of the these points provides us with values for σ and τ (Fig. 8), however the form of such function makes it hard to find the absolute minimum with a least-squares method. To solve such problem, we can see that as n grows, the exponential term decays with characteristic time τ, and the function takes a linear form:

$$(51)\qquad 2\sigma^2\left[n + \tau\left(e^{-\frac{n}{\tau}} - 1\right)\right] \to_{n \gg \tau} 2\sigma^2\left[n - \tau\right]$$

which allows us to fit the parameters straightforwardly with a linear regression.

The criterion chosen to determine the lower cutoff for the points used consists in increasing the lower bound for the fitting procedure until the relative error of the fitted function evaluated at the lower bound (with respect to the real data) is lower than 1%. We see that this condition is already verified starting from the 12th year, so we will be dropping only 12 points of the 39 available. The results of the regression are:

$$(52)\qquad \sigma = 0.109 \pm 0.002\ \mathrm{y}^{-1/2}$$

$$(53)\qquad \tau = 0.607^{+0.4}_{-0.5}\ \mathrm{y}$$

The error of τ has been expressed in such way because we observed that the dynamics of the Z_p's does not really change by choosing for τ any value between 0.1 and 1. In Fig. 8 we see that the linear regression is very close to the exact function comprehensive of the exponential term, so the approximation we performed for the calibration is consistent with the results.

5.3. $\bar{\mu}$ calibration. The parameter $\bar{\mu}$ quantifies the average growth rate of all the commodities. A correct calibration of such parameter should therefore allow to reproduce quantitatively the growth observed after 38 years of time evolution.

In this scenario it is convenient to use the *logarithmic growth* we used before to define yearly returns (it is addictive). The growth of a a commodity p in an time interval $[T_1, T_2]$ will be defined as:

$$(54)\qquad \lambda_{p,T_2-T_1} = \frac{1}{T_2 - T_1}\log\frac{Z_{p,T_2}}{Z_{p,T_1}}$$

The average growth of the real data in the period $[0, 38]$ will therefore be:

$$(55)\qquad \lambda_{38}^{real} = \frac{1}{38 \cdot 226}\sum_{p \in \text{goods}} \log\frac{Z_{p,38}^{real}}{Z_{p,0}^{real}} = 0.0889\ \mathrm{y}^{-1}$$

Our task is to reproduce correctly this quantity when simulating the evolution of the data, so we will need to understand the intrinsic growth of the network in the

absence of any deterministic drifts. For this reason we define the growth deprived of the contribution coming from the inflation as:

$$c_{38} = \lambda_{38} - \frac{1}{38}\int_0^{38} I_t \, \mathrm{d}t \tag{56}$$

which, for the real data, is $c_{38}^{real} = 9.77 \times 10^{-3}\ \mathrm{y}^{-1}$.

Using the auxiliary functions we introduced for the calibration of the parameter G we see that growth we defined is simply the average of the function f over the products

$$c_{38} = \langle f_p(0, 38) \rangle \tag{57}$$

Evaluating this quantity on the real data we find indeed 55.

Given Eq. 38 the following relation holds:

$$c_{38}^{real} = G \left\langle g_p^{real}(0, 38) \right\rangle + \bar{\mu} \tag{58}$$

For a correct estimation of $\bar{\mu}$, we first proceed with simulations in which we use the calibrated values of the parameters G, σ and τ and we artificially set $\bar{\mu} = 0$. Introducing the $\sim$ superscript to characterize such quantities, the equation they must satisfy will be the following:

$$\tilde{f}_p(0, 38) = G\tilde{g}_p(0, 38) + \frac{1}{38}\int_0^{38} \eta_t \, \mathrm{d}t \tag{59}$$

which substituted in $\tilde{c}_{38} = \left\langle \tilde{f}_p(0, 38) \right\rangle$ yield

$$\tilde{c}_{38} = G \left\langle \tilde{g}_i(0, 38) \right\rangle \tag{60}$$

Taking the difference between Eq. 58 and 60 we find that the following relation must hold for $\bar{\mu}$:

$$\bar{\mu} = c_{38}^{real} - \tilde{c}_{38} + G\left(\langle \tilde{g}_p(0, 38) \rangle - \left\langle g_p^{real}(0, 38) \right\rangle\right) \tag{61}$$

The term inside the parenthesis is negligible because depends on the differences of the ratios between $Z_{p,t}^{real}$ and $Z_{p,t}^{tilde}$ which have very large values, so finally we find that a good approximation of $\bar{\mu}$ is

$$\bar{\mu} = c_{38}^{real} - \tilde{c}_{38} \tag{62}$$

We performed 100000 simulations for the estimation of $\tilde{c}_{38}$ with $\bar{\mu}$ set artificially equal to zero. Averaging the results yields

$$\tilde{c}_{38} = (3.46 \pm 0.05) \times 10^{-3}\ \mathrm{y}^{-1} \tag{63}$$

which in the end yields

$$\bar{\mu} = (6.31 \pm 0.05) \times 10^{-3}\ \mathrm{y}^{-1} \tag{64}$$

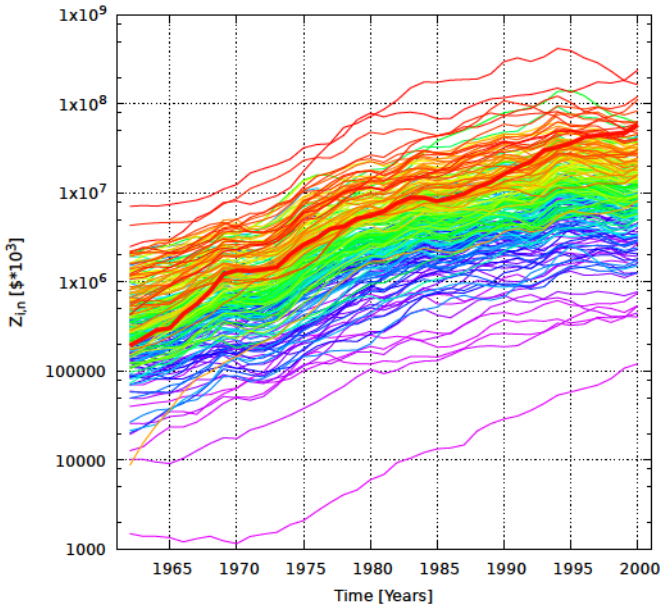


FIGURE 9. Simulation of the evolution of global exports starting from the real initial condition of 1962. The color scheme is assigned as in Fig. 1, with the thicker line referring to electronic commodities. The values used for the parameters are those obtained with the calibration procedure: $G = 0.042$ y^{-1}, $\bar{\mu} = 6.31 \times 10^{-3}$ y^{-1}, $\sigma = 0.109$ y$^{-1/2}$ and $\tau = 0.607$ y.

6. Results

The first result obtained with the model is obviously a synthetic reproduction of the historical data depicted in Fig. 1. Starting from the real Z_p of 1962 and using the calibrated values for the parameters G, $\bar{\mu}$, σ and τ and the inflation $I(t)$ we can simulate the evolution of the exports for 38 years, until the last year observed (2000).

In Fig. 9 we show the results obtained with a simulation with integration step $\Delta t = 1/1000$: we immediately see that not only the rainbow effect has been reproduced, but also that the goods that in the initial years were out of ranking (e. g. electronic commodities) reach the right position in the final year, undergoing an anomalous growth compared to the average (this result will be formally quantified in the following section).

6.1. Ranking reproduction. One way to validate quantitatively the convergence of the simulated rankings to the real one in the final years is to evaluate yearly the *Spearman rank corelator* r_s [**61**]. Such coefficient evaluates the correlation between the rankings of two sets of data instead of the actual values of the variables (as with Pearson's coefficient [**62**]). Spearman's correlator assumes values between -1 and +1:

- $r_s = +1$ indicates a perfect positive correlation, meaning that the two rankings coincide;
- $r_s = -1$ indicates a perfect negative correlation which happens if the two rankings are one the opposite of the other;

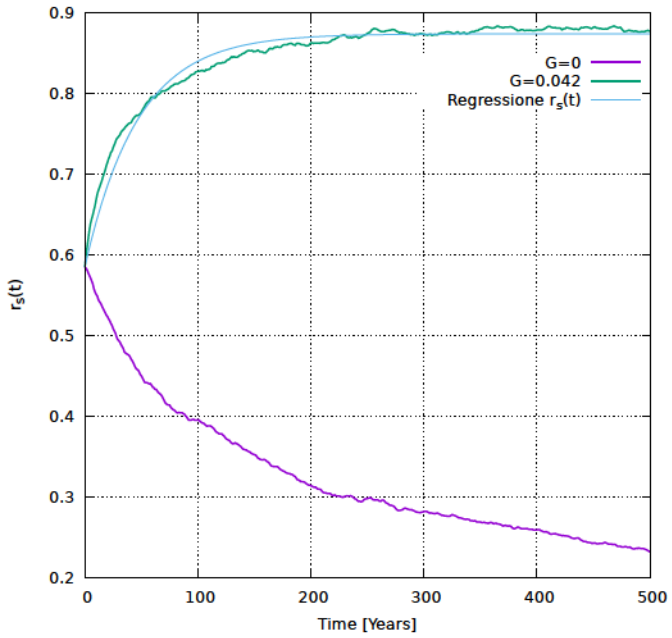


FIGURE 10. Evolution in time of the Spearman rank correlator. Both curves (the green one with the calibrated value for G, the purple one with $G = 0$) have been obtained averaging 100 simulation evolved for a time period of 500 years. Single simulations fluctuate around their respective curves. The light blue line represents the exponential regression that provides the characteristic time $\tau_s \simeq 47$ years.

- $r_s = 0$ means that there is no correlation at all.

if the rankings are made of non-equal integer numbers, given a set of N elements $\{X_i\}_{i=0\ldots N}$ and $\{Y_i\}_{i=0\ldots N}$ the correlator is defined as:

$$r_s = 1 - \frac{6 \sum_i \left(\mathrm{rg}\left(X_i\right) - \mathrm{rg}\left(Y_i\right)\right)^2}{n\left(n^2 - 1\right)} \tag{65}$$

where rg () is the ranking of the element within its set.

We proceed simulating synthetic evolutions of the data starting from the real initial conditions and we compare the yearly ranking $z_p^{simul}(t)$ with the real one z_p^{real} (which is evaluated in the last decade of historical data as in 8).

As we can clearly see in Fig. 10 the correlator tends exponentially in time towards a positive correlation (fluctuations are the reason why the asymptotic value stays below 1). This proves that our model is able to reproduce correctly also the evolution of the goods which underwent an anomalous growth. In the same figure we also reported the evolution of the correlator in simulations in which we artificially set $G = 0$: it is clear how this scenario leads to the total absence of correlation among the two systems. This fact shows the importance of the role played by the transfer matrix $J_{pp'}$ and provides ulterior validation to the choices made in the construction of our model..

An analysis of the curve that arises from the evolution of r_s provides us with a sort of characteristic time τ_s which can be interpreted, from an economical perspective, as the time the global economy needs to re-establish the correct ranking

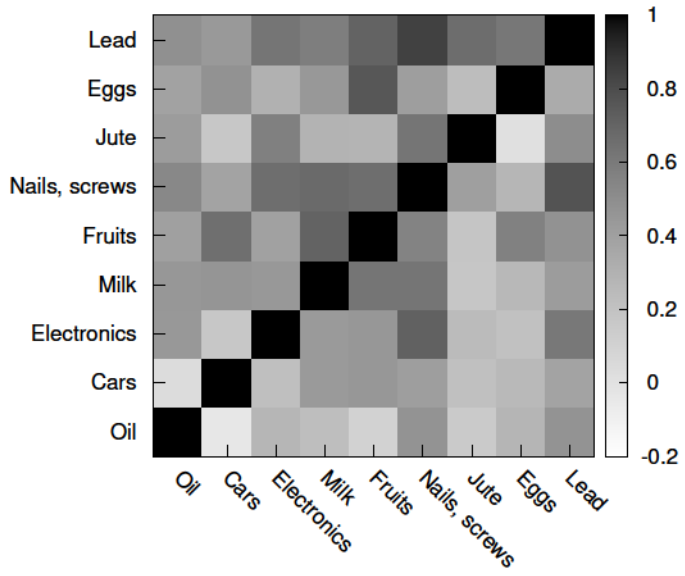


FIGURE 11. Comparisons between the correlators obtained form the real data (above the diagonal) and the synthetic ones (below the diagonal), obtained averaging 20 synthetic histories obtained with real initial conditions of 1962 and with the values for the parameters obtained in the calibration procedure. Commodities were chosen to be equally representative of all the set (3 with high, 3 with average and 3 with low ranking positions).

after that, for instance, a worldwide crisis occurs. An exponential fit of the form:

$$r_s(t) - r_s(\infty) = [r_s(0) - r_s(\infty)]\, e^{-t/\tau_S} \tag{66}$$

provides us with the following characteristic time:

$$\tau_s = 46.7 \pm 0.5 \text{ y}$$

A more in-depth analysis, shows that given the same initial conditions we find a characteristic time that is usually inversely proportional to G (keeping the values of the other parameters fixed).

7. Reproduction of the correlation matrix

Here we present another result that validates the choices we made in the construction of the correlation matrix $c_{pp'}$ defined in 13. Exploiting the simulations we performed, we evaluated in retrospect a synthetic analogous matrix $c_{pp'}^{simul}$. In Fig. 11 we illustrate a comparison between the matrix obtained from historical data and the one obtained from the simulations. The comparison has been carried out over a subset of 9 representative goods (3 among the top rankings, 3 in the middle ones and 3 in the lower positions).

As we can see the symmetry of the plot highlights how the simulated data are consistent with the model we built also from this perspective: $c_{pp'}^{simul}$ captures the same correlations that were present in the original data.

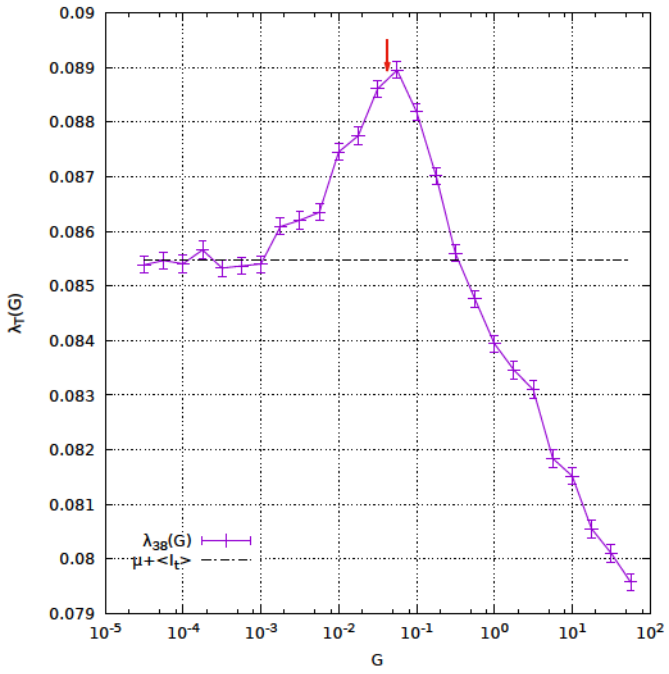


FIGURE 12. Average growth rate λ_{38} vs. G. The black dashed line indicates the growth coming from deterministic drifts $\bar{\mu} + \sum_{i=1}^{38} I_t/38$, while the arrow points to the calibrated value of $G = 0.042 \pm 0.001$ on the real data. 10000 simulations were performed for each point in the plot.

8. Optimal growth

One important feature that characterizes our model is that the growth we observe is not entirely given by the deterministic contribution coming from $\mu_t = \bar{\mu} + I_t$. The multiplicative nature of the terms composing the equation might allow positive stochastic fluctuations to spread locally more efficiently rather than unfavorable ones. Such feature is strictly connected with the explore-exploit trade-off dilemma presented in [**51**], in which one has to decide whether to exploit a favorable local fluctuation rather than redirecting its resources towards new and unexplored venues.

After defining the following average growth rate in the period 1962-2000:

$$\lambda_{38} = \frac{1}{38 \cdot 226} \sum_{p \in \text{goods}} \log \frac{Z_{p,38}}{Z_{p,0}} \tag{67}$$

it is straightforward to try to understand how varying the coupling constant G that regulates the intensity of the transfer of resources might affect such growth. Therefore we proceed simulating data with our model starting from the same initial condition of the year 1962 and keeping the values of $\bar{\mu}$, σ and τ fixed to the values obtained in the calibration procedure, but varying the value of G to see how this affects the resulting average growth rate.

The results we obtained are shown in Fig. 12: for very small values of G the growth is simply determined by deterministic contributions, in fact in such limit

Eq. 5 becomes

$$\frac{\partial Z_{p,t}}{\partial t} \xrightarrow[G\to 0]{} (\eta_{p,t} + \bar{\mu} + I_t)\, Z_{p,t} \tag{68}$$

which integrated between 0 and 38 yields

$$\log \frac{Z_{p,38}}{Z_{p,0}} = \int_0^{38} (\eta_{p,t} + \bar{\mu} + I_t)\, \mathrm{d}t \tag{69}$$

Averaging over the p's the stochtic component $\eta_{p,t}$ disappears and we are left with

$$\lambda_{38}|_{G=0} = \bar{\mu} + \frac{1}{38}\sum_{t=0}^{38} I_t \simeq 0.0855 \text{ y}^{-1} \tag{70}$$

which coincides with the deterministic drift comprehensive of the average inflation rate in time.

Increasing the value of G we observe instead a progressive increase of the growth rate until a maximum value is reached, after which the curve changes its trend reaching values that are even below the line given by the deterministic component. Such behavior is caused by the specificity of the initial conditions, which imply big amount of transfers of resources in the initial stages of the evolution which would not happen in the case of steady state initial conditions, as we'll see in the following section.

A peculiar fact that arises from such analysis is that the calibrated value $G \simeq 0.04 \text{ y}^{-1}$ on the real data is really close the the point of the maximum of the growth curve: this indicates that the global network somehow was able to self-regulate itself to the best possible conditions (from a 39 years growth perspective).

Keeping into account the average weight of the links of the transfer matrix $\sum_{i\neq j} J_{ij}/N$, the calibrated value of G corresponds to an average 2% transfer of resources for each product every year.

9. Growth for $T \gg \tau_s$

The analysis presented in the previous section studied the growth over the 38 years period following 1962. We can try and see what happens for longer time simulations in which $T \gg \tau_s$, using as inflation the average value $\sum_{t=0}^{38} I_t/38 \simeq 0.079 \text{ y}^{-1}$.

In Fig. 13 we show (in green) the curve $\lambda_{400}(G)$ starting from the real 1962 initial conditions. The first thing we notice is that the *bell*-shape became both wider and higher, allowing for a general better growth with respect to the 38 year case. We also notice that the G coordinate of the maximum is one order of magnitude bigger than previously, so the real calibrated value of G is further from the maximum in a long-time perspective.

Another fact that arises from this analysis with $T = 400$ y is that for large values of G the growth does not go below the line of the deterministic component, as in the 39 years growth. This effect is common in studies of networks starting from steady state initial conditions. In fact we expect that, if no relevant perturbations occur, the network will stabilize in the long time to a state close to the steady one.

Now we look for the steady state conditions of this network in order to simulate the growth that comes by starting from such condition. Of course we'll need to set the deterministic drift μ_t equal to 0 in the dynamics Eq. 5:

$$\frac{\partial \log Z_{p,t}}{\partial t} = \sum_{p'\neq p} \left(J_{pp'} \frac{Z_{p',t}}{Z_{p,t}} - J_{p'p} \right) + \eta_{p,t} \tag{71}$$

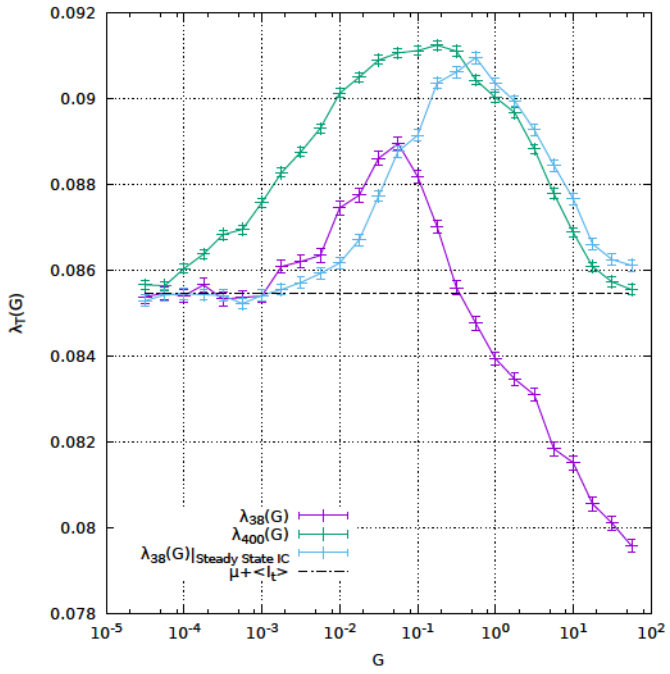


FIGURE 13. Comparison between the average growth rate λ_T vs. G for $T = 38$ y (purple) and $T = 400$ y (green) with 1962 real initial conditions and for $T = 38$ y with steady state initial conditions (light blue). The black dashed line represents the growth coming from the deterministic drift $\bar{\mu} + \sum_{i=1}^{38} I_t/38$.

Averaging on the products the stochastic term $\eta_{p,t}$ disappears, yielding

$$\left\langle \frac{\partial}{\partial t} \log Z_{p,t} \right\rangle = \left\langle \sum_{p' \neq p} \left(J_{pp'} \frac{Z_{p',t}}{Z_{p,t}} - J_{p'p} \right) \right\rangle \tag{72}$$

On the left hand side we have the derivative of the average growth (non-normalized) which is equal to zero in the steady state. Imposing this condition we find that the steady state is given by:

$$Z_{p,t} = \frac{\sum_{p \neq p'} J_{pp'} Z_{p',t}}{\sum_{p \neq p'} J_{p'p}} \tag{73}$$

We can find a set of $Z_{p,t}$'s that satisfy the above relation starting from the initial condition and imposing the iterative scheme

$$Z_{i,t} = \frac{\sum_{p \neq p'} J_{pp'} Z_{p',t-1}}{\sum_{p \neq p'} J_{p'p}} \tag{74}$$

and iterating until the fixed point is reached. Starting from such conditions we studied again the growth λ_{38} as a function of G, obtaining the light-blue curve shown in Fig. 13. Here we verified the idea that a long-time evolution approximates the effect of a steady-state regime: as in the case $T \gg \tau_s$ we find that the maximum's G coordinate is one order of magnitude greater with respect to the one calibrated over real data, and also that for large values of G the curve does not go below the line defined by the deterministic drift.

The reason why the two curves differ is probably given by the relation of inverse proportionality that connects G and τ_s: for small G's a time of $T = 400$ y is not

big enough for this kind of analysis. On the other hand, for large values of G (corresponding to smaller values of τ_s) we find that the the two curves are in good agreement.

10. Discussion

In this chapter we illustrated the construction of an autonomous dynamical model describing the evolution of the value of product exports within a global-economy complex network. Minimality has been an obliged feature in order to highlight cooperative endogenous mechanisms underlying economic growth. Specifically, in the equations we singled out a deterministic growth term including inflationary contributions, a stochastic one representing the alternance of favorable and unfavorable conditions to the development of each export class, and finally a network of value-transfers among products. This last ingredient has been successfully identified primarily through a ranking-based criterion. The stochastic component is characterized both by time-correlations associated to the duration of economic trends and by product cross-correlations reflecting, e.g., the existence of economy sectors in which several exports are simultaneously involved. In spite of its parsimonious character, the model provides realistic estimates of fluctuation properties, characteristic response times, and average growth rates. It also provides an evaluation of the average percentage of value transfer for the exports. Importantly, through the distinction between $\bar{\mu}$ and the empirical growth rate, it characterizes which part of the growth can be ascribed to transfer mechanisms determined by investments and structural interdependences.

Here we also could verify that the global growth complies with the typical conditions of an explore-exploit problem. The optimal solution for such a problem depends on both structural interdependences among products and strategic investment choices. Noteworthy, the calibrated network couplings realize close-to-optimal conditions for maximal average growth in the period covered by data. Thus, for the prevailing correlated noise conditions, the network transfer rates appear to be self-organized towards a close to optimal solution of the explore-exploit dilemma.

CHAPTER 3

Growth dynamics and complexity of national economies

1. Introduction

A fundamental problem in the study of economic growth is the quantitative assessment of the effect that the variety and quality of goods produced by a country has on its overall productivity [**63, 64**]. This assessment faces the extra difficulty that the productivity depends also on nontradable capabilities and on intangible assets. An advance in this field was recently made within the approach to economic complexity. This consisted in the proposal of a measure of diversity in the productions of a country which takes into account their degree of specialization, as deducible from a comparative analysis at global level [**7**]. An idea at the basis of this approach [**65, 8, 66, 67**] is that the productive basket of all countries, if properly analysed, should supply also most of the information encoded in assets like education, quality of life, technological sophistication, or institutions.

In the work published in [**2**] we managed to push this view further, considering data concerning yearly exports of all countries, retrieved from the global trade network, as a complete set of coordinates suitable for a closed description of their evolution in time. As we show below, supplying a dynamic model of all these data allows to make contact with aspects of the economic complexity of the nations, which seem to be hardly revealable by other approaches. At the same time, comparisons of the parameters quantifying these aspects in the dynamic description, can be made with a novel measure of economic complexity extracted directly and reliably from the data whose evolution is described by the model.

Inspired by the success of relatively more simple descriptions used in various fields, like population and evolutionary dynamics [**68**], portfolio strategies [**53**] or interface growth [**69**], we formulate stochastic differential equations for the evolution of the exports of the various countries. The possibility of synthetic simulations opens the way to counterfactual or path dependence analyses. For example, the equations generate through noise effects alternances of periods of favourable and unfavourable conditions for each export, and the occurrence of contextual transfers of resources between different exports, can lead to the result of enhancing or depressing average growth. This mechanism gives rise to explore-exploit tradeoff alternatives in portfolio optimization [**53**]. Since the model includes a coupling parameter which tunes simultaneously all the rates of resource transfers between different exports, we can estimate as a function of this coupling the potential average growth which could have occurred in a past period, and compare it with the historical one. This is illustrated in the Results Section below for the case of the USA, China and Russia. The counterfactual analysis relative to the period 1995-2015 shows in each case how pronounced are the maxima of average growth and how far they fall from the historically calibrated values of the transfer coupling. This provides an indication of the growth potential of the countries associated with variations of the average transfer rate.

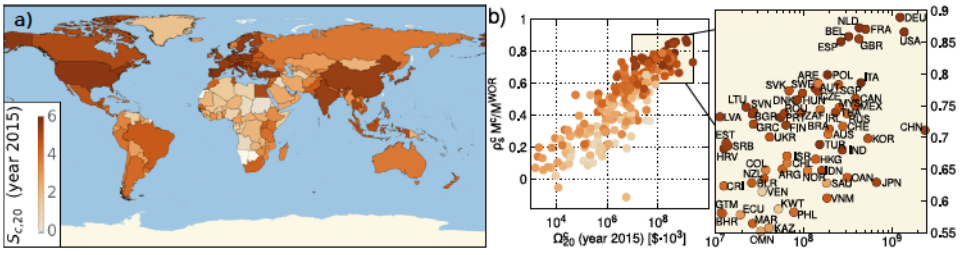


FIGURE 1. Panel (a): 223 countries colored with a palette related to the entropic measure S_c, computed for 2015. Few countries with no data are white. Panel (b): This scatter plot reveals that the total export Ω_{20}^c and the normalized Spearman's correlator $\rho_s^c M^c/M^{WOR}$ discussed in the text are strictly correlated. The ranking of most advanced countries is close to that valid for globally aggregated exports. The color scheme is the same as the one used in the Panel (a) and shows that S_c is clearly increasing with increasing Spearman's correlator.

The model presented in the previous chapter and published in [1] dealt with the simpler, but similar, problem of the time evolution of exports aggregated among all countries at global level. The similarity is due to the circumstance that, at both aggregate and single developed nation level, data show a progressive enforcement and stabilization of a peculiar way in which resources are distributed in different productions. This tendency towards a similar relative distribution of resources in the baskets should be regarded as a main effect of the reciprocal influences of different national economies in the global market. As shown in Fig. 1(b), especially for the most developed countries, the rankings associated with these distributions appear very close to that valid after global aggregation. These distributions in fact constitute a basic ingredient for the construction of our dynamics.

The precise distributions of resources into different exports are not taken into account by approaches which attribute to each country the role of exporter of a given product simply on the basis of sharp threshold criteria. Such criteria have been extensively adopted recently in the construction of economic complexity measures [**7, 65, 8, 66, 67**] The structure of our dynamical model of export growth suggests to look for complexity measures which fully take into account the precise relative weights of the different exports in the national baskets. In development economics measures of diversity of productions using these weights as arguments of Shannon entropy functions are already known [**70, 71**]. A key result we provide here is the construction of an iterative, safely and rapidly convergent scheme for a consistent simultaneous evaluation of the diversification of national exports and the specialization of products. The starting inputs in the iterative scheme are Shannon entropy indicators of bare diversity and specialization. An achievement in this context is anticipated in Fig. 1(a), where we report in a world map our estimated entropy measure of economic complexity for 223 countries referring to historical data of 2015. The analysis of the various parameters and indexes of our model is considerably helped by comparisons with this novel entropic measure and $\mathrm{GDP_{pc}}$ data.

2. Original databases

2.1. Export Data. The data used for quantifying the products exported by each country are extracted from the BACI database [**72**]. This database, redacted

TABLE 1. Yearly values (in percentage) of the World aggregate CPI.

year	$I(t)$	year	$I(t)$	year	$I(t)$	year	$I(t)$
1996	5.55	2001	3.59	2006	2.63	2011	2.86
1997	4.69	2002	2.71	2007	2.50	2012	2.24
1998	4.11	2003	2.42	2008	3.68	2013	1.61
1999	3.54	2004	2.35	2009	0.51	2014	1.73
2000	3.95	2005	2.60	2010	1.84	2015	0.59

by the CEPII Research Institute, is in turn built upon the COMTRADE database (freely accessible at the UN COMTRADE website [**47**]), and consists in a revision of the latter, in which the authors check and rectify the amounts declared by the exporting and importing countries to make them consistent with each other. All the prices are expressed in terms of current US dollars (thus, justifying the insertion of an inflation term in the dynamic model). The BACII database is a 4 dimensional panel that spans the following axes:

(1) Time: data are reported on an annual basis (starting from 1995) and are constantly updated with the latest data available. At the time we started performing our analysis the database covered a 21 years range, with the latest data being those of 2015.
(2) Product: The products are classified according to a 6 digit code (the Harmonized System 2007 [**73**]) which consists in roughly 5000 different categories of products.
(3) Exporting country: about 223 countries are present in the database. Data are organized at a country-to-country level, so that one can know how much of a single product has been exported from any country to any another.
(4) Same for importing countries.

2.2. Gross Domestic Product. The values of the Gross Domestic Product Per Capita ($\mathrm{GDP_{pc}}$), have been extracted from a database [**74**] of the United Nations Statistics Division (UNSD). This database spans the period 1970-2015 and quantities are expressed in current US dollars (same as for the BACII database).

2.3. Consumer Price Index. Finally the Consumer Price Index (CPI) values were given by the Organisation for Economic Co-operation and Development (OECD) [**59**]. In table 1 we report the values used in our analysis, which correspond to a weighted aggregate at a global level of all the CPIs of every country present in the OECD databases.

3. Data aggregation

The Harmonized System classification allows us to organize the data in less specific categories by simply aggregating products sharing the most significant digits. Our analysis was performed over 1238 "macro" categories built with 4 digit data for exports of each country. Another aggregation has been carried out over the importing countries axis, in order to obtain the $Z^c_{p,n}$ representing the total amount of product p exported by country c in the year n. The entropic measures construction (see section 10) was carried out on the resulting dataset. In order to apply our dynamic model we had to perform a further selection in the $Z^c_{p,n}$ panel. With data coarse-grained to a 4 digit level, it may still happen that some entries become zero. In principle this should indicate that a certain country stops suddenly to export a product. However, such occurrence is more likely related to

an intrinsic problem of missing information in the original database: the data we work on are aggregated, so, if a product disappears, it means mean that the country has ceased to produce a whole macro category of products in just one year, which is unlikely to happen. Moreover these zeroes appear mainly when dealing with underdeveloped countries. So, this suggests that the problem lies in the redaction of the database itself. Having a $Z^c_{p,n} = 0$ is an issue because it produces a singularity when inserted in Eq. 122. So, when this happens, we need to remove the entire time-series associated with that product. We decided to proceed with the calibration only for those countries who kept more than 80% of their exports after these removals. So, the dynamic model was applied to 131 countries in total.

4. Distributions of exports from individual countries

For each country and each product we consider the yearly exports realized from 1995 to 2015 (see SI). Limiting here examination to 131 countries whose export data do not show too many interruptions in the whole period, we call $Z^c_{p,n}$ the total value (in thousands of US-dollars) of the product category p ($p = 1, 2, \ldots, M^c$) exported in the year n ($n = 0, 1, \ldots, T$ with $T = 20$) by country c ($c = 1, 2, \ldots, N = 131$). The number of exported products, M^c, varies from country to country, but for the most developed economies $M^c \lesssim 1238 \equiv M^{WOR}$. An interesting result that emerged in ref. [1] is that, the exports aggregated at global level, $Z^{WOR}_{p,n} = \sum_c Z^c_{p,n}$, are sorted in value according to a ranking which is maintained and slowly stabilizes, up to slight fluctuations, in the years. Its existence and the fact that similar stable rankings are approached by the exports of each individual country, especially the most developed, is a key feature of the organization of economies. For each country, we decide to assign a reference ranking of product p based on the fraction of the value of product p over the whole country export, $\Omega^c_n \equiv \sum_{p=1}^{M^c} Z^c_{p,n}$, averaged over the last 5 years:

$$z^c_p \equiv \frac{1}{5} \sum_{n=16}^{20} \frac{Z^c_{p,n}}{\Omega^c_n} \tag{75}$$

In order to compare the ranking of the z^c_p of country c to that of the exports aggregated at global level in year n, $z^{WOR}_{p,n}$, we evaluate the Spearman's rank correlation coefficient [**61**] ρ^c_s between the former and the latter set deprived of the goods that are not exported by country c. Since an extensive mismatch between the two sets could yield a misleadingly high coefficient, we further multiply the Spearman's correlation by M^c/M^{WOR}. Fig. 1(b) reports this quantity coupled with the total export Ω^c_n for $n = 20$ (year 2015). It shows that, indeed, developed countries share a common ranking structure, close to that of globally aggregated exports ($\rho^c_s \sim 1$). Moreover, with lower total export the correlator decreases.

In Fig. 2, we show the time series of the products exported by the USA (a), China (b) and Russia (c). The wavelengths of the colors used for each product p are proportional to the corresponding z^c_p: the rainbow image perceived in the figures confirms the existence and stability of the ranking of product values. Exceptions are of course present, corresponding to goods changing their rank in the period. Furthermore, the rainbow effect in the graph is slightly more pronounced for the USA compared to China. This is a sign of the fact that China's growing economy is still reorganizing its internal structure. Russia instead shows a wider spread of the product export values, indicating that in its case there is a certain number of products whose impact on the country's economy is low and a relatively smaller number of products with a high impact, like coal, petroleum, natural gas and minerals. Fig. 2 also shows that China's exports increased much faster than those

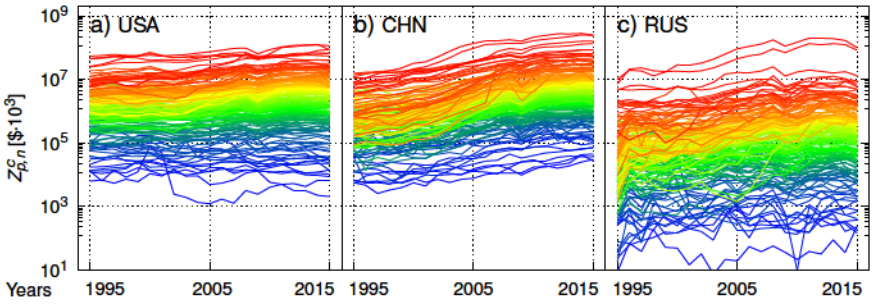


FIGURE 2. Interpolations of yearly values of products exported by the USA (a), China (b) and Russia (c) from 1995 to 2015. The wavelength of the color for each group of products is proportional to the corresponding z^c_p. The relative stability of products rankings in value over time causes marked rainbow effects in all images. For better visualization, the data have been further coarse-grained to roughly 200 categories of products, corresponding to 3 digits of the HS codifications.

of the USA. In fact, defining the average growth rate over time T for a country c as

$$\lambda^c_T \equiv \frac{1}{M^c \cdot T} \sum_{p=1}^{M^c} \log \left[\frac{Z^c_{p,T}}{Z^c_{p,0}} \right] \tag{76}$$

one gets $\lambda^{USA}_T = 2.1\%$ and $\lambda^{CHN}_T = 11.2\%$. Russia experienced average export growth comparable to China $\lambda^{RUS}_T = 10.5\%$.

5. Model of Export Dynamics for individual countries

Since yearly export records result from variations in much shorter periods, we set up equations in continuous time t measured in year units ($t = 0$ corresponds to the end of 1995): the value of product p exported by country c in the year preceding time t is indicated with $Z^c_p(t)$, with $0 \leq t \leq T$, (T corresponds to 2015) Thus, $Z^c_{p,n}$ gives a discrete representation of this function of time in 21 points.

We write a stochastic system of equations for the $Z^c_p(t)$'s in the following form.

$$\begin{aligned} \partial_t Z^c_p(t) &= \sum_{p' \neq p} \left[J^c_{pp'} Z^c_{p'}(t) - J^c_{p'p} Z^c_p(t) \right] + \left[\eta^c_p(t) + \mu^c(t) \right] Z^c_p(t) \\ &= \sum_{p'} A^c_{pp'} Z^c_{p'}(t) + \left[\eta^c_p(t) + \mu^c(t) \right] Z^c_p(t) \end{aligned} \tag{77}$$

where η^c is a colored zero-mean Gaussian noise. The first term in $Z^c_p(t)$ on the rhs of Eq. 122 would enter, in somewhat simplified form, in a geometric Brownian motion description of asset dynamics familiar from the standard model of finance [**75, 53**].

In the following sections we provide an exhaustive description of the dynamics described by this system of equations in order to give a better understanding of the meaning of the parameters and of the model ingredients.

5.1. Correlation matrix. In the basket of a country, this matrix plays the important role of quantifying the correlations between the increments of each pair of exported products. The first step in building such matrix is to evaluate the

yearly logarithmic returns

$$R^c_{p,n} = \log\left[\frac{Z^c_{p,n}}{Z^c_{p,n-1}}\right] \tag{78}$$

which we standardize referring to averages over all the years (indicated by $\langle A_n\rangle = \Sigma_n A_n/T$ for a quantity A_n).

$$r^c_{p,n} = \frac{R^c_{p,n} - \langle R^c_{p,n}\rangle}{\sqrt{\langle \left(R^c_{p,n}\right)^2\rangle - \langle R^c_{p,n}\rangle^2}} \tag{79}$$

The correlation between the products p and p' is thus defined as

$$c^c_{pp'} \equiv \langle r^c_{p,n} r^c_{p',n}\rangle = \frac{1}{T}\sum_{n=1}^{T} r^c_{p,n} r^c_{p',n} \tag{80}$$

where $T = 20$ represents the last available year for the returns (2015 in this case). As we will see in the following sections, the correlation matrix will play an important role in the construction of both the matrix ruling the transfers between different productions and the colored noise.

5.2. Matrix $J^c_{pp'}$. The matrix $J^c_{pp'}$ is constructed exploiting the correlation matrix as a sort of "inverse distance" (see Tinbergen's gravity law [**76**]): the shorter this "distance", the larger the rate at which the transfer of resources between two products occurs. In the structure of the transfer term of the equations, the gravity law is present also because the transfer term is proportional to both $Z^c_{p'}$ and z^c_p. Indeed, we write:

$$J^c_{pp'} = G^c z^c_p |c^c_{pp'}| \tag{81}$$

where G^c is a coupling multiplicative constant that regulates the magnitude of these transfer processes, and z^c_p is the fraction of total export for product p averaged the last five years:

$$z^c_p = \frac{1}{5}\sum_{n=16}^{20} \frac{Z^c_{p,n}}{\sum_{p'} Z^c_{p',n}} \tag{82}$$

This is the quantity used to discuss the ranking in value of exports in the main text. The transition matrix $\mathbf{A}^c$, with elements $A^c_{pp'} = J^c_{pp'}$ for $p' \neq p$ and $A^c_{pp} \equiv -\sum_{p'\neq p} J^c_{p'p}$, establishes and maintains the observed ranking of the various $Z^c_p(t)$'s once we choose $J^c_{pp'} \propto z^c_p$. Indeed, with such a choice, z^c_p is in the kernel of $\mathbf{A}^c$ and, in the absence of noise terms, the solutions of Eq. 122 tend to a long-time attractor in which the Z^c_p's are in the kernel of $\mathbf{A}^c$ itself [**77**] and are thus proportional to the z^c_p's.

5.3. Noise. The noise that defines the stochastic dynamics of Eq. 122 needs to take into account that fluctuations in the growth conditions are correlated both in time and between different products. The former correlations require to define η^c as a colored noise with characteristic time τ^c and variance σ^c, while the latter will imply that we can't use a Kronecker delta for the correlation among noises for different products. Indeed, fluctuations affecting a certain product will be felt by other products that are "spatially" close. For these reasons, we define η^c as a colored Gaussian noise with mean and correlation

$$\langle \eta^c_p(t))\rangle = 0 \tag{83}$$

$$\langle \eta^c_p(t_1)\, \eta^c_{p'}(t_2)\rangle = c^c_{pp'}\, \frac{(\sigma^c)^2}{\tau^c}\, \mathrm{e}^{-|t_1-t_2|/\tau^c} \tag{84}$$

τ^c represents a typical opportunity/crisis period duration [51], while σ^c regulates the magnitude of the fluctuations. Here the averages indicated by $\langle . \rangle$ have the usual meaning in stochastic differential calculus. $c^c_{pp'}$ appears in the noise correlator because the noise is directly responsible for the return correlations recorded historically.

5.4. Drift. The term $\mu^c(t)$ represents a deterministic drift, accounting for the average percentual input of resources in the export production in a given country in the absence of mutual influences. The time dependence comes from the inclusion of the inflationary effects to which the exports are exposed every year (as explained in section 2) Thus, we write

$$\mu^c(t) = \bar{\mu}^{\,c} + I(t) \tag{85}$$

in order to separate the constant, average contribution to the drift $\bar{\mu}^c$ from the inflationary one $I(t)$, which can be read as an yearly step-wise function whose values are taken from the OECD [59] and reported in table 1.

6. Calibration procedure

The values of the free parameters are determined by best fits of the historical records. We notice that all the parameters to be fitted in the model, and rates $J^c_{pp'}$, are assumed to be time independent, in a sort of mean field spirit. This is justified by the considerable complexity of the problem. Introducing a time-dependence for the $J^c_{pp'}$ in particular, would in principle allow to explore optimization strategies of the average growth more realistic than those one can test upon tuning only the constant value of G^c for the whole period, as discussed in the next Section. Indication that such dependence on time could be a realistic feature of a more sophisticated model is provided by the often large variances displayed by the yearly records whose average yields our matrix $c^c_{pp'}$. In spite of this, we take these averages as time-independent matrix elements, in a spirit which is not far from that of a recent study of the multi-layered network structure underlying financial and macroeconomic dynamics [78, 79]. When dealing with the data of certain countries we find that the calibration of these parameters reveals the presence of a noise that cannot be fully reproduced by the dynamics of Eq. 122 (see SI for more details). This occurs especially in the case of less developed countries since when a country is not driving in the global economy, the fluctuations of its productions can be strongly affected by the dynamics of other countries, an effect not explicitly included in our model. For this reason, in order to properly calibrate the fluctuation parameters σ^c and τ^c, we need to clean the data by removing these external-noise effects. This suggests us to introduce an additional, fifth parameter, σ^c_0, which in turn becomes an interesting indicator of the role of a country in the World economy scene, by quantifying its sensibility to other countries' influence.

From this point on, we will drop in this section the $\cdot^c$ to avoid redundancy in the notations. The calibration procedure we follow is similar to that presented in ref. [1], with the main difference lying in the presence of the new background noise σ_0. Dividing Eq. 122 by $Z_p(t)$ and integrating in the time interval $[n_1, n_2]$ (n_1 and n_2 integers), we obtain

$$f_p(n_1, n_2) = G g_p(n_1, n_2) + \bar{\mu} + \frac{1}{n_2 - n_1} \int_{n_1}^{n_2} \eta_p(t)\, \mathrm{d}t \tag{86}$$

where we introduced the functions

$$f_p(n_1, n_2) = \frac{1}{n_2 - n_1} \left(\log \frac{Z_{p,n_1}}{Z_{p,n_2}} + \sum_{n=n_1}^{n_2} I_n \right) \tag{87}$$

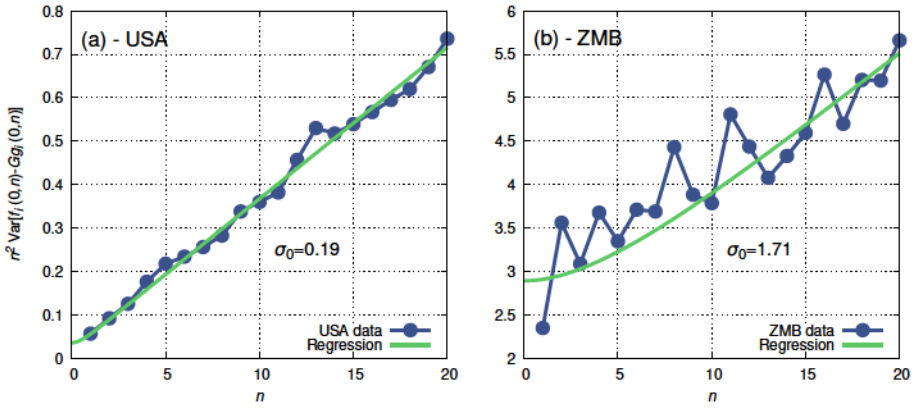


FIGURE 3. An offset noise term characterize the networks of underdeveloped countries. Here we report the quantity $n^2\mathrm{Var}\left[f_p(0,n) - Gg_p(0,n)\right]$ evaluated from the data of a driving country in the economical scene (USA, panel (a)) and for an underdeveloped country (Zambia, panel (b)). It is evident that an offset term arises for the latter country, justifying the introduction of the additional parameter σ_0.

$$g_p(n_1, n_2) = \frac{1}{n_2 - n_1} \sum_{\substack{n=n_1 \\ p' \neq p}}^{n_2-1} \left[\frac{z_p \left| c_{pp'} \right|}{2} \left(\frac{Z_{p',n}}{Z_{p,n}} + \frac{Z_{p',n+1}}{Z_{p,n+1}} \right) - z_j \left| c_{pp'} \right| \right] \tag{88}$$

which are estimated approximately by sums in place of integrals because of the discrete nature of the data at our disposal. We immediately see that Eq. 86 establishes a linear relation between such functions, with G representing the slope. Therefore we perform a robust linear regression of the scatter plot f vs. g to estimate the value of G. The intercept obtained with the regression is a first estimate of $\bar{\mu}$ (the stochastic term here averages to 0), but a more accurate derivation can be performed after calibrating the parameters σ and τ. In order to do that, we first need to rearrange Eq. 86 by moving to the left side the Gg_p term. We then evaluate the variance of the resulting equation. By setting $n_2 = n$ and $n_1 = 0$ we obtain:

$$v(n) := n^2\mathrm{Var}\left[f_i(0,n) - Gg_i(0,n)\right] = 2\sigma^2 \left(n + \tau \left(e^{-n/\tau} - 1 \right) \right) \tag{89}$$

When dealing with developed countries this relation is usually quite solid, but for the rest of the cases the empirical plots show that there is a missing offset in the equation (see Fig. 3). Such effect is imputable to the fact that the countries are not a closed system and are subject to what happens outside of their borders: the less a country is driving in the global economy, the more its trades will be affected by fluctuations of the countries it relates to. In order to correctly calibrate the parameters σ and τ of a country, we thus need to clean the data by introducing an offset noise term σ_0, which in turn will be an indicator of the extra fluctuations the country's exports are exposed to. Eq. 89 becomes then:

$$v(n) = 2\sigma^2 \left(n + \tau \left(e^{-n/\tau} - 1 \right) \right) + \sigma_0^2 \tag{90}$$

The exponential term decays quite fast (ex post analysis shows that $\langle \tau \rangle_c = 2.8 \pm 0.2$ y). So, by neglecting the first years terms, we are able to perform a linear regression: the slope $m = 2\sigma^2$ can be directly inverted to obtain the value of σ,

while the intercept equation $q = -2\sigma^2\tau + \sigma_0^2$ needs to be coupled with another condition in order to determine the remaining two parameters. For this reason we integrate Eq. 91 in the full time range $[0, T]$ obtaining:

$$V = \frac{1}{2}\sum_{n=0}^{T-1}[v(n) + v(n+1)] = 2\sigma^2\left(\frac{T^2}{2} - \tau T + \tau^2\right) + T\sigma_0^2 \tag{91}$$

where we disregarded the exponential term $e^{-T/\tau}$ obtained after the integration. Inverting the system of equations finally yields

$$\tau = \sqrt{\frac{V - Tq}{2\sigma^2} - \frac{T^2}{2}}\ ,\ \sigma_0^2 = q + 2\sigma^2\tau \tag{92}$$

We calibrate the last remaining parameter $\bar{\mu}$ through repeated synthetic simulations (see below) of the system of equations deprived of the deterministic drift $\bar{\mu}$ itself. In this case the average growth reads $\lambda_T^* = h_T(G) + \sum_{t=0}^{T} I_t$. Thus, one can find the value of $\bar{\mu}$ that reproduces correctly the growth from the historical data by difference:

$$\bar{\mu} = \lambda_T - \lambda_T^* \tag{93}$$

7. Integration

The stochastic differential Eq. 122 involves a colored Gaussian noise that evolves according to

$$\eta_p(t) = \rho\eta_p(t - \mathrm{d}t) + \sqrt{1-\rho^2}\frac{\sigma}{\sqrt{\tau}}\xi_p(t) \tag{94}$$

where $\rho \equiv e^{-\,\mathrm{d}t/\tau}$ and $\xi_p(t)$ is a zero-mean Gaussian noise with correlation $\langle \xi_p(t)\xi_{p'}(t')\rangle = c_{pp'}\delta(t - t')$. In order to obtain a noise with this properties we perform the $\mathbf{C} = \mathbf{L}\mathbf{D}\mathbf{L}^T$ Cholesky decomposition (see for instance [80]) of the matrix $\mathbf{C} \equiv c_{pp'}$, which can be applied to any symmetric matrix and does not require the evaluation of any square roots in the diagonal terms (which may become a problem when dealing with large matrices whit relatively small entries). Thus, we are able to apply the decomposition to a vector of independent Gaussian noises $\vec{\tilde{\xi}}$ to generate $\vec{\xi} = \mathbf{L}\mathbf{D}^{1/2}\vec{\tilde{\xi}}$. By substituting the correlated noise in Eq. 122 and discretizing we obtain

$$\Delta Z_{p,t} = a_p(t, \vec{Z}_t)\,\mathrm{d}t + b_p(t, \vec{Z}_t)\Delta W_{p,t} \tag{95}$$

where we introduced the Wiener processes $\Delta W_{p,t} = \sqrt{\Delta t}\xi_{p,t}$ and the two coefficients

$$a_p(t, \vec{Z}_t) = \left[\sum_{p' \neq p}\left(J_{pp'}\frac{Z_{p',t}}{Z_{p,t}} - J_{p'p}\right) + \rho\eta_{p,t-\Delta t} + \bar{\mu} + I_t\right] Z_{p,t} \tag{96}$$

$$b_p(t, \vec{Z}_t) = \sqrt{1-\rho^2}\frac{\sigma}{\sqrt{\tau}}\sqrt{\Delta t}Z_{p,t} \tag{97}$$

With the assumption $\Delta t \ll \tau$ different integration prescriptions yield the same results: we chose a second order Runge-Kutta scheme for the Itô prescription, suited for systems of equations and characterized by both strong and weak convergence of order 1 [81]. If we partition the time interval $[0, T]$ in L sub-intervals each of length Δt, the integration approximation $\{\vec{w}_l\}_{l=1\dots L}$ estimated over such mesh is

given by

$$\begin{cases} \vec{w}_0 = \vec{Z}_0 \\ \vec{w}_{l+1} = \vec{w}_l + \frac{\vec{K}_{1,l}+\vec{K}_{2,l}}{2} \\ \vec{K}_{1,l} = \vec{a}(t_l, \vec{w}_l)\Delta t + \mathbf{B_1}\vec{b}(t_l, \vec{w}_l) \\ \vec{K}_{2,l} = \vec{a}(t_{l+1}, \vec{w}_l + \vec{K}_{1,l})\Delta t + \mathbf{B_2}\vec{b}(t_{l+1}, \vec{w}_l + \vec{K}_{1,l}) \end{cases} \tag{98}$$

Here $\mathbf{B_{1,2}}$ are two diagonal matrices containing the Wiener processes associated with the noise $\Delta W_{p,n} \equiv \sqrt{\Delta T}(\xi_{p,n+1} - \xi_{p,n})$, which have entries

$$\begin{aligned} (\mathbf{B_1})_{pp'} &= \delta_{pp'}\left(\Delta W_{p,n} - S_{p,n}\sqrt{\Delta t}\right) \\ (\mathbf{B_2})_{pp'} &= \delta_{pp'}\left(\Delta W_{p,n} + S_{p,n}\sqrt{\Delta t}\right) \end{aligned} \tag{99}$$

where $S_{p,n}$ are random variables that can assume values ± 1 with equal probability $p = 1/2$. In all the numerical integrations performed we set $\Delta t = 0.01$, satisfying the condition $\Delta t \ll \tau$.

8. Growth

Concerning the average growth introduced in Eq. 121: one would like to determine the separate contributions of the terms in Eq. 122. To this purpose one can divide Eq. 122 by $Z_p^c(t)$, integrate it in the interval $[0, T]$ and take the average over the set of products. This allows to obtain:

$$\lambda_T^c = h_T^c(G^c) + \bar{\mu}^c + \frac{1}{T}\int_0^T I(t)\,\mathrm{d}t \tag{100}$$

where h_T^c results from integration of the transfer terms. This term depends strongly on G^c and can be estimated by discrete summations. Thus, as anticipated, part of the growth is directly associated with the cooperative transfer terms in the equations. Of course, the effect of transfers on growth depends also on the fluctuations caused by the colored noise. The combination of the two factors may realize conditions in which the positive trends of some productions are optimally exploited to increase the growth. So, besides determining G^c and $h_T^c(G_c)$ on the basis of the historical data, it makes sense to simulate histories in which, e.g., G_c is varied, while keeping historical initial conditions and other parameters fixed at the calibrated values. Thus, one can deduce the effect that a variation of G^c could have had on the average growth. This is interesting and not devoid of prescriptive value, since G_c can be in principle partly controlled, e.g., by regulations and investment policies.

In Fig.4, we report such plots for λ_{20}^c of the USA (a), China (b) and Russia (c) indicating also the values of G^c, (abscissa pointed by the red arrow in panels a,b,c), of λ_{20}^c, (y-coordinate of red arrow tip), and of $\bar{\mu}^c$ plus average inflation rate (horizontal black dashed line) as determined by calibration on historical data. The agreement of the arrow tips with simulation data points confirm the consistency of the calibrated model dynamics. The distance between the level of the black dashed line and the data points give a measure of the contribution to the growth to be ascribed to the transfer terms in the equations. This contribution appears positive in the historical cases, but becomes always negative for sufficiently high G^c values. Remarkably, for the USA the contribution of the transfer term is very small, and not susceptible of sensible increments upon an increase of G^c. The situation is very different for China and Russia, where we recognize a much larger contribution of the transfers and, most important, a substantial increment of the growth with larger G^c's. Fig. 4(d) reports the plots for the three countries after subtraction of the average deterministic contribution.

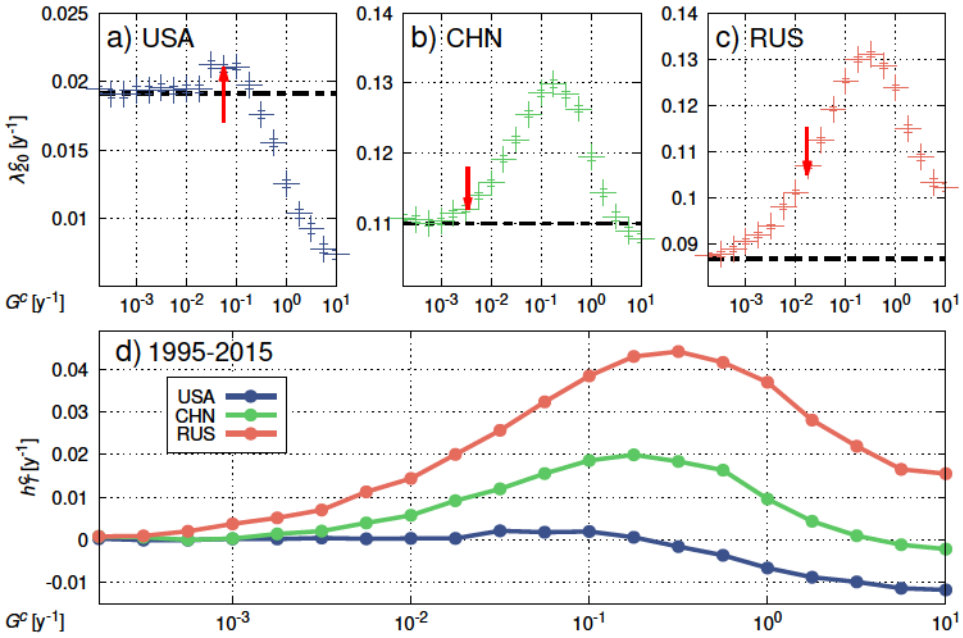


FIGURE 4. Panels (a), (b) and (c) report the average growth $\lambda^c_{20}(G)$ of the USA, China and Russia, respectively. They are obtained with simulations for different values of the transfer parameters G^c. Red arrows point to the calibrated values of G^c and the growth estimated from historical data. The black dashed lines represent the deterministic drift contributions to the average growth. $\bar{\mu}^c + \frac{1}{T}\int_0^T I(t)\,dt$. Panel (d) shows the previous curves deprived of the drift, in order to confront the growth contributes h^c_{20} of cooperative nature.

9. Entropic measures of diversity and specialization

In order to further clarify the meaning of the model parameters we define a quantity with the same scope as that of the Economic Complexity Index (ECI) [**65**] or Fitness [**8, 82**]. To be directly comparable with intensive indicators, like GDP per capita, it should be a function of the relative weights of the exports in the distribution of each basket, expressed as $s_{c,p} \equiv Z^c_p/\Omega^c$, for product p and country c (for simplicity we drop the year index n in this and the following section). The Shannon's entropy function already in use in development economics as a variety or diversity indicator is a natural candidate [**70, 71**]. This indicator increases with both the number and the evenness of products shares. It has the advantage of being independent of the detailed structure and number of intermediate stages leading to the final repartition of resources [**9**], and thus is quite robust. So, as a first step we define a country's production diversity as $S^{(0)}_c = -\sum_p s_{c,p} \log s_{c,p}$. Within this analysis we can analyze all the countries in the database regardless of time interruptions of the export data, so $N = 223$. An entropic indicator taking into account the specialization of the exports can also be defined as $Q^{(0)}_p = \log N + \sum_c q_{c,p} \log q_{c,p}$, where $q_{c,p} \equiv Z^c_p/Z^{WOR}_p$ are the shares of country c evaluated with respect to the overall global export of product p. Here we exploit the fact that the maximum Shannon entropy of a product is $\log(N)$. Thus, $Q^{(0)}_p$ increases with the entropy of the product being more distant from this maximum. As such, it weigths

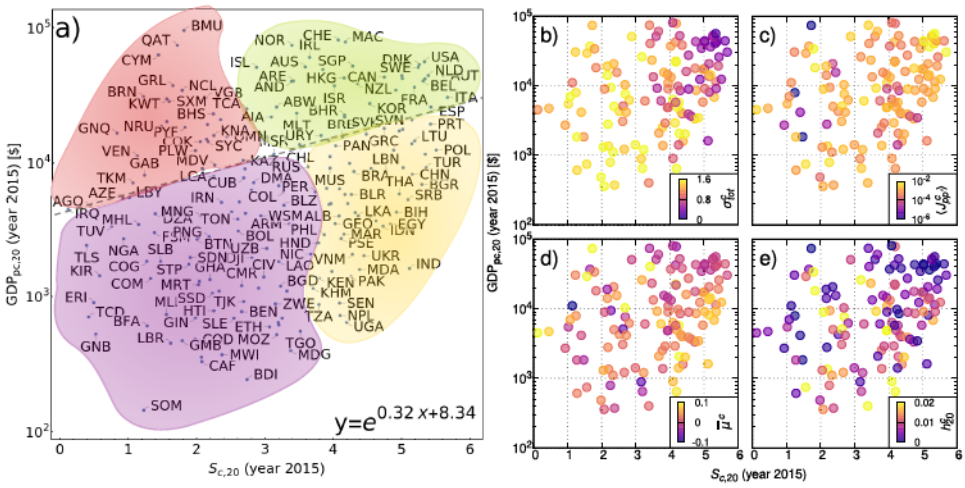


FIGURE 5. Panel (a): S_c and $\mathrm{GDP_{pc}}$ for 2015 allow to distinguish among developed (green), emerging (yellow), underdeveloped (purple) and economically risky (red) countries. In panel (b),(c),(d) and (e) the colors used are representative, respectively, of: the overall (inner and outer) fluctuations of a country σ^c_{tot}, the average transfer $\langle J^c_{pp'}\rangle$, the drift $\bar{\mu}^c$ and the cooperative growth term h^c_{20}. The parameters, together with the entropic complexity, help us in assessing the correct level of development of a country.

the degree of specialization of product p. So far $S_c^{(0)}$ depends only on the shares in value of the products produced by the country, and not on their specialization. Analogously, $Q_p^{(0)}$ does not depend on how developed are the various countries exporting product p. We can introduce such dependences by reweighting the shares entering each of the two quantities through factors depending on the other. So, for example in the case of the product diversity the idea is to start weighting each product p by Z_p^c multiplied by a quality factor $Q_p^{(0)}$, and to modify consistently the normalization $\tilde{\Omega}^c$. In this way each product enters in the entropy function with the importance deriving from its specialization at global level. For consistency such reweighting should be iterated until a fixed point is reached. The iteration step from stage $k-1$ to k reads:

$$\text{(101)} \qquad \begin{cases} S_c^{(k)} = -\sum_p \dfrac{Z_p^c Q_p^{(k-1)}}{\sum_{p'} Z_{p'}^c Q_{p'}^{(k-1)}} \log\left[\dfrac{Z_p^c Q_p^{(k-1)}}{\sum_{p'} Z_{p'}^c Q_{p'}^{(k-1)}}\right] \\ Q_p^{(k)} = \log(N) + \sum_c \dfrac{Z_p^c S_c^{(k-1)}}{\sum_{c'} Z_p^{c'} S_{c'}^{(k-1)}} \log\left[\dfrac{Z_p^c S_c^{(k-1)}}{\sum_{c'} Z_p^{c'} S_{c'}^{(k-1)}}\right] \end{cases}$$

The algorithm has a unique fixed point, S_c and Q_p to which both $S_c^{(k)}$ and $Q_p^{(k)}$ converge rapidly, as we tested numerically. This fixed point yields our consistent entropic measures of productions diversity and products specialization. In Fig. 1b the colors of points representing the various countries give a measure of how large S_c is. We find high values for the most developed countries at the upper right corner of the box. In general, moving vertically in the box one finds a gradient of S_c, which should be expected because the correlation with the ranking of globally aggregated exports increases. Significant is also the gradient for horizontal moves. However, some countries to the right, with a larger total export, have in this case lower S_c

than those to the left, because their exports are less balanced, and dominated, e.g., by petroleum (like Saudi Arabia, Venezuela, Kuwait etc...). A comparison between our entropic measure and the Fitness F_c defined in ref. [8] is reported 7, where we show that a relation $F_c = e^{\text{const}\cdot S_c}$ is very plausible (evaluating the Spearman correlation index between F_c and S_c yields a value $\rho_s \simeq 95\%$).

Our definitions of S_c and Q_p differ from those of the analogous quantities constructed within the economic complexity approach in a basic aspect: while in these approaches the quantities are defined directly as weighted averages, in our case weights are used to renormalize arguments to insert in indicators related to the Shannon entropy function. The use of this function introduces an essential ingredient of non-linearity[8] in the iterations and guarantees at the same time global stability and smoothness of S_c and Q_p. Thus, it provides also a way out of some mathematical pathologies which occasionally affect the previous algorithms[83]. S_c can also be straightforwardly evaluated at different levels of fine graining of the export data (number of digits used for coding the products), allowing consistent quantification of inter- and intra-sectoral contributions to diversification[71]. Finally, another advantage of our entropic measures is the fact that they make use of the full information contained in the Z_p^c's values, without resorting to binarizations through threshold criteria based on Revealed Comparative Advantage RCA [10, 11].

10. Analysis of the Entropic scheme

10.1. Construction of the algorithm. To simplify notations here the subscript n is omitted, thus, our considerations will refer to data for a generic year. The starting point that led to the construction of the algorithm for the entropies presented in the previous section, is the idea to use Shannon's entropy [9] as measure of the diversity in the basket of the exported products of a country [71]. Here the role of probabilities is played by the shares $s_{c,p} \equiv Z_p^c / \sum_{p'} Z_{p'}^c$ of the total export by the country in a certain year. Therefore a first, bare measure of diversity will be:

$$S_c^{(0)} = -\sum_{p=1}^{M} \frac{Z_p^c}{\sum_{p'} Z_{p'}^c} \log \left[\frac{Z_p^c}{\sum_{p'} Z_{p'}^c} \right] \tag{102}$$

The more a basket is equally partitioned among the products, the higher will be the value of this entropy. The ideal (unrealistic) case is therefore the one in which every product has a share $1/M$, which gives the upper bound limit value $\log M$ to this entropy. On the other hand, when all the exports are concentrated in a restricted niche of products (like in oil-exporting countries) the entropy will be small, with lower bound value 0 corresponding to the case in which only one product is exported. We can define an analogous quantity for the products, in order to understand which products are more common, and therefore easy to produce, and which are more rare. The shares are evaluated in this case over the total global export of each product, as $q_{c,p} \equiv Z_p^c / \sum_{c'} Z_p^{c'}$, Hence we define:

$$Q_p^{(0)} = \log(N) + \sum_{c=1}^{N} \frac{Z_p^c}{\sum_{c'} Z_p^{c'}} \log \left[\frac{Z_p^c}{\sum_{c'} Z_p^{c'}} \right] \tag{103}$$

so that in this way Q will be higher for the most specialized products and lower for the most common ones. The next step in creating the algorithm is to use S and Q respectively as a weight in valuating more refined entropies: the new weighted shares will therefore be:

$$s'_{c,p} = \frac{Z_p^c Q_p^{(0)}}{\sum_{p'} Z_{p'}^c Q_{p'}^{(0)}} \ , \ q'_{c,p} = \frac{Z_p^c S_c^{(0)}}{\sum_{c'} Z_p^{c'} S_{c'}^{(0)}} \tag{104}$$

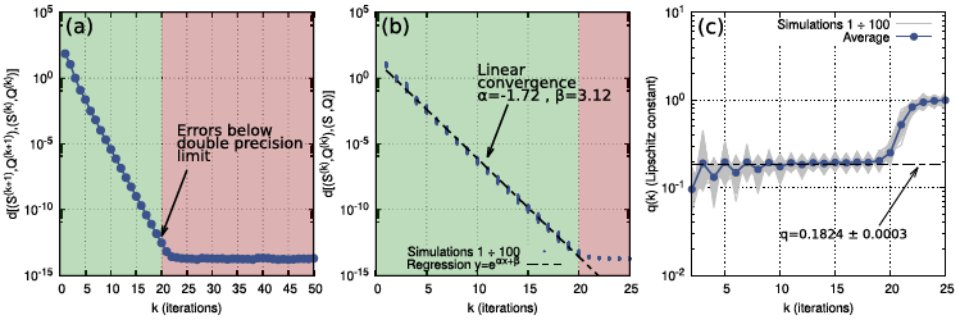


FIGURE 6. Panel (a): Euclidean distance d evaluated between consecutive steps of the iterative algorithm of the diversities $S_c^{(k)}$ and specializations $Q_p^{(k)}$. The distance decreases exponentially fast until the value $d \sim 10^{-28}$ (around iteration number 20) where the numerical precision of the computer reaches its limit. Panel (b): Distance of iterations from the fixed point of convervence of 100 simulations starting from completely random initial conditions (blue dots). The black dashed line represent the exponential regression $y = e^{\alpha x+\beta}$ in the region where the double precision limit is not exceeded. Panel (c): Valuation of the Lipschitz constant as the ratios of distances from the fixed point of consecutive iterations. The 100 iterations starting from random initial conditions are colored in grey, whereas in blue we see the average performed over such data sample. The dashed line represents the estimation of the Lipschitz constant obtained by the regression illustrated in the previous panel.

Using these shares as entries, we valuate the new entropic diversities $S_c^{(1)}$ and specializations $Q_p^{(1)}$. We then naturally iterate the procedure to further refine our measures. The algorithm at the k-th steps takes the form:

$$S_c^{(k)} = -\sum_p \frac{Z_p^c Q_p^{(k-1)}}{\sum_{p'} Z_{p'}^c Q_{p'}^{(k-1)}} \log\left[\frac{Z_p^c Q_p^{(k-1)}}{\sum_{p'} Z_{p'}^c Q_{p'}^{(k-1)}}\right] \tag{105}$$

$$Q_p^{(k)} = \log(N) + \sum_c \frac{Z_p^c S^{(k-1)}}{\sum_{c'} Z_p^{c'} S_{c'}^{(k-1)}} \log\left[\frac{Z_p^c S_c^{(k-1)}}{\sum_{c'} Z_p^{c'} S_{c'}^{(k-1)}}\right] \tag{106}$$

which is the one presented in the previous section.

10.2. Convergence. In panel (a) of Fig. 6 we show the euclidean distance evaluated between consecutive iterations, which is defined as

$$d\left[\left(S^{(k+1)}, Q^{(k+1)}\right), \left(S^{(k)}, Q^{(k)}\right)\right] = \left(\sum_c (S_c^{(k+1)} - S_c^{(k)})^2 + \sum_p (Q_p^{(k+1)} - Q_p^{(k)})^2\right)^{1/2} \tag{107}$$

where $S^{(k)}$ and $Q^{(k)}$ are the vectors with components $\{S_c^{(k)}\}_{c=1\ldots N}$ and $\{Q_p^{(k)}\}_{p=1\ldots M}$ respectively. As we can see the distance decreases exponentially fast with a constant rate, until the differences between the iterations go below the double precision limit. In order to understand if the fixed point (S_c, Q_p) is unique (and therefore if the algorithm is globally convergent), we run the iterative algorithm with random initial conditions chosen with a uniform distribution in the range $[0, \log M]$ for the S_c's

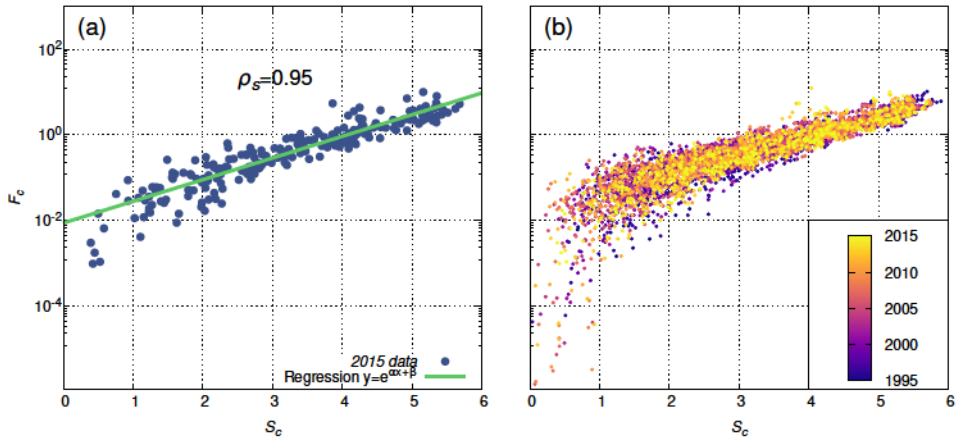


FIGURE 7. Panel (a): Scatter plot comparing the Fitness F and the Entropic Diversity S measures of the countries in the year 2015. A strong positive correlation is found, indicating that an exponential relation holds (green line). Panel (b): Scatter plot of panel (a) containing all the data of every year in the period 1995-2015, each colored with colors following the palette.

and $[0, \log N]$ for the Q_p's. Then, for every iteration k we determined the distance from the fixed point and found that the relation

$$d\left[\left(S^{(k)}, Q^{(k)}\right), (S, Q)\right] = e^{\alpha k + \beta} \tag{108}$$

holds with coefficients $\alpha = -1.720 \pm 0.001$ and $\beta = 3.12 \pm 0.02$, found by fitting the points obtained with 100 different random initial conditions (see panel (b) of Fig. 6). With this parameters we were therefore able to estimate the Lipschitz constant [84], defined as

$$q \equiv \lim_{k\to\infty} \frac{d\left[\left(S^{(k+1)}, Q^{(k+1)}\right), (S, Q)\right]}{d\left[\left(S^{(k)}, Q^{(k)}\right), (S, Q)\right]} \tag{109}$$

which turns out to be

$$q = e^{\alpha} = 0.1824 \pm 0.0003 \tag{110}$$

These estimates suggest to classify the algorithm as globally convergent with a linear rate of convergence ($q < 1$ means that the map associated with the algorithm is a contraction). In panel (c) of Fig. 6 we report the ratios of distances from the fixed points of consecutive iterations, and confront them with the value of the Lipschitz constant q obtained with the previous regression.

11. Comparison with ECI and Fitness

The Fitness algorithm was introduced by Pietronero's group in [8] as an improvement to Hidalgo and Housmann's ECI indicator [7], which was characterized by a fundamental problem in its original formulation. In this section we will introduce both algorithms and highlight in detail the issues and concerns that lead us to the construction of our entropic measure.

11.1. Revealed compared advantage and binarization. One key difference between our measure and both Fitness and ECI is that they are not built upon the

matrix of exports Z_p^c but rather on a binarization of it, obtained using the Revealed Comparative Advantage (RCA) [10], defined as

$$\mathrm{RCA}_p^c = \frac{\frac{Z_p^c}{\sum_{p'} Z_{p'}^c}}{\frac{\sum_{c'} Z_p^{c'}}{Z_p^c}} \tag{111}$$

Basically RCA is larger than one when the share of exports of country on a given product is larger than the share of that product on the global trade, therefore it is used to set a threshold for a countries export. When RCA_p^c is greater or equal to 1 country c is considered an exporter of product p, while when $\mathrm{RCA}_p^c < 1$ that country is not an effective exporter of that product.

Along this idea, a binarized version of the matrix of the Z's is performed:

$$M_p^c = \begin{cases} 1 & \text{if } \mathrm{RCA}_p^c \geq 1 \\ 0 & \text{if } \mathrm{RCA}_p^c < 1 \end{cases} \tag{112}$$

Such matrix will serve as the main ingredient in the iterative schemes. Such choice exposes the schemes to two main concerns: first, there is undeniably an enormous loss of information that comes with the binarization procedure. We throw away all the (continuous) information contained in the matrix Z_p^c in exchange for a discrete "yes-or-no" indicator. Secondly, there is a arbitrariness introduced with the choice of the hard threshold, which, taken in such an early stage of the analysis, is going to introduce a bias in every outcoming result.

11.2. ECI. The *method of refletions* presented in [7] is extremely simple, and in its simplicity lies its inconsistency. It consists in iteratively calculating the average value of the previous-level properties of a node's neighbors. Starting from the zeroth level

$$\begin{cases} k_c^{(0)} = \sum_{p'} M_{p'}^c \\ k_p^{(0)} = \sum_{c'} M_p^{c'} \end{cases} \tag{113}$$

which respectively represent the diversification of a country (number of products exported) and the ubiquity of a product (number of countries exporting a given product), the iterative scheme for $n \geq 1$ is defined as follows.

$$\begin{cases} k_c^{(n)} = \frac{1}{k_c^{(0)}} \sum_{p'} M_{p'}^c k_{p'}^{(n-1)} \\ k_p^{(n)} = \frac{1}{k_p^{(0)}} \sum_{c'} M_p^{c'} k_{c'}^{(n-1)} \end{cases} \tag{114}$$

The authors, at the time of the publication, didn't seem to notice that this scheme, iterated, lead to a trivial solution in which all k's are equivalent with one another. In fact, if we join the equation of the system 114 we get:

$$k_c^{(N)} = \frac{1}{k_c^{(0)}} \sum_{p',c'} \frac{M_{p'}^c M_{p'}^{c'} k_{c'}^{(n-2)}}{k_{p'}^{(0)}} \tag{115}$$

which, for large n is equivalent to solving the eigenvalue problem

$$k_c = \sum_{c'} M'^{cc'} k_{c'} \tag{116}$$

where the matrix

$$M'^{cc'} = \frac{1}{k_c^{(0)}} \sum_{p'} \frac{M_{p'}^c M_{p'}^{c'}}{k_{p'}^{(0)}} = \frac{\sum_{p'} M_{p'}^c M_{p'}^{c'}}{\sum_{p',c'} M_{p'}^c M_{p'}^{c'}} \tag{117}$$

has largest eigenvalue equal to 1 (because $\sum_{c'} M'^{cc'} = 1$, as in a transition probability matrix). The scheme therefore converges to an uninformative eigenvector whose entries are all the same.

The authors later noticed such inconsistency, and later, in the book [**65**], went around the issue by simply choosing to look at the eigenvector associated with the second largest eigenvalue, which was named the Economic Complexity Index (ECI). Since M' is the analogous to a transitional probability matrix, this choice corresponds in using the eigenvactor containing the speed of convergence to the steady state for a Markov jump process, therefore such choice seems rather unclear and unjustified.

11.3. Fitness. Pietronero's work published in [**8**] revealed the issue presented above, and presented a solution to such problem which consisted in introducing non-linearity in the iterative scheme. He defines the fitness of a country F and the complexity of a product Q as the quantities obtained through the following iterative scheme:

$$\begin{cases} \tilde{F}_c^{(n)} = & \sum_{p'} M_{p'}^c Q_{p'}^{(n-1)} \\ \tilde{Q}_p^{(n)} = & \frac{1}{\sum_{c'} M_p^{c'} \frac{1}{F_{c'}^{(n-1)}}} \end{cases} \rightarrow \begin{cases} F_c^{*(n)} = & \frac{\tilde{F}_c^{(n)}}{\left\langle \tilde{F}_c^{(n)} \right\rangle_c} \\ Q_p^{*(n)} = & \frac{\tilde{Q}_p^{(n)}}{\left\langle \tilde{Q}_p^{(n)} \right\rangle_p} \end{cases} \tag{118}$$

where the brackets $\langle \cdot \rangle_x$ indicates an arithmetic average over the ensemble x.

With such choice linearity is broken, therefore the issue that characterized the ECI scheme is not present anymore. However, the choice of using a reciprocal of a reciprocal in the definition of the complexity Q^* rises some convergence issue [1] in some occasions, as pointed out in [**83**]. This happens, for instance in the cases in which a country with F^* close to zero happens to be considered an exporter for a certain product p: this automatically causes the complexity Q^* of such product to go to zero, regardless of any other contribution to the complexity that might have come from other countries.

Here we see how the RCA plays an important role in these scenarios, since choosing a different threshold might have avoided these situations, or also given rise to other divergences. This two main concerns were the reason why we decided to find an alternative to the Fitness, and we found in Shannon's entropy the answer.

11.4. Comparison. In Fig. 7 we compare our indicators with the current state of the art in terms of complexity indicators, represented by the fitness F introduced in [**8**]. As we can clearly see from panels (a) and (b), we have an exceptionally high correlation between F and S: evaluating the Spearman correlator yearly we obtain on average a value of $\langle \rho_s(S, F) \rangle_t = 0.95$, suggesting the existence of a monotonic relation between the two quantities. Such a strong correlation was to be expected, since the algorithm of Fitness has the main purpose of rewarding countries exporting specialized products and data show that countries exporting such products are characterized by high degree of diversification. So, it is not surprising that rewarding in first place number of exports and evenness of relative shares, as in our case, can lead to a substantial agreement. An exponential relation of the form $F \propto e^{\alpha S}$ is found to be quite plausible, with α assuming values close to 1.

[1] The non convergence of the algorithm is masked for two reasons. First, if we were to define the quantity *simplicity* $S = 1/Q$ we would see that it is allowed to diverge towards infinity. Second, both F^* and Q^* follow Pareto-like distribution, therefore using Euclidian metrics doesn't allow to capture this hidden divergence.

When comparing our measure of specialization of the products Q_p with its analog computed with the Fitness algorithm, we still see a positive correlation but definitely lower ($\langle \rho_s(Q, Q^*) \rangle_t = 0.42$ on average).

We suspect that this is probably due to the stability issues that arise because of the nonlinear structure of Q^*, as it was recently pointed out in [**83**]. This suspect is also supported by the fact that the correlation between F and F^* is manifestly worse in the low fitness region, where we find the countries most affected by the convergence issue. We therefore need to further investigate this aspect to better understand how the two specialization measures are related.

12. Calibrated model parameters

Entropic complexity and GDPpc provide an obvious benchmark plane to assess levels of development and stability of countries, even if , of course, one could consider also other quantities, like the eigenvector centrality determined recently in the network study of Ref. [**78, 79**]. Fig. 5(a) reports positions on this plane of 223 countries in 2015. An exponential fit (grey dotted line) shows that $\mathrm{GDP_{pc}}$ increases on average with S_c, although a precise relation does not hold. The four colored regions are consistent with the model parameters comparisons reported in Fig. 5(b)-(e). The boundaries are drawn based on the exponential fit of y vs x indicated in the figure (horizontal), and on a reproduction in our $(S_c, \mathrm{GDP_{pc}})$ plane of a similar plot in the plane $(F_c, \mathrm{GDP_{pc}})$ of Ref. [**82**]. In Fig. 5(b) the 131 representative points of the calibrated countries are colored from a palette defined by the values of the total noise amplitude, $\sigma^c_{tot} = \sqrt{(\sigma^c)^2 \tau^c + (\sigma^c_0)^2}$. The noise strength is rather low in the upper right corner (developed countries), and very large in the lower left one (risky economies). The vertical gradient is present also on the right side, where, however, some developing countries, like China and India, show relatively low noise effects. Panel (c) reports the average annual transfer rate of resources between exports, $\langle J^c_{pp'} \rangle = \sum_{pp'} J^c_{pp'}/M^c$, an indicator of the flexibility of the economy. This is always large for developed countries. One finds exceptionally high values also for some less developed countries. Panel (d) reports $\bar{\mu}^c$, the average annual rate of resources input. This is only one of the contributions to the slope seen, e.g., in the time series in Fig. 2. It is rather low for countries in the upper right corner. This should not surprise, since $\bar{\mu}^c$ is not a usual growth indicator, like $\mathrm{GDP_{pc}}$, but rather quantifies the average rate of investment of resources for the growth net of inflation and cooperativity effects. It is much higher in the lower right corner, where emerging economies with high S_c are located. On the left vertical side countries with low entropic complexity have a mixed behavior, signaling that these are dynamically far from stationarity. The low value of $\bar{\mu}^c$ for developed economies means that, when the country approaches a state with high entropy and high $\mathrm{GDP_{pc}}$, fresh resources for growth start diminishing. A similar effect has been clearly observed in ref. [**1**] for exports aggregated at global level over a period of 39 years. The indicator h^c_T is reported in panel (e), showing that developed economies have a relatively low contribution to growth coming from cooperative effects. This is consistent with the fact that for these countries the margins of optimization of growth upon variation of G^c are rather narrow, as seen, e. g., in the counterfactual analysis for the USA (Fig. 4(a)). In particular h^c_T remains very small compared to the total drift inclusive of inflation. As already remarked above, h^c_T depends on combined effects of the transfers and of noise. Higher values of h^c_T are found for countries with low $\mathrm{GDP_{pc}}$ and medium S_c. Thus, relatively low transfer rates combined with high levels of noise can produce relevant cooperative growth effects.

13. Discussion

The data driven approach to growth proposed here has features and implications worth of further investigation in both development economics and network theory. Our dynamic model provides novel insight into the complexity of economic systems and into the role played by cooperative effects in growth. The idea that the productive output of a country, if weighted in terms of competitiveness, allows to fully account for endowments [**7, 8**], has been extended to the dynamic context. Indeed, the basic assumption made here is that, for each country, the export basket structure with its fluctuations should also be sufficient to determine the statistics of its evolution in time. The model embodies explicit interaction effects between productions which are found to be an essential ingredient of growth dynamics, but are generally not addressed in the economic literature. Such effects play a key role in enhancing or depressing the growth due to their interplay with the variability of market conditions. The fact that the long term dynamics is closely controlled by the relative distribution of resources in the export basket of each country suggests that interventions aimed at increasing efficiency of an economic system should stably modify the structure of its basket, making it as similar as possible to those of the most efficient economies. The effects of such modification should be testable by simulations of our model.

Besides inspiring the dynamic model construction, the focus on the distribution of resources in different exports suggested a novel way to obtain measures of complexity based on the full information content of the data. We showed here that a tool to estimate variety familiar in development economics, namely the Shannon entropy function [**71**], can be used to evaluate the diversity of productions consistently with the specialization of products. The resulting entropic measures, S_c and Q_p, do not rely on RCA threshold criteria [**10, 11**] and are nicely convergent and stable. In the broader context of the theory of networks, the use made here of the Shannon entropy function to extract these measures is fully original with respect to previous applications aimed at characterizing topological heterogeneity [**85**], or at assessing the statistical significance of monopartite projections of bipartite networks [**86**]. Our results open a novel, entropy based way to explore the properties of bipartite networks so frequently met in the real world [**87**].

Our entropic measures are derived directly from those distributions identified here as the main generators of dynamics. This strongly supports the expectation that these measures should embody essential information concerning economic growth and should be good candidates as appropriate collective variables for describing this phenomenon in low dimensionality spaces [**82**]. Thus, the nexus explored in this work is deep and promises to further improve our understanding of growth complexity.

CHAPTER 4

Analysis of the G7 countries

1. Introduction

The NBER-UN database [49] we used to build the time series of the exports aggregated at a global level contains also data specific to single countries. In this chapter we will talk about the analysis we carried out in the case of specific countries (published in [3]) following the lines of the work done for the global aggregated export.

When aggregating yearly exports at the global level, for each product one gets a resultant coinciding with that obtained by aggregating imports. A country, instead, is a subset of a larger system in which trades can be carried out in both directions: imports and exports. The NBER-UN database contains both this kinds of data, reporting yearly the amount of commodities that have been imported from and exported towards the rest of the work, therefore we need to decide which data we need to use in our analysis. The choice of exports is, of course, appropriate if, for example, the aim is to investigate the productive potential of the countries in accordance with general ideas of the Economic Complexity approach [7, 8]. We performed this kind of analysis on a more recent database, which was presented in chapter 3. Here, more consistently with the previous investigation of aggregated exports at a global level, we consider, for each country, the basket reporting the sum of yearly exports and imports for each product. This choice is consistent with the aim of describing the growth of the economies, which is related to the increments of both imports and exports. We limit here our consideration to the countries of G7 (Canada, France, Germany, Italy, Japan, the United Kingdom, and the USA) and, as a benchmark, we consider the global aggregate exports realized by all countries worldwide.

2. Data construction

We proceed with the construction of the dataset of a single country c. For a given product p, the NBER-UN database contains both the amount of import $Z^{c,I}_{p,n}$ and export $Z^{c,E}_{p,n}$ for any given year n in the time period 1962-2000. In Fig. 1 it is shown the case of the USA as an illustrative example. The same coarse graining procedure to 3 digits of the SITC4 classification has been carried out, reducing the number of classes to 202 in the case of the import and 221 for the exports.

Unfortunately with this kind of analysis we come across some issues related to the data: in some occasion, certain commodities do not present any record for a given year. In such cases one needs to remove the good in order to avoid zeroes which cannot be taken into account by the dynamical evolution. This problem however is quite rare in the case of the G7 countries, and it is further reduced by the aggregation imoprt+export that we operate to construct the final time series of the country:

$$Z^{c}_{p,n} = Z^{c,I}_{p,n} + Z^{c,E}_{p,n} \tag{119}$$

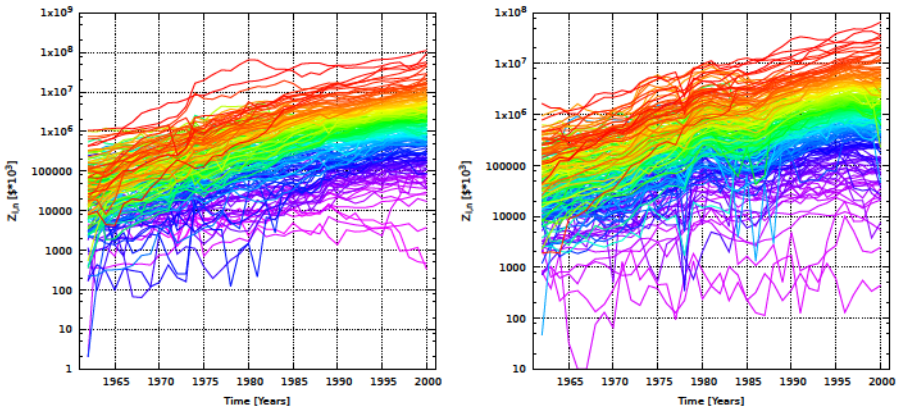


FIGURE 1. Comparison between the import (left) and export (right) time series of the USA in the time period 1962-2000. The color's wavelength is proportional to the ranking of each product.

3. Model

Data for yearly exports and imports were taken from the international trade data furnished by the National Bureau of Economic Research [**46**] and cover a period of 39 years from 1962 to 2000. The products are classified on the basis of the Standardized International Trade Code at 3-digit level (SITC-3) and trades are reported in US-dollars. We limit here our consideration to the countries of G7 (Canada, France, Germany, Italy, Japan, the United Kingdom, and the USA) [1] and, as a benchmark, we consider the global aggregate exports realized by all countries worldwide. We denote as $Z^c_{p,n}$ the total value (in thousands of current US-dollars) of the product category p ($p = 1, 2, \ldots, M^c$) traded (import plus export) in the year n ($n = 0, 1, \ldots, T$ with $T = 38$) by country (world) c. The number of products, M^c, varies from country to country and is reported in table 1. When referring to worldwide aggregated data, the superscript c is implicitly assumed to mean world.

As already noticed in previous papers [**1, 2**], a key feature of the organization of economies consist in the fact that the various products, besides showing an approximate average exponential growth, have a rather stable ranking over the whole considered period. In fact, if we assign to each product, p, of a given country a color whose wavelength is proportional to the fraction of the value of product p over the whole country basket, $\Omega^c_n \equiv \sum_{p=1}^{M^c} Z^c_{p,n}$, averaged over the last 10 years covered by the database,

$$z^c_p \equiv \frac{1}{10} \sum_{n=29}^{38} \frac{Z^c_{p,n}}{\Omega^c_n} \,, \tag{120}$$

a rainbow image is perceived when plotting the exports as a function of time (see Figure 2).

The interpolations shown in Figure 2 resemble those of geometric Brownian motions: on average they show similar drifts and fluctuations keeping amplitude constant on a logarithmic scale. This suggests to define average growth as [**51**]:

[1] Note that imports and exports for each one of these countries have been recorded as imports from and export towards the rest of the world.

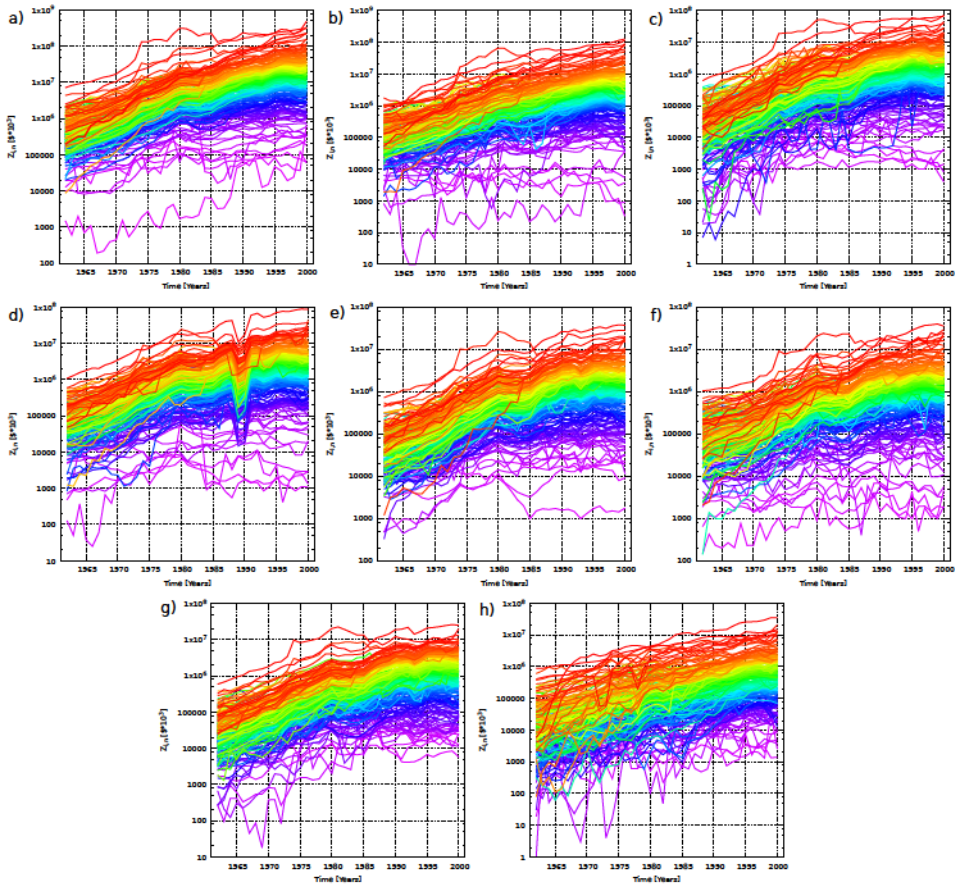


FIGURE 2. Time evolution of the trades $Z^c_{p,t}$ of the World (**a**), the USA (**b**), Japan (**c**), Germany (**d**), France (**e**), the UK (**f**), Italy (**g**) and Canada (**h**). The wavelengths of the colors used are directly proportional to z^c_p.

$$\lambda^c_T = \frac{1}{T \cdot M^c} \sum_{p=1}^{M^c} \log \left[\frac{Z^c_{p,T}}{Z^c_{p,0}} \right] \tag{121}$$

In Table 1 (second column) we report the estimated growth for the world and for each country. As we can see Canada and Japan are the countries who experienced the largest average growth, while the United Kingdom is the country that grew the least among the G7.

Since yearly trade records result from variations in much shorter periods, we switch to continuous time t in year units (with $t = 0$ corresponding to 1962): the trade values in the year preceding time t is now indicated with $Z^c_p(t)$. Thus, $Z^c_{p,n}$ gives a discrete representation of $Z^c_p(t)$ with a point for every year in the database. Taking inspiration from a former paper of Goudré *et al.* [**51**] and following Ref. [1] the minimal model describing how the values of various trades evolve in time can be written as a set of Stochastic Differential Equations (SDE):

$$\partial_t Z^c_p(t) = \left[\eta^c_p(t) + \mu^c(t) \right] Z^c_p(t) + \sum_{p' \neq p} \left[J^c_{pp'} Z^c_{p'}(t) - J^c_{p'p} Z^c_p(t) \right] , \tag{122}$$

TABLE 1. Values of the model parameters for the world and each country member of the G7. Errors of the calibration procedure are also reported. Exception is made for τ^c for which the errors obtained from the calibration procedure are not very meaningful since even a great change (up to 80%) in its value influences very little the results and the other calibrated parameters.

	M^c	λ_T^c [y^{-1}]	G^c [y^{-1}]	$\bar{\mu}^c$ [y^{-1}]	σ^c [y$^{-1/2}$]	τ^c [y]	h_T^c [y^{-1}]
World	226	0.089	0.042 ± 0.001	$(6.31 \pm 0.05) \times 10^{-3}$	0.109 ± 0.002	0.607	$(3.46 \pm 0.05) \times 10^{-3}$
USA	224	0.086	0.034 ± 0.001	$(2.25 \pm 0.07) \times 10^{-3}$	0.153 ± 0.005	2.64	$(4.75 \pm 0.07) \times 10^{-3}$
Japan	220	0.109	0.011 ± 0.001	$(2.33 \pm 0.01) \times 10^{-2}$	0.203 ± 0.009	5.5	$(6.43 \pm 0.09) \times 10^{-3}$
Germany	224	0.086	0.023 ± 0.001	$(4.28 \pm 0.04) \times 10^{-3}$	0.137 ± 0.001	1.67	$(2.79 \pm 0.04) \times 10^{-3}$
France	222	0.087	0.042 ± 0.001	$(3.53 \pm 0.07) \times 10^{-3}$	0.128 ± 0.001	1.70	$(4.55 \pm 0.06) \times 10^{-3}$
UK	221	0.072	0.035 ± 0.001	$(-1.192 \pm 0.001) \times 10^{-2}$	0.18 ± 0.01	12.7	$(4.82 \pm 0.07) \times 10^{-3}$
Italy	221	0.092	0.068 ± 0.002	$(7.67 \pm 0.09) \times 10^{-3}$	0.112 ± 0.004	0.09	$(5.88 \pm 0.11) \times 10^{-3}$
Canada	219	0.100	0.007 ± 0.001	$(1.40 \pm 0.01) \times 10^{-2}$	0.188 ± 0.001	1.02	$(7.02 \pm 0.10) \times 10^{-3}$

where $\mu^c(t)$ represents a deterministic drift, accounting for the average growth of the exports (including the inflationary one) and $\eta_p^c(t)$ is a multiplicative noise representing the variability of conditions faced by different products at different times. Such variable conditions may depend on many complex causes which can be both "internal" to the country itself and "external", i.e. due to dynamics in countries with which goods are traded. Finally, the coupling terms $J_{pp'}^c$ describe the shift of resources from production j to production i.

Stochastic differential equations similar to (122), were used to describe other problems like population and evolutionary dynamics [**68, 88**], portfolio strategies [**53**], interface growth [**69**] or optimal pinning of vortices by random defects in materials [**56, 55**]. An interesting aspect of Eq. (122) is the fact that, if the noise $\eta_p^c(t)$ is correlated in time, as it is natural to expect, its combination with the coupling terms proportional to $J_{pp'}^c$ may induce a nonzero average growth even in the absence of deterministic trends driving the single productions [**51**].

The deterministic drift $\mu^c(t)$, represents the average growth of the import plus exports in a given country in the absence of contributions generated by the interplay between the correlated noise and transfer terms. The time dependency comes from the need to include inflationary effects due to the fact that import-export values are expressed in the current currency of the specific year. Thus, we write

$$\mu^c(t) = \bar{\mu}^c + I(t) \,, \tag{123}$$

in order to separate the real average contribution to the drift, $\bar{\mu}^c$, from the inflationary one $I(t)$, which can be read as an yearly step-wise function whose values are taken from the Organization for Economic Cooperation and Development (OECD) [**59**] and reported in table 1.

To take into account the variability of conditions to which the production and acquisition of a given product are subjected, Eq. (122) contains a multiplicative noise, η_p^c, that, together with the deterministic drift, lets the quantities $Z_p^c(t)$ perform a sort of geometric Brownian motion. As we mention this noise has to be correlated in time and our choice, similar to what has been done in Ref. [**51**], is for an exponential correlation with a characteristic time τ^c which represents the typical duration of opportunity/crisis periods. Since the various products are reasonably clustered in sets facing similar external conditions, we also introduce a correlation among products. Thus, the noise has zero average, $\langle \eta_p^c(t) \rangle = 0$, and its

correlator reads

$$\langle \eta_p^c(t_1)\, \eta_{p'}^c(t_2)\rangle = c_{pp'}^c\, \frac{(\sigma^c)^2}{\tau^c}\, \mathrm{e}^{-|t_1-t_2|/\tau^c}\ , \tag{124}$$

where σ^c weights the importance of the stochastic part of the dynamics and $c_{pp'}^c$ are the elements of the correlation matrix constructed from the variations at equal times of the different exports in the database [1] (see Material and Methods section for further details).

Finally, we put the coupling terms equal to:

$$J_{pp'}^c = G^c z_p^c |c_{pp'}^c|\ . \tag{125}$$

where $c_{pp'}^c$ are again the elements of the correlation matrix entering in the noise and G^c is a coupling constant that regulates the magnitude of the transfer of resources among different products. By considering the factor $|c_{pp'}^c|$ as a proxy for inverse distance, Eq. (125) is consistent with the gravity law, often used for estimating transfer rates in economics [**76, 49**]. The proportionality of $J_{pp'}^c$ to z_p^c guarantees that, for $t \to \infty$, $Z_p^c(t) \sim z_p^c$.

The network specified by the $J_{pp'}^c$'s is deeply different from other networks proposed in the economic complexity literature as that based on products similarities [**11**]. In the latter, for instance, *apples* and *pears* are strongly connected because they need the same infrastructures and undergo similar production processes. Such kind of similarities are indirectly present in our case as a secondary effect inherent to the correlation matrix, which in turn also takes into account the fact that a fluctuation in the trade of *oil* at a certain time is likely going to affect the production of *apples*, *pears* and many other products. However, an even more important feature of the matrix of transfer rates is the proportionality $J_{pp'}^c \propto z_p^c$, which strongly weights the influence of each single product on the global dynamics. Indeed, if a given product p' experiences favorable conditions for growth, in force of this proportionality, part of its extra gain tends to be mostly redistributed towards nodes with larger z_p^c. A proper estimate of how effective this mechanism is, requires to take into account also proper notions of centrality of the directed network. The next section is partly devoted to such discussion. At the same time, given the proportionality of $J_{pp'}^c$ to z_p^c, the structure of our network is such that the most traded products are also the most central nodes (see Figure 3).

4. Results

4.1. Matrix of transfer rates. For each country and for the World, the directed network having products as nodes and links is quite complex. Figure 3 gives a partial and undirected visual representation of the network structure in the case the USA and the world. The links $\tilde{J}_{pp'}^c$ of the undirected graphs U^c and U^{WOR} are built with a maximum criterion: $\tilde{J}_{pp'}^c = \max\{J_{pp'}^c, J_{p'p}^c\}$. At a first glance, the qualitative structure of the two networks appears very similar: there are few nodes (about 2% of the total number) that are central (degreewise) in the graphical representation, with the remaining nodes being connected almost exclusively to the more central ones in a fashion similar to scale-free networks [**54, 89**]. In both cases, the central nodes are given by the same categories of goods, namely oil, cars, machinery and electronics related products. Finally, the USA seem to have an economy thoroughly dominated by machinery and electronics, and, quite surprisingly, the oil related nodes are substantially smaller than in the case of the world network.

A more quantitative analysis to highlight these analogies and differences can be performed by determining a suitable node centrality measure. The $J_{pp'}^c$ matrix has strictly positive real entries, that can be considered the adjacency matrix of a full

TABLE 2. Top 10 products in the world and USA $J^c_{pp'}$ networks with highest authority score. The value of z^c_p (defined in 120) together with its ranking is also reported.

World					
Auth. rank	Auth. score	z^c_p rank	z^c_p	SITC-3	Commodity Description
1	1	2	0.052	781	Passenger motor vehicles (excluding buses)
2	0.976	1	0.057	333	Crude petroleum and oils obtained from bituminous minerals
3	0.792	3	0.038	776	Thermionic, microcircuits, transistors, valves, etc
4	0.459	6	0.025	784	Motor vehicle parts and accessories
5	0.440	4	0.031	752	Automatic data processing machines and units thereof
6	0.432	10	0.016	778	Electrical machinery and apparatus
7	0.423	5	0.026	764	Telecommunication equipment, parts and accessories
8	0.408	13	0.015	641	Paper and paperboard
9	0.407	9	0.016	541	Medicinal and pharmaceutical products
10	0.405	12	0.015	583	Polymerization and copolymerization products
USA					
Auth. rank	Auth. score	z^c_p rank	z^c_p	SITC-3	Commodity Description
1	1	2	0.054	776	Thermionic, microcircuits, transistors, valves, etc
2	0.697	4	0.041	333	Crude petroleum and oils obtained from bituminous minerals
3	0.570	3	0.049	752	Automatic data processing machines and units thereof
4	0.564	1	0.069	781	Passenger motor vehicles (excluding buses)
5	0.331	7	0.031	764	Telecommunication equipment, parts and accessories
6	0.330	6	0.035	784	Motor vehicle parts and accessories
7	0.322	10	0.018	778	Electrical machinery and apparatus
8	0.307	9	0.020	874	Measuring, checking, analysis, controlling instruments, parts
9	0.292	5	0.036	792	Aircraft and associated equipment, and parts thereof
10	0.279	14	0.014	641	Paper and paperboard

directed weighted graph. To properly exploit such feature we chose to make use of the Kleinberg's authority score [**90**] In Table 2 we show the top 10 commodities with respect to Kleinberg's authority: as we can see, 8 out of 10 products are present in both charts and only swap position in ranks, confirming that the two networks are very similar to each other. It is interesting to notice that the ranking of products in terms of authority score does not differ very much from that in terms of z^c_p. So, the structure of our network is such that the most traded products are also the most central nodes (see Figure 3). Interestingly, Sharma et al. have recently shown that a very similar structure arises in the financial network at sectoral level by using a methodology based on multi-layered networks [**78, 79**]: in fact, their results show that there exists a one-to-one mapping between the economic size of the sectors and their centrality in the corresponding financial network.

In Figure 4 we plotted the empirical survival distribution function (ESDF) of the authority scores for the networks of the World and of all the G7 countries. One notices that the world's network has two nodes with a very high value (781 and 333) while the USA have only one (776). Another interesting feature pointed out by Figure 4 is the initial exponential decay of all the authority scores ESDF. This is another property that our product networks have in common with scale-free networks. All the networks of the G7 countries present a distribution steeper than the world's one. Taking inspiration from work related to the vulnerability of networks [**91**], steepness can be interpreted as an alternative instability indicator [**92**], since concentrating high values of centrality in few products, will result in an exposure of the country to major risks, in the eventuality of a crisis striking such sectors (negative trends will spread very easily to the rest of the nodes). These scenarios could be in principle tested with our dynamical model, through simulations of hypothetical setbacks of high-centrality products.

4.2. Calibrated parameters. For every G7 country, we performed a calibration procedure in order to determine the values of the model parameters that best fit

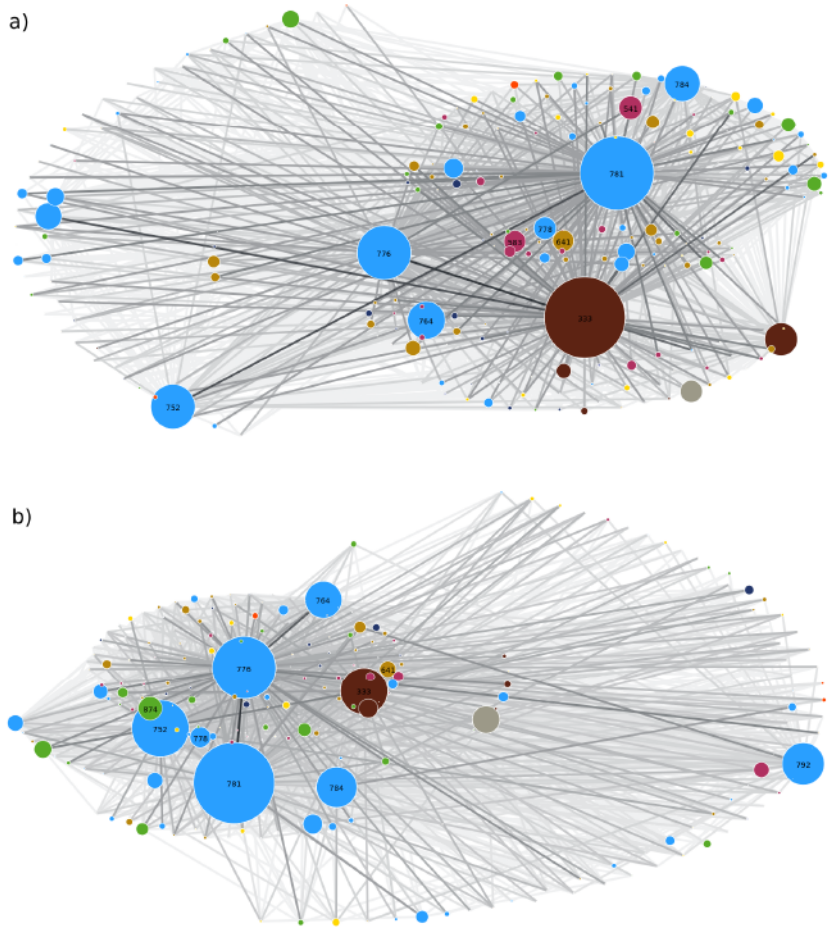


FIGURE 3. Representation of undirected graphs U^{WOR} (panel a) and U^{USA} (panel b) associated with the corresponding $\tilde{J}^c_{pp'} = \max\{J^c_{pp'}, J^c_{p'p}\}$ matrices. Since U^c are fully connected graphs with approximately $M^c(M^c-1) \sim 5 \cdot 10^4$ edges, here we reported only the 10% of the strongest links, which in turn are colored with a palette that is lighter for the weakest among these. The size of the nodes is directly proportional to the value of the ranking z^c_p associated with the product p (the SITC code is highlighted for the most central nodes, see Table 2), while the colors are representative of the macro-category of products illustrated in the legend.

the historical data. In Table 1 we show such values with the associated errors. We note that σ^c assumes values in the interval $[0.1, 0.2]$ $\mathrm{y}^{-1/2}$, with Canada and Japan showing the highest ones, and the world the smallest. The latter feature has to be expected since multiplicative noise should get reduced by the aggregation process.

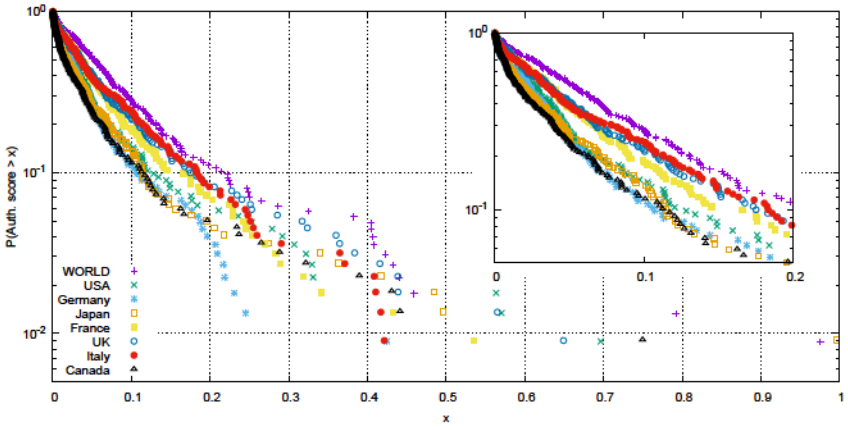


FIGURE 4. Empirical survival distribution function (ESDF) of the products' authority valuated over the graphs associated with the $J^c_{pp'}$ network of the world and the countries member of the G7. For the majority of the less central products the trend follows an exponentially decaying law, with all the countries showing a curve slightly steeper than the one of the world. Top ranked products of the world and the USA are reported in table 2.

The values of the parameter G^c of all the countries have comparable magnitudes. In the next section we will investigate more accurately the role played by such parameter in determining the overall average growth λ^c_T of every country.

All the values of $\bar{\mu}^c$, are of comparable magnitude and are substantially lower than the contribution which can be ascribed to average inflation. This is approximately around 0.08 y^{-1} (see Table 1 in chapter 2). Quite interesting is the case of the UK, the only country showing a negative value for the deterministic drift $\bar{\mu}^c$: this is probably due to the fact that the UK, in the second half of the twentieth century, did not manage to fully exploit the relationship with its trade partners as well as the other countries did. Nevertheless, its inner trade network is still good because it shows an average growth similar to that of the other countries.

4.3. Counterfactual analysis and optimization. Our model allows to distinguish three distinct contributions to the overall growth. In fact, if we integrate Eq. 122 and average over the products we obtain

$$\lambda^c_T = h^c_T(G^c) + \bar{\mu}^c + \frac{1}{T}\int_0^T I(t)\,\mathrm{d}t \tag{126}$$

On top of the two deterministic contributions due to $\bar{\mu}^c$ and the average inflation, we have the term $h^c_T(G^c)$. This results from the integration of the transfer terms and can be estimated from historical data by discrete summations. Its magnitude depends sensibly on G^c, and on the interplay that the resource transfers have with the fluctuations determined by multiplicative noise. Indeed, favorable stochastic fluctuations at a local level, if properly exploited, can spread globally to the rest of the network more efficiently than unfavorable ones. In portfolio optimization, this fact leads to the explore-exploit dilemma [**51**] of deciding whether to exploit a local opportunity (of amplitude σ^c and expected duration τ^c) or to move towards possibilities offered by other nodes in the network by transferring a percentage of the local investment (at a transfer rate speed controlled by G^c).

It is therefore interesting to study and quantify what the overall growth of the network would have been if we varied the coupling constant of the transfers rate G^c. Such *counterfactual analysis*, produces the plots reported in Figure 5: in panel (**a**) we show, for every G7 country and the world, the dependence of λ_T^c on G^c (in logarithmic scale). These results are obtained by simulating many times the evolution of every network at fixed, historically calibrated values of $\bar{\mu}^c$, σ^c and τ^c, and varying only G^c. For extremely low values of G^c the growth is exclusively determined given by the deterministic drifts, and eq. 122 reads:

$$\frac{\partial Z_p^c}{\partial t} = \left(\eta_p^c(t) + \bar{\mu} + I(t)\right) Z_p^c(t). \tag{127}$$

Once integrated in the interval $[0,T]$ and averaged over the products this equation yields the relation $\lambda_T^c = \bar{\mu}^c + \frac{1}{T}\int_0^T I(t)\ \mathrm{d}t$. For increasing values of G^c we see that every country produces the same kind of curve: λ_T^c rises until it reaches a peak for some specific value of G^c, and then starts to fall for large values. Almost every country (the exception is Canada) has a curve that for extremely large values of G^c goes beneath the plateau defined by the deterministic drift: this means that in conditions of frenetic transfers the contribution to the growth h_T^c is negative.

The arrows in the plots indicate the coordinates of G^c and λ_T^c determined from the historical data (values shown in table 1). In order to better compare the intrinsic growth of the network, in panel (**b**) of Figure 5 we plot the same curves deprived of their drift contributes.

Clearly, every network is characterized by different peak amplitudes, and by different locations of the peaks with respect to the historically calibrated G^c. For two countries (namely Canada and Japan) the plots indicate a much higher, unexpressed growth potential compared to that of the other countries. Indeed, an even mild increase of G^c for them would have produced substantial extra growth. These two countries are also those with the highest historical h_T^c, and this is in agreement with the fact that Japan and Canada have been two emerging economies in the second half of the twentieth century that reached a well established position nowadays (indeed they became G7 members). Looking at the rainbow plots in Figure 2 we also see that for these countries numerous products were out of rank in 1962 and only through the later evolution reached a presumably more stable position in the year 2000. This is in agreement with results obtained for the calibration of the parameter σ^c: as we can see in Table 1 and already remarked above, Canada and Japan are the countries with the two highest σ^c among those observed. Having a great amplitude of fluctuations, together with a good organization of the transfer rates, can lead to a very high overall growth.

For the other countries, we find that the peak has an amplitude of magnitude comparable to that of the world (just slightly bigger). We also see that the values of the calibrated parameter G^c are pretty close to that of the world, while those of Canada and Japan are almost one order of magnitude smaller. Moreover, we find that the value of the historically calibrated G^c is rather close to the location of the maximum (as for the world). All these hints tell us that these other countries have an economy which is much more similar to that of the world, because they have been well established since the beginning of the analysis in 1962, while Canada and Japan, as previously stated, underwent big radical changes in this 39 years period that led them to become leading economies in the world scenario.

Finally, we observe that a common feature shared by all countries is that the corresponding calibrated value of G^c is always on the left side of the peak. This is an indication of a conservative character of their economies, in the sense that

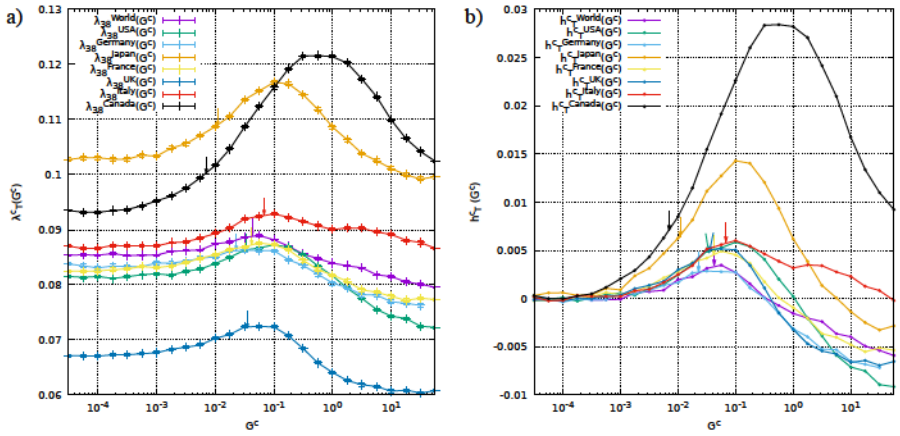


FIGURE 5. (a) Study of the overall growth λ_T dependence on the parameter G^c. Arrows indicate, for every country (world), the coordinates of the real historical data. (b) Same curves of the previous panel deprived of the deterministic drifts (average inflation and $\bar{\mu}^c$), hence showing the contribution to the growth h^c_T given by the cooperative effects of the dynamical network.

these countries prefer the safety of exploiting the resources rather than exploring new directions of investment through transfers.

5. Discussion

A dynamic description of the evolution in time of the bipartite network connecting countries to the traded goods can be a key to identify interesting complexity features of the economies and of the products. The choice of undertaking such modelization amounts to pushing further some basic points of view of the economic complexity approach [**7, 65**]: besides assuming that export/import panels should be sufficient to take into account intangible factors of the economies, it is postulated that specifying the composition of the trade baskets should be fully sufficient to account for their time evolution up to the information contained in the deterministic drift and the noise. The dynamical insight one can gain is independent from, and complementary to that provided by analyses of the Fitness complexity approach [**8, 66, 67, 82**]. Here we tried to give an account of these features and of the perspectives they are opening, by considering the application of our stochastic differential system of equations to the case of the G7 countries and to the data aggregated for the whole world. The relevant emerging aspects are related to the novel network structure underlying the dynamics of transfers between different productions and to the possibility of performing synthetic simulations and counterfactual analyses. The preliminary results shown should provide a clear indication of the information one can expect to extract by a careful analysis of the networks and of the model dynamics.

CHAPTER 5

Growth of entropy production in system out of equilibrium

1. Introduction

In the previous chapters we treated a typical non-equilibrium process (growth) in the specificity of a macroeconomic context. Our treatment of the dynamics we observed, based on a stochastic description, led to identify Shannon and Boltzmann's entropy function [**9**] as a fundamental ingredient to establish rankings for both countries and products, allowing us to accomplish the main goal of the whole field of Economic Complexity.

Entropy and non-equilibrium are recurrent keywords in the most recent developments of statistical mechanics. The entropy production in systems out of equilibrium is expected to undergo fluctuations crucial for issues like the interpretation of micro- or nano-manipulation experiments or the functioning of molecular motors. At stationarity these fluctuations are controlled by Cramer's functions, which, in the case of Markovian dynamics, have been shown to obey functional relations analogous to those valid for deterministic dynamics [**14, 15, 16, 17**]. The fluctuation theorems implied by these relations are of key importance for the above applications and for the whole field of stochastic thermodynamics.

Any Markovian description assumes a limited resolving power in distinguishing between different states of the system. Since this power is determined by the experimental apparatus, it is a fundamental problem to establish how the entropy production possibly changes in a coarse-grained description in which a reduced number of states accounts for the use of instruments with lower resolution. This is a very difficult issue since coarse-graining of any Markovian model necessarily produces memory effects which hinder the analysis. A pioneering attempt to address such problem performed numerically coarse-grainings consisting of merging into single entities clusters of states connected by fast transition rates [**20, 21**]. A wide gap between these rates and the other time scales in the problem allowed to approximate the coarse-grained dynamics by a Markov jump process, on the basis of which one could define an entropy production. However, exact insight into entropy production for coarse-grained dynamics and the possible validity for it of the fluctuation theorem is missing. Further motivation to investigate such issue comes from the fact that recently coarse-graining of dynamical data to make them consistent with Markov jump processes was even proposed as a strategy to detect entropy production in active matter systems [**18, 19**]. In these studies, however, the possible role played by memory effects was not considered.

In this chapter we show how Markov jump processes can be exactly coarse-grained, and elucidate the non Markovian nature of the resulting dynamics. We also prove that the entropy production rate at stationarity and its fluctuations remain invariant. Such invariance thus qualifies as the guiding principle to set up a correct coarse-graining in non-equilibrium statistical mechanics, excluding the possibility that a reduction of the number of states of the system could cause deviations from the fluctuation relations for the coarse-grained entropy.

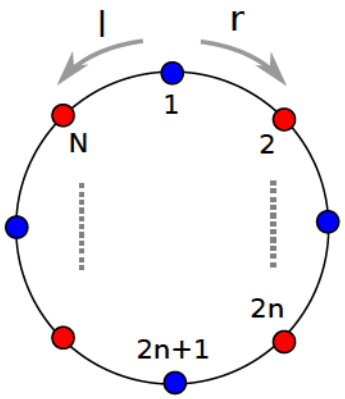


FIGURE 1. Sketch illustrating the jump model characterized by r and l jump rates. For treatment simplicity the total number of initial sites N is chosen to be even, but this is not a strict requirement.

2. Periodic linear network with N states

We start considering jump processes on networks in which symmetry and homogeneity allow analytic treatment with a limited number of equivalent states. A paradigmatic example (depicted in Fig. 1) is the periodic linear chain of N states for which we can write:

$$[r+l+\partial_t]\,P_i(S,t) = rP_{i-1}(S-\log(r/l),t) + lP_{i+1}(S+\log(r/l,t) \tag{128}$$

where $P_i(S,t)$ is the probability that at time t the system is in state i with total entropy S accumulated along the trajectory in the interval $(0,t)$. The positive rates r/l refer to jumps from a state to the right/left neighbor, with entropy increments $\log(r/l)/\log(l/r)$. Based on Eq. 128, for $P_i(t) = \Sigma_S P_i(S,t)$ holds the usual Master equation, for which a unique stationary solution, reached from arbitrary initial conditions, is guaranteed by rigorous theorems. Periodicity ($N+1=1$) allows dissipation and probability current flow. Indicating by $Q(S,t) = \Sigma_i P_i(S,t)$ the probability of having produced a total entropy S at time t, the corresponding scaling cumulant generating function (SCGF) ε can be extracted from the generating function

$$G(\lambda,t) = \sum_S e^{\lambda S} Q(S,t) \sim_{t\to\infty} e^{\varepsilon(\lambda,r,l)t} \tag{129}$$

by solving a first order differential equation implied by Eq. 128 for G. This yields

$$\varepsilon(\lambda,r,l) = re^{\lambda \log r/l} + le^{-\lambda \log(r/l)} - (r+l). \tag{130}$$

Deriving ε with respect to λ at $\lambda = 0$ one obtains the scaled moments of S and, upon Legendre-Fenchel transforming, the Cramer's function $I(\sigma,r,l) = \sup_{\lambda\in\mathbb{R}}[\sigma\lambda - \varepsilon(\lambda,r,l)]$, with $\sigma = \frac{S}{t}$.

Coarse-graining can be performed by, e.g., eliminating from the system in Eq. 128 all even states (assuming even $N \geq 4$). This is done exactly by switching to Fourier transforms in time for the $P_i(S,t)$'s:

$$[r+l+i\omega]\,\tilde{P}_i(S,\omega) = r\tilde{P}_{i-1}(S-\log(r/l),\omega) + l\tilde{P}_{i+1}(S+\log(r/l,\omega) \tag{131}$$

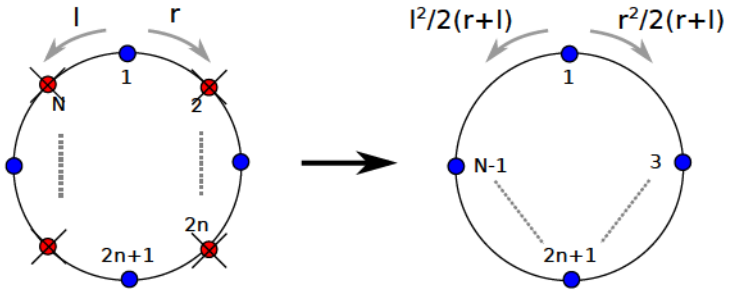


FIGURE 2. Illustration of the coarse-graining procedure of Eq. 128 in which the odd sites are decimated. The resulting system is not Markovian jump process because of memory effects, but jump rates can still be defined.

where $\tilde{P}_i(S,\omega) = \int_{\mathbb{R}} P_i(S,t)e^{i\omega t}\,\mathrm{d}t$. Expanding the right hand terms as function of their neighbors $i \pm 2$ we get :

$$\left[-\tfrac{1}{2(r+l)}\omega^2 + i\omega\right]\tilde{P}_i(S,\omega) = \tfrac{r^2}{2(r+l)}\tilde{P}_{i-2}(S-2\log(r/l),\omega) + \\ +\tfrac{l^2}{2(r+l)}\tilde{P}_{i+2}(S+2\log(r/l),\omega) - \tfrac{r^2+l^2}{2(r+l)}\tilde{P}_i(S,\omega) \tag{132}$$

and upon anti-transforming the odd (as well as the even) P_i's satisfy

$$\left[\tfrac{1}{2(r+l)}\partial_t^2 + \partial_t\right]P_i(S,t) = \tfrac{r^2}{2(r+l)}P_{i-2}(S-2\log(r/l),t) + \\ +\tfrac{l^2}{2(r+l)}P_{i+2}(S+2\log(r/l),t) - \tfrac{r^2+l^2}{2(r+l)}P_i(S,t) \tag{133}$$

In this relatively simple case, an implicit uniform rescaling by a factor 2 of all the surviving $P_i(t) = \Sigma_{S'}P_i(S',t)$'s is sufficient to maintain for them the meaning of probabilities and to get consistency of the coarse-grained stationary state with the original one. However, the $P_i(t)$'s do not satisfy a Master equation anymore. Indeed, summing both sides of Eq. 2 over S' yields an equation with both ∂_t and ∂_t^2. This memory effect is best analysed by performing coarse-graining directly on the trajectories of the original, undecimated process. Recording only the first times of arrival to even states, the distribution of the waiting times separating two successive arrivals can be evaluated exactly and shows clear non-exponential, non-Markovian behavior. In Figs. 4 and 5 we plot this distribution and compare it with the histogram obtained by simulations. The analysis shows that the coefficients on the r.h.s. of Eq. 133, while losing the character of elements of a Markovian statistical matrix, still represent rates of (right or left) jumps to nearest neighbor even sites: the coarse-grained trajectories can be generated by allowing the particle to jump on nearest neighbor even sites with probabilities proportional to r^2 and l^2, after waiting times extracted from the distribution discussed above.

Based on Eq. 133, putting $Q'(S',t) = \Sigma_i P_i(S',t)$ one can write a differential equation of the second order in time for $G'(\lambda,t) = \sum_{S'} e^{\lambda S'}Q'(S',t)$, the generating function for the entropy production rate along coarse-grained trajectories (see SI):

$$\left[\partial_t^2 + 2(r+l)\partial_t + r^2 + l^2\right]G'(\lambda,t) = \\ = \left[r^2 e^{2\lambda\log(r/l)} + l^2 e^{-2\lambda\log(r/l)}\right]G'(\lambda,t) \tag{134}$$

From the long t behavior of G' we extract eventually ε' for the coarse-grained entropy production and find $\varepsilon'(\lambda,r,l) = \varepsilon\,(\lambda,r,l)$. Thus, the Cramer's function remains the same as that of the original process satisfying the fluctuation theorem: $\varepsilon'(\lambda-1,r,l) = \varepsilon'(-\lambda,r,l)$.

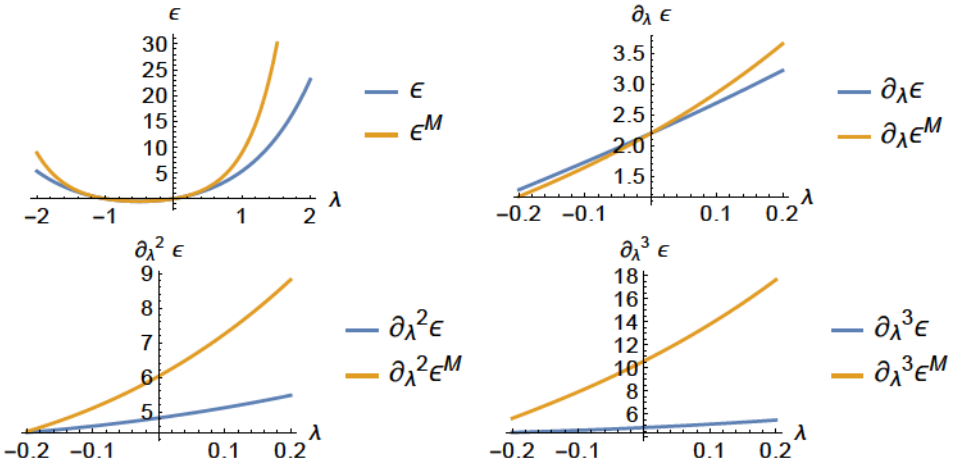


FIGURE 3. Comparison between the scaled cumulants generating functions ε and ε^M of the model described by Eq. 128 and its coarse grained and markovianized version, for $r = 3$ and $l = 1$. Values of the derivatives in $\lambda = 0$ represent the value of scaled cumulants.

Eqs. 133 and 134, being exact, suggest to enforce Markovianity of the coarse-grained dynamics by just dropping from them the terms with second order time derivative terms. The result is an approximate SCGF $\varepsilon^M(\lambda, r, l) = \varepsilon(\lambda, \frac{r^2}{2(r+l)}, \frac{l^2}{2(r+l)})$, which leaves invariant the average rate $\sigma = (r - l)\log(r/l)$, while producing larger higher moments (Fig. 3). This is the effect of replacing the proper distribution of waiting times by a Markovian (exponential) one consistent with the coarse grained transition rates. The invariance of the average σ is guaranteed by the fact that the waiting time distribution with memory has the same mean as the exponential one of the approximate Markovian process.

We can also imagine that in the linear network nearest neighbor links with fast transition rates (R and L) alternate with links with slow rates (r and l, $r \ll R$, $l \ll L$). Our decimation in which only odd or even sites survive is a way of realizing exactly a merging of each cluster of two states connected by fast rates, into a single state, the type of coarse-graining performed approximately in Ref. [**20**]. Computing along lines similar to those above the SCGF of the coarse-grained model with memory we find that the Cramer's function remains invariant also in this case. We can also coarse-grain networks with higher coordination. An example is the triangular one mapped onto hexagonal, based on a sort of star-triangle relation. Adopting torus topology for periodicity, we can apply, e.g., a drift along one of the circumferences of the torus. Also in these cases proper account of the memory effects leads to equations which yield invariant SCGF's. All these scenarios will be treated extensively in the sections that follow.

2.1. Simulations. We can simulate the original process using the Gillespie algorithm [**93**], which we can summarize in the following scheme:

$$\begin{cases} t_i &= t_{i-1} + \zeta_i \\ x_i &= x_{i-1} + \eta_i(r, l) \end{cases} \tag{135}$$

where

$$\zeta_i \sim \mathrm{Exp}(\tau) = \frac{1}{\tau} e^{-t/\tau} \ , \ \tau = \frac{1}{r + l} \tag{136}$$

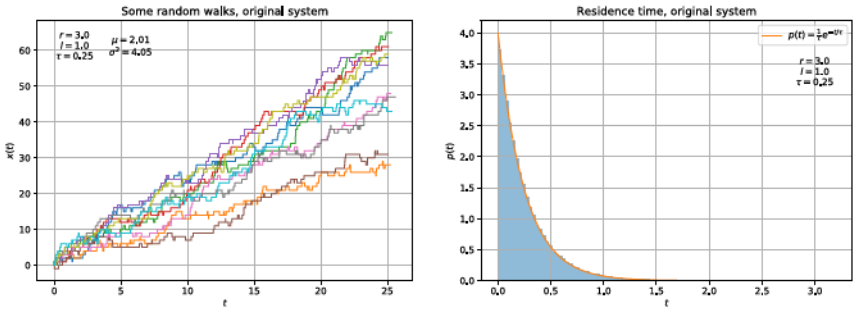


FIGURE 4. Some random walks and distribution of residence times of the original process. The parameters used are: $r = 3$, $l = 1$ and the distance between sites a is set equal to the unit.

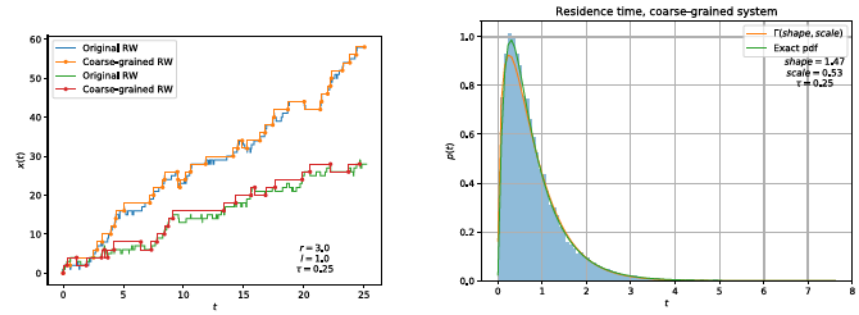


FIGURE 5. Some random walks and distribution of residence times of the coarse-grained process associated to the original one illustrated in Fig. 4. In the plot of the residence times distribution two analytic curves are shown: the green one is the exact distribution derived in Eq. 148. The orange one is a Gamma distribution with parameters (shape and scale) obtained from mean and variance of data. It's an approximation, but it's quite good.

and $\eta_i(r, l)$ is equal to +1 with probability $p = r/(r + l)$ and -1 with probability $q = l/(r + l)$. In Fig. 4 we see an example of such process simulated with rates $r = 3$ and $l = 1$. We see that the scaled cumulants of the process are in agreement with the analytic predictions.

Now we can coarse grain such process by, for instance eliding all the odd sites. The resulting process is shown in Fig. 5. We immediately see that an important effect introduced by the memory effects (the $\alpha\partial_t^2$ term) can be easily seen by looking at the residence times distribution, which is obviously far from an exponential one. Another indicator that the process is not Markovian comes also from the cumulants: if we were given directly the coarse-grained process and we evaluated the scaled cumulants we would see that they are not consistent with that of a Markov process (which have the form $a^n(r + (-1)^n l)$).

Now we derive the distribution of the residence times (RT) of the coarse-grained process (CGP). The characteristic function associated with an exponential probability distribution $p(T = t, \tau)$ is:

$$\varphi_T(\omega) = (1 - i\tau\omega)^{-1} \tag{137}$$

The RT of the CGP is a sum of a random number of exponentially distributed times. If, for instance, one of the rates is zero, we would find that this random number is fixed to 2, because we always proceed in one direction. In this case we would get that the characteristic function of the new RT is

$$\varphi_{T=T_1+T_2}(\omega) = \varphi_{T_1}(\omega) * \varphi_{T=T_2}(\omega) = (1 - i\tau\omega)^{-2} = \Gamma(2,\tau) \tag{138}$$

where Γ represents the Gamma distribution. Unfortunately in the general case the number of steps needed to perform a coarse-grained step may vary. We need to understand how.

We'll call a "successful" event one in which two consequent steps in the same direction happen, which would result in a coarse-grained step. We call it p. An unsuccessful one consists in going opposite directions in two consequent steps. We call it q. The probability these two event happen is:

$$p = \frac{r^2 + l^2}{(r+l)^2} \ , \ q = \frac{2rl}{(r+l)^2} \tag{139}$$

Therefore the characteristic function of the CG RT is:

$$\varphi_T(\omega) = p\varphi_{T_{1,2}}(\omega)\varphi_{T2,2}(\omega) + \tag{140}$$

$$+ pq\varphi_{T_{1,4}}(\omega)\varphi_{T2,4}\varphi_{T_{3,4}}(\omega)\varphi_{T4,4}(\omega) + \tag{141}$$

$$+ pq^2\varphi_{T_{1,6}}(\omega)\varphi_{T2,6}\varphi_{T_{3,6}}(\omega)\varphi_{T4,6}(\omega)\varphi_{T_{5,6}}(\omega)\varphi_{T6,6}(\omega) + ... \tag{142}$$

$$.... + pq^n \prod_{k=1}^{n+2} \varphi_{T_{k,n+2}}(\omega) + ... \tag{143}$$

where we have $\sum_{k=1}^{n+2} T_{k,n+2} = T$ for every $n \in \mathbb{N}$. Since all the functions are the same (the $T_{k,n+2}$ it's just a reminder for the associated pdf) we can call them simply $\varphi(\omega)$ and write:

$$\varphi_T(\omega) = p\varphi(\omega)^2 \sum_{k=0}^{\infty} \left(q\varphi(\omega)^2\right)^k \tag{144}$$

$$= \frac{p\varphi(\omega)^2}{1 - q\varphi(\omega)^2} \tag{145}$$

$$= \frac{p(1 - i\tau\omega)^{-2}}{1 - q(1 - i\tau\omega)^{-2}} \tag{146}$$

Finally we can find the (unique) probability function associated with the characteristic one with an inverse Fourier transform

$$p(t = T) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\omega t} \varphi_T(\omega) \ dt \tag{147}$$

and we finally find (we also write p and q in terms of r and l):

$$p(t, r, l) = \frac{l^2 + r^2}{2\sqrt{2lr}} e^{-x(\sqrt{2lr}+l+r)} \left(e^{2x\sqrt{2lr}} - 1\right) \tag{148}$$

In Fig. 5 we see that the simulations coincide with this analytic derivation. We want also to stress the fact that this procedure can be used to simulate straightforwardly a process with memory: we just need to follow the scheme of the Gillespie algorithm with the exception that now we extract the residence times from this distribution.

Another thing we can do is evaluate the CG rates from the CG trajectories. If we count the total number of right (and left) jumps and divide them by the total evolution time we obtain a value in complete agreement with the expected values $r^2/2(r+l)$ and $l^2/2(r+l)$. So in the end we are entitled to call these rates as well as those of the original model.

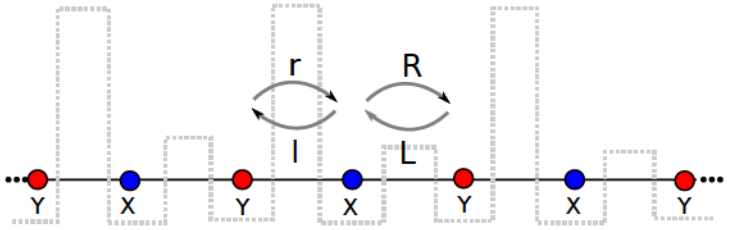


FIGURE 6. Sketch of the jump model described by Eq. 149. The gray line in the background depicts the barriers that cause the difference between the two pairs of rates (here we have $R, L \gg r, l$).

3. Clusters scenario

In literature [20, 21] one finds works dealing with coarse graining of systems with fast and slow sites. Here we study the simplest model possible involving one fast and one slow site alternating each other over a periodic lattice. This is achieved through the use of two pairs of small and big rates respectively (rates are inversely proportional to the time spent in the site).

3.1. Setup. We are dealing with a one-dimensional lattice as before, only that now we have two pair of rates (R, l) and (r, L) characterizing odd and even sites respectively (see Fig. 6). If we assume that $R \gg r$ and $L \gg l$ we'll have that transitions between $2k+1$ and $2k+2$ happen much more frequently than transitions between $2k$ and $2k+1$. The system has therefore 2 states (odds and even), and the probability density will have to satisfy the following equations:

(149)
$$\begin{cases} \partial_t P_i^X(S,t) &= rP_{i-1}^Y(S-\log(r/l),t) + LP_{i+1}^Y(S+\log(R/L),t) - (R+l)P_i^X(S,t) \\ \partial_t P_i^Y(S,t) &= RP_{i-1}^X(S-\log(R/L),t) + lP_{i+1}^X(S+\log(r/l),t) - (r+L)P_i^Y(S,t) \end{cases}$$

3.2. Statistic. In order to find the statistic associated with this 2-state markovian model we build the transition rate matrix:

(150)
$$\mathbb{W} = \begin{pmatrix} -(R+l) & r+L \\ R+l & -(r+L) \end{pmatrix}$$

The associated tilted matrix is:

(151)
$$\mathbb{W}_\lambda = \begin{pmatrix} -(R+l) & re^{\lambda\log(r/l)} + Le^{-\lambda\log(R/L)} \\ Re^{\lambda\log(R/L)} + le^{-\lambda\log(r/l)} & -(r+L) \end{pmatrix}$$

The SCGF coincides with the largest eigenvalue of this tilted matrix. It has the form:

(152)
$$\varepsilon(\lambda) = \frac{1}{2}\left(L+R+l+r - \sqrt{(L+R+l+r)^2 + 4(1-e^{-\lambda\log\frac{Rr}{Ll}})(Rre^{+\lambda\log\frac{Rr}{Ll}} - Ll)}\right)$$

Deriving in $\lambda = 0$ provides us with the scaled cumulants for the entropy production:

(153)
$$\sigma = \frac{\langle S\rangle}{t} = \frac{Rr - Ll}{R+L+r+l}\log\frac{Rr}{Ll}$$

(154)
$$\frac{\mathrm{Var}[S]}{t} = \left[\frac{Rr+Ll}{R+L+r+l} - 2\frac{(Rr+Ll)^2}{(R+L+r+l)^3}\right]\left(\log\frac{Rr}{Ll}\right)^2$$

which coincide with those of the original model for $R = r$ and $L = l$.

3.3. Coarse-Graining. In the Fourier space the system of the P's becomes:

$$\begin{cases} (R+l-i\omega)\tilde{P}_i^X(S,t) &= r\tilde{P}_{i-1}^Y(S-1,t) + L\tilde{P}_{i+1}^Y(S+1,t) \\ (r+L-i\omega)\tilde{P}_i^Y(S,t) &= R\tilde{P}_{i-1}^X(S-1,t) + l\tilde{P}_{i+1}^X(S+1,t) \end{cases} \tag{155}$$

We can exploit the fact that the sites alternate each other to substitute the second equation (evaluated in $i \pm 1$) in the right members of the first equation. Going back to the original time-space we obtain the following equation for both the odd and even sites (we drop the $\cdot^{X,Y}$ superscript):

$$\left[\frac{1}{R+L+r+l}\partial_t^2 + \partial_t\right] P_i(S,t) = \frac{Rr}{R+L+r+l}P_{i-2}(S - \log\frac{Rr}{Ll}, t) + \\ + \frac{Ll}{R+L+r+l}P_{i+2}(S + \log\frac{Rr}{Ll}, t) - \frac{Rr+Ll}{R+L+r+l}P_i(S,t) \tag{156}$$

We can exploit the results obtained in the 1-state case to extract the SCGF provided by Eq. 156. We find that that the following polynomial equation must hold:

$$\varepsilon'(\lambda)^2 + (R+L+r+l)\varepsilon'(\lambda) = Rre^{\lambda \log\frac{Rr}{Ll}} + Lle^{-\lambda\frac{Rr}{Ll}} - Rr + Ll \tag{157}$$

The dominant solution is equivalent to the original SCGF of Eq. 152, therefore also in this case we find $\varepsilon'(\lambda) = \varepsilon(\lambda)$.

3.4. Markovianization. One could try to quantify the errors introduced by neglecting the memory term in the coarse grained model. If we omit the 2nd derivative in 156 and map $i \pm 2 \to i \pm 1$ we obtain the ordinary ME:

$$\partial_t P_i(S,t) = \frac{Rr}{R+L+r+l}P_{i-1}(S - \log\frac{Rr}{Ll}, t) + \\ + \frac{Ll}{R+L+r+l}P_{i+1}(S + \log\frac{Rr}{Ll}, t) - \frac{Rr+Ll}{R+L+r+l}P_i(S,t) \tag{158}$$

which is associated with the following SCGF for the entropy production:

$$\varepsilon^M(\lambda) = \frac{Rr}{R+L+r+l}e^{\lambda\log\frac{Rr}{Ll}} + \frac{Ll}{R+L+r+l}e^{-\lambda\log\frac{Rr}{Ll}} - \frac{Rr+Ll}{R+L+r+l} \tag{159}$$

The statistics associated with this process is:

$$\sigma^M = \frac{\langle S^M\rangle}{t} = \frac{Rr-Ll}{R+L+r+l}\log\frac{Rr}{Ll} \tag{160}$$

$$\frac{\mathrm{Var}[S^M]}{t} = \frac{Rr+Ll}{R+L+r+l}\left(\log\frac{Rr}{Ll}\right)^2 \tag{161}$$

So the error introduced by this approximation affects the variance, and can be measured as the difference between the two:

$$\Delta = \frac{\mathrm{Var}[S] - \mathrm{Var}[S^M]}{t} = -2\frac{(Rr-Ll)^2}{(R+L+r+l)^3}\left(\log\frac{Rr}{Ll}\right)^2 \tag{162}$$

One thing we notice is that this difference is always negative, so neglecting memory effects has always as consequence the observation of a higher scaled variance ($\Delta < 0$). Now let's assume that the rates take the form $R, L = Ce^{\pm af/2}$ and $r, l = ce^{\pm af/2}$, so they all satisfy LDB. Let's also assume that we are in a regime in which $C \gg c$. One finds that the correction in this regime is:

$$\Delta = -8a^2f^2\frac{C^2c^2(e^{af}-e^{-af})^2}{(C+c)^3(e^{af/2}+e^{-af/2})^3} \propto \frac{c^2}{C} \tag{163}$$

Therefore it is one order smaller with respect to C compared to $\mathrm{Var}[S^M]/t \propto c$. We can conclude that in this limit the correction provided by keeping into account

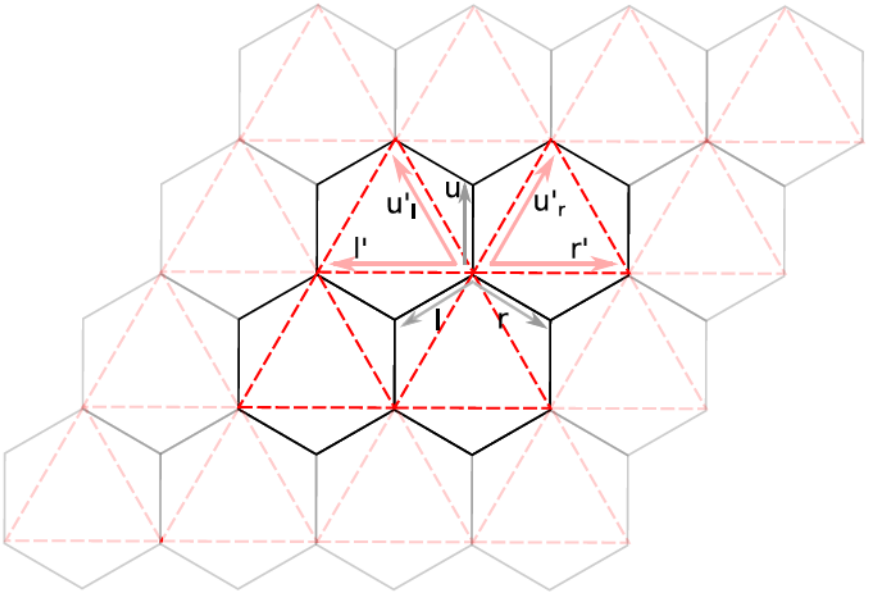


FIGURE 7. Sketch for the tar triangle model. Black lines represent the initial hexagonal structure (with size of length a), while the red lines the resulting one after the decimation process. Periodicity is assumed in both horizontal and vertical directions.

memory is negligible, on the other hand it when $C = c$ it assumes the highest value possible.

4. Star-Triangle

This is a 2-dimensional model in which we are able to decimate sites without introducing additional hopping neighbors. We start from a regular hexagonal 2D lattice (composed of 2 different states) with sides of length a as in Fi. 7, in which we decimate alternatively the sites of each hexagon. The resulting lattice consists of a 1-state model in which each site has 6 neighbors equally distant and uniformly distributed around it (a regular triangular lattice). We write the equations for the most general case, in which there are 2 rates for each direction: $u$up and $d$own for vertical jumps, $r_{u,d}$ for diagonal hops to the right and $l_{u,d}$ to the left. In this way one is therefore allowed to introduce biases in both horizontal and vertical directions.

4.1. Setup. The master equation associated with the original hexagonal model is the following:

(164)
$$\begin{cases} [r_d + l_d + u + \partial_t]\, P^X_{i,j}(S,t) & = r_u P^Y_{i-1,j}(S - \log\frac{r_u}{l_d}, t) + l_u P^Y_{i+1,j}(S - \log\frac{l_u}{r_d}, t) + d P^Y_{i,j+1}(S - \log\frac{d}{u}, t) \\ [r_u + l_u + d + \partial_t]\, P^Y_{i,j}(S,t) & = r_d P^X_{i-1,j}(S - \log\frac{r_d}{l_u}, t) + l_d P^X_{i+1,j}(S - \log\frac{l_d}{r_u}, t) + u P^X_{i,j-1}(S - \log\frac{u}{d}, t) \end{cases}$$

This case is similar to that of the cluster model. The transition matrix associated with this process is:

(165)
$$\mathbb{W} = \begin{pmatrix} -(r_d + l_d + u) & r_u + l_u + d \\ r_d + l_d + u & -(r_u + l_u + d) \end{pmatrix}$$

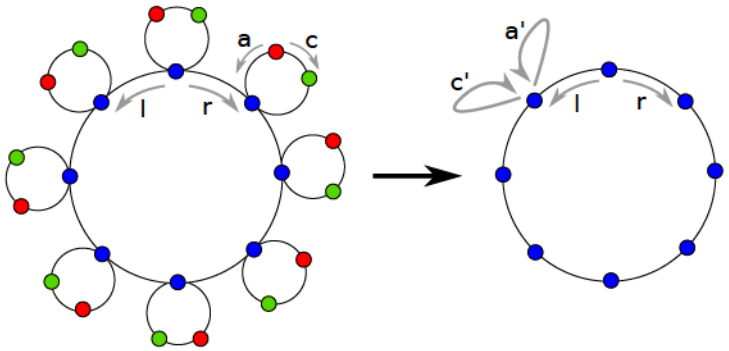


FIGURE 8. Decimation of secondary loops in the jump model described by Eq. 169. Only the X sites (blue color) composing the system survive the procedure. The resulting system is characterized by self-loops of transition rates $a' = \frac{a^3}{3(a^2+ac+c^2)}$ and $c' = \frac{c^3}{3(a^2+ac+c^2)}$.

while associated tilted matrix is:

(166)
$$\mathbb{W}_\lambda = \begin{pmatrix} -(r_d + l_d + u) & r_u e^{\lambda \log \frac{r_u}{l_d}} + l_u e^{\lambda \log \frac{l_u}{r_d}} + d e^{\lambda \log \frac{d}{u}} \\ r_d e^{\lambda \log \frac{r_d}{l_u}} + l_d e^{\lambda \log \frac{l_d}{r_u}} + u e^{\lambda \log \frac{u}{d}} & -(r_u + l_u + d) \end{pmatrix}$$

The SCGF coincides with the largest eigenvalue of the tilted matrix. It can be computated easly but has a complex form. A check shows that in the case $d = u$, $r_d = r_u = r$ and $l_d = l_u = l$ it has the form:

(167)
$$\varepsilon(\lambda, r, l) = r e^{\lambda \log r/l} + l e^{-\lambda \log(r/l)} - (r + l)$$

as expected.

4.2. Coarse graining. For the coarse graining we can proceed analogously to the cluster model and reduce the system to a single equation:

(168)
$$\begin{aligned} \left[r_u(r_d + u) + l_u(l_d + u) + d(r_d + l_d) + (r_u + l_u + r_d + l_d + u + d)\partial_t - \partial_t^2\right] P^X_{i,j}(S,t) = \\ = r_u[r_d P^X_{i-2,j}(S - \log \frac{r_d r_u}{l_u l_d}, t) + u P^X_{i-1,j-1}(S - \log \frac{u r_u}{d l_d}, t)] + \\ + l_u[l_d P^X_{i+2,j}(S - \log \frac{l_d l_u}{r_u r_d}, t) + u P^X_{i+1,j-1}(S - \log \frac{u l_u}{d r_d}, t)] + \\ + d[r_d P^X_{i-1,j+1}(S - \log \frac{d r_d}{u l_u}, t) + l_d P^X_{i+1,j+1}(S - \log \frac{d l_d}{u r_u}, t)] \end{aligned}$$

The SCGf associated with this coarse-grained model is found to coincide with the one of the original model.

5. Secondary loops

A fundamental problem left is to study situations in which there is more than one loop with circulating current in the network. In the presence of secondary loops it has been suggested that if coarse-graining removes them, entropy production could be altered. This seems plausible if one regards also the coarse-grained dynamics as Markovian. Indeed, in Markov jump processes the entropy production can be directly ascribed to the current flowing in all the loops. Below we show that, if memory effects are exactly taken into account invariance of the SCGF still holds.

5.1. Setup. Let us consider the network in Fig. 8. It consists of a main chain of N (X) states analogous to that in Eq. 128, with right/left jump rates r/l. Secondary 3-state loops are further attached to every X state at position i along the chain, by connecting it to a Y and a Z state. Rates c and a apply to clockwise and anticlockwise transitions within the loops, respectively. The states Y and Z of each loop i are those we want to decimate, leaving only the X states. Before decimation, upon summing over the index i specifying different X, Y, and Z states, equations equivalent to Eq. 1 above, one gets:

$$\begin{cases} [r+l+a+c+\partial_t]P_X(S,t) & = rP_X(S-\log\frac{r}{l},t)+lP_X(S-\log\frac{l}{r},t)+ \\ & \quad +cP_Y(S-\log\frac{c}{a},t)+aP_Z(,S-\log\frac{a}{c},t) \\ [a+c+\partial_t]P_Y(S,t) & = aP_X(S-\log\frac{a}{c},t)+cP_Z(S-\log\frac{c}{a},t) \\ [a+c+\partial_t]P_Z(S,t) & = cP_X(S-\log\frac{c}{a},t)+aP_Z(S-\log\frac{a}{c},t) \end{cases} \tag{169}$$

where, e.g., $P_X(S,t)=\Sigma_i P_{X,i}(S,t)$ is the probability that a trajectory ends at time t at a generic X state with total entropy production S. The SCGF for the entropy production $\varepsilon(\lambda,r,l,c,a)$ of this system can be found by evaluating the dominant eigenvalue of the tilted transition matrix $\mathbb{W}_\lambda$, whose elements are defined by:

$$(\mathbb{W}_\lambda)_{ij}=\sum_\rho W_{ij}^\rho e^{\lambda\log\frac{W_{ij}^\rho}{W_{ji}^\rho}} \tag{170}$$

where i,j run over X, Y, and Z, and W_{ij}^ρ is a rate of transition $j\to i$, with ρ specifying the type of transition (r, l, a or c in the different cases) (SI). By deriving the SCGF in $\lambda=0$ we obtain the average entropy production rate:

$$\sigma=\frac{r-l}{3}\log\frac{r}{l}+(c-a)\log\frac{a}{c} \tag{171}$$

The first terms represent the contribution given by jumps on the main chain, while the second one is relative to the transitions occurring within the secondary loops.

5.2. Decimation. Now we want to decimate the vertices outside the main loop, e.g the Y and Z kind of sites, therefore we exploit the last two equation in order to find an expression of P_Y and P_Z in terms of just P_X. Going in the Fourier space and setting $\alpha=a/(i\omega+a+c)$ and $\gamma=c/(i\omega+a+c)$ we get the system:

$$\begin{cases} P_Y(I,S,\omega) & = \alpha P_X(I,S-\log\frac{a}{c},\omega)+\gamma P_Z(I,S-\log\frac{c}{a},\omega) \\ P_Z(I,S,\omega) & = \gamma P_X(I,S-\log\frac{c}{a},\omega)+\alpha P_Y(I,S-\log\frac{a}{c},\omega) \end{cases} \tag{172}$$

which solved for the S values that appear in the first equation becomes:

$$\begin{cases} P_Y(I,S-\log\frac{c}{a},\omega) & = \frac{\gamma}{1-\alpha\gamma}\left[\alpha P_X(I,S,t)+\gamma P_X(I,S-3\log\frac{c}{a},t)\right]= \\ & = \frac{a(i\omega+a+c)P_X(i,\sigma,\omega)+c^2P_X(i,\sigma-3\log\frac{c}{a},\omega)}{-\omega^2+2i\omega(a+c)+(a^2+ac+c^2)} \\ P_Z(I,S-\log\frac{a}{c},\omega) & = \frac{\alpha}{1-\alpha\gamma}\left[\alpha P_X(I,S-3\log\frac{a}{c},\omega)+\gamma P_X(I,S,\omega)\right]= \\ & = \frac{a^2P_X(I,S-3\log\frac{a}{c},\omega)+c(i\omega+a+c)P_X(I,S,\omega)}{-\omega^2+2i\omega(a+c)+(a^2+ac+c^2)} \end{cases} \tag{173}$$

Upon substituting in the original system and antitransforming we finally find a third-order differential equations in time for $P_X(S',t)$, which plays here the same

role as $Q'(S',t)$ in the linear chain case:

(174)
$$\begin{aligned}\left[\left(3(a^2+ac+c^2)+2(r+l)(a+c)\right)\partial_t+(3(a+c)+r+l)\partial_t^2+\partial_t^3\right]P_X(I,S,t)=\\ =\left[r(a^2+ac+c^2)+2r(a+c)\partial_t+r\partial_t^2\right]P_X(I-1,S-\log\frac{r}{l},t)+\\ +\left[l(a^2+ac+c^2)+2l(a+c)\partial_t+l\partial_t^2\right]P_X(I+1,S-\log\frac{l}{r},t)+\\ +c^3P_X(I,S'-3\log\frac{c}{a},t)+a^3P_X(I,S-3\log\frac{a}{c},t)+\\ -\left[a^3+c^3+(r+l)(a^2+ac+c^2)\right]P_X(I,S,t)\end{aligned}$$

Here if we sum over all the entropies $S\in[-\infty,+\infty]$ we find the equation for the probability of the total displacement $P_X(I,t)$: here we see that the terms a^3 and c^3 will disappear, because they contribute only to the entropy and not to the displacement. For the generating function $G'(\lambda,t)=\Sigma_S e^{\lambda S'}P_X(S,t)$ of the coarse-grained network we eventually obtain:

$$\left[\alpha+\beta\partial_t+\gamma\partial_t^2+\partial_t^3\right]G'(\lambda,t)=0 \tag{175}$$

where α, β and γ are functions of the jump rates and of λ

(176)
$$\begin{aligned}\alpha&=r(a^2+ac+c^2)e^{\lambda\log r/l}+l(a^2+ac+c^2)e^{-\lambda\log r/l}-(r+l)(a^2+ac+c^2)\\ &\quad+a^3+c^3+a^3e^{3\lambda\log a/c}+c^3e^{-3\lambda\log a/c}\\ \beta&=2r(a+c)e^{\lambda\log r/l}+2l(a+c)e^{-\lambda\log r/l}-\left(3(a^2+ac+c^2)+2(r+l)(a+c)\right)\\ \gamma&=re^{\lambda\log r/l}+le^{-\lambda\log r/l}-(3(a+c)+r+l)\end{aligned}$$

In the limit $t\to\infty$ we get $G'\sim e^{t\varepsilon'}$, where ε' is the dominant root of the polynomial characteristic equation associated with Eq. 175. Remarkably ε' coincides with the SCGF $\varepsilon(\lambda,r,l,c,a)$ of the original process. So, also in this case $\varepsilon'=\varepsilon$ (SI), maintaining validity of the fluctuation theorem. We remark that in this example the presence of secondary loops leads to terms in the equation for G' which represent entropy gains originating from P_X's computed in values of S which are not shifted: these contributions take into account what the secondary loops were producing in the original, full description. Such terms were not present in the linear network case and are precisely those which allow to abtain $\varepsilon'=\varepsilon$ in the presence of secondary loops. In the linear chain example only loss terms associated to Q''s in unshifted S were present.

5.3. Markovianization. Also in this example the invariance of the entropy production fluctuations is clearly guaranteed by memory terms that make Eq. 175 incompatible with a Markov jump description. At variance with the linear chain case, here we have not only higher order time derivatives, but also first order time-derivatives of probabilities evaluated in shifted values of produced entropy. Indeed, the full expression of the coefficient β presented in Eq. 176 is:

$$\begin{aligned}\beta\quad&=3(a^2+ac+c^2)+2(r+l)(a+c)+\\ &\quad-2(a+c)\left(re^{\lambda\log r/l}+le^{-\lambda\log r/l}\right)\end{aligned} \tag{177}$$

where the second line corresponds to terms $\partial_t Q(S\pm\log r/l,t)$ in the equation for the probability associated to Eq.175. If, like in the linear chain case, we want to reduce to Markov jump the coarse-grained process, these terms must not be dropped altogether, but just kept at unshiftet S values. In this way, which amounts to disregard only discrete gradient contributions in S, the average entropy production rate σ remains invariant. Such procedure will be presented in detailed in the

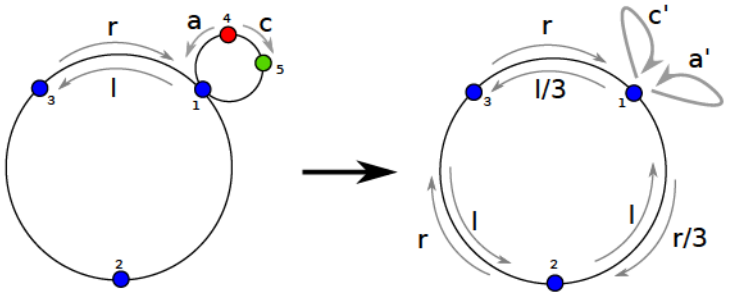


FIGURE 9. Sketch of the jump model described by Eq. 179. The coarse grainining procedure performed on the system leaves only the sites 1,2 and 3 (blue color) composing the main loop. In the resulting system the rates concerning the site 1 change because of the new steady state probability, and moreover additional self-loops of transition rates $a' = \frac{a^3}{3(a^2+ac+c^2)}$ and $c' = \frac{c^3}{3(a^2+ac+c^2)}$ arise always for this site.

following section with the aid of two emblematic examples here we anticipate, the resulting Master Equation of such procedure for this specific model: The resulting Master Equation will be:

$$\begin{aligned}(178)\quad &\left[\frac{a^3+c^3}{3(a^2+ac+c^2)} + (r+l) + \partial_t\right] P_X(I,S,t) = \\ &= rP_X(I-1, S-\log\frac{r}{l}, t) + lP_X(I+1, S-\log\frac{l}{r}, t) + \\ &+ \frac{c^3}{3(a^2+ac+c^2)} P_X(I, S-3\log\frac{c}{a}, t) + \\ &+ \frac{a^3}{3(a^2+ac+c^2)} P_X(I, S-3\log\frac{a}{c}, t) +\end{aligned}$$

6. Coarse graining of general markovian networks

The work done so far addressed cases in which the symmetry of the problem allowed a straightforward analytic approach. But what happens when we have to deal with a general markovian network? The novel ingredient that we must introduce to deal with this cases is the re-normalization of probabilities after the decimation procedure: if we decimate some states, their probabilities must be redistributed to the closest neighboring sites that survived the decimation. Note that this should have be done also in the symmetrical models, but in these cases the renormalization coefficient is the same for all the probabilities, therefore such effect is not visible in the end. Below we will study two representative cases.

6.1. One main loop - one secondary loop. We build a model similar to that of the secondary loops presented in the previous section, with the difference being in the fact that we use only 5 sites: the three sites 1,2,3 compose the main loop, while the secondary loop shares the site number 1 with the main loop and is therefore composed of the sites 1,4,5 (see Fig. 10). The rates in the main loop are r and l, while in the secondary loop we have a and c. The equation for the original

system is:

$$
(179)\quad \begin{cases} [r+l+a+c+\partial_t]P_1(I,S,t) & = rP_3(I-1,S-\log\frac{r}{l},t)+lP_2(I+1,S-\log\frac{l}{r},t)+ \\ & \quad +cP_5(I,S-\log\frac{c}{a},t)+aP_4(i,I,S-\log\frac{a}{c},t) \\ [r+l+\partial_t]P_2(I,S,t) & = rP_1(I-1,S-\log\frac{r}{l},t)+lP_3(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_3(I,S,t) & = rP_2(I-1,S-\log\frac{r}{l},t)+lP_1(I+1,S-\log\frac{l}{r},t) \\ [a+c+\partial_t]P_4(I,S,t) & = aP_5(I,S-\log\frac{a}{c},t)+cP_1(I,S-\log\frac{c}{a},t) \\ [a+c+\partial_t]P_5(I,S,t) & = cP_4(I,S-\log\frac{c}{a},t)+aP_1(I,S-\log\frac{a}{c},t) \end{cases}
$$

Evaluating the SS shows that all sites are equivalent, with $P_i = 1/5 \ \forall \ i$. The average entropy production is:

$$
(180)\qquad \sigma = \frac{3}{5}\left[(r-l)\log(r/l)+(a-c)\log(a/c)\right]
$$

If we decimate the sites 4 and 5 we eliminate the secondary loop of the system. The equations describing this decimated system are:

(181)

$$
\begin{cases} \left[\left(3(a^2+ac+c^2)+2(r+l)(a+c)\right)\partial_t+(3(a+c)+r+l)\partial_t^2+\partial_t^3\right]P_1(I,S,t) = \\ \qquad = \left[r(a^2+ac+c^2)+2r(a+c)\partial_t+r\partial_t^2\right]P_3(I-1,S-\log\frac{r}{l},t)+ \\ \qquad +\left[l(a^2+ac+c^2)+2l(a+c)\partial_t+l\partial_t^2\right]P_2(I+1,S-\log\frac{l}{r},t)+ \\ \qquad +c^3P_1(I,S-3\log\frac{c}{a},t)+a^3P_1(I,S-3\log\frac{a}{c},t)+ \\ \qquad -\left[a^3+c^3+(r+l)(a^2+ac+c^2)\right]P_1(I,S,t) \\ [r+l+\partial_t]P_2(I,S,t) = rP_1(I-1,S-\log\frac{r}{l},t)+lP_3(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_3(I,S,t) = rP_2(I-1,S-\log\frac{r}{l},t)+lP_1(I+1,S-\log\frac{l}{r},t) \\ (-\partial_t^2+2\partial_t(a+c)+(a^2+ac+c^2))P_5(I,S-\log\frac{c}{a},t) = a(\partial_t+a+c)P_1(i,S,t)+c^2P_1(i,S-3\log\frac{c}{a},t) \\ (-\partial_t^2+2\partial_t(a+c)+(a^2+ac+c^2))P_4(I,S-\log\frac{a}{c},t) = a^2P_1(I,S-3\log\frac{a}{c},t)+c(\partial_t+a+c)P_1(I,S,t) \end{cases}
$$

Now we can drop the states 4 and 5 and reduce the system. In this coarse grained system the SS probabilities of the states 2 and 3 stay unaltered, while the state 1 will "absorb" those of states 4 and 5. Therefore we will have:

$$
(182)\qquad P_1^{CG} = 3P_1
$$

$$
(183)\qquad P_2^{CG} = P_2
$$

$$
(184)\qquad P_3^{CG} = P_3
$$

How can we derive this result in a formal way? We can say that decimating 4 and 5 is equivalent to force the associated probabilities P_4 and P_5 to be zero. If we perform this operation, we are now left with probabilities that do not add up to one. That's because when we decimate a site, its probability is redistributed to the sites connected to him with a coefficient that is proportional to the exiting rate towards such site. This is can be explained in terms of transitional probability. If we decimate the site j, the probability of being in the state i in the CG model is $P_i + \pi_{ij}P_j$, where $\pi_{ij} = W_{ij}/\sum_k W_{kj}$ is the transitional probability of going from state j to state i in the original model. Formally:

$$
(185)\qquad P_i^{CG} = P_i + \pi_{ij}P_j \sim_{\text{SS}} P_i + \pi_{ij}\frac{P_j^{\text{stat}}}{P_i^{\text{stat}}}P_i = P_i\left(1+\pi_{ij}\frac{P_j^{\text{stat}}}{P_i^{\text{stat}}}\right)
$$

where the $\sim$ relation is justified because in the SS we need this equality to hold, and this is possible only if this relation is satisfied also throughout all the evolution, since the rates do not depend on time.

If we want to decimate two sites simultaneously (e.g. j and k) the formula then becomes:

$$P_i^{CG} = P_i + \pi_{ij}P_j + \pi_{ik}P_k + \pi_{kj}P_j + \pi_{jk}P_k \sim_{\rm SS} \tag{186}$$

$$\sim_{\rm SS} \frac{P_i}{P_i^{\rm stat}}\left(P_i^{\rm stat} + \pi_{ij}P_j^{\rm stat} + \pi_{ik}P_k^{\rm stat} + \pi_{kj}P_j^{\rm stat} + \pi_{jk}P_k^{\rm stat}\right) = \tag{187}$$

$$= \frac{P_i}{P_i^{\rm stat}}\left(P_i^{\rm stat} + (\pi_{ij}+\pi_{kj})P_j^{\rm stat} + (\pi_{ik}+\pi_{jk})P_k^{\rm stat}\right) \tag{188}$$

So in general if we want to eliminate every site with index $j \in J$ we have:

$$P_i^{CG} = \frac{P_i}{P_i^{\rm stat}}\left(P_i^{\rm stat} + \sum_{j\in J}\left(\pi_{ij} + \sum_{k\in J\ ,\ k\neq j}\pi_{kj}\right)P_j^{\rm stat}\right) \tag{189}$$

If we apply such procedure in this specific model we get by decimating simultaneously sites 4 and 5:

$$P_1^{CG} = P_1\left(1 + \frac{a+c}{a+c} + \frac{a+c}{a+c}\right) = 3P_1 \tag{190}$$

$$P_2^{CG} = P_2 \tag{191}$$

$$P_3^{CG} = P_3 \tag{192}$$

$$P_4^{CG} = 0 \tag{193}$$

$$P_5^{CG} = 0 \tag{194}$$

Imposing the condition that they sum to 1 we find the correct steady state. Substituting and dropping the CG superscript we find:

(195)
$$\begin{cases} \left[\left(3(a^2+ac+c^2)+2(r+l)(a+c)\right)\partial_t + (3(a+c)+r+l)\partial_t^2 + \partial_t^3\right]\frac{1}{3}P_1(I,S,t) = \\ \qquad = \left[r(a^2+ac+c^2) + 2r(a+c)\partial_t + r\partial_t^2\right]P_3(I-1, S-\log\frac{r}{l}, t) + \\ \qquad + \left[l(a^2+ac+c^2) + 2l(a+c)\partial_t + l\partial_t^2\right]P_2(I+1, S-\log\frac{l}{r}, t) + \\ \qquad + c^3\frac{1}{3}P_1(I, S-3\log\frac{c}{a}, t) + a^3\frac{1}{3}P_1(I, S-3\log\frac{a}{c}, t) + \\ \qquad - \left[a^3+c^3+(r+l)(a^2+ac+c^2)\right]\frac{1}{3}P_1(I,S,t) \\ [r+l+\partial_t]P_2(I,S,t) = \frac{r}{3}P_1(I-1, S-\log\frac{r}{l}, t) + lP_3(I+1, S-\log\frac{l}{r}, t) \\ [r+l+\partial_t]P_3(I,S,t) = rP_2(I-1, S-\log\frac{r}{l}, t) + \frac{l}{3}P_1(I+1, S-\log\frac{l}{r}, t) \end{cases}$$

In the SS we find $P_2 = P_3 = 1/5$ and $P_1 = 3/5$. We can Markovianize the system by neglecting higher order time derivatives and expanding the shifted ones. We find:

(196)
$$\begin{cases} \left[\frac{a^3+c^3}{3(a^2+ac+c^2)} + \frac{r+l}{3} + \partial_t\right]P_1(I,S,t) = rP_3(I-1, S-\log\frac{r}{l}, t) + lP_2(I+1, S-\log\frac{l}{r}, t) + \\ \qquad + \frac{1}{3(a^2+ac+c^2)}\left(c^3P_1(I, S-3\log\frac{c}{a}, t) + a^3P_1(I, S-3\log\frac{a}{c}, t)\right) \\ [r+l+\partial_t]P_2(I,S,t) = \frac{r}{3}P_1(I-1, S-\log\frac{r}{l}, t) + lP_3(I+1, S-\log\frac{l}{r}, t) \\ [r+l+\partial_t]P_3(I,S,t) = rP_2(I-1, S-\log\frac{r}{l}, t) + \frac{l}{3}P_1(I+1, S-\log\frac{l}{r}, t) \end{cases}$$

Evaluating the entropy production of this CGM system we find an agreement with the original one σ.

One thing we should notice here, is that since the sites are not equivalent in the SS anymore the entropy production contributions coming from transitions among these states do not have the form $\log(W_{ji}/W_{ij})$ but they keep into account the SS probability:

$$\Delta S = \log\frac{P_i^{SS}W_{ji}}{P_j^{SS}W_{ij}} \tag{197}$$

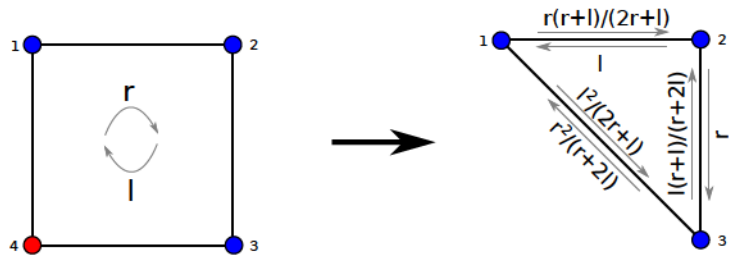


FIGURE 10. Sketch of the jump model described by Eq. 198. The coarse grainining procedure performed on the system decimates the 4th state leading to a triangle loop system with memory. The rates of the coarse grained system are relative to the process obtained with the markovianization procedure.

7. Square loop to triangle loop

In this scenario we are dealing with the simple case of a loop composed by four sites in which we want to decimate one single site, therefore obtaining a loop with a triangle shape (Fig. 10). This case is emblematic because it's a simple example in which the eliminated site is "redistributed" among two sites. If we assume that all the sites are connected with r and l rates, the original system is described by the following ME:

$$\begin{cases} [r+l+\partial_t]P_1(I,S,t) &= rP_4(I-1,S-\log\frac{r}{l},t)+lP_2(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_2(I,S,t) &= rP_1(I-1,S-\log\frac{r}{l},t)+lP_3(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_3(I,S,t) &= rP_2(I-1,S-\log\frac{r}{l},t)+lP_4(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_4(I,S,t) &= rP_3(I-1,S-\log\frac{r}{l},t)+lP_1(I+1,S-\log\frac{l}{r},t) \end{cases} \tag{198}$$

All sites are equivalent, so in the SS we find $P_i^{SS}=1/4$ for every i. The average entropy production is easily found to be $\sigma=(r-l)\log r/l$. Decimating the site number 4 with the usual procedure yields the following system:

$$\begin{cases} [r^2+rl+l^2+2(r+l)\partial_t+\partial_t^2]P_1(I,S,t) &= r^2P_3(I-2,S-2\log\frac{r}{l},t)+l(r+l+\partial_t)P_2(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_2(I,S,t) &= rP_1(I-1,S-\log\frac{r}{l},t)+lP_3(I+1,S-\log\frac{l}{r},t) \\ [r^2+rl+l^2+2(r+l)\partial_t+\partial_t^2]P_3(I,S,t) &= r(r+l+\partial_t)P_2(I-1,S-\log\frac{r}{l},t)+l^2P_3(I+2,S-2\log\frac{l}{r},t) \\ [r+l+\partial_t]P_4(I,S,t) &= rP_3(I-1,S-\log\frac{r}{l},t)+lP_1(I+1,S-\log\frac{l}{r},t) \end{cases} \tag{199}$$

Now if decimate definitively the site 4, we have to renormalize the probabilities as in the previous exercise.

$$P_1^{CG}=P_1\left(1+\frac{r}{r+l}\right) \tag{200}$$

$$P_2^{CG}=P_2 \tag{201}$$

$$P_3^{CG}=P_3\left(1+\frac{l}{r+l}\right) \tag{202}$$

$$P_4^{CG}=0 \tag{203}$$

Inverting these relations and substituting them in the system yields:

$$\begin{cases} [r^2+rl+l^2+2(r+l)\partial_t+\partial_t^2]\frac{r+l}{2r+l}P_1(I,S,t) &= r^2\frac{r+l}{r+2l}P_3(I-2,S-2\log\frac{r}{l},t)+l(r+l+\partial_t)P_2(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_2(I,S,t) &= r\frac{r+l}{2r+l}P_1(I-1,S-\log\frac{r}{l},t)+l\frac{r+l}{r+2l}P_3(I+1,S-\log\frac{l}{r},t) \\ [r^2+rl+l^2+2(r+l)\partial_t+\partial_t^2]\frac{r+l}{r+2l}P_3(I,S,t) &= r(r+l+\partial_t)P_2(I-1,S-\log\frac{r}{l},t)+l^2\frac{r+l}{2r+l}P_1(I+2,S-2\log\frac{l}{r},t) \end{cases} \tag{204}$$

Now we proceed with the markovianization of the system. We drop higher order derivatives and we expand shifted first order derivatives. Since now sites are

inequivalent, when we expand these shifted derivatives we need to take into account of the proportionality relations that hold between them in the SS. Namely:

$$\partial_t P_j \longrightarrow \frac{P_j^{SS}}{P_i^{SS}} \partial_t P_i \tag{205}$$

Performing these operations finally yields:

(206)
$$\begin{cases} [r^2+rl+l^2+2(r+l)\partial_t]\frac{1}{2r+l}P_1(I,S,t) & = r^2\frac{1}{r+2l}P_3(I-2,S-2\log\frac{r}{l},t)+lP_2(I+1,S-\log\frac{l}{r},t)+\frac{l}{r+2l}\partial_t P_1(I,S,t) \\ [r+l+\partial_t]P_2(I,S,t) & = r\frac{r+l}{2r+l}P_1(I-1,S-\log\frac{r}{l},t)+l\frac{r+l}{r+2l}P_3(I+1,S-\log\frac{l}{r},t) \\ [r^2+rl+l^2+2(r+l)\partial_t]\frac{1}{r+2l}P_3(I,S,t) & = rP_2(I-1,S-\log\frac{r}{l},t)+\frac{r}{r+2l}\partial_t P_3(I,S,t)+l^2\frac{1}{2r+l}P_1(I+2,S-2\log\frac{l}{r},t) \end{cases}$$

And upon rearranging we get:

(207)
$$\begin{cases} [r^2+rl+l^2+(2r+l)\partial_t]\frac{1}{2r+l}P_1(I,S,t) & = r^2\frac{1}{r+2l}P_3(I-2,S-2\log\frac{r}{l},t)+lP_2(I+1,S-\log\frac{l}{r},t) \\ [r+l+\partial_t]P_2(I,S,t) & = r\frac{r+l}{2r+l}P_1(I-1,S-\log\frac{r}{l},t)+l\frac{r+l}{r+2l}P_3(I+1,S-\log\frac{l}{r},t) \\ [r^2+rl+l^2+(r+2l)\partial_t]\frac{1}{r+2l}P_3(I,S,t) & = rP_2(I-1,S-\log\frac{r}{l},t)+l^2\frac{1}{2r+l}P_1(I+2,S-2\log\frac{l}{r},t) \end{cases}$$

A check shows that: the probability is conserved, the SS is the one expected, and finally that the average entropy production coincides with that of the original system, so it seems we did everything correctly.

8. Discussion

A key advantage of our apporoach is the fact that we eliminate states based on the exact equations for the entropy production: this is why we get rigorous control of the arising memory effects. The analytical results presented above, partly inspired by the renormalization group approach to critical phenomena, were relatively simplified by the choice of networks with a limited number of equivalent states. Our calculations of course become more complicated, when dealing with less symmetric networks and/or decimation patterns. Our coarse-graining strategy is based on the general idea of eliminating from the equations governing entropy production a number of network nodes, leaving to a subset of the surviving states the role of representing also the states dropped from the equations. In the presence of inequivalences among states, this implies the necessity of correcting the coarse-grained equations by taking into account that, at stationarity, states in charge of representing the eliminated network sites should account also for their probabilities. Such operation reduced to perform uniform, irrelevant rescalings of the state occupation probabilities in the examples discussed above. We have also discussed two illustrative examples of how one has to proceed with situations in which inequivalent states survive coarse-graining. In the first we discussed a single site on a chain which replaces a whole loop, while in the second two sites share the duty of representing an erased intermedied one. In both cases the probabilities representing some of the surviving states have to be rescaled in order to maintain consistency with the stationary state as recorded along the coarse-grained trajectory.

We also learned that it is possible to convert into Markovian the process resulting from coarse-graining, without altering the average rate of entropy production. For general networks our findings suggest the possibility that an empirical coarse-graining performed directly on the trajectories could lead to enough information to set up a Markovian counterpart of the process with the same average rate of entropy production σ. It would be sufficient to identify from the statistics over coarse-grained trajectories the various rates which contribute to entropy creation, like we could do in the examples above.

In summary, this work shows, in the context of Markov jump processes, that an exact coarse-graining, taking into account memory effects, is possible. This coarse-graining is such to guarantee invariance of the Cramer's function controlling

entropy fluctuations at stationarity. This provides a strong, previously absent, argument in favor of the general validity and applicability of fluctuation theorems. Our results are also representative enough to suggest what theory should pursue when addressing coarse-graining in more general contexts.

CHAPTER 6

Negative differential mobility

1. Introduction

The work presented in the previous chapter followed a line of research that was originally inspired by the seemingly unrelated topic of Negative Differential Mobility (NDM). This is the phenomenon where increasing a force leads to a decrease of mobility due to, for example, trapping or crowding. It has been observed in several experiments on electronic properties of materials [**30, 31, 32, 33**]. It can also occur in gel electrophoresis of polymers [**34, 35, 36, 37**] and chemical reaction networks [**38**].

Negative differential mobility is often studied theoretically in driven lattice gas models [**22, 23, 24, 25, 26, 27, 28, 29**]. Since there exists at this moment no general theory for systems out of equilibrium, it is usual to study this particular and many other nonequilibrium phenomena using specifically designed Markov jump models. The dynamics of these processes follows once the rates are defined. To do this one can use a general principle such as local detailed balance but this is in general not enough to fix the rates.

For example, the biased random walk is a popular model for describing one-dimensional transport. Local detailed balance fixes the ratio of the rates for the particle to hop to the right, p, and left, q, as $p/q = e^F$ (where we put $k_B T$ and the distance between neighbouring sites equal to one). Here F is a force that, for example, is produced by an electric field acting on the particle when it is charged or is the force exerted by an optical tweezer when it represents a molecular motor. In the latter case, it has been found that it is necessary to introduce a load distribution factor θ in the jump rates, $p \sim e^{\theta F}$ and $q \sim e^{(\theta-1)F}$, in order to get agreement with experiments [**94, 95, 96**]. This load factor takes into account the change in the microscopic free energy landscape of the motor due to the force. The necessity of this factor cannot be obtained from local detailed balance alone but can be deduced by taking into account information on a more microscopic scale.

This example shows that an investigation of how a microscopic description can lead under coarse graining to a jump process description is in principle needed to assess the precise choice of transition rates. However, without undertaking such an ambitious program, it is important to explore the physical implications that different choices of rates can have by analysing specific prototype models. Such analysis reveals particularly useful when focusing on non equilibrium phenomena whose occurence sensibly depends on these choices.

This was the first step we took in trying to understand to which extent these jump models can be considered a faithful representation of a real world scenario. The world *real*, in this case, is a synonym of Langevin dynamics, since this is as close as we can get to a continuous stochastic system that can model what we actually observe in a physics experiment. In the end the problem we are dealing with is the coarse-graining of observables in systems out of equilibrium: bridging the gap between a continuous and a mesoscopic world is our final goal, but such task is extremely ambitious and therefore requires some intermediate steps.

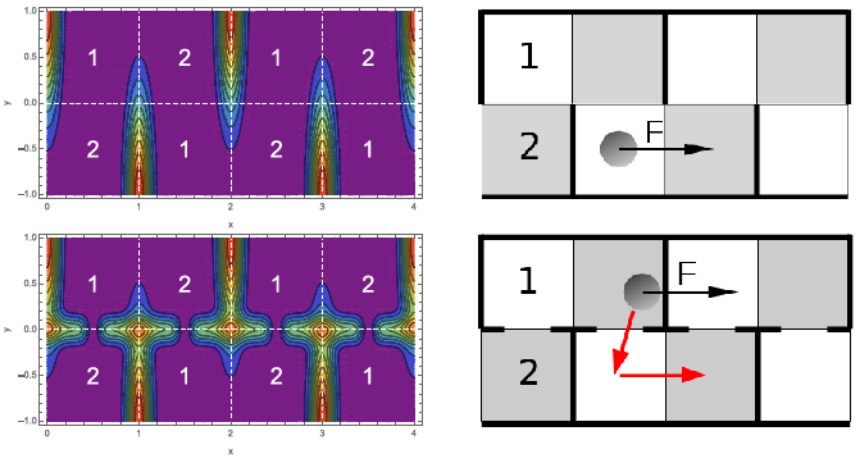


FIGURE 1. The left-hand side shows contour plots of two energy landscapes V_1 (top) and V_2 (bottom). These are coarse grained to the discretised version on the right-hand side where tick lines are walls that correspond to high energy barriers.

Here we perform such an analysis in models for NDM. We introduce a few exactly solvable one-particle Markov models which differ in their transition rates and study how these different rates influence the presence or absence of NDM. The choice of the rates is motivated by qualitative considerations of the way in which different free energy landscapes could influence the coarse grained description. Various landscapes lead to different rates in the coarse grained Markov model.

We also go beyond the average current and study higher (scaled) cumulants and the whole probability distribution of the current using large deviation theory [**97**]. In this way our work is also a first step in extending the large deviation theory for currents in exclusion processes [**98, 99, 100**] to lattice gases showing NDM.

This chapter is organised as follows. In section 2, we introduce our Markov jump models. In section 3, we review the large deviation approach to current fluctuations. In section 4, we apply this theory and discuss the difference between the models with and without NDM at the level of current distributions. Finally, in section V we present our conclusions.

2. Models

We begin by introducing three related Markov jump models in which a particle can hop between the sites of a lattice with two lanes, see Fig. 1. The jump models are considered to be discretisations of an underlying microscopic model in which a particle performs Brownian motion in an energy landscape $V(x, y) - Fx$ where F is the force that puts the system out of equilibrium.

To be more specific, in Fig. 1, left-hand side, we show contour plots of two energy landscapes that on a coarse grained scale correspond to the jump models shown in the right-hand side. High barriers in the energy landscape are considered as walls in the discretised version. In the potential V_1 (upper figure) there are only barriers and walls that are oriented perpendicular to the force. For V_2 (lower figure) the landscape is more complex and the walls have the shape of a letter T.

The different Markov models are distinguished by different jump rates between neighbouring sites. We denote by k^+ and k^- the rates to move in the direction of the field and against the field. As soon as $F \neq 0$, cells with a wall on the right are not equivalent with those where the wall is on the left. Therefore, the jump rates

Model	k^+	k^-	k^0
A	$e^{F/2}$	$e^{-F/2}$	1
B	$e^{F/2}$	$e^{-F/2}$	$1/\cosh(F/2)$
C	$e^{\theta F}$	$e^{(\theta-1)F}$	$1/\cosh(F/2)$

TABLE 1. Jump rates of the particle in the three models.

in the direction perpendicular to the force can be different if one goes from a cell with the wall on the left to one with a wall on the right or vice versa. We call the associated rates u^+ and u^- respectively.

We will from now assume that local detailed balance holds even though it is nontrivial to establish a precise connection between its form at a microscopic, e.g. Langevin description, and that in the associated jump rate model.

The rates to jump between sites i and j then obey

$$\frac{k(i \to j)}{k(j \to i)} = \exp[E_i - E_j + F\Delta x] \tag{208}$$

where E_i is the energy when the particle is at site i and Δx is the size of the x-component of the jump (which can be $-1, 0$ or 1). Here we will assume that all E_i are equal. This can be the case if we identify the position of the particle in the discrete model with the location of the minima in the potential $V(x)$. Local detailed balance then implies $k^+/k^- = e^F$ and $k^0 \equiv u^+ = u^-$.

Given local detailed balance, there is still a large freedom in the choice of the rates. The rate k^0 can be F-dependent or not. When adding a positive (negative) force, a particle in a cell left (right) of a wall is trapped and first has to change lane in order to move forward. If there is no wall between the two lanes, we do not expect the corresponding rates to depend significantly on F (Fig. 1, top). In the opposite case (Fig. 1, bottom), the particle in the underlying model first has to move against the force to cross the barrier so that we define the corresponding rate k^0 to be equal to $1/\cosh(F/2)$.

Furthermore we distinguish between models where walls can be passed or not. We introduce an extra (small) parameter ϵ to describe this effect such that the rates for crossing a wall are given by $k_W^{\pm} = \epsilon k^{\pm}$. This parameter could be given in terms of a Kramers' rate $\epsilon \sim e^{-\Delta E}$ where ΔE is the height of the barrier.

Finally, in order to describe the effect of wall crossing in even more detail, we also allow the possibility of a load-distribution number θ ($\theta - 1$) multiplying F in k^+ (k^-). As mentioned in the introduction, this takes into account the shift in the location and the height of the energy barriers in the presence of a force. These shifts lead to a modification in Kramers' rate for passing an energy barrier which is captured by the load factor.

With these considerations in mind we define three models. In model A, k^0 does not depend on F whereas in the other two it does. In model B, $k^0 = 1/\cosh(F/2)$. In these two models, we do not include a load factor, implying $\theta = 1/2$. Finally, in model C, the load factor is an extra parameter. In Table I, we give a summary of the rates for the various models.

With the introduction of periodic boundaries and using the symmetries of the lattice all our models are two state Markov chains. We denote state 1 (2) as the state in which the particle is to the right (left) of a vertical wall. The probability to be in state i at time t, $P_i(t)$, evolves according to the master equation

$$\frac{dP_i}{dt} = \sum_{j=1,2} M_{ij} P_j(t) \tag{209}$$

A straightforward calculation shows that the generator M is given by

$$M = \begin{pmatrix} -k^0 - k^+ - k_W^- & k^0 + k^- + k_W^+ \\ k^0 + k^+ + k_W^- & -k^0 - k^- - k_W^+ \end{pmatrix} \tag{210}$$

The stationary state $P^\star$ is the eigenvector with eigenvalue 0 of M. The average current in the stationary state is given by

$$J(F) = (k^+ - k_W^-)P^\star(1) - (k^- - k_W^+)P^\star(2) \tag{211}$$

We mention here that another popular choice has been to take model A and normalise all rates by a factor $2(1+\cosh F)$ [**22, 25, 27**]. From the master equation (209), it follows that this can be seen as a rescaling of time to a new force dependent time $t'(F) = t/[2(1+\cosh F)]$. Since for F large, the rates in this model become constant or go to zero, the current at most goes to a constant. Therefore within such a model it is not possible to obtain current-force relations which increase with F for F sufficiently large.

We will also be interested in the entropy production rate $\sigma(F)$. A general expression for this quantity for a Markov chain is well known [**101**]. In the present case it reduces to

$$\begin{aligned} \sigma(F) &= [(k^+ \ln(k^+/k^-) + k_W^- \ln(k_W^-/k_W^+)]P^\star(1) \\ &+ [(k^- \ln(k^-/k^+) + k_W^+ \ln(k_W^+/k_W^-)]P^\star(2) \end{aligned} \tag{212}$$

Using detailed balance and (211) this can be simplified to

$$\sigma(F) = J(F)F \tag{213}$$

as could be expected from non equilibrium thermodynamics.

3. The current large deviation function

In recent years, much interest has been devoted to go beyond the average current and investigate current fluctuations in various non-equilibrium systems such as exclusion processes [**98, 99, 100**]. Fluctuations of the current away from its average value are characterised in terms of a large deviation function $I(Y)$ which plays a role analogous to that of entropy in equilibrium statistical mechanics. Equivalently, one can describe these fluctuations in terms of the scaled cumulant generating function (SCGF) which is like a non equilibrium free energy.

Here we collect a few basic results of the theory of large deviations in continuous time Markov processes. For more details we refer to [**98, 17**].

In the large deviation approach to current fluctuations, one introduces a stochastic variable $Q(t)$ which increases (decreases) by one each time the particle makes a step in the direction of (against) the field. For t large, the probability density that $Q(t)/t$ equals Y is then proportional to $e^{-tI(Y)}$. Here $I(Y)$ is the large deviation function (LDF, also called rate function) which is zero at the average current $J(F)$ and positive otherwise.

Alternatively, one can introduce the SCGF as

$$\lambda(s) = \lim_{t\to\infty} \frac{1}{t} \ln\langle e^{sQ(t)}\rangle \tag{214}$$

where the average is taken over all realisations of the process. The large deviation function and the SCGF are connected through a Legendre-Fenchel transformation

$$I(Y) = \sup_s [Ys - \lambda(s)] \tag{215}$$

The scaled cumulants of the current are obtained by taking derivatives of $\lambda(s)$ with respect to s. One has for the first three cumulants

$$J(F) = \lim_{t\to\infty}\frac{1}{t}\langle Q(t)\rangle = \frac{\partial\lambda(s)}{\partial s}(s=0) \tag{216}$$

$$\Delta(F) \equiv \lim_{t\to\infty}\frac{1}{t}[\langle Q(t)^2\rangle - \langle Q(t)\rangle^2] = \frac{\partial^2\lambda(s)}{\partial s^2}(s=0) \tag{217}$$

$$\chi(F) \equiv \lim_{t\to\infty}\frac{1}{t}[\langle (Q(t)-\langle Q(t)\rangle)^3\rangle] = \frac{\partial^3\lambda(s)}{\partial s^3}(s=0) \tag{218}$$

For convenience we will refer to the second and third scaled cumulant as the variance and the asymmetry, even though the standard definition of the variance includes an extra factor t.

It is known [**17**] that the SCGF equals the largest eigenvalue of a tilted generator $M(s)$ which for the models considered in this paper has the form

$$M(s) = \begin{pmatrix} -k^0 - k^+ - k_W^- & k^0 + k^- e^{-s} + k_W^+ e^s \\ k^0 + k^+ e^s + k_W^- e^{-s} & -k^0 - k^- - k_W^+ \end{pmatrix} \tag{219}$$

Here off-diagonal elements that correspond with a step in the direction of the force are multiplied by e^s while those associated to a step against the force get a factor e^{-s}.

We have calculated the SCGF for our three models from an exact diagonalisation of the tilted generator. The average, the variance and the asymmetry of the current can then be obtained by simple derivation.

The average current (216) and the variance (217) appear in the thermodynamic uncertainty relation (TUR) [**102, 103, 104**]

$$\frac{J(F)^2}{\Delta(F)} \leq \frac{\sigma}{2k_B} \tag{220}$$

This relation implies that in order to decrease the variance of the current, more entropy has to be produced.

More recently, also a so called kinetic uncertainty relation (KUR) was derived [**105**]. It is expressed in terms of the dynamical activity $K(t)$ which, for a given realisation of a jump process, equals the number of transitions made up to time t. This dynamic activity, also called frenesy, takes into account non-dissipative aspects of non equilibrium systems and is needed to characterise the physics of systems far from equilibrium [**106, 107**]. The kinetic uncertainty relation is

$$\frac{J(F)^2}{\Delta(F)} \leq \kappa(F) \tag{221}$$

where $\kappa(F) = \lim_{t\to\infty}\langle K(t)\rangle/t$, i.e. the average activity per unit time. The KUR might give a better bound on current fluctuations when the system is far from equilibrium [**105**], though in a chemical reaction network the reverse situation has been observed [**38**].

In our models we have

$$\kappa(F) = (k^+ + k^0 + k_W^-)P^\star(1) + (k^- + k^0 + k_W^+)P^\star(2) \tag{222}$$

Finally, we have also determined the rate function $I(Y)$ using (215). For given Y we have to determine $s^\star(Y)$ which satisfies

$$Y = \frac{\partial\lambda(s)}{\partial s}(s^\star) \tag{223}$$

Then $I(Y) = Ys^\star(Y) - \lambda(s^\star(Y))$. In practice it is more convenient to vary s and make a parametric plot of the curve $\{\frac{\partial\lambda(s)}{\partial s}(s), \frac{\partial\lambda(s)}{\partial s}(s)s - \lambda(s)\}$ [**97**].

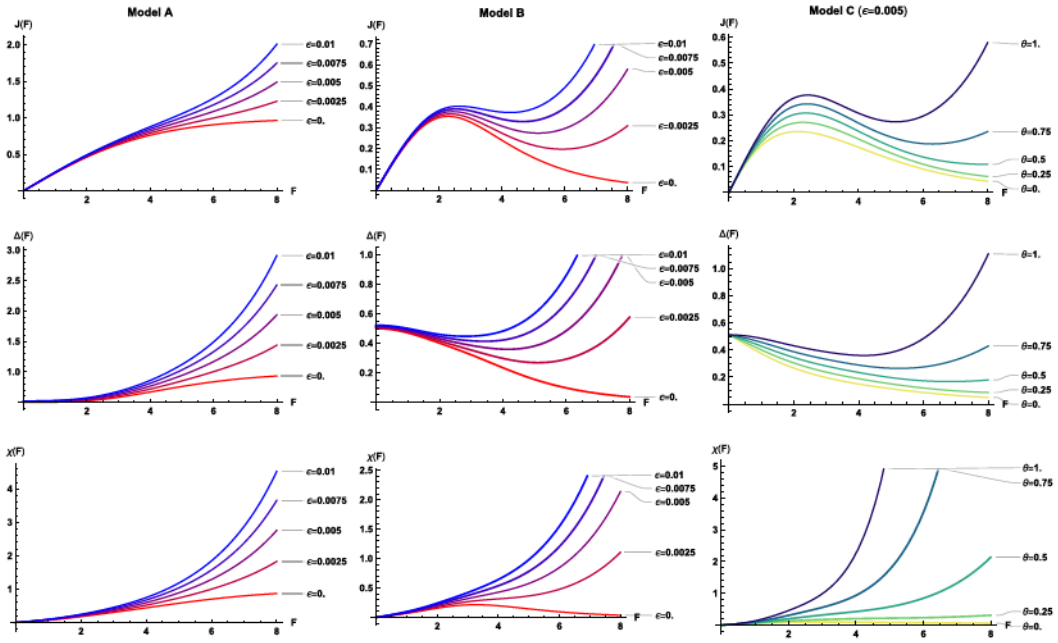


FIGURE 2. Average current (top), variance (middle) and asymmetry (bottom) as a function of F for various ϵ values. The left (middle) panel shows the result for model A(B). On the right we show the same quantities for model C at fixed value of $\epsilon = 0.005$ as a function of the load factor θ.

4. Results

In Fig. 3 we plot the average current (top), the variance (middle) and the asymmetry (bottom) as a function of F for the three models. For model A and B (left and middle panel) the results are given for different ϵ-values, while for model C (right panel) results are shown for different θ-values and at $\epsilon = 0.005$.

We find that model A does not show NDM whereas the other two models do. There is also NDM in these models in a regime of negative F (not shown). This is because they have an ($F \leftrightarrow -F$, left $\leftrightarrow$ right) symmetry.

Notice that for model B and for $\epsilon = 0$ the average current goes to zero for large F and that NDM is therefore present for all forces above a critical one. If the walls can be surpassed and ε is not too big, NDM is present in a finite interval of forces, i.e. for $F_-(\epsilon) \leq F \leq F_+(\epsilon)$ after which the current increases again. For ϵ above a critical value $\epsilon_c = 0.015$, NDM disappears. In Fig. 3, we present a plot of $F_-(\epsilon)$ and $F_+(\epsilon)$ in model B.

Interestingly, the observed behaviour of the current in model B and C at $\varepsilon \neq 0$ is qualitatively similar to the behaviour seen in some experiments on hopping conductivity [**30, 31, 32, 33**]. In all of these, it has been found that the current as a function of the field shows NDM for a certain range of forces, after which it increases again.

If we normalise all the rates of model A with a factor $2(1+\cosh F)$ [**22, 25, 27**] one finds that also model A shows NDM. On the other hand, we then find that the current goes to a constant at large force which is not what one would expect physically for a particle in a flow or for the hopping conductivity of an electron.

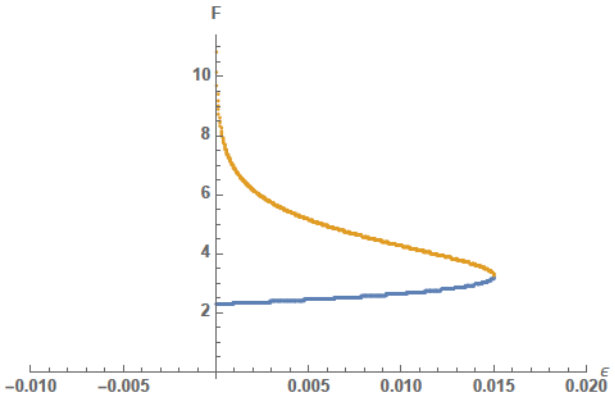


FIGURE 3. This figure shows, for model B, the regime in which there is negative differential mobility. It is bounded by the forces $F_-(\epsilon)$ (blue, lower curve) and $F_+(\epsilon)$ (orange, upper curve). Above $\epsilon = 0.0015$, NDM disappears.

We now look whether signs of NDM can also be seen in the fluctuations of the current. In model A, the variance is monotonically increasing both as a function of F and as a function of ε. In contrast, in the models showing NDM, the variance goes to zero at large F for $\varepsilon = 0$ while it goes through a minimum if $\varepsilon \neq 0$. Hence in both cases there is a regime in which $\partial\Delta(F)/\partial F < 0$. This phenomenon can be called negative differential variance.

Also the asymmetry is observed to behave in a similar way: in the models with NDM there is a regime where $\partial\,\chi(F)/\partial F < 0$. Such a regime is absent in model A. Notice however that this behaviour is only observed for ϵ and $\theta - 1/2$ very small.

Alternatively, the same conclusions can be reached by plotting the rate function $I(Y)$. In Fig. 4 we plot this function for the three models for various forces and ϵ or θ (model C). When the average current increases (decreases) the minimum in $I(Y)$ shifts to the right (left). Similarly, an increase (decrease) in $\Delta(F)$ makes $I(Y)$ broader (smaller) while an increase (decrease) in $\chi(F)$ tilts the function more to the right (left).

Finally we look at the uncertainty relations (220) and (221). In Fig. 5 we plot, for models A and B, $\sigma(F)/2$ and $\kappa(F)$ at $\epsilon = 0$ and $\epsilon = 0.005$. In the same plot we also show $J(F)^2/\Delta J$. We see that the thermodynamic uncertainty relation becomes almost an equality for small forces, while it gives a strict upper bound at large forces. For larger forces, where it can be argued that the system is further from equilibrium, the KUR gives a better upper bound. This result holds independently of the behaviour of the entropy production at large F. The same conclusion was found to be valid for the other models and other parameter values.

5. Discussion

In this chapter we showed the investigation we carried out on current fluctuations in some Markov chain models that were constructed to give a coarse grained description of the Brownian motion of a particle in an energy landscape. We have chosen the rates in such a way that they give an average current that behaves similarly to the one expected in the microscopic model. While the different models all obey the same local detailed balance we have seen that choosing different rates can lead to the absence or presence of negative differential mobility. If we include more parameters in the rates, such as the possibility to cross a wall or a load factor,

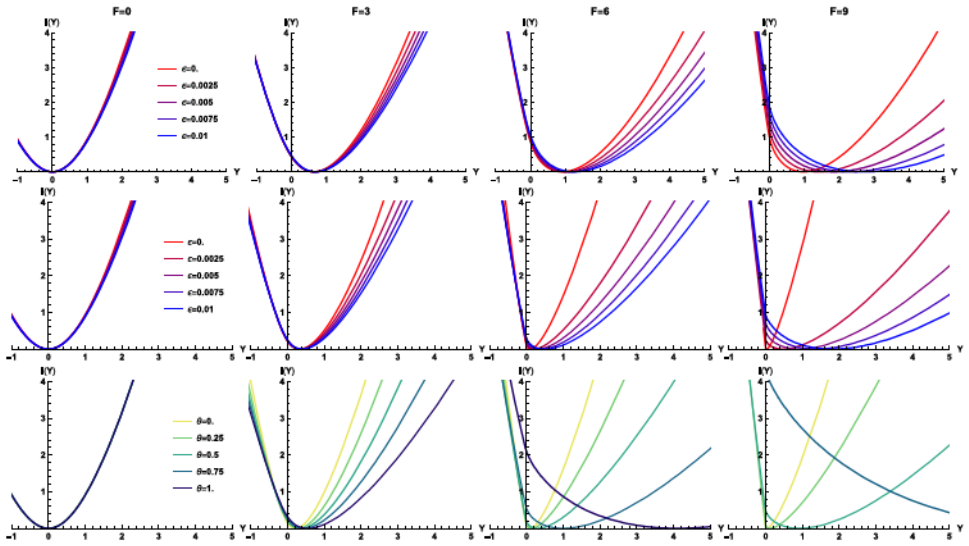


FIGURE 4. Large deviation function $I(Y)$ in model A (top) and B (middle) at $F = 0, 3, 6$ and 9 for different ϵ-values. The bottom figure shows $I(Y)$ in model C as a function of θ at $\epsilon = 0.005$.

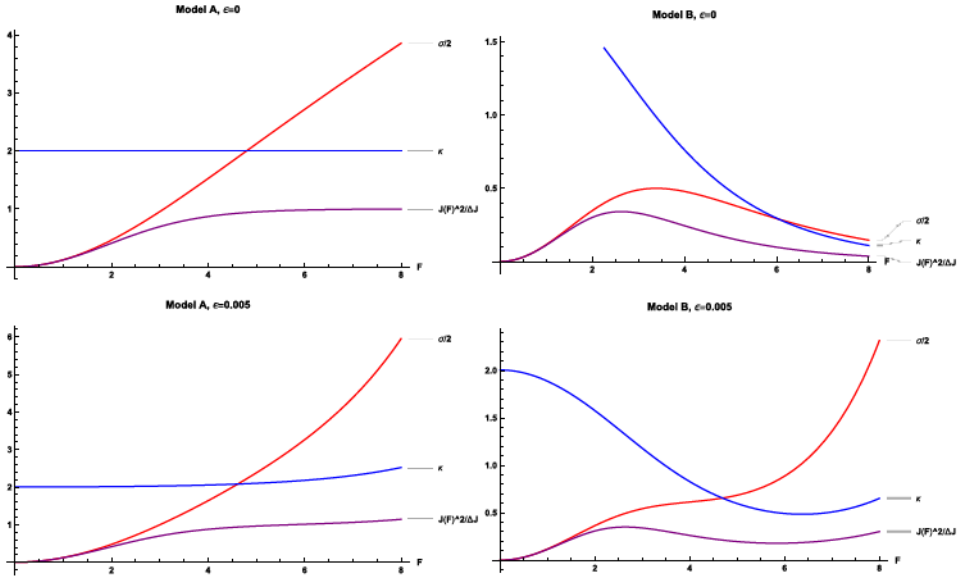


FIGURE 5. Entropy production rate $\sigma(F)/2$, average activity rate $\kappa(F)$ and $J(F)^2/\Delta J$ as a function of F in model A and model B for $\epsilon = 0$ and $\epsilon = 0.005$.

we can obtain a current-force relation that qualitatively resembles that observed in experimental systems.

Besides average currents we also investigated the current fluctuations. We have seen that in models with NDM, and for $\epsilon = 0$ also the variance and the asymmetry of the current have a maximum which for the variance is at $F = 0$. When $\varepsilon \neq 0$

the behaviour becomes more interesting so that for some parameter values the variance and the asymmetry go through a minimum as a function of F. This can be compared with lattice gas models of NDM. In these one considers one or more force driven particles in the presence of passively diffusing particles that do not feel any force. The latter can act as traps and cause NDM [**27, 28**]. It has been found that for such a model the variance of the current goes through a maximum [**29**]. It would be interesting to see whether with a suitable modification of the lattice gas model one could also find regimes in which the average current and its fluctuations show both a maximum and a minimum.

In this work, the relation between the microscopic and the mesoscopic model was made at a heuristic level. It would be of great interest to make a more detailed mapping between the two levels of description. One can for example ask how the rates should be chosen such that the models at different scales have the same current and the same diffusion constant. Or one can require that the entropy production at the two levels stays the same. These questions have been studied in an approximate way for Markov chain models in which a clear separation of timescales can be made between fast and slow variables [**20, 21**]. But that work leaves open the question on what happens if one coarse grains a Markov chain where no clear separation of time scales is present, or when one coarse grains from a microscopic diffusion process to a mesoscopic Markov model. We are currently performing studies in this direction, which will help us in approaching and tackling the general problem of coarse graining of stochastic systems across different formalisms.

Aknowledgments

TO DO...

APPENDIX A

Network model of conviction-driven social segregation

During my PhD course I had also the chance to explore fields not directly linked to the main research topics presented in this thesis. One peculiarity of the physics complex systems is in fact the possibility of exploiting the same tools to perform analyses on completely unrelated topics. In my case, a statistical mechanics approach revealed to be very effective in the study of the phenomenon of social segregation. In this appendix I'll present the work (published in [**5**]) that I was able to carry out as a side project collaborating with a heterogeneous group of people external to my research group.

1. Introduction

Social segregation is a primary problem for our well-being, and for the policy-making of our governments. The most basic questions regarding social segregation concern its quantification, and the prediction and prevention of its onset and its outcomes. Attempts to approach the problem from a quantitative viewpoint date back to the late 1960s, with a model proposed by the economist Thomas C. Schelling [**39, 40**]. In this model, individuals are embedded in a two-dimensional lattice, and are characterized by a threshold "tolerance" to other individual opinions. This model naturally attracted the attention of statistical physics because of its analogy with Blume-Emery-Griffiths and Potts models, and more in general with binary mixtures and interfacial dynamics. It shows a complex phase diagram, including threshold phenomena (phase transitions) where opinions separate spatially and may form patterns [**41, 42, 43, 44**]. Schelling's model demonstrates that even mild preferences for a set of agents for defining themselves as a local minority can produce strong spatial segregation patterns, challenging the common view that discrimination is a necessary condition for segregation.

While spatial "steric" interactions and dimensionality are very important in Schelling's model, human interactions can in most cases be described as network-like [**108, 109, 110, 111, 112**]. In a situation with (nearly) immutable convictions and limited tolerance to other opinions, individuals sharing the same conviction might find themselves severed from society even if their potential for social interaction is not limited by spatial constraints. Such a situation is very dangerous for society, for the danger of triggering self-propelled distortions of reality shared between many individuals. For example, this is particularly relevant in the on-line world of social networks. The diffusion of on-line non-intermediated unverified and polarized contents and the spread of misinformation is becoming a pressing problem for our society. One of the most relevant driving forces has been recognised as the echo-chamber effect [**113, 114, 115**]. It consists in the formation of segregated clusters of users who share some strong common opinions, increasingly reinforcing these ideas and thus becoming impenetrable to news diverging from their point of view.

Thus, another possible approach (relatively less explored) may attempt to describe segregation using opinion-based network models, such as the voter model [**116,**

117, 118]. The complex networks literature provides many examples of segregation in the structure of relationships (from school friendship to value- and belief-oriented partitioning) empirical data [**119, 120**]. However, the literature on complex networks models focuses mostly on how opinion dynamics is shaped by network-like human interactions, i.e., on how individuals change their mind based the opinions of others [**121, 117, 118**]. Such a framework is not well-suited to describe segregation, where precisely the opposite occurs, i.e., human interactions change following stable "opinions", or other more general individual-specific factors (as it happens in Schelling's model). Indeed, some of these factors may be very strongly rooted in individuals, such as convictions, religious and cultural factors, and even immutable physical or racial features. A comparatively smaller thread of studies [**122, 116, 123, 124**] has considered the coevolution of network connections and opinions. In such models, individuals can both change their mind and change their connections, and segregated states can emerge, depending on the intrinsic time scales of these processes [**122, 123**]. However, the conditions for reaching segregated states are not the main focus of these investigations, which are typically focused on the conditions for reaching consensus. In order to understand the factors leading to segregated states, it is important to address the case where node attributes (convictions) are persistent.

There is very little work in the literature addressing such situation on networks. A fairly recent study [**125**], considered the emergence of segregation in a social network by a model with continuous opinions and an individual "aversion bias" favoring the severing of connections with increasing difference of opinions, in favor of random rewiring. They proved the existence of attractor steady states with given segregation levels that are independent of initial conditions, and characterized the time scales of convergence to these states. However, this study did not address the possibility and existence of the threshold phenomena that are ubiquitious in Schelling's model. Such phenomena are important to address, as argued in the previous paragraphs.

Here, we define an alternative model of segregation on networks based on *discrete* convictions, and we study it through analytical calculations and direct simulation. In our model, individuals may choose to follow other individuals based on sharing the same conviction, or based on their popularity (regardless of conviction). The trade-off between these two moves defines a transition between a well-mixed and a segregated state. A threshold parameter, analogous (but not equivalent) to the "tolerance" parameter in Schelling's model, weighs the two different possible choices. We analyze this model in the case of binary states of the agents (two possible convictions, such as Democrats and Republicans), and we are able to fully characterize the conditions for the emergence of phase transitions the relaxation time scales of the system in the segregated and non-segregated phases. Importantly, in order for transitions to exist, the conviction move has to occur on the same time scale of the popularity move, regardless of the size of the community being segregated. Finally, we show that minority convictions segregate more easily, and we characterize this phenomenon quantitatively.

2. Definition of the model

Our model describes a social network as a directed graph where individuals (nodes) follow other individual's opinions by sending directed edges to their corresponding nodes. The initial condition is a random directed graph $G_0(N, m, h)$ made of $N \in \mathbb{N}$ nodes. Each node has fixed outdegree $m \in \mathbb{N}$ (relaxing this assumption to fluctuating outdegrees does affect the results, see below). A fraction $h \in [0; 1]$ of individuals hold a certain conviction, which we identify with the color *red* (as

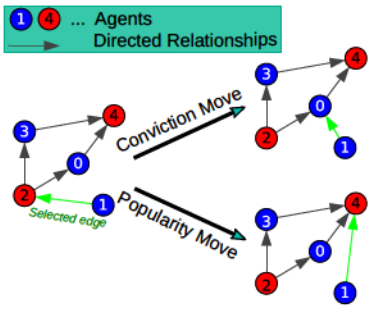


FIGURE 1. Illustration of the action of the model basic moves. Nodes represent agents and colors represent convictions. Edges represent directed social connections (A follows B if an edge is sent from A to B). The selected edge to be removed is in both cases $e_{1\to 2}$. In a conviction move, the new target can be chosen only among the blue nodes (in the sketch this move creates the edge $e_{1\to 0}$), while in a popularity move the new target can be chosen regardless of its opinion, so that every node with an in-degree greater than 0 is a potential candidate (in the sketch this moves creates the edge $e_{1\to 4}$).

opposed to the probability $1-h$ of holding the opposite conviction, i.e. being colored in *blue*). The total number of edges $M = N \cdot m$ defines the size of our system. The graph is constructed through the associated adjacency matrix by filling randomly with m ones the matrix rows of a zero matrix (we exclude the matrix diagonal elements which would indicate self-edges). As a consequence of this construction procedure, the in-degrees follow a Poisson distribution with average value m (as in an Erdõs-Rényi random graph [**126**]).

The network evolves at *fixed* conviction, by choosing at each step one of two possible rewiring moves (Fig. 1) accordingly to the choice parameter $\varphi \in [0;1]$:

- with probability φ a *conviction move* chooses randomly one among all the edges $e_{i\to j}$ between two nodes holding different convictions (which we will call "heterogeneous" edges), deletes, chooses uniformly a new target node k holding the same conviction as i and creates a new "homogeneous" edge $e_{i\to k}$;
- alternatively, with probability $1-\varphi$, a *popularity move* which chooses randomly one edge $e_{i\to j}$ among all the edges of the network, deletes it, and creates a new edge $e_{i\to k}$ with a target k chosen among all the nodes with a preferential attachment criterion, i.e. with a probability equal to the in-degree of the target node normalized by the total number of edges M.

It is important to underline the fact that the opinion move selects the edge to be removed in the basket of the heterogeneous edges. As it will be more clear in the following, this choice is essential in order to obtain a threshold phenomenon for segregation.

We quantify the segregation using as order parameter the total number of homogeneous edges connecting nodes with the same conviction. The order parameter corresponds to observables commonly used in the literature, and roughly quantifies the links that need to be severed to separate the two communities. In real-world situation the structure within the communities that form can vary. Our calculations indicate that this structure is not relevant for our model (see below). In the initial

condition ($t = 0$), and for M sufficiently large, the densities of the four different kinds of edges (red to red, blue to blue, red to blue and blue to red) are:

$$\begin{aligned} e_0(rr) &= h^2 \\ e_0(bb) &= (1-h)^2 \\ e_0(rb) &= e_0(br) = h(1-h) \quad . \end{aligned} \tag{224}$$

More in general, for every step $t > 0$, the link densities are functions of this parameter order parameter. Indeed, since $\Omega_t := M(e_t(rr) + e_t(bb))$, one has

$$\begin{aligned} e_t(rr) &= \frac{h^2}{h^2 + (1-h)^2} \frac{\Omega_t}{M} \\ e_t(bb) &= \frac{(1-h)^2}{h^2 + (1-h)^2} \frac{\Omega_t}{M} \\ e_t(rb) &= e_0(br) = \frac{M - \Omega_t}{2M} \quad . \end{aligned} \tag{225}$$

We define a segregated phase as a state where, for large networks, typically all the heterogeneous edges disappear, leaving the network with only edges between like-minded nodes, characterized by a saturation of the order parameter to the maximum value $\Omega_t = M$.

3. Results

3.1. A transition to a segregated state emerges at a critical point. By construction of the model dynamics, conviction moves favor the transition to a segregated phase, while popularity moves try to reestablish the disorder and will also affect the in-degree distribution. Moreover, we expect networks characterized by asymmetric densities of opinions ($h \neq 1/2$) to reach a segregated phase more easily.

Starting by the same initial random graph G_0, we evolved the network for different values of φ and at each step we recorded the order parameter $\Omega_t(\varphi)$, starting from initial conditions with $\Omega_0 = 1/2$ for $h = 1/2$ (Fig. 2a), representing the fraction of homogeneous edges (connecting individuals with equal convictions). For low values of φ, the system does not segregate, but they reach a balance between popularity- and conviction-based moves. As the value of φ increases, conviction-based moves become increasingly dominant, and the steady-state value of the order parameter increases until it reaches the maximum possible value M, indicating that typically the number of heterogeneous edges is negligible compared to the total number of edges, and the system reaches a segregated phase. This behavior suggests the existence of a critical value φ_c of the choice parameter, above which the steady state of the network is always in a segregated phase.

In order to find the critical value of the choice parameter analytically, we used a mean-field approach, based on an estimate of the average variation $\Delta\Omega_t$ at every step. Conviction moves increase Ω_t by 1, while popularity moves might act differently depending on the probability of picking an edge of a certain kind, and also on the kind of the new edge created. The resulting mean-field equation is

$$\Delta \langle \Omega_t(\varphi, h) \rangle = \underbrace{\vartheta\varphi}_{\text{conv. move}} + \underbrace{(1-\varphi)\left[\vartheta p_t^+(h) - p_t^-(h)\right]}_{\text{pop. move}} , \tag{226}$$

where the Heaviside step function $\vartheta := \theta(M - \Omega_t)$ excludes forbidden moves once the segregation state is reached, while $p_t^{\pm}(h)$ are the probabilities of respectively increasing and decreasing the order parameter with a popularity move.

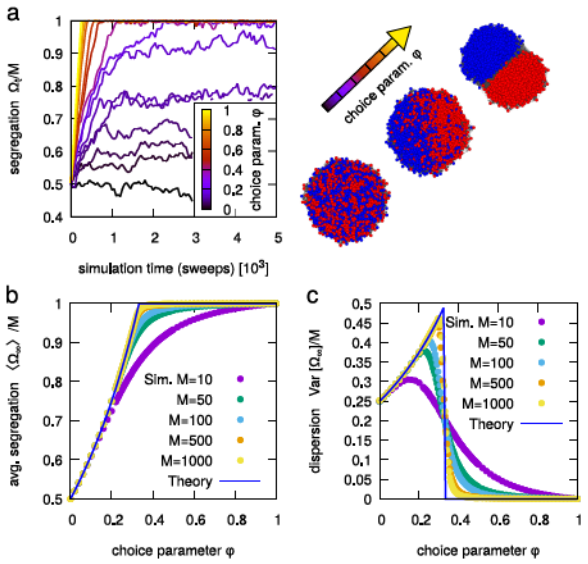


FIGURE 2. A threshold phenomenon to a segregated state appears for a critical value of the choice parameter φ_c. **a)** Evolution of the fraction of homogeneous links. The plot shows the order parameter normalized by the total number of edges M plotted against sweeps. The curves are obtained by simulating the evolution of the same initial random graph $G_0(N = 500, m = 5, h = 1/2)$ for different values of φ. For low φ, the long-time value of $\Omega_\infty(\varphi)$ relaxes to a steady state where the edges connecting nodes with different colors fluctuate around a finite value, while as φ grows, it reaches one (a segragated state) in a finite time. The right-hand panel shows some illustrative simulation snapshots, where the network is visualized with a spring model based on shared links. **b)** Plot of the mean order parameter at steady state versus the choice parameter φ comparing the analytical results (solid line) of Eq. 230 with numerical simulations for different sizes of the network M (symbols). This analysis supports a segregation transition for $\varphi_c = 1/3$ (for $h = 1/2$). **c)** Fluctuations scale linearly with the size of the system. Plot of the dispersion of the order parameter from the simulations in panel b (symbols). As the size of the network grows, the variability across realizations peaks around the critical value $\varphi_c = 1/3$ reflecting the prediction of Eq. 235 (solid line).

In the continuum time limit, and for $h = 1/2$ (for a more general derivation for every $h \in [0; 1]$ see section 5.1) Eq. (226) gives the following differential equation for the average value of the order parameter

$$\partial_t \langle \Omega_t(\varphi) \rangle = \vartheta \frac{1+\varphi}{2} - (1-\varphi)\frac{1+\vartheta}{2}\frac{\langle \Omega_t(\varphi)\rangle}{M} \,. \tag{227}$$

This equation can be explicitly integrated (for $\varphi \neq 1$), yielding the time dependence for the average value of the order parameter,

$$\frac{\langle \Omega_t(\varphi)\rangle}{M} = \left[\left(1 - \frac{1}{2}\vartheta\right) - \frac{\vartheta}{1+\vartheta}\frac{1+\varphi}{1-\varphi}\right] e^{-(1-\varphi)\frac{1+\vartheta}{2M}t} + \\ + \frac{\vartheta}{1+\vartheta}\frac{1+\varphi}{1-\varphi} \,. \tag{228}$$

In the pre-segregation regime (where $\Omega_t < M$ and therefore $\vartheta = 1$) the relaxation is then exponential with characteristic time

$$\tau_\Omega = \frac{M}{1-\varphi}. \tag{229}$$

Hence, the asymptotic value

$$\frac{\langle \Omega_\infty(\varphi) \rangle}{M} = \min_{\varphi \in [0;1)} \left\{ 1, \frac{1+\varphi}{2(1-\varphi)} \right\} \tag{230}$$

will be reached for times $t \gg \tau_\Omega$. Fig. 2b compares this prediction with direct simulations. The model behaves as expected already for relatively small-sized networks ($M = 100$) and gradually moves towards the predicted curve as the size of the system grows. The relaxation time scales agree with the theoretical predictions, which can be used as criteria for stationarity (Figure 9). By setting $\langle \Omega_\infty(\varphi) \rangle = 1$ in Eq. 230 and solving for φ one finds the critical value of the choice parameter at which the transition occurs, which for $h = 1/2$ is $\varphi_c = 1/3$. This transition has a clear similarity with second order phase transitions [**127**] , because of a discontinuity in the first derivative of Ω_t with respect to φ. The analogy identifies the order parameter Ω with the magnetization, while the role of the temperature is played here by the choice parameter φ.

The fluctuations of the order parameter also characterize the transition. These can be estimated by the second cumulant moment $\mathrm{Var}[\Omega_\infty(\varphi)]$. A peak in amplitude of the fluctuations at the critical value φ_c should signal the transition. In the social segregation interpretation, this means that the transition to a segregated state is also marked by sudden growth and shrinkage of its connections to the rest of the world. In order to access the fluctuations analytically, we explicitly considered the master equation [**128**]). Calling $P_t(\Omega)$ the probability of having Ω homogeneous edges at time t the master equation is defined as

$$\partial_t P_t(\Omega) = \sum_{\Omega' \neq \Omega} W(\Omega|\Omega') P_t(\Omega') - W(\Omega'|\Omega) P_t(\Omega) \ , \tag{231}$$

where $W(\Omega|\Omega')$ are the transition rates of moving from a network with Ω' homogeneous edges to a network of Ω edges, which for our system (always in the case of $h = 1/2$) is

$$\begin{aligned} W(\Omega|\Omega') = \delta_{\Omega',\Omega-1} \left[\varphi + (1-\varphi) \frac{M-\Omega'}{2M} \right] + \\ + \delta_{\Omega',\Omega+1} (1-\varphi) \frac{\Omega'}{2M} + \delta_{\Omega',\Omega} \frac{1-\varphi}{2} \ . \end{aligned} \tag{232}$$

In the above equation, the first row describes the contribution of both the opinion and popularity moves to an increase in Ω, while the second row describes the contributions of the popularity move to respectively decrease and keep unaltered the order parameter. Then we define the factorial moment generating function

$$G(s,t) = \sum_{\Omega=0}^{M} s^\Omega P_t(\Omega) \ , \tag{233}$$

where $s \in \mathbb{R}$ is the dual parameter of Ω. Combining Eqs. (231) and (233) (see section 5.2) yields the following partial differential equation,

$$\partial_t G(s,t) = G(s,t) \frac{1+\varphi}{2} (s-1) + \partial_s G(s,t) \frac{1-\varphi}{2M} \left(1 - s^2\right) \ . \tag{234}$$

By evaluating $\partial_s^n [\partial_t G(s,t)|_{s=1}]$ for every $n \in \mathbb{N}$ we obtain a closed system of time-only differential equations giving the exact dynamics (including the transient phase) of all the factorial moments. The first factorial moment coincides with the average,

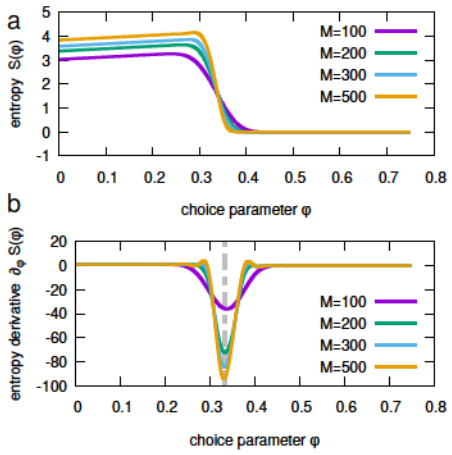


FIGURE 3. Entropy is characterized by a discontinuity in correspondence with the critical value of the choice parameter φ_c. A) The entropy of the system as a function of the order parameter φ for different system size M. B) Its derivative with respect to φ_c. The dashed gray line represent the predicted critical threshold $\varphi_c = 1/3$.

so we find again Eq. 227, whereas the second factorial moment gives $\langle \Omega_t^2 \rangle$ and hence the variance. Taking the long-time limit we obtain an analytical expression for the fluctuations

$$\frac{\mathrm{Var}[\Omega_\infty(\varphi)]}{M} = \begin{cases} \frac{1+\varphi}{4(1-\varphi)} & \text{for } \varphi \leq 1/3 \\ 0 & \text{for } \varphi > 1/3 \end{cases} \tag{235}$$

Fig. 2c shows that as the size of M (number of edges) of the network grows, the simulations tend to agree with this large-M prediction, showing a behavior that resembles that of the susceptibility in second-order phase transitions, with fluctuations amplitude scaling linearly in M.

By means of the generating function formalism, we can go further and calculate exactly the stationary solution of the Master Equation (231) with transition rates given by Eq. (232). The resulting stationary probability function P_{stat} is (see section 5.3 for detailed calculations):

$$P_{\mathrm{stat}}(\Omega) = \frac{2^{-\frac{M(\varphi+1)}{1-\varphi}} \left(\frac{M(\varphi+1)}{1-\varphi}\right)^{(\Omega)}}{\Omega!}, \tag{236}$$

where $x^{(\Omega)}$ is the factorial power of x and it is given by $\frac{\Gamma(x+1)}{\Gamma(-\Omega+x+1)}$. From Eq. (236) we can then define the entropy of the system $S(\varphi) = -\sum_{\Omega=0}^{M\to\infty} P_{\mathrm{stat}}(\Omega) \log[P_{\mathrm{stat}}(\Omega)]$ and its derivative with respect to the choice parameter φ. As Figure 3 shows, by plotting $S(\varphi)$ and $\partial_\varphi S(\varphi)$ we can effectively see that the system undergoes a genuine phase transition.

3.2. Overlap of time scales is necessary for a segregation transition to exist. We now discuss more in detail an essential ingredient for a segregation sharp transition to exist, the fact that the conviction move occurs on the same time scale of the popularity move, regardless of the size of heteorogeneous edges in the system. In other words, the conviction move is realized at each step with

probability φ drawing directly from the basket of heterogeneous edges in order to observe the transition.

We can understand this result by considering a similar model in which the opinion move is, for instance, defined as follows. Select an edge randomly among all the M edges of the network (rather then from the basket of the heterogeneous ones) and if the edge is heterogeneous execute the conviction move, otherwise leave the network unaltered and move on by executing a new step. In this model the mean-field equation, Eq. 226 will take an additional term representing the heterogeneous edge density multiplying the conviction move term,

$$\Delta \langle \Omega_t \rangle = \underbrace{\vartheta \varphi \frac{M - \langle \Omega_t \rangle}{M}}_{\text{op. move variant}} + (1 - \varphi) \left[\vartheta p_t^+(h) - p_t^-(h) \right] . \tag{237}$$

The critical value φ_c is found setting $\Delta \langle \Omega_t \rangle$ to zero and the average value of the order parameter saturates to its maximum value M. Substituting these quantities one immediately finds that the contribution of the opinion move disappears, leaving us with the equation $(1 - \varphi_c) \left[\vartheta p_t^+(h) - p_t^-(h) \right] = 0$ which has the only trivial solution $\varphi_c = 1$ (that represents a model in which only opinion based move are executed). In other words, a segregated phase is found only in the trivial case where the agents only choose their connections by conviction.

This analysis also gives a general condition for the existence of a transition, which is that the conviction move has to be such that the multiplicative factor introduced in the opinion move term in Eq. (237) translates into a function $f(\Omega_t)$ characterized by the condition $f(M) \neq 0$). A possible justification for this forcing in the opinion move can be found by considering some realistic situations characterized by a segregation phenomenon driven by strong convictions (ethnicity, political orientation, religious beliefs, etc.). If an agent is left only with opposite minded neighbors, it is likely going to be the first one to decide to sever a connection and rewire with someone with the same conviction. For this reason, we believe that direct targeting of heterogeneous connection in an environment of strong convictions might be a realistic assumption. We also note that the mean-field equations apply for fluctuating outdegrees, as long as the fluctuations are controlled, so that this behavior is more general than the fixed outdegree model that we consider here. Simulations with fluctuating outdegrees fully support this statement (Figure 10). We expect that in a realistic situation the outdegree (number of individuals that one individual follows) is much more compact than the in degree (number of followers).

3.3. The popularity move broadens the in-degree distribution in the unsegregated phase, but does not affect the transition point. We proceed by considering the role of the popularity move in setting the in-degree distribution and in the segregation transition. The initial random graph $G_0(N, m, h)$ has by definition Poisson-distributed in-degrees k_{in} for large N, with a mean equal do the fixed outdegree of every node of the network m. As the network evolves, the distribution of the in-degrees changes at each popularity move, because the most popular nodes are more likely to be chosen as a target for the newly created edges. This determines a departure from the initial distribution towards heavier-tailed distributions, in analogy with the "rich gets richer" principle that usually characterizes social networks [**116**]. In order to properly characterize this behavior evaluated the empirical survival distribution function (ESDF) of the in-degree distributions of evolved graphs G_t for different values of the choice parameter. The ESDF indicates the probability of observing a node i with in-degree $k_{in}(i)$ greater

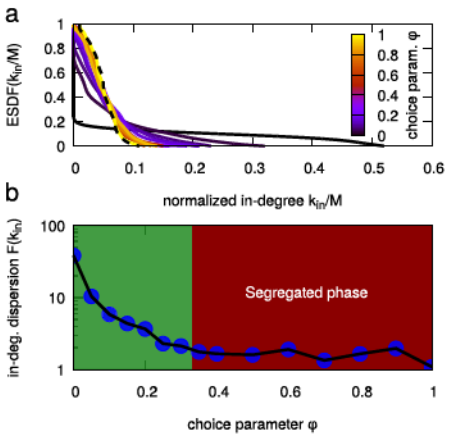


FIGURE 4. Preferential attachment from the popularity move broadens hte in-degree distribution. **a)** Empirical survival distribution function (ESDF) of the in-degree distributions of networks evolved for different values of φ. The plot was obtained by evolving an initial random graph $G_0(N = 100, m = 5, h = 1/2)$ for $t = 10^6$ steps (the in-degrees are normalized with respect to the total number of edges $M = 500$). The broadening of the distribution indicates the increasing presence of bigger attractors in the evolved networks. **b)** Two different trends for the Fano factor of the in-degrees are observed in the regions below and above the segregation transition. The plot reports the Fano factor of the in-degrees distributions shown in panel a versus the choice parameter φ. In the region above the critical value of the choice parameter $\varphi_c = 1/3$ the deviation from a Poisson distribution ($F(k_{in}) = 1$) is small, while the unsegregated region shows a super-exponential departure (the vertical axis is in log-scale) towards larger dispersions as φ decreases.

then a certain value k_{in}, and is defined as

$$\text{ESDF}(k_{in}) = \frac{1}{M}\sum_{i=0}^{M}\theta(k_{in} - k_{in}(i)) \ , \tag{238}$$

Fig. 4a shows that when $\varphi = 1$ the initial distribution is unaltered (the dashed line represents the distribution for the initial random graph G_0), but as φ decreases the in-degree distributions take increasingly heavier tails.

The same phenomenon can be quantified by a single broadness parameter such as the Fano factor of the in-degrees $F(k_{in})$, defined as

$$F(k_{in}) = \frac{\text{Var}[k_{in}]}{\langle k_{in}\rangle} \ . \tag{239}$$

This parameter is 1 for a Poisson distribution, whereas greater values indicate larger dispersion. Fig. 4 shows this parameter plotted as a function of the choice parameter φ. The Fano Factor increases as popularity-based moves become more probable (as φ goes to zero). Moreover two different trends appear to characterize the region below and above the critical value $\varphi_c = 1/3$.

Finally, although we found that popularity-based rewiring increases the dispersion of social connections in the unsegregated regime, this preferential attachment ingredient

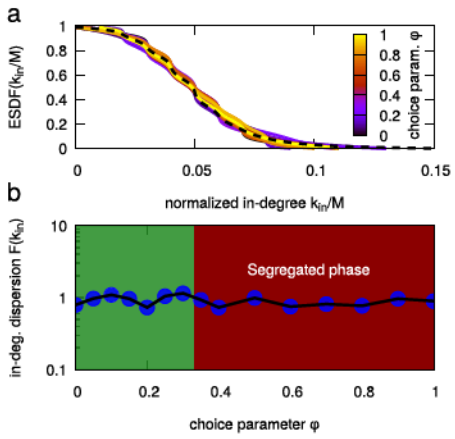


FIGURE 5. Skewed node popularity does not affect segregation. **ab)** Same plots as Fig. 4, for a model in which the popularity move is changed with a rewiring on a uniformly chosen random node. This model shows the same phase transition as the original one (and in particular the plots in Fig. 2bc are identical), but the transition is not accompanied by changes in node degree.

does not affect the segregation transition in any way, as we have verified by substituting popularity-based rewiring with random rewiring in our simulations (Fig. 5). Although one may expect that the presence of popular individuals may help avoiding the emergence of segregation due to their capacity of attracting new nodes regardless of their opinion, this does not happen in this model. The reason is easily understood from Eq. (226) and (237), which govern the dynamics of the order parameter, where it is clear that the in-degree distribution never comes into play.

3.4. Minority convictions segregate more easily. The results presented up to this point were obtained under the hypothesis of equally represented convictions condition ($h = 1/2$). A more generic case describes minority versus majority convictions, characterized by different values of h. The differences from the symmetric case concern both the characteristic time τ_Ω needed to reach the steady state and the critical value φ_c at which the transition to a segregated phase occurs.

In order to study this asymmetric situation we write a mean-field equation valid for every value of $h \in [0,1]$. Starting from Eq. 226, we just need to specify how the terms $p_t^{\pm}(h)$ depend on h (see section 5.1),

$$p_t^+(h) = \frac{M - \langle \Omega_t \rangle}{2M}$$
$$p_t^-(h) = \frac{h(1-h)}{h^2 + (1-h)^2} \frac{\langle \Omega_t \rangle}{M} \,. \tag{240}$$

The resulting mean-field equation can be integrated in the continuum limit as in the symmetric case $h = 1/2$, yielding the dynamics of the average value of the order parameter. The critical value φ_c on the asymmetry h is obtained again by imposing the segregation regime conditions $\Delta \langle \Omega_t \rangle = 0$ and $\langle \Omega_t \rangle = M$. Solving for φ gives

$$\varphi_c(h) = \frac{h(1-h)}{1 - h(1-h)} \tag{241}$$

for the critical value. This relation satisfies the red-blue symmetry $\varphi_c(h) = \varphi_c(1-h)$ with maximum value $\varphi_c(1/2) = 1/3$ (as in Eq. 230) for the symmetric case.

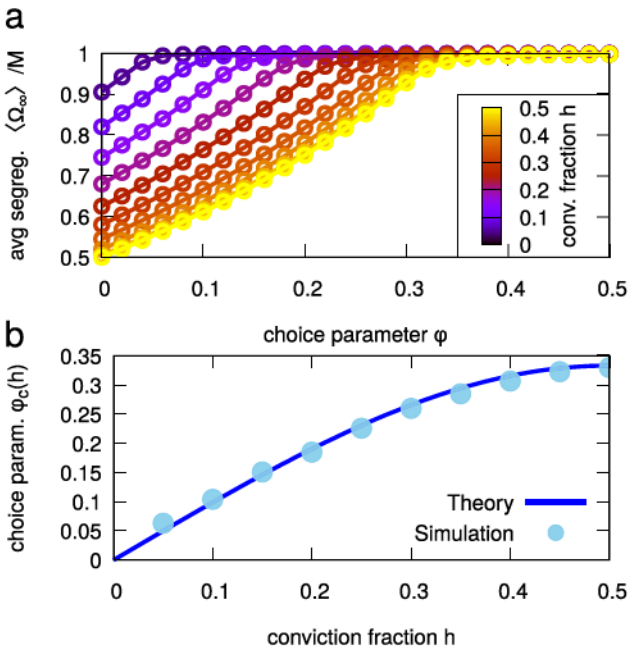


FIGURE 6. Minority convictions tend to segregate more easily. **a)** Average value of the order parameter $\langle \Omega_t \rangle$ in networks evolved from initial networks $G_0(100, 5, h)$ for different values of $h \leq 1/2$ (the results for $h > 1/2$ are the same due to the symmetry $h \to 1 - h$). As the the density of nodes holding a certain conviction decreases, the networks will reach a segregated phase for lower values of φ. **b)** Simulations confirm the analytical prediction for the critical points of the model. The critical points (symbols) are ectracted from the curves in panel a, for different values of h, and compared with the prediction described by Eq. 241 (solid line).

Fig. 6b compares the predicted critical point from Eq. 241 to simulations of evolved networks for different values of h Fig. 6a. This analysis shows that a situation characterized by a minority conviction favors segregation for lower values of the choice parameter, indicating that the symmetric situation is the one in which segregation can be more easily avoided (the situation is analogous to the miscibility gap for phase segregation in a binary mixture).

The characteristic duration of the transient before a steady state is reached is also affected by the presence of a minority conviction. The solution of the mean-field equation gives

$$\tau_\Omega(h) = \frac{2M\left[h^2 + (1-h)^2\right]}{1-\varphi}\,, \tag{242}$$

i.e., the characteristic relaxation time will increase for asymmetric convictions. This time scale is important in cases where the segregation dynamics competes with the spreading of consensus [**122, 123**].

3.5. Scale-invariance close to the transition. The limit of large system size, $M \to \infty$, is better analyzed in terms of a finite-size scaling ansatz, typical of

critical phenomena [**129, 130**]. We define the normalized choice parameter

$$t = \frac{\varphi - \varphi_c}{\varphi_c} \,. \tag{243}$$

and the intensive order parameter

$$m = \frac{M - \Omega_\infty}{M} \tag{244}$$

so that

$$\langle m \rangle = 1 - \frac{\langle \Omega_\infty \rangle}{M} = \langle \frac{M - \Omega_\infty}{M} \rangle \tag{245}$$

and we assume that $\langle m \rangle$, which in principle depends on both M and t separately, is an homogeneous function of t and a suitable power of M, that is

$$\langle m \rangle = |t|^\beta \tilde{f}_1(M^y t) \tag{246}$$

in the large (small) M (t) limit with $M^y t$ fixed. y and β are exponents that are expected to be independent of the microscopic details of the dynamical model, characterizing the transition point, while f is a scaling function, which might depend on the model specificities. Since we expect that m is non-zero (zero) for $t < 0$ ($t > 0$) the scaling function f should behave asymptotically as

$$\lim_{x \to +\infty} \tilde{f}_1(x) = 0, \qquad \lim_{x \to -\infty} \tilde{f}_1(x) = \text{constant} > 0 \tag{247}$$

In order to estimate the two scaling exponents β and y, we plot $m|t|^{-\beta}$ versus $M^y t$ and determine the exponents so that the best collapse of the different curves is obtained. Indeed one should obtain a different curve for each value of M as t varies and this is what we observe for generic pair β and y. However for $\beta = 1$ and $y = 1/2$ the various curves collapse in a range of $x \equiv M^y t$ that increases as M becomes larger and larger as Fig.7, panel (a), shows.

The same analysis leads to the following scaling ansatz for the variance of m (corresponding to $\mathrm{Var}[\Omega_\infty]/M^2$) in terms of the original extensive order parameter):

$$\mathrm{Var}[m] = t^2 \tilde{f}_2(M^{1/2} t) \tag{248}$$

and the corresponding collapse is shown in Fig.7, panel (b). Both scaling Eqs.(246) and (248) are captured by the more general scaling ansatz of the distribution function of m

$$P(m, t, M) = |t|^{-1} \tilde{P}(m t^{-1},\ M^{1/2} t) \tag{249}$$

3.6. A model with pure intra-specific aversion leads to an equivalent segregation threshold behavior. Motivated by the literature on segregation models based on aversion between unlike individuals [**39, 125**], we asked whether the same threshold phenomenon observed in our model could be present in case of conviction moves that were based purely on aversion bias.

To this end, we defined a variant of our model where the conviction move (with probability φ) chooses randomly one heterogeneous edge, between two nodes holding different convictions and rewires it to a random node. In this variant, the popularity move (with probability $1 - \varphi$ at each step) remains the same. Under this variant, Eq. (226) becomes

$$\Delta \langle \Omega_t(\varphi, h) \rangle = \underbrace{\vartheta \frac{\varphi}{2}}_{\text{conv. move}} + \underbrace{(1 - \varphi) \left[\vartheta p_t^+(h) - p_t^-(h) \right]}_{\text{pop. move}} , \tag{250}$$

immediately leading to the expression,

$$\frac{\langle \Omega_\infty(\varphi) \rangle}{M} = \min_{\varphi \in [0;1)} \left\{ 1, \frac{1}{2(1 - \varphi)} \right\} \tag{251}$$

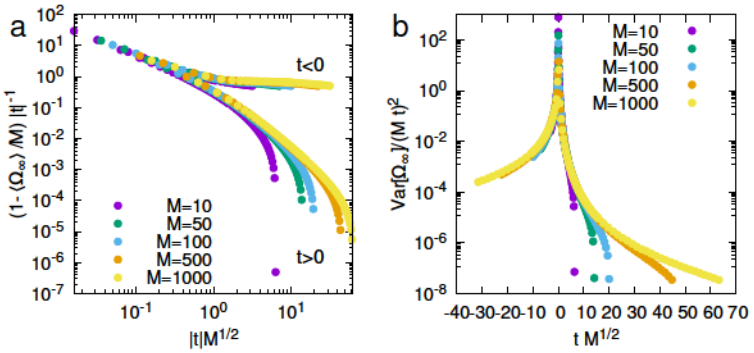


FIGURE 7. The fraction of homogeneous edges and its variance obey scaling. a) Scaling collapse for the fraction of homogenous edges. b) Scaling collapse for the variance. The x and y axes of both plots compare the functions predicted by Eqs. 246 and 248. The symbols correspond to data points from simulations at different network size above and below the segragation transition point.

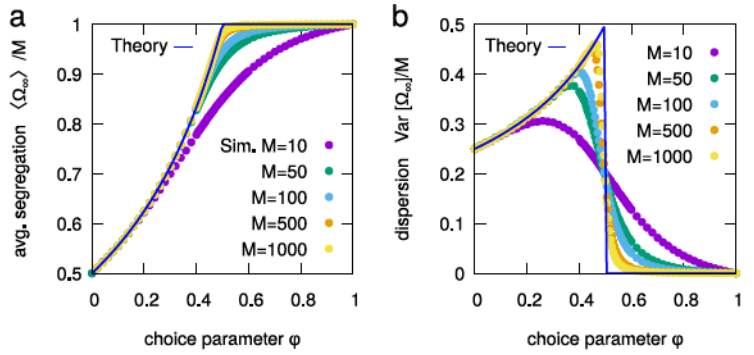


FIGURE 8. The sudden transition to a segregated state remains in a model with aversion bias only. a) Mean order parameter at steady state versus the choice parameter φ comparing theory (solid line) with numerical simulations for different sizes of the network M (symbols). This analysis supports a segregation transition for $\varphi_c = 1/2$ (for $h = 1/2$). b) The dispersion of the order parameter (symbols) shows the same behavior as the standard model (compare with Fig 2).

for the mean fraction of heterogeneous edges.

By setting $\langle \Omega_\infty(\varphi) \rangle = 1$ in Eq. 251 and solving for φ one finds again the critical value, which for $h = 1/2$ is $\varphi_c = 1/2$. An analogous reasoning can be followed for solving for the higher moments of the distribution of Ω. Fig. 8 shows that direct simulations of the aversion bias model are fully in line with these theoretical predictions. Thus, we conclude that aversion alone is sufficient to produce a sudden segregation threshold.

4. Discussion and Conclusions

Social segregation is ubiquitous in our society, and manifests itself as fragmentation of social networks at all scales, in countries, cities, schools, firms, governmental agencies, etc. Its consequences may lead to a wide range of nefastous phenomena ranging from inefficient planning to war. It is driven by diverse and enormously complex sociological, cultural, environmental and economic dilemmas, which are unlikely to be solved in the near future. However, since the pioneering work of Schelling there is increasing agreement that there may be common quantitative traits in the "macroscopic" dynamics of segregation that emerge from this complexity [**40, 41, 42, 125, 131**]. A quantitative understanding of the consequences of such simple features on the dynamics of a social network may be important to develop efficient estimators to be used in real-life examples to detect and prevent segregation phenomena.

The framework developed here shows that complete segregation in a network setting without any spatial aspects can emerge as a threshold phenomenon that corresponds to a genuine phase transition. Close to such transition point, small perturbations of the system can cause very large rearrangements in the state. Importantly, we have shown that such transition point is scale invariant, hence "universal" in the statistical physics sense. This supports the hypothesis that close to this critical point more detailed descriptions of social interactions are not necessary, since a wide class of models may behave similarly. Our model is different from standard Schelling models, as we put ourselves in the conditions where spatial structure is not relevant (which is applicable for example to well-mixed social structures and on-line social connections), and more similar to the model of ref. [**123**], which was built to explore the relative roles of conviction rewirings with opinion dynamics, and did not characterize the conditions leading to a segregated state in the limit where convictions dominate.

By studying the model in general setting, we have shown that the competition of time scales between different rewring moves is essential in determining the possibility of having a phase transition (a phenomenon that one would want to avoid in a real situation) *versus* a crossover to the segregated state. In empirical cases of social dynamics, not much is known about the time scales of the different rewiring moves and there is no *a priori* reason why a model should take one assumption or the other. It would be interesting to quantify this trends in empirical data. Potentially the real-world values of these time scales could vary form case to case, and bridge between these extreme-case scenarios. Additionally, their ratio of the two time scales is a parameter that may be acted upon.

We can also parallel this model with available physical models for the separation of phases and mixtures. For example, binary mixtures can be described in a coarse-grained way as a set of particles of two kinds filling a cubic lattice, with an energy cost for particles of one kind sitting next to particles of the other kind. This system (equivalent to an Ising model) shows a spatial phase separation when temperature is lowered. Contrary to this case, in our model set on a network a concept of distance is missing, since all individuals can potentially interact with any other agent in each move. However, we can parallel our results to a variant of the above model where instead of the usual "local" fraction of lattice sites occupied by each kind of particle, we write the free energy in terms of the parameter used here, i.e., the fraction of homogeneous edges $e_h = -\Omega/M$. The energetic term is simply $-\chi e_h$. In order to write the entropy, we consider the network as a gas of edges formed by connecting nodes. We compute the number of ways to assign Ω edges out of M, considering that each edge is spurious if two colors of the same kind are selected. The resulting free energy is $\beta F = e_h \log(e_h) + (1 - e_h) \log(1 - e_h) - e_h \chi$. Minimizing this free

energy and comparing with the equations governing our model shows that they are different, and our model cannot be reconducted to this simple case. The question remains open on whether there is a simple equilibrium model recapitulating the phase-separation behavior shown by our segregation model.

Segregation in social networks may be driven by both homophyly (the choice of social interactions with like individuals) and aversion. These ingredients are mixed in different proportion in the existing literature. Our basic model contains both, since in the conviction-based rewirings interactions between dissimilar partners are rewired in favor of homogeneous ones. Schelling's model [**39**] shows that aversion from dissimilar network partners alone, coupled with a random selection of new partners, may be sufficient to induce segregation. Our analysis of a model variant where the conviction-based rewiring is based on pure aversion supports this conclusion. Indeed, this variant shows the same type of threshold phenomenon, in full quantitative agreement with the main model. The (expected) quantitative change is that in the case of pure aversion the transition point is shifted to higher values of the choice parameter φ, compared to the case where both aversion and homophyly are in place.

Overall, our analysis supports the conclusion that whether conviction-based rewiring is based on aversion or homophyly is not a key ingredient for the existence of a segregation threshold. Instead, the important feature to determine a threshold phenomenon for segregation is that the the conviction-based rewiring of the network (based on aversion or homophyly, or both) occurs on the same time scale of the popularity-based rewirings (i.e. the establishment of social interactions that are non-discriminant). In the alternative scenario in which, e.g., each kind of rewiring occurs proportionally to the number of extant interactions, segregation occurs smoothly. In such situation, at all levels of the bias in establishing interactions (quantified by the choice parameter φ) the network maintains a finite fraction of interactions between dissimilar individuals.

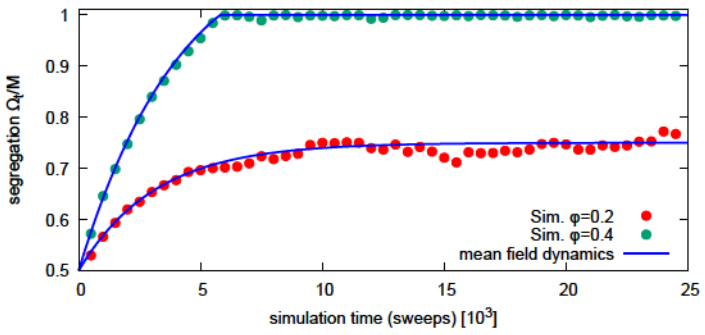


FIGURE 9. Relaxation dynamics and criterion for equilibrium. The plot shows, for a starting graph G_0 ($N = 500, m = 5, h = 1/2$), a time series of Ω_t/M in the transient regime for two different values of φ, above and below the critical value. The mean-field predictions (solid lines) perfectly agree with simulations. Thus, the relaxation times of the solution of the time-dependent mean-field equations provide a good criterion for stationarity.

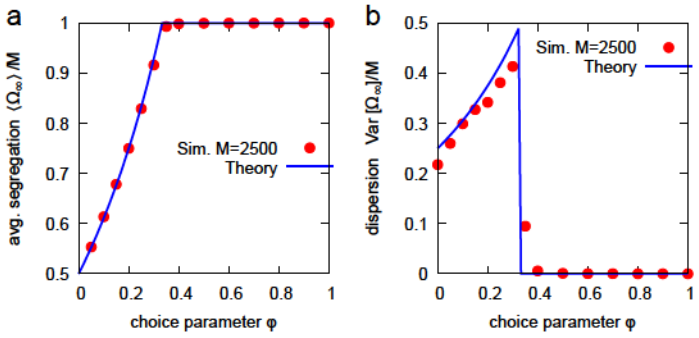


FIGURE 10. The results hold for fluctuating outdegrees. The plots show mean and dispersion of the order parameter, obtained with an initial Erdos-Renyi network with 250 nodes and 2500 directed edges.

5. Analytical calculations and additional figures

This section presents in further detail the two different methods used to derive the analytic expressions for the cumulants of the order parameter (namely equations 230 and 235).

5.1. Mean-field approach. As previously explained, the mean-field approach consists in quantifying the average variation of the order parameter at every step of the dynamics, which resulted in equation 226. The meaning of the terms of such equation have already been discussed, here we will present the more general derivation of the contributions $p_t^{\pm}(h)$ for every $h \in [0,1]$, which will yield the more general solution of equation 228 for different densities of colored nodes.

The terms $p_t^{\pm}(h)$ represent the probabilities of, respectively, increasing and decreasing the order parameter Ω when a popularity move is performed:

$$\begin{aligned} p_t^+(h) &= \text{Prob}\left[e_t(rb) \to e_t(rr)\right] + \text{Prob}\left[e_t(br) \to e_t(bb)\right] \\ p_t^-(h) &= \text{Prob}\left[e_t(rr) \to e_t(rb)\right] + \text{Prob}\left[e_t(bb) \to e_t(br)\right] \end{aligned} \tag{252}$$

which are found to be

$$\begin{aligned} p_t^+(h) &= \frac{M - \langle \Omega_t(\varphi,h)\rangle}{2M} \\ p_t^-(h) &= \frac{\langle \Omega_t(\varphi,h)\rangle}{M}\frac{h(1-h)}{h^2+(1-h)^2} \end{aligned} \quad . \tag{253}$$

By substituting these coefficients in equation 228 and taking the continuous-time limit we obtain the following differential equation,

$$\partial_t \langle \Omega_t(\varphi,h)\rangle = \vartheta\frac{1+\varphi}{2} - (1-\varphi)\frac{2h(1-h)(1-\vartheta)+\vartheta}{2\left(h^2+(1-h)^2\right)}\frac{\langle \Omega_t(\varphi,h)\rangle}{M} \quad , \tag{254}$$

which can be explicitly integrated in time (for $\varphi \neq 1$), yielding

$$\begin{aligned} \frac{\langle \Omega_t(\varphi,h)\rangle}{M} = \left[1+\vartheta\left(\frac{\langle \Omega_0(\varphi,h)\rangle}{M}-1\right) - \vartheta\frac{1+\varphi}{2}\alpha(\varphi,h)\right] \cdot \\ \cdot e^{-\frac{t}{\alpha(\varphi,h)}} + \vartheta\frac{1+\varphi}{2}\alpha(\varphi,h) \quad , \end{aligned} \tag{255}$$

where the initial condition is

$$\frac{\langle \Omega_0(\varphi,h)\rangle}{M} = e_0(rr) + e_0(bb) = h^2 + (1-h)^2 \tag{256}$$

and the coefficient α is

$$\alpha(\varphi, h) = \frac{2(h^2 + (1-h)^2)}{(1-\varphi)(2h(1-h)(1-\vartheta)+\vartheta)} \,. \tag{257}$$

If we evaluate this coefficient in the unsegregated phase (where $\vartheta \equiv 1$), we obtain the characteristic time of the transient phase, which is

$$\tau(\varphi, h) = \frac{2(h^2 + (1-h)^2)}{1-\varphi} \tag{258}$$

Taking the limit $t \to \infty$ of equation 255 yields the steady-state solution of the order parameter, which for every $\varphi \in [0,1)$ and $h \in [0,1]$ is,

$$\frac{\langle \Omega_t(\varphi, h) \rangle}{M} = \min\left\{1, \frac{1+\varphi}{1-\varphi}\left(h^2 + (1-h)^2\right)\right\} \,. \tag{259}$$

Fig. 6 shows the phase diagram for $\langle \Omega_t \rangle$, which is in agreement with the fact that the critical value of the choice parameter φ_c becomes lower as we move away from the symmetric nodes density given by $h = 1/2$ (discussed in section 3.4).

5.2. Master equation and moment-generating function approach. This section treats in further detail the derivation of a generic factorial moment of the order parameter Ω. Substituting the rates 232 in the master equation 231 one gets,

$$\begin{aligned} \partial_t P_t(\Omega) = P_t(\Omega-1)\left[\varphi + (1-\varphi)\frac{M-\Omega+1}{2M}\right] + \\ + P_t(\Omega+1)(1-\varphi)\frac{\Omega+1}{2M} - P_t(\Omega)\frac{1+\varphi}{2} \end{aligned} \tag{260}$$

In order to find a differential equation for the FMGF 233 we first multiply by s^Ω both sides of equation 260, and then we sum over the order parameter Ω itself. The probabilities $P_t(\Omega)$ are obviously defined only for $\Omega \in [0, M]$, so we need to explicitly set $P_t(\Omega) \equiv 0$ when Ω is outside that range. This notation has a practical advantage that allows us to extend the summation over Ω from the range $[0, M]$ to the range $[-1, M+1]$. This frees from border-term issues when re-indexing the summation for the terms on the right side. To evaluate the contribution with the $P_t(\Omega - 1)$ coefficient, we set $\Omega' = \Omega - 1$ and obtain

$$\begin{aligned} \sum_{\Omega'=-2}^{M} s^{\Omega'+1} P_t(\Omega')\left[\frac{1+\varphi}{2} - \frac{\Omega'}{2M}\right] = \\ = \left[s\frac{1+\varphi}{2} - \frac{1-\varphi}{2M}s^2\partial_s\right] G(s,t) \,, \end{aligned} \tag{261}$$

where we introduced a derivative in s in order to eliminate the multiplicative Ω' in the summation. The same trick can be used for the $P_t(\Omega+1)$ term (this time we set $\Omega' = \Omega + 1$):

$$\sum_{\Omega'=0}^{M+2} s^{\Omega'-1}(1-\varphi)P(\Omega')\frac{\Omega'}{M} = \frac{1-\varphi}{2M}\partial_s G(s,t) \tag{262}$$

Finally, the $P_t(\Omega)$ term does not require any re-indexing and immediately yields $G(s,t)(1+\varphi)/2$. Putting all the pieces together we finally find the desired equation 234 for the dynamics of the FMGF.

Equation 234 is a partial differential equation that contains derivatives both in s and t. Since we are only interested in finding the moments of the equation, we can avoid solving it explicitly: if we evaluate $\partial_s^n[\partial_t G(s,t)|_{s=1}]$ for every $n \in \mathbb{N}$ we

obtain a closed system of time-only differential equations for the dynamics of the moments. In fact we can easily see that

$$\partial_s^n G(s,t)|_{s=1} = \langle \frac{\Omega!}{(\Omega-n)!} \rangle \tag{263}$$

For $n = 1$, we are evaluating the first factorial moment, which coincides with the average. A straightforward calculation shows that we obtain precisely equation 227 (in the unsegregated phase with $\vartheta \equiv 1$). For $n = 2$, we find the equation of the second factorial moment $\langle \Omega(\Omega-1) \rangle = \langle \Omega^2 \rangle - \langle \Omega \rangle$, which reads

$$\partial_t \langle \Omega^2 \rangle - \partial_t \langle \Omega \rangle = -2\frac{1-\varphi}{M}\langle \Omega^2 \rangle + \left(1+\varphi+\frac{1-\varphi}{M}\right)\langle \Omega \rangle \tag{264}$$

By evaluating the steady-state solution ($\partial_t \langle \Omega^2 \rangle = \partial_t \langle \Omega \rangle = 0$) of this equation and substituting the steady-state form of $\langle \Omega \rangle$, we find the steady-state equation of $\langle \Omega^2 \rangle$, which in turn gives us the variance

$$\mathrm{Var}\,[\Omega] = \langle \Omega^2 \rangle - \langle \Omega \rangle^2 = \frac{1+\varphi}{4(1-\varphi)} \ . \tag{265}$$

This equation coincides with the one presented in equation 235 (in the unsegregated phase).

5.3. Full Stationary Solution. Starting from the Master Equation (260) we can write the full equation for the Generating Function $G(s,t)$

$$\partial_t G(s,t) = aG(s,t)(s-1) + b\partial_s G(s,t)\left(1-s^2\right) \ , \tag{266}$$

where $a = \frac{1+\varphi}{2}$ and $b = \frac{1-\varphi}{2M}$; M is the total number of links. We assume the initial condition ($P(\Omega, t=0) = \delta_{\Omega - M/2}$ and thus we have $G(s,0) = s^{M/2}$. Additionally, the normalisation condition fixes $G(1) = 1$.

The stationary solution for Eq. (266) is simple to find by solving directly the PDE, and leads to

$$G(s) = \left(\frac{1+s}{2}\right)^{a/b} \ . \tag{267}$$

In order to solve the full transient of the PDE (266) we use the so-called method of characteristics. Setting $f(s) = -b\left(1-s^2\right)$, then Eq. (266) corresponds to the following system of differential equations:

$$\dot{s}(t) = f(s) \tag{268}$$

$$\frac{d}{dt}G(s(t),t) = a(s(t)-1)G(s(t),t) \ . \tag{269}$$

Eq. (268) leads to the integral equation $-\int_{s(0)}^{s} \frac{dz}{1-z^2}\,dz = \int_0^t b dt$ where s it evaluated at a final time t, i,e, $s(t) = s$. Solving this equation leads to

$$s = \frac{\mathrm{Cosh}(bt)s(0) - \mathrm{Sinh}(bt)}{\mathrm{Cosh}(bt) - s(0)\mathrm{Sinh}(bt)} \tag{270}$$

and

$$s(0) = \frac{s\mathrm{Cosh}(bt) + \mathrm{Sinh}(bt)}{\mathrm{Cosh}(bt) + s\mathrm{Sinh}(bt)} \ . \tag{271}$$

Finally, performing the integral $\int_{s(0)}^{s} a(s(\tau)-1)d\tau$ we find

$$G(s,t) = e^{-at}(\mathrm{Cosh}(bt) + s\mathrm{Sinh}(bt))^{a/b} \cdot \left(\frac{s\mathrm{Cosh}(bt) + \mathrm{Sinh}(bt)}{\mathrm{Cosh}(bt) + s\mathrm{Sinh}(bt)}\right)^{M/2} \ . \tag{272}$$

In the limit t$\to \infty$, this expression gives the stationary solution Eq. (267). Expanding this in series around $s = 0$, and matching term by term, one can find

the transient solution $P(\Omega, t)$. In fact, we have that $G(s,t) = P(0,t) + sP(1,t) + ... + s^M P(M,t)$ and $G(0,t) = P(0,t)$. Expanding the steady state solution of $G(s)$ in series around $s = 0$, we obtain $G(s) = \sum_{\Omega=0}^{M} \frac{\partial_s^\Omega G(s)_{|s=0}}{\Omega!} s^\Omega$ leading to Eq. (236) in the main text. We highlight that Eq. (236) only holds for $\varphi \in [0, \varphi_c)$.

Bibliography

[1] Michele Caraglio, Fulvio Baldovin, and Attilio L Stella. Export dynamics as an optimal growth problem in the network of global economy. *Scientific reports*, 6, 2016.

[2] Gianluca Teza, Michele Caraglio, and Attilio L Stella. Growth dynamics and complexity of national economies in the global trade network. *Scientific reports*, 8(1):15230, 2018.

[3] Gianluca Teza, Michele Caraglio, and Attilio Stella. Data driven approach to the dynamics of import and export of g7 countries. *Entropy*, 20(10):735, 2018.

[4] Gianluca Teza, Stefano Iubini, Marco Baiesi, Attilio L Stella, and Carlo Vanderzande. Rate dependence of current and fluctuations in jump models with negative differential mobility. *arXiv preprint arXiv:1904.05241*, 2019.

[5] Gianluca Teza, Samir Suweis, Marco Gherardi, Amos Maritan, and Marco Cosentino Lagomarsino. Network model of conviction-driven social segregation. *Physical Review E*, 99(3):032310, 2019.

[6] Philip W Anderson. More is different. *Science*, 177(4047):393–396, 1972.

[7] C.A. Hidalgo and R. Hausmann. The building blocks of economic complexity. *Proc. Natl. Acad. Soc. USA*, 106:10570, 2009.

[8] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero. A new metrics for countries' fitness and products' complexity. *Sci. Rep.*, 2:723, 2012.

[9] C.E. Shannon. A mathematical theory of communication. *The Bell System Tech. Journal*, 27:379, 1948.

[10] Bela Balassa. Comparative advantage in manufactured goods: a reappraisal. *The Review of Economics and Statistics*, pages 315–319, 1986.

[11] C.A. Hidalgo, B. Klinger, A.-L. Barabási, and R. Hausmann. The product space conditions the development of nations. *Science*, 317:482, 2007.

[12] Virginia Domínguez-García and Miguel A Munoz. Ranking species in mutualistic networks. *Scientific reports*, 5:8182, 2015.

[13] Luke R Thompson, Jon G Sanders, Daniel McDonald, Amnon Amir, Joshua Ladau, Kenneth J Locey, Robert J Prill, Anupriya Tripathi, Sean M Gibbons, Gail Ackermann, et al. A communal catalogue reveals earth's multiscale microbial diversity. *Nature*, 551(7681):457, 2017.

[14] Giovanni Gallavotti. Chaotic dynamics, fluctuations, nonequilibrium ensembles. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 8(2):384–392, 1998.

[15] Giovanni Gallavotti and Ezechiel Godert David Cohen. Dynamical ensembles in stationary states. *Journal of Statistical Physics*, 80(5-6):931–970, 1995.

[16] Giovanni Gallavotti. New methods in nonequilibrium gases and fluids. *Open Systems & Information Dynamics*, 6(2):101–136, 1999.

[17] Joel L Lebowitz and Herbert Spohn. A gallavotti–cohen-type symmetry in the large deviation functional for stochastic dynamics. *Journal of Statistical Physics*, 95(1-2):333–365, 1999.

[18] Christopher Battle, Chase P Broedersz, Nikta Fakhri, Veikko F Geyer, Jonathon Howard, Christoph F Schmidt, and Fred C MacKintosh. Broken detailed balance at mesoscopic scales in active biological systems. *Science*, 352(6285):604–607, 2016.

[19] J Gladrow, N Fakhri, FC MacKintosh, CF Schmidt, and CP Broedersz. Broken detailed balance of filament dynamics in active networks. *Physical review letters*, 116(24):248301, 2016.

[20] Saar Rahav and Christopher Jarzynski. Fluctuation relations and coarse-graining. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(09):P09012, 2007.

[21] A Puglisi, S Pigolotti, Lamberto Rondoni, and A Vulpiani. Entropy production and coarse graining in markov processes. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(05):P05015, 2010.

[22] RKP Zia, Eigil Luxhøj Præstgaard, and OG Mouritsen. Getting more from pushing less: Negative specific heat and conductivity in nonequilibrium steady states. *American Journal of Physics*, 70(4):384–392, 2002.

[23] Bart Cleuren and Christian Van den Broeck. Brownian motion with absolute negative mobility. *Physical Review E*, 67(5):055101, 2003.
[24] Pulak K Ghosh, Peter Hänggi, Fabio Marchesoni, and Franco Nori. Giant negative mobility of janus particles in a corrugated channel. *Physical Review E*, 89(6):062115, 2014.
[25] Sebastian Leitmann and Thomas Franosch. Nonlinear response in the driven lattice lorentz gas. *Physical Review Letters*, 111(19):190603, 2013.
[26] Urna Basu and Christian Maes. Mobility transition in a dynamic environment. *Journal of Physics A: Mathematical and Theoretical*, 47(25):255003, 2014.
[27] O Bénichou, P Illien, G Oshanin, A Sarracino, and R Voituriez. Microscopic theory for negative differential mobility in crowded environments. *Physical Review Letters*, 113(26):268002, 2014.
[28] Marco Baiesi, Attilio L Stella, and Carlo Vanderzande. Role of trapping and crowding as sources of negative differential mobility. *Physical Review E*, 92(4):042121, 2015.
[29] Pierre Illien, Olivier Bénichou, Gleb Oshanin, Alessandro Sarracino, and Raphaël Voituriez. Nonequilibrium fluctuations and enhanced diffusion of a driven particle in a dense environment. *Physical Review Letters*, 120(20):200606, 2018.
[30] DI Aladashvili, ZA Adamiya, KG Lavdovskii, EI Levin, and BI Shklovskii. Negative differential resistance in the hopping conductivity region in silicon. *JETP Letters*, 47(8):466–469, 1988.
[31] Cheng Cen, Stefan Thiel, Jochen Mannhart, and Jeremy Levy. Oxide nanoelectronics on demand. *Science*, 323(5917):1026–1030, 2009.
[32] Xiao-Fei Li, Lingling Liu, Qing Yan, Qin-Kun Li, Yunxiang Wang, Mingsen Deng, and Qi Qiu. Strong current polarization and perfect negative differential resistance in few-fen 4-embedded zigzag graphene nanoribbons. *Physical Chemistry Chemical Physics*, 19(4):2674–2678, 2017.
[33] Viet Hung Nguyen, Jérôme Saint-Martin, Damien Querlioz, Fulvio Mazzamuto, Arnaud Bournel, Yann-Michel Niquet, and Philippe Dollfus. Bandgap nanoengineering of graphene tunnel diodes and tunnel transistors to control the negative differential resistance. *Journal of Computational Electronics*, 12(2):85–93, 2013.
[34] Björn Åkerman and Kenneth D Cole. Electrophoretic capture of circular dna in gels. *Electrophoresis*, 23(16):2549–2561, 2002.
[35] Davide Michieletto, Marco Baiesi, Enzo Orlandini, and Matthew S Turner. Rings in random environments: sensing disorder through topology. *Soft Matter*, 11(6):1100–1106, 2015.
[36] Davide Michieletto, Davide Marenduzzo, and Enzo Orlandini. Topological patterns in two-dimensional gel electrophoresis of dna knots. *Proceedings of the National Academy of Sciences*, 112(40):E5471–E5477, 2015.
[37] Stefano Iubini, Enzo Orlandini, Davide Michieletto, and Marco Baiesi. Topological sieving of rings according to their rigidity. *ACS Macro Letters*, 7(12):1408–1412, 2018.
[38] Gianmaria Falasco, Tommaso Cossetto, Emanuele Penocchio, and Massimiliano Esposito. Negative differential response in chemical reactions. *arXiv preprint arXiv:1812.11245*, 2018.
[39] Thomas C Schelling. Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186, 1971.
[40] Thomas C Schelling. Models of segregation. *The American Economic Review*, 59(2):488–493, 1969.
[41] Luca Dall'Asta, Claudio Castellano, and Matteo Marsili. Statistical physics of the schelling model of segregation. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(07):L07002, 2008.
[42] Laetitia Gauvin, Jean-Pierre Nadal, and Jean Vannimenus. Schelling segregation in an open city: a kinetically constrained blume-emery-griffiths spin-1 system. *Physical Review E*, 81(6):066120, 2010.
[43] Tim Rogers and Alan J McKane. Jamming and pattern formation in models of segregation. *Physical Review E*, 85(4):041136, 2012.
[44] Laetitia Gauvin, Jean Vannimenus, and J-P Nadal. Phase diagram of a schelling segregation model. *The European Physical Journal B-Condensed Matter and Complex Systems*, 70(2):293–304, 2009.
[45] Thierry Biben and Jean-Pierre Hansen. Phase separation of asymmetric binary hard-sphere fluids. *Physical review letters*, 66(17):2215, 1991.
[46] R.C. Feenstra, R.E. Lipsey, H. Deng, A.C. Ma, and H. Mo. World trade flows: 1962-2000. *NBER Working Paper Series*, 11040:1, 2005.
[47] UN. Commodity trade statistics database. *URL: https://comtrade.un.org*, 2017.
[48] UN. Provisional center product classification. *URL: http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=14*, 1991.

[49] R. Feenstra. *Advanced International Trade: Theory and Evidence*. Princeton, NJ: Princeton University Press, 2004.

[50] Jean-Philippe Bouchaud and Marc Mézard. Wealth condensation in a simple model of economy. *Physica A: Statistical Mechanics and its Applications*, 282(3-4):536–545, 2000.

[51] T. Goudreé, A. Dobrinevski, and J.-P. Bouchaud. Explore or exploit? a generic model and an exactly solvable case. *Phys. Rev. Lett.*, 112:050602, 2014.

[52] D.R. Nelson and N.M. Shnerb. Non-hermitian localization and population biology. *Phys. Rev. E*, 58:1383, 1998.

[53] J.-P. Bouchaud and M. Potters. *Theory of Financial Risks and Derivative Pricing*. Cambridge University Press, Cambridge, 2000.

[54] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509, 1999.

[55] T. Halpin-Healy and Y.-C. Zhang. Kinetic roughening phenomena, stochastic growth, directed polymers and all that. aspects of multidisciplinary statistical mechanics. *Phys. Rep.*, 254:215, 1995.

[56] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.*, 56:889, 1986.

[57] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

[58] Matteo Ciccarelli and Benoit Mojon. Global inflation. *The Review of Economics and Statistics*, 92(3):524–535, 2010.

[59] OECD. Inflation (cpi) (indicator). *doi: 10.1787/eee82e6e-en, URL : https://data.oecd.org/price/inflation-cpi.html (Accessed on 22 November 2017)*, 2017.

[60] WorldBank. Gross domestic product (gdp) (indicator). *URL : https://data.worldbank.org (Accessed on 22 November 2017)*, 2017.

[61] R. Hogg and A.T. Craig. *Introduction to Mathematical Statistics*. New York: Macmillan, 1995.

[62] Karl Pearson. Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London. A*, 185:71–110, 1894.

[63] G.M. Grossman and E. Helpman. Quality ladders in the theory of growth. *Rev. Econ. Stud.*, 58:43, 1991.

[64] P. Aghion and P.W. Howitt. *Quality ladders in the theory of growth*. MIT Press, Cambridge, MA, 1998.

[65] Ricardo Hausmann, César A Hidalgo, Sebastián Bustos, Michele Coscia, Alexander Simoes, and Muhammed A Yildirim. *The atlas of economic complexity: Mapping paths to prosperity*. Mit Press, 2014.

[66] G. Caldarelli, M. Cristelli, A. Gabrielli, L. Pietronero, A. Scala, and A. Tacchella. A network analysis of countries' export flows: Firm grounds for the building blocks of the economy. *PLos ONE*, 7(10):e47278, 2012.

[67] M. Cristelli, A. Gabrielli, A. Tacchella, G. Caldarelli, and L. Pietronero. Measuring the intangibles: a metric for the economic complexity of countries and products. *PLos ONE*, 8(8):e70726, 2013.

[68] D.R. Nelson and N.M. Shnerb. Non-hermitian localization and population biology. *Phys. Rev. E*, 58:1383, 1998.

[69] A.-L. Barabási and H.E. Stanley. *Fractal Concepts in Surface Growth*. Cambridge University Press, Cambridge, 1995.

[70] Alexis P Jacquemin and Charles H Berry. Entropy measure of diversification and corporate growth. *The journal of industrial economics*, pages 359–369, 1979.

[71] Pier Paolo Saviotti and Koen Frenken. Export variety and the economic performance of countries. *Journal of Evolutionary Economics*, 18(2):201–218, Apr 2008.

[72] Guillaume Gaulier and Soledad Zignago. Baci: International trade database at the product-level. the 1994-2007 version. Working Papers 2010-23, CEPII, 2010.

[73] WCO. Harmonic system nomenclature. *URL: http://www.wcoomd.org*, 2007.

[74] UNSD. Per capita gdp at current prices in us dollars. *URL: https://unstats.un.org/unsd/snaama/dnlList.asp*, 2015.

[75] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[76] J. Tinbergen. An analysis of world trade flows. In J. Tinbergen, editor, *Shaping the World Economy: Suggestions for an International Economic Policy*. New York: The Twentieth Century Fund, 1962.

[77] D. Glasser, D. Horn, and R. Meidan. Properties of certain zero column-sum matrices with applications to the optimization of chemical reactors. *J. Math. Anal. Appl.*, 73:315, 1980.

[78] Kiran Sharma, Balagopal Gopalakrishnan, Anindya S Chakrabarti, and Anirban Chakraborti. Financial fluctuations anchored to economic fundamentals: A mesoscopic network approach. *Scientific reports*, 7(1):8055, 2017.

[79] Kiran Sharma, Anindya S Chakrabarti, and Anirban Chakraborti. Multi-layered network structure: Relationship between financial and macroeconomic dynamics. *arXiv preprint arXiv:1805.06829*, 2018.

[80] David S Watkins. Fundamentals of matrix computations john wiley & sons. *Inc., New York, NY*, 1991.

[81] AJ Roberts. Modify the improved euler scheme to integrate stochastic differential equations. *arXiv preprint arXiv:1210.0933*, 2012.

[82] M. Cristelli, A. Tacchella, and L. Pietronero. The heterogeneous dynamic of economic complexity. *PLos ONE*, 10(2):e0117174, 2015.

[83] Greg Morrison, Sergey V. Buldyrev, Michele Imbruno, Omar Alonso Doria Arrieta, Armando Rungi, Massimo Riccaboni, and Fabio Pammolli. On economic complexity and the fitness of nations. *Scientific Reports*, 7(1), nov 2017.

[84] Lawrence Craig Evans and Ronald F Gariepy. *Measure theory and fine properties of functions*. Chapman and Hall/CRC, 2015.

[85] Meng Cai, Ying Cui, and H Eugene Stanley. Analysis and evaluation of the entropy indices of a static network structure. *Scientific reports*, 7(1):9340, 2017.

[86] Fabio Saracco, Mika J Straka, Riccardo Di Clemente, Andrea Gabrielli, Guido Caldarelli, and Tiziano Squartini. Inferring monopartite projections of bipartite networks: an entropy-based approach. *New Journal of Physics*, 19(5):053022, 2017.

[87] Mark Newman. *Networks: an introduction*. Oxford university press, 2010.

[88] Thomas Gueudré and David G. Martin. Optimal growth entails risky localization in population dynamics. *EPL (Europhysics Letters)*, 121(6):68005, 2018.

[89] G. Caldarelli. *Scale-Free Networks*. Oxford University Press, 2007.

[90] Jon M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632, 1999.

[91] Luca Dall'Asta, Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. Vulnerability of weighted networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(04):P04006, 2006.

[92] Joël Cariolle and Michaël Goujon. MEASURING MACROECONOMIC INSTABILITY: A CRITICAL SURVEY ILLUSTRATED WITH EXPORTS SERIES. *Journal of Economic Surveys*, 29(1):1–26, jul 2013.

[93] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[94] Anatoly B Kolomeisky and Michael E Fisher. Molecular motors: a theorist's perspective. *Annu. Rev. Phys. Chem.*, 58:675–695, 2007.

[95] AWC Lau, David Lacoste, and Kirone Mallick. Nonequilibrium fluctuations and mechanochemical couplings of a molecular motor. *Physical Review Letters*, 99(15):158102, 2007.

[96] Anatoly B Kolomeisky. Motor proteins and molecular motors: how to operate machines at the nanoscale. *Journal of Physics: Condensed Matter*, 25(46):463101, 2013.

[97] Hugo Touchette. The large deviation approach to statistical mechanics. *Physics Reports*, 478(1-3):1–69, 2009.

[98] Bernard Derrida. Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07023, 2007.

[99] Sylvain Prolhac. Tree structures for the current fluctuations in the exclusion process. *Journal of Physics A: Mathematical and Theoretical*, 43(10):105002, 2010.

[100] Mieke Gorissen, Alexandre Lazarescu, Kirone Mallick, and Carlo Vanderzande. Exact current statistics of the asymmetric simple exclusion process with open boundaries. *Physical Review Letters*, 109(17):170601, 2012.

[101] Jürgen Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern Physics*, 48(4):571, 1976.

[102] Andre C Barato and Udo Seifert. Thermodynamic uncertainty relation for biomolecular processes. *Physical Review Letters*, 114(15):158101, 2015.

[103] Karel Proesmans and Christian Van den Broeck. Discrete-time thermodynamic uncertainty relation. *EPL (Europhysics Letters)*, 119(2):20001, 2017.

[104] Todd R Gingrich, Jordan M Horowitz, Nikolay Perunov, and Jeremy L England. Dissipation bounds all steady-state current fluctuations. *Physical Review Letters*, 116(12):120601, 2016.

[105] Ivan Di Terlizzi and Marco Baiesi. Kinetic uncertainty relation. *Journal of Physics A: Mathematical and Theoretical*, 52(2):02LT03, 2018.

[106] Christian Maes. *Non-dissipative effects in nonequilibrium systems*. Springer, 2018.
[107] Marco Baiesi, Christian Maes, and Bram Wynants. Fluctuations and response of nonequilibrium states. *Physical Review Letters*, 103(1):010602, 2009.
[108] Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572, 2002.
[109] Duncan J Watts and Steven H Strogatz. Collective dynamics of'small-world'networks. *nature*, 393(6684):440, 1998.
[110] Luis A Nunes Amaral, Antonio Scala, Marc Barthelemy, and H Eugene Stanley. Classes of small-world networks. *Proceedings of the national academy of sciences*, 97(21):11149–11152, 2000.
[111] M Barthélemy. Crossover from spatial to scale-free networks. *Europhys. lett*, 63:915–921, 2003.
[112] Marc Barthélemy. Spatial networks. *Physics Reports*, 499(1):1–101, 2011.
[113] Alina Sirbu, Vittorio Loreto, Vito D P Servedio, and Francesca Tria. Opinion dynamics with disagreement and modulated information. *Journal of Statistical Physics*, page 1–20, 2013.
[114] Fabiana Zollo, Petra Kralj Novak, Michela Del Vicario, Alessandro Bessi, Igor Mozetič, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Emotional dynamics in the age of misinformation. *PloS one*, 10:e0138740, 2015.
[115] Michela Del Vicario, Gianna Vivaldo, Alessandro Bessi, Fabiana Zollo, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports*, 6:37825, December 2016.
[116] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591, 2009.
[117] Vishal Sood and Sidney Redner. Voter model on heterogeneous graphs. *Physical review letters*, 94(17):178701, 2005.
[118] S Suweis, E Bertuzzo, Lorenzo Mari, I Rodriguez-Iturbe, A Maritan, and A Rinaldo. On species persistence-time distributions. *Journal of theoretical biology*, 303:15–24, 2012.
[119] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002.
[120] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.
[121] Eli Ben-Naim, Laurent Frachebourg, and Paul L Krapivsky. Coarsening and persistence in the voter model. *Physical Review E*, 53(4):3078, 1996.
[122] Petter Holme and Mark EJ Newman. Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E*, 74(5):056108, 2006.
[123] Richard Durrett, James P Gleeson, Alun L Lloyd, Peter J Mucha, Feng Shi, David Sivakoff, Joshua ES Socolar, and Chris Varghese. Graph fission in an evolving voter model. *Proceedings of the National Academy of Sciences*, 2012.
[124] Byungjoon Min and Maxi San Miguel. Fragmentation transitions in a coevolving nonlinear voter model. *Scientific Reports*, 7(1):12864, 2017.
[125] Adam Douglas Henry, Paweł Prałat, and Cun-Quan Zhang. Emergence of segregation in evolving social networks. *Proceedings of the National Academy of Sciences*, 108(21):8605–8610, 2011.
[126] Paul Erdos and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.
[127] LD Landau and EM Lifshitz. Statistical physics, vol. 5. *Course of theoretical physics*, 30, 1980.
[128] Crispin Gardiner. Stochastic methods. *Springer Series in Synergetics (Springer-Verlag, Berlin, 2009)*, 1985.
[129] Michael E Fisher. The theory of equilibrium critical phenomena. *Reports on progress in physics*, 30(2):615, 1967.
[130] F.J.W. Hahne. *Critical Phenomena: Proceedings of the Summer School Held at the University of Stellenbosch, South Africa January 18–29, 1982*. Lecture Notes in Physics. Springer Berlin Heidelberg, 2006.
[131] Amaud Banos. Network effects in schelling's model of segregation: New evidence from agent-based simulation. *Environment and Planning B: Planning and Design*, 39(2):393–405, 2012.